Navigating the Virtual Environment Using Microsoft Kinect

Xiaochen Yang  Lichuan Pan

Honor Code
We, Xiaochen Yang and Lichuan Pan, pledge our honor that we have neither given nor received aid on this work.

Abstract—In this paper, we present a simple way “Walking in place” to navigate a virtual environment using Microsoft Kinect. Our system can be re-implemented in relatively low-cost. In order to evaluate this navigation method, we also use Xbox game controller as another navigation method. A fun finding red balls experiment is proposed to evaluate these two methods. Ten subjects participated this experiment with different background and pre-experience. The result contradicts our original hypothesis that Kinect perform better than Xbox game controller after statistical analysis. The possible reasons are discussed in the paper.

Keywords—Virtual environment; Kinect; head-mounted display; subjects; evaluate

1. INTRODUCTION
Virtual Environment (VE) is using computer-generated environment to represent real world scenes or objects, or imagination. It usually has an interactive system which will give users an illusion of displacement to another location. With the rapid development in entertainment, Virtual Environment has a great impact on people’s daily life. With the great success of the film Avatar, other science and motion pictures, and many other similar products in recent years, people are more shocked by this technique and willing to know more. Also, VE is widely used in heritage and archaeology, Virtual Reality (VR) reconstruction, radio, fine art, games, music, therapeutic applications and training, etc. It is implemented in manufacturing and urban design.

VE navigation is important in electronic games as well as in combat training and real flight simulation in the Air Force. Finding methods for spatial virtual environment navigation is a challenging and important problem that has received significant attention. Methods that explored in previous works include joystick navigation, walking-in-place (virtual walking) and push-button-fly [5], WIP–Wii [1], Walk in Virtual Reality profile [2] and WIP [4].

Many researches are interested in navigating VE using Kinect because of its real time skeletal tracking. Such kind of works includes navigation, but its VE is not that large [4] and detail tracking such as hand tracking [3].

This paper works on one aspect of VE navigation, targeting at electronic game area. Actually, VE navigation is not only important in electronic games, but also very helpful in other areas, such as soldier training and aviation simulation. We use Microsoft Kinect to recognize human gestures in order to represent the real actions in the VE navigation. In our design, there are two navigation modes: Walk and Fly, at changeable speeds. With the fast-speed Fly mode, the user can navigate in a large-scale outdoors environment and rapidly reach a specific 3D area, while with the slow Walk and Fly speed, the user can enjoy the beautiful scenes of the VE. This system basically can be low-cost re-implemented if the HMD’s price can further be reduced (depends on the brand of HMD). After the usability test, users generally think this VE navigation system is fun and on the other hand, something can be applied to improve the user experience.

However, at the same time we found the motion sickness due to VE navigation wearing a HMD and the relatively high navigation speed. So we guess the next-generation HMD should be developed further to adapt fast advancing speed of VE development (fortunately already there are researchers working on this challenges). Because nowadays more and more people are talking about true space travelling, not just astronauts, such kind of simulation is necessary and highly beneficial to people.

The rest of the paper is organized as follows: Section 2 describes the system description, which covers the system work flow. Section 3 presents apparatus / methods used. Section 4 shows results of the experiments and statistical analysis. Section 5 gives our discussion of this system and Section 6 presents the conclusions.
2. SYSTEM DESCRIPTION

The work flow of our system is showed in Figure 1: Microsoft Kinect captures the single-user’s gestures, and then sends the gesture data to FAAST in order to get recognized. After the process of gesture data, FAAST sends out control data to Vizard to control the VE navigation. The virtual scene is showed both in two monitors and in HMD.

Fig. 1. Single-user VE navigation system using Microsoft Kinect

2.1 Hardware

Our system consists of several simple, common devices (see Figure 2), one head-mounted display (HMD), one Microsoft XBOX 360 game controller, one Microsoft Kinect, and a PC desktop with two LCD monitors. These are all the devices we need which will not cost too much to get all of them. Our system is easy to set up with low cost.

Fig. 2. The overview of the system

A head-mounted display (HMD) [7] is used to show the virtual environment in stereopsis. In Figure 2, the two LCD monitors provides two offset image separately to the left and right view of the HMD. In one of our experiments, the user needs to wear the HMD, and what they see is exactly the same as it shows on two LCD monitors. The difference is, wearing HMD makes user feel more stereoscopic than seeing the monitors.

The XBOX game controller is one way we use to navigate the virtual environment. When subjects use this game controller to navigate the VE, they do not wear HMD. The lacks of stereopsis may influence their performance.

Microsoft Kinect [8] is novel game controller presented by Microsoft in November 2010. It has two 3D depth sensors and one RGB camera. We use its skeleton tracking feature to capture user’s gesture, without the need to use a game controller. Standing in front of the Kinect, user can perform Walk in place (WIP), which moves the main-view in the virtual environment, but not in the real world.

2.2 Software

In our system, the software plays an important role. Mainly we use two kinds of software to realize our work. FAAST (Flexible Action and Articulated Skeleton Toolkit) [9], developed by University of Southern California, is used to recognize user’s body posture and specific gestures and emulate keyboard input triggered for our program to navigate the virtual environment.

Vizard [10] is a Python-based integrated development environment, which provides a lot of features that facilitate the development of virtual application. Our main program is written in Python using WordViz as the IDE.

2.3 Virtual Scene Navigation

The virtual scene we use to navigate is a modern city with several skyscrapers in the middle. The real city texture images have been applied on the 3D models, making the whole city look more realistic. In this large virtual environment, we basically applied two ways to navigate: the Walking mode and Flying mode.

Walking Mode

In the walking mode, user is required to stand in front of the Kinect. User is moving forward in the virtual city by applying walk in place in the real world. Turning is captured via Kinect by lifting their arms up. Lifting left arm makes left turning, and similarly, right turning is applied by right arm lifting. Acceleration in moving forward is performed by lean user’s body a bit forward. Also, lifting left hand over head brings user back to the original place in the virtual environment.

Flying Mode

Since this is a pretty big virtual scene, it is more wonderful to fly among those high buildings instead of
just walking through them. In order to turn on the flying mode, user should perform a superman posture, that is, user should lift his right arm straight up over the head, and put his left hand on the waist. When user starts flying, he can put his left hand down, and only use right arm to control the direction: lift up and down, move left and right. Similar as walking mode, flying speed can be accelerated by leaning the body forward. Landing is also realized in our program. Whenever user wants to land, he just needs to straight up both arms over his hand, then the walking mode is re-entered until the superman pose is applied.

3. Methods

3.1 Participants
Ten subjects between the ages of 20 and 30 participated in the experiment. Among them, six are males and four are females. They are all graduate students from Vanderbilt University but with different majors. The female subjects do not have much pre-experience of navigating a virtual environment compared to the male subjects.

3.2 Experimental Design
Each of the 10 participates explored the virtual environment under two different ways, XBOX game controller plus HMD navigation and body gesture navigation using Kinect plus HMD. Ten subjects are divided into two groups, and each group has three males and two females.

The task is to explore the virtual city environment by finding five red balls among forty randomly generated balls. Figure 3 shows the bird-view of this whole virtual city environment. The red area is the place where forty balls are randomly generated. From these forty balls, randomly pick five balls and change their color to red. The task of this experiment is to find all the red balls using as less time as they can. Subject is placed at the center of the red area, initially. In order to find all the targeted balls, they probably should traverse all the roads in this environment. The red balls are pretty obvious among ordinary balls, which can be seen in a distance place. When subject come near to the red ball, the ball will turn blue and translucent gradually (see Figure 4). At the same time, the number of red balls that have been found shows in the left-up corner of the screen.

3.3 Apparatus
The order of test method is different; one group takes the game controller navigation with HMD experiment first, and then finishes the Kinect plus HMD exploration, while the other group applies the opposite order.

4. Results
The subjects’ mean tct (task complete time) of navigating the virtual environment via two different ways is showed in Figure 6. A repeated measure ANOVA is performed on the mean tct under the two
conditions, Xbox game controller with HMD and Kinect with HMD with a significant difference between them (p < 0.05). The Xbox game controller with HMD condition is the faster one 96s, and the slower one is Kinect with HMD condition 123s.

Fig. 5. Task complete time of two methods.

We perform a two-way ANOVA for comparing the means of the task complete time of two methods using MATLAB statistic toolbox and calculate the p value (p = [0.0269 0.0035 0.9363]). The vector p shows the p-values for the two methods, 0.0269, the ten individual subjects, 0.0035, and the interaction between method and individual. These values show that both method and individual affect the task complete time, but there is no evidence of a synergistic effect of the two.

We are also interested in the learning effects of the two trials. In order to compare two trials, we average the task complete time of two methods for each trial. The overall performance of two trials is plotted in Figure 7, and mean time and standard error of two trials is showed in Figure 8. The average task complete time of first trial is about 121s and the second trial is only 98s.

Fig. 6. 1. Xbox game controller + HMD 2. Kinect + HMD. Red line is the median tct value (second). Edges of the box indicate 25-th and 75-th percentiles.

Fig. 7. Task complete time of two trials.

Fig. 8. 1. Trial 1. 2. Trial 2. Red line is the median tct value (second). Edges of the box indicate 25-th and 75-th percentiles.

5. DISCUSSION

Our original hypothesis is that using Kinect to navigate the virtual environment should perform better than Xbox game controller. However, the experiment result contradicts our hypothesis. From the performance of two methods, we found that subjects do better using Xbox game controller. Also, in our questionnaire after experiment, nine of ten people prefer Xbox game controller rather than Kinect to navigate the virtual environment, and think it is more enjoyable. We want to find the reason why the performance is so different. Again, in the questionnaire, we may grasp some aspects of the reason. One of our subjects wrote that “HMD is not suitable for the one with glasses, and I struggled to see clearly and feel dizzy.” The background survey shows that almost all of them do not have the previous experience to wear a HMD, and some of them really feel dizzy the first time they wear it. However, there are still some people who feel good with HMD. Why they still take longer time with Kinect navigation than using an
Xbox game controller? One possible factor is that when they wear the HMD and walk in the place, sometimes they turn their body a little bit and then they are not facing Kinect directly. This may affect the sensitivity of capturing skeleton gesture by Kinect. Another possible reason is that they feel more comfortable using Xbox game controller since almost everyone used it before (or something similar to it). Even though they may still feel a little dizzy, they are so familiar with the game controller, so they complete the task fast. The last possible reason is that using body gesture to navigate a virtual environment is not so controllable. Every subject is different with each other in height, weight, the length of arm, etc. The action they perform is also different with each other. Sometimes Kinect cannot recognize a certain gesture when a subject does not perform properly. There is also a factor that affects the test is that random. Someone may be very lucky that he choose a way that five red balls happened to be on the way. Then he can finish the task without traversing all the roads. However, others may really need to go through all the roads in order to find all the red balls. Since all the red balls are randomly generated on the roads, it may have some cases that one trial was completed in relatively short time.

We also found that subjects perform better in the second trial than the first trial. Nine of ten did better in their second trial. The possible reason is that subjects are more familiar with the system when they did the test again. Another interesting result is that male subjects perform better than female subjects. Three of four female subjects do not like any video or PC game, and all of them do not have much experience in playing 3D games. However, almost all the boys played 3D games before, some of them are quite familiar with them, such as Counter-Strike, Warcraft, etc. Only a few minutes’ instructions are needed to train a male subject. It takes longer time to explain to female subjects and practice the system before the experiment.

During the experiment, we found that some subjects prefer to use Fly mode. Their task complete time is shorter than others since Fly mode can provide a fast moving forward speed and a better view of the whole scene, which makes it easier to find all the red balls. An effective strategy can also be very helpful. For example, when one subject arrive the cross, he usually look around to check if there is any red ball in other directions before turning.

6. CONCLUSIONS

In this paper we presented an interactive system using Microsoft Kinect to navigate the virtual environment. We design an experiment to evaluate the performance of Kinect compared with Xbox game controller. The result shows that subjects performed better using game controller instead of Kinect, and we discussed the possible reasons of this.

Our future work focuses on improving the virtual scene, and adjusts Kinect to make it more controllable.

REFERENCES


