Using Whole-Body Orientation for Virtual Reality Interaction

Vitor A.M. Jorge, Juan M.T. Ibiapina, Luis F.M.S. Silva, Anderson Maciel, Luciana P. Nedel

Instituto de Informática
Universidade Federal do Rio Grande do Sul (UFRGS)

{vamjorge, jmtibiapina, luisfmssilva, amaciel, nedel}@inf.ufrgs.br

Abstract

When interacting in the real world, it is common to remember locations of objects based on our own body previous locations and postures. In this paper we discuss the benefits of the whole-body awareness in 3D interactive applications. We propose a technique for navigation and selection in 3D environments which uses an optical tracking system to provide whole-body position and orientation. We explore the peephole metaphor with a tablet PC to artificially enlarge the display and interaction area. In the application implemented to prove the proposed concepts, the tablet is held by the participant who moves it around and points it in any direction for visualization and interaction.

1. Introduction

When exploring a 3D virtual environment with a mouse and keyboard, users easily become disoriented. Jacob et al. [4] remind us that we are not guided purely by visual cues when moving in the real world. We are also guided by some understanding of the surrounding environment, of our body and the presence of other people, as well as by some sense about physics. Thinking of navigation in real world is relatively easy as our complex biological system provides important information to aid in self orientation. The human vision is capable of providing stereo perspective views of the world, giving a notion of position and distance from visible objects. The labyrinth provides the information about up, down and balance. Our sense of touch makes us aware of obstacles when we are in direct contact, even when we cannot see them. To finish, we dispose of a sense of position and orientation which makes us know, all the time, where our limbs and other body part are, which is called proprioception [2].

However, when the focus passes from the real world into a virtual world, and one starts to interact with a virtual environment using conventional interfaces, all corporal cues vanish. This often reduces body sensation and causes disorientation. Despite the fact that there is a 3D view, normally there is no information to guide us other than what we see on the display. In the best cases a stereo view is available, but generally only two-dimensional mini-maps of the environment and the keys we press to change what we see are provided. To illustrate that, think of playing a first person shooting game in a regular personal computer. There is a strong dissociation between vision and movement, i.e., the virtual eye/camera position and orientation in space is controlled by key pressing, while in fact we are still sitting there in front of the screen.

In the present work we explore human orientation capabilities without relying only on the sense of vision. We propose to do so using the history one has about their body postures while moving in the real world. This is done by implementing the peephole metaphor using a tablet PC as a window to the virtual world, artificially enlarging the display and interaction area. The tablet is held by the participant, who moves it around and point it in any direction

Figure 1. The pictures illustrate the use of whole body awareness as a tool to aid the users to keep themselves oriented during the exploration of a virtual world

for visualization and interaction, as shown in Figure 1. Depending on user orientation and position, the display shows different portions of the environment and the system allows interaction with objects in that specific portion of the space.

2. Related work

When researchers first introduced techniques for navigation in virtual reality (VR) they naturally proposed the creation, in the virtual environment, of virtual navigation tools analog to real ones. This includes the use of maps, compasses, etc. Darken and Sibert [3] created a toolset based on this principle and studied how their tools influence the navigation behavior of some subjects. One of their conclusions is that people need cues, e.g. visual and audible information, which can be combined to make targets easier to find.

Experimental results reported by Wartenberg et al. [12] indicate that whole-body movement information - vestibular and proprioceptive signals, motor efferent commands - are decisive in grabbing correct directional information while navigating through the real space. They showed that all senses contribute to spatial orientation, and proposed a simple task in order to analyze and separate elementary spatial information about distances and directions.

Van Rhijn and Mulder [11] show that the movement performed by users in the virtual environment should correspond to the actual movement in the real world, in order to be intuitive or natural. Mine et al. [7] present a technique to instantly bring objects in reach so that users can manipulate them using proprioceptive cues. It shows significant improvements in positional accuracy to objects held in one's hand in relation to those fixed in space. These works are reinforced by Rohs et al. [9] through 2D map exploration using arm/hand movement and target localization, hence requiring spatial memorization. In the work of Bakker et al. [1], the participant was immersed in a virtual forest and asked to turn specific angles using different interaction techniques. Best results are reached when the kinesthetic feedback is present, that is, when the participant moved his legs to turn around.

While previous research works explore body awareness in specific situations, either limited to two dimensions or to a reduced number of DOF, in this work we propose a general interaction case in 3D with 6 DOFs.

3. Design and Implementation

Peephole interaction has been implemented with somewhat different strategies by several authors [13, 6, 10]. A peephole generally occurs when a spatially aware display is moved and reveals different parts of a virtual environment – as a virtual window – showing parts of objects or areas too large to be seen at once.

In this work we propose a peephole implementation where a tablet PC is used as a window to a virtual world. Through the tablet display, it is possible to inspect and interact within a virtual world. When the user walks or turns around, the position and orientation of his/her head are tracked using an optical system. As a consequence, images exhibited in the tablet display change accordingly. Figure 2 illustrates the system setup.

In order to obtain spatial awareness, we are using BraTrack [8], a precise and low-cost marker-based commercial optical tracking system. The system is composed by a set of off-the-shelf USB cameras (two in this setup) customized with electronic boards that provide flash strobes using a huge number of infrared LEDs. Trackable artifacts are built using sets of reflective spheres. Images are acquired in a synchronized way by the camera modules at

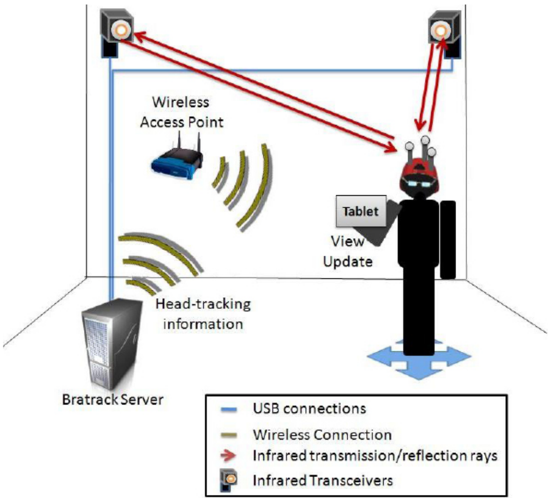

Figure 2. Overview of the system setup

60 Hz. A 2D pre-processing identifies infrared artifacts in each camera, and a 3D reconstruction module finds the 3D position and orientation of each artifact that is visible in more than one camera. In case there is occlusion, the system cannot provide accurate tracking.

Such occlusion limitation of the device is the main reason why we decided to attach the artifact to the head instead of the tablet PC, as many projects in handheld augmented reality do. While it would be great to track both the head and the tablet, the same occlusion problem would arise. Therefore, we track only the head of the user and assume that the distance between the head and the tablet PC is fixed. In a future version of this system, the tablet PC should also be tracked.

Aiming at tracking the head of the user – and capture the position and orientation of the virtual camera – an artifact was attached on a hat, as shown in Figure 1. In such a way, the head movement controls the virtual camera position and orientation in space. The motion tracked by BraTrack is sent to the tablet PC by a wireless connection between BraTrack server and the tablet (see Figure 2).

Regarding the display, we have chosen to use a tablet PC, because changing it into a peephole device is straightforward. To do that, we adopted the fixed-cursor design. With this design, the person sees the virtual environment through the tablet display with a fixed crosshair positioned at the center of the screen. Thus, moving the device means aiming, and objects placed under the crosshair can be selected with a click. Since both hands were used to hold the device – still not as light as wished, we needed some strategy to allow the user to point and click objects without the stylus, keyboard or mouse. To do so,

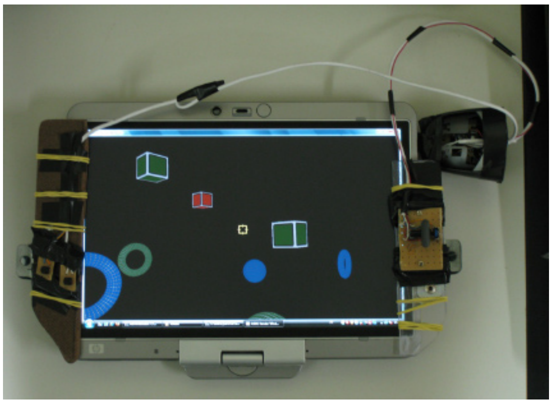

Figure 3. Detailed view of the tablet PC is presented with mouse buttons attached at the left and the wheel attached at the right side.

we attached two mouse buttons on the left of the tablet and the mouse wheel on the right to chose among menu options (see Figure 3).

4. Other Tracking Strategies Tested

Before the choice for the BraTrack system, we also tested the use of other two solutions for 3D tracking: the Augmented Reality Tool Kit (ARToolkit) [5], a very popular low-cost academic optical tracker that works with any type of webcam and flat binary markers detected by computer vision; and the Nintendo Wii® remote controller, which allows for immersive interaction metaphors in home environments, with a software API[1].

The main reason why we did not keep ARToolkit is that it is extremely sensible to brightness changes. Another problem with the toolkit was the noise. The matrix generated by the system and passed to our application at each frame had variations that sometimes led to severe discontinuities (shaking or flicking) in movement. In order to reduce that, we had to interpolate matrices between frames to generate a more stable movement, but users reported response delays. Besides, even in a light controlled environment, when the marker could effectively be detected, regular web-cams show a limited field of view. Due to these issues, the use of Augmented Reality Tool Kit was discarded.

Then we tested the Wiimote. Despite the fact that the system responded very precisely, the implementation efforts required were

[1] http://www.cs.cmu.edu/ johnny/projects/wii/

considerably higher. Additionally, the problem of having a very limited field of view persisted; occlusions were too much frequent. Due to the time required to customize the interface, and the occlusion problem, this implementation strategy was also abandoned. However, using more than one Wii controller to enlarge the tracked area could be a good implementation choice for someone that does not have a more suitable tracking system available.

Finally, the BraTrack system was tested and chosen for several reasons. It is faster. While the ARToolkit works at a frame rate of 30fps, BraTrack works at 60fps, which minimizes latency problems. Also, BraTrack uses two cameras for tracking, instead of one. Then, the range of movement extends to a greater area than the ones provided by the two previous systems. This does not invalidate the previous choices, especially if one wishes to use them with more than one camera/device at once, but that would require further customization effort, and tracking itself is not the main focus of this work.

5. Results

With the purpose of observing the possible advantages of an interaction technique coherent with the real space, and to test the use of body mnemonics and intuition in a navigation task, a simple pair-matching memory game application was implemented. In the game, the participant has to navigate through the virtual environment and correctly match four pairs of objects, just as in a classic memory game. Initially, objects are shown as question marks. When the participant selects a question mark, the real shape of the hidden object is shown. Once two objects are selected, if they have the same shape and color, they disappear after 2 seconds and a green check mark is shown on the bottom-right of the screen. If they are different, they turn back into question marks after 3 seconds becoming hidden again and a red "X" is shown, indicating the error. This process is repeated until there are no objects left and the game ends.

We observed the time and the number of clicks needed to complete the game; the path followed, and the subjective impressions of some users about the experience of using the peephole display compared with the mouse and keyboard.

We analyzed the individual performance of each subject in both tasks. We noticed that people with game oriented behavior are extremely skilled with mouse and keyboard, reaching a much higher performance in this modality comparing to their non-game-oriented fellows. At the same time, we observed that non-game-oriented people always had their best performance with the peephole. More than that, their delta between mouse and peephole is often higher with the peephole than the delta of the gamers is with the mouse and keyboard. This suggests that the subjects profile plays an important role when it comes to choosing the best interaction technique.

6. Conclusion and Future Work

In order to verify the advantages of the whole-body awareness in 3D interactive environments, we proposed a technique for navigation and selection in virtual environments using the position and orientation of the user´s body. Results obtained are not yet statistically significant, but preliminary tests presented very promising results, mainly with users that are not familiar with 3D interaction.

Concerning our future work, we plan to perform a formal evaluation of our system with subjects with different profiles, including people without previous experience on heavy computer interaction. This should provide a reasonable comparison, avoiding experienced users addicted to mouse use. Another important open question is measure. What is the level of body awareness considered to be good? How about environment awareness? How is it measured with standard interfaces? Such questions are valid and need to be further investigated.

7. Acknowledgments

We would like to thank Hewlett-Packard for the tablet-PC, MAMS Tecnologia for the BRATRACK system, and the volunteers who tested the prototype. This work was partially supported by CNPq-Brazil.

8. References

[1] N. Bakker, P. Werkhoven, and P. Passenier. Aiding orientation performance in virtual environments with proprioceptive feedback. *Virtual Reality Annual International Symposium*, 1998. Proceedings., IEEE 1998, pages 28–33, 18-18 1998.

[2] K. R. Boff, L. Kaufman, and J. P. Thomas, editors. *Handbook of Perception and Human Performance:*

Sensory Processes and Perception, volume 1. Wiley-Interscience, April 1986.

[3] R. P. Darken and J. L. Sibert. A toolset for navigation in virtual environments. In *UIST '93: Proceedings of the 6th annual ACM symposium on User interface software and technology*, pages 157–165, New York, NY, USA, 1993. ACM.

[4] R. J. Jacob, A. Girouard, L. M. Hirshfield, M. S. Horn, O. Shaer, E. T. Solovey, and J. Zigelbaum. Reality-based interaction: a framework for post-wimp interfaces. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 201– 210, New York, NY, USA, 2008. ACM.

[5] H. Kato and M. Billinghurst. Marker tracking and hmd calibration for a video-based augmented reality conferencing system. *Augmented Reality, 1999. (IWAR '99) Proceedings. 2nd IEEE and ACM International Workshop on*, pages 85– 94, 1999.

[6] S. Mehra, P. Werkhoven, and M. Worring. Navigating on handheld displays: Dynamic versus static peephole navigation. *ACM Trans. Comput.-Hum. Interact.*, 13(4):448–457, 2006.

[7] M. R. Mine, J. Frederick P. Brooks, and C. H. Sequin. Moving objects in space: exploiting proprioception in virtualenvironment interaction. In *SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 19–26, New York, NY, USA, 1997. ACM Press/Addison-Wesley Publishing Co.

[8] F. Pinto, A. Buaes, D. Francio, A. Binotto, and P. Santos. Bratrack: a low-cost marker-based optical stereo tracking system. In *SIGGRAPH '08: ACM SIGGRAPH 2008 posters*, pages 1–1, New York, NY, USA, 2008. ACM.

[9] M. Rohs, J. Sch¨oning, M. Raubal, G. Essl, and A. Kr¨uger. Map navigation with mobile devices: virtual versus physical movement with and without visual context. In *ICMI '07: Proceedings of the 9th international conference on Multimodal interfaces*, pages 146–153, New York, NY, USA, 2007. ACM.

[10] Z. Szalav´ari and M. Gervautz. Using the personal interaction panel for 3d interaction. In *Proceedings of the Conference on Latest Results in Information Technology*, pages 3–6, 1997.

[11] A. van Rhijn and J. D. Mulder. Spatial input device structure and bimanual object manipulation in virtual environments. In *VRST '06: Proceedings of the ACM symposium on Virtual reality software and technology*, pages 51–60, New York, NY, USA, 2006. ACM.

[12] F. Wartenberg, M. May, and P. Péruch. *Spatial Cognition*, volume 1404/1998 of *Lecture Notes in Computer Science*. Springer, Berlin / Heidelberg, January 1998.

[13] K.-P. Yee. Peephole displays: pen interaction on spatially aware handheld computers. In *CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 1–8, New York, NY, USA, 2003. ACM.