3D Input for 3D Worlds

Sreeram Sreedharan[1], Edmund S. Zurita[2], Beryl Plimmer[3]
Department of Electrical and Computer Engineering[1&2], Department of Computer Science[3]
University of Auckland, Auckland, New Zealand
sreeram.sreedharan@gmail.com, ezur001@gmail.com, beryl@cs.auckland.ac.nz

ABSTRACT
Virtual Worlds present a 3D space to the user. However, input devices are typically 2D. This unnatural mapping reduces the engagement of the experience. We are exploring using Wii controllers to provide 3D gesture-based input to the 3D virtual world, Second Life. By evaluating its usability, we found that gesture-based interfaces are appealing and natural for hand gestures such as wave but difficult to map to facial expressions.

ACM Classification Keywords
H5.2. Information interfaces and presentation (e.g., HCI): User Interfaces.

General Terms
Design, Human Factors.

Author Keywords
3-dimensional, input device, virtual environments, Nintendo Wii controller, Second Life, social interaction, usability

1. INTRODUCTION
We are developing a 3-dimensional (3D) gesture-based interface for a 3D environment to investigate its potential to provide more a immersive experience. The Nintendo Wii controller (also known as the Wii Remote or Wiimote, see Figure 1) is the 3Dinput device we are using. It can detect motion and direction information, allowing users more natural interaction with virtual environments such as Second Life[4] (Figure 2).

Over recent years, virtual environments have grown in popularity. These 3Dvirtual worlds have developed into effective tools for social interaction, becoming more realistic not only in terms of appearance, but also in the types of inter-personal interactions available.

OzCHI 2007 Proceedings, ISBN 978-1-59593-872-5

Figure 1: The Nintendo Wii Controller

Second Life is the virtual environment we are using to investigate the gesture interface. It is an online community where users are represented with user-created personas called 'avatars'. Linden Labs' Second Life currently has around 5 million registered accounts [5]. Users interact using a combination of text messages, to act as the verbal component of conversation, and gestures performed by the avatar to act as the non-verbal component.

However, avatar gestures are currently activated by using text commands. For example, by typing '/hey' in the chat text box, the avatar performs a waving action with its hand (see Figure 2). Second life is typical of 3Dvirtual environments; the vast majority of 3Dactions are performed using 2-dimensional input devices such as the keyboard and mouse.

With the release of Nintendo's Wii gaming console, the popularity of 3Duser input has also increased by providing a gesture interface to a mainstream audience. The Nintendo Wii controller has a built-in accelerometer to measure accelerations in the X, Y, and Z axes, an infra-red sensor for accurate pointing, various buttons, and Bluetooth for communication with a Wii console or a computer[7].

However, while 3Dinput has proven to be an effective means of interaction for games on the Nintendo Wii, its efficacy is yet to be formally determined in the context of online social interaction within a virtual environment. Our focus with this work is on the user experience of using a hand-held device and gestures to interact with a 3D environment

2. RELATED WORK

While 3D gesture input for navigation of virtual environments has been explored in a number of different types of systems there are few reports of the usability of these interfaces. The usability is complex and we have drawn on reports of these experiences for our prototype.

Research into the usability of 3D interfaces for virtual reality environments has been conducted by Cabral et al.[1]. They performed experiments involving the virtual reality environments CAVE and Powerwall. From the researchers' observations of user preferences, it was found that users generally preferred a lack of physical stress over system responsiveness when manipulating in-world objects. Meanwhile, users generally desired freedom of movement for in-world navigation, placing great importance on making interaction with the virtual environment as immersive as possible.

Another study on the usability of gesture interfaces was made by Kwon and Gross [3] who combined the use of sensors on the body with motion capture cameras for a motion training application. Usability tests of their system found users often formed the opinion that a gesture-based interface made the application more appealing and encouraged the users to concentrate more on the tasks to be achieved.

There is little existing literature to suggest how one should design a 3D gesture interface. We based our approach towards designing a logical and ergonomic interface on guidelines offered by Nielsen et al.[6]. Their suggested steps are: First, determine the desired functions of the gesture interface. Second, study the actions potential users would use to perform each gesture. Third, use that information to determine the gesture vocabulary or the system. Finally, benchmark the gesture interface using usability tests.

Figure 2: Avatar performing a 'Hey' gesture in Second Life

3. GESTURE INTERFACE DESIGN

We determined the key functions required by selecting commonly used in-game gestures performed by avatars. Of these, half were hand gestures, and the other half were not.

We also included navigational functions as part of our study since we are determining the overall usability of the Wii controller in Second Life. The next step is to collect a set of intuitive gestures from users which map to these functions.

3.1. User Study

This part of our user study was designed to determine initial user perceptions and reactions to the Wii controller and gesture-based interfaces and to determine a set of default gestures [6]

We interviewed ten people of varying cultural and socio-economic backgrounds. All were proficient computer users very familiar with the keyboard and mouse. However, not all had used a Wii controller before. Each was provided a Wii controller and asked to demonstrate how they would perform the following actions using the Wii controller: Walk, run, fly, look, move the camera, select an in-world object.

This was done to determine how users would expect to navigate through a virtual environment when given the Wii controller as an interaction device. The actions which we requested users to perform using the Wii controller correspond directly to tasks which must be performed by the user when navigating through Second Life.

Then participants were shown videos of a Second Life avatar performing each of the following gestures and asked how they would perform them using the Wii controller: Hey, laugh, point me, point you, yes/agree, no/disagree.

We could then extract common gestures to help us determine the default gesture vocabulary through the observations we made during the study.

3.2. User Study Observations

We found that participants showed a preference for the 2-dimensional interfaces for navigational actions, often using buttons on the controller instead of performing a gesture with the Wii controller.

For the second half of the interview, participants tended to mimic the avatar's actions. Participants used actions which very closely resembled the avatar's actions for all hand gestures. Even for avatar gestures that tended to not use the hands, many participants attempted to mimic the avatar.

For the 'Yes' and 'No' gestures, only two participants preferred using buttons. One participant positioned the Wii controller next to his head to capture head movement while the rest used up-and-down movement of the hand for yes, and side-to-side movement for no.

The ‘Laugh’ gesture especially does not map so well to the Wii controller, however one participant still attempted to mimic the avatar.

Of the gestures which we asked participants to perform, only five were common across most participants. These five gestures are: Hey, Yes, No, Point at me, Point at you as shown in Fig. 3.

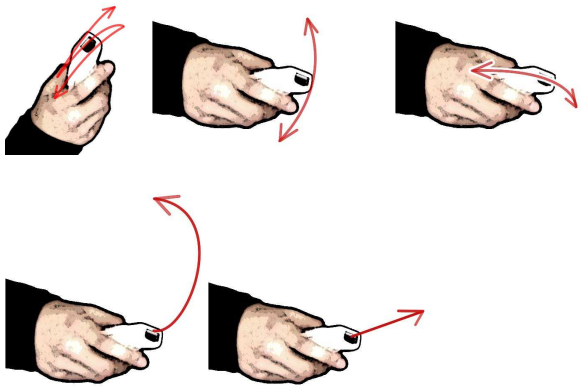

Figure 3: 5 gestures used in our study. From top-left to bottom-right: Hey, Yes, No, Point at you, Point at me.

4. Implementation

We implemented a prototype gesture interface called “Silverbolt”. It performs three main functions: record, recognize and perform.

4.1. Record

Accelerometer data from the Wii controller is transmitted to the computer via Bluetooth. This data is collected and then stored in an array-list data structure which we can then process.

4.2. Recognize

Our system implemented a gesture recognition system using a variation on Rubine’s Algorithm[8]. This has been used in other system requiring recognition capabilities (for example Inkkit [2]). It employees hidden markov models and a set of gesture features to statistically determine the most likely gesture match.

Rubine’s algorithm is trained by providing it with a set of example gestures for each class (type) of expected gesture. Our system can be trained by adding recorded gestures to a particular gesture class. A major decision when implementing this algorithm is the features to select.

The features we arrived at heuristically were:

- The duration of a gesture.
- The starting orientation of the gesture.
- The amplitudes of acceleration values in the X, Y and Z directions.
- The number of turning points in the acceleration vs. time graph of this data in the X, Y and Z directions.

Rubine’s algorithm calculates a score for each gesture class representing how well an input gesture matches a given gesture class. The maximum score represents the gesture class which the input gesture most closely matches. This application provides an interface in which users may create, train and recognize gestures.

4.3. Perform

The current implementation of “Silverbolt” performs actions within Second Life via the Windows API by first recognizing the input gesture and then converting it to the appropriate Second Life instruction.

5. USABILITY TESTING

Our second user study was designed to determine the efficiency of Silverbolt as well as to obtain users’ opinions of the system in comparison to a keyboard and mouse. The range of Second Life avatar gestures used in the study consisted of: /hey, /yes, /no, /pointyou, /pointme.

Users were required to navigate through a course we devised performing seven tasks as they went. The user’s avatar was guided through the course and given the in-world challenges by another avatar controlled by the researcher. The tasks included answering trivial questions such as “is the sky blue?” or “who is holding the Wiimote?” One question required a ‘hey’ response, three were yes/no questions and three were poinyou/pointme. Ten participants were asked to perform the test using both a keyboard and mouse interface and our Silverbolt interface. Before the Silverbolt test participants trained the recognizer with four examples of each gesture. After both tests the participants were asked to rate Silverbolt in comparision to the keyboard and mouse on: their enjoyment, ease of use and overall rating and preference on a 10 point Likert scale (1 worse, 5 same, 10 better).

The results suggest that users generally preferred the Silverbolt interface over keyboard and mouse. Participants found the gesture interface more enjoyable to use, giving the system an average rating of 6.9. We also found that participants perceived the system a little easier to use than the standard keyboard and mouse interface on average an average rating of 5.75. Finally, participants gave the gesture recognition system a higher overall rating of 6.35. Furthermore, 70.0% of participants would rather use the Wii controller interface over the keyboard interface.

As far as the recognition rates for our gesture recognition system are concerned, 92.9% of all questions were answered correctly using Silverbolt, however only 72.9% of all questions were answered correctly on the first attempt.

This compares to 90.0% of questions answered on the first attempt using the keyboard and mouse interface. Most keyboard mistakes made were typing errors often caused by unexpected auto-completion within the Second Life chat-box.

Problems with recognition affected one participant's opinions negatively, rating the system 1/10 in all aspects. It was observed that the participant's Wii controller gestures became more and more erratic from subsequent recognition failures. These erratic movements were noticeably different to the gestures the participant performed during the training stage.

Interestingly, another participant faced similar difficulties. However, even with the participant's limited success in completing different tasks, the participant still rated the system 7/10 in terms of enjoyment, 8/10 in terms of ease of use and would use the Wii controller interface over the keyboard interface. The participant made the comment that the gestures used in the gesture-based interface were a lot easier to remember than the text commands provided by Second Life.

Some participants made the comment that by using our gesture interface more they would be able to perform better in our testing course. The effect of user experience and training on the usability of our system would be interesting to investigate in a longer term study.

6. CONCLUSIONS

We found that users are currently unfamiliar with 3D gesture interfaces, often devising ergonomically inappropriate gestures or resorting to 2-dimensional aspects of the Wii controller such as buttons. It is possible to set an intuitive gesture vocabulary for hand gestures such as wave using user feedback. The wide variety of responses for facial gestures (look etc) suggests that these do not map naturally to a hand gesture. For gross body gestures (run etc) users preferred the Wii buttons and joystick, we did not investigate whether this is a result of previous experience or whether buttons and joysticks are inherently more suited to this type of action.

We also found that the Silverbolt 3D gesture interface provides users with a more enjoyable and intuitive user interface for Second Life. However, our recognition accuracy needs to be increased to improve the experience of users of this system and better facilitate further study.

7. FUTURE WORK

Implementation of more sophisticated gesture recognition algorithms would improve the user experience. Studies can be done using multiple algorithms and the results compared to remove any skewing of results caused by algorithm choice.

Currently our work dealt mainly with the social interaction aspects of user interaction in Second Life. A logical next step would be to determine the usability of this 3D gesture based system for other tasks which need to be performed such as interaction with in-world objects.

Larger scale and longitudinal studies are needed to fully determine the usability of a Wiimote-style interface for virtual environments and negate the novelty effect.

7. ACKNOWLEDGMENTS

We would like to thank John Corey from the University of Auckland Centre for Software Innovation for providing us with the idea of using Second Life and the Wiili.org community for providing useful information and resources which helped us during the implementation of our project.

REFERENCES

[1] Cabral, M. C., Morimoto, C. H., & Zuffo, M. K. (2005). *On the usability of gesture interfaces in virtual reality environments*. Paper presented at the Proceedings of the 2005 Latin American conference on Human-computer interaction.

[2] Chung, R., Mirica, P., & Plimmer, B. (2005). *InkKit: a generic design tool for the tablet PC*. Paper presented at the Proceedings of the 6th ACM SIGCHI New Zealand chapter's international conference on Computer-human interaction: making CHI natural.

[3] Kwon, D. Y., & Gross, M. (2005). *Combining body sensors and visual sensors for motion training*. Paper presented at the Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology.

[4] Linden Research Inc. (2007). Second Life.

[5] Linden Research Inc. Second Life: Your World. Your Imagination. Retrieved 1 April, 2007, from http://www.secondlife.com

[6] Nielsen M., S. M., Moeslund T.B., Granum. E. (2004). A Procedure for Developing Intuitive and Ergonomic Gesture Interfaces for HCI. In *Gesture-Based Communication in Human-Computer Interaction* (Vol. 2915/2004, pp. 409-420): Springer Berlin / Heidelberg.

[7] Nintendo. (2007). In Depth Regional Wii Coverage. Retrieved 23 March, 2007, from http://wii.nintendo.com/controller.jsp

[8] Rubine, D. (1991). *Specifying gestures by example*. Paper presented at the Proceedings of the 18th annual conference on Computer graphics and interactive techniques