Comparing low cost input devices for interacting with 3D Virtual Environments

C. Ardito, P. Buono, M. F. Costabile, R. Lanzilotti and A. L. Simeone
Dipartimento di Informatica, Università di Bari, Italy
{ardito, buono, costabile, lanzilotti, simeone}@di.uniba.it

Abstract — **Interaction with 3D Virtual Environments has always suffered from a lack of widely available and low cost input devices. Recently, thanks to the diffusion of gaming systems such as the Microsoft XBox 360 or the Nintendo Wii, new input devices are on the market at a relatively cheap price. This paper describes a study whose aim is to compare input devices in order to identify effective alternatives for the mouse and keyboard in such settings where their use is not advisable or feasible, e.g. museums and other public areas. This study has been carried out using a 3D Virtual Environment in which the participants were required to perform three canonical 3D interaction tasks. Two different groups participated to the test: the first group was involved in a pilot study to check the test environment. The second group performed the test.**

Keywords — **3D virtual environments, human-computer interaction, input devices, user studies.**

I. INTRODUCTION

Interacting in a Virtual Environment (VE) often requires devices specifically built for that purpose, such as gloves, wands and other 6 DOFs devices. This equipment is generally expensive and, in most cases, it is exclusively confined to university laboratories and industry professionals. The average consumer has been confined, until very recently, to the classic mouse and keyboard, with occasional special purpose devices (joysticks, driving wheels, etc). There is undoubtedly a vast gap between input devices used by researchers and those used by consumers, which is hindering the spread of 3D User Interfaces as valid or even better alternatives of traditional interfaces. In fact, there is no universally accepted Virtual Reality framework, with common reference devices adopted because of their overall performance and efficiency. Devices are usually chosen due to their cost or attractiveness, rather than their usability and comfort [1].

There are certain scenarios in which mouse and keyboard are unpractical. Indeed, in order to be comfortably used, they do require a surface where they can be put on and a seat for the user. Public areas like museums or airport/train terminals cannot employ this kind of set-up for a variety of reasons, especially for space requirements. Other devices must be employed if we want users to interact more actively with 3D environments in these settings. Also, new interaction metaphors could be devised for these devices. Research on these topics will provide useful hints to the designers of tomorrow's input devices.

Thanks to the investments of big companies such as Microsoft, Sony and Nintendo, many different input devices are now available to a vast majority of people. We analyzed them and, by comparing their cost and the easiness to set them up within a desktop PC environment, we chose the Microsoft Xbox 360 gamepad and the Nintendo Wii Remote. In this paper we describe the experimental study we designed to compare these devices to the mouse and keyboard to see how well they performed for interacting in a 3D VE.

Two different groups participated to the test. Afterwards, they were interviewed about their experience. They had to complete standard 3D interaction tasks such as translating and rotating an object. The first group was involved in a pilot study to check the test environment. The second group performed the actual test.

This paper is organized as follows: in the next section the related work is briefly discussed with reference to similar studies to ours. In Section 3, a brief overview of the two chosen devices is given. Section 4 presents the test environment and Section 5 reports the experimental study. Section 6 concludes the paper.

II. RELATED WORK

Testing and developing interaction techniques has always been a mainstay of the research within the field of 3D User Interfaces. Studies such as [2],[3] helped define and identify the most basic interaction techniques for manipulating objects in the environment: ray-casting, 3D cursors, fixed or extensible virtual hand representations [4], etc. These techniques were also compared and evaluated in [5]. Bowman et al. propose a formal framework for testing and comparing different approaches of various interaction techniques in VEs [6]. A notable study by Hinckley et al tests different rotation techniques and compared the results obtained with different devices [7]. In [8], a study evaluated how various tasks are performed with 2D, 3D and hybrid interfaces, which employs 2D interaction techniques for tasks that are typically 2D in nature, such as writing or selecting from a menu, and 3D interaction techniques for typical 3D tasks such as selection and manipulation. The authors showed that the best approach was obtained using interfaces that

---

Partial support for this research was provided by Italian MIUR (grant "CHAT") and by EU and Regione Puglia under grant DIPIS.

properly match in dimension the task objective. With the same spirit, we wanted to evaluate the possibility of successfully using low cost devices for interaction in 3D environments.

The Wii remote (or Wiimote) has been recently used to control the visualization of medical data [9]; the authors use the Wiimote's motion sensing capability in two different modes: pointing and manipulation. In the first mode, the Wiimote is used to control a virtual pointer, whereas in the other mode it is used to move or rotate the 3D volume.

The Wiimote, in recent times, has gained a lot of interest from enthusiasts as well as researchers: the work Johnny Lee did with a Wiimote is a popular example [10]. In his work Lee shows how a Wiimote can be reverse engineered to be used as a tool for finger tracking or as a drawing pen. By placing the infrared sensors on glasses, it can be used to render view-angle dependant scenes so that they may be used to simulate the changing parallax and the field of view.

In [11] an Augmented Reality mobile phone is used to manipulate virtual objects. Isokoski and Martin [12] performed an experiment similar to ours in nature by comparing the use in First Person Shooter games of different input devices for the task of aiming, namely: a wheel mouse, a track mouse, an XBox controller, and mouse and keyboard. The results of their study are similar to ours, since they show that, among the considered devices, the mouse is still the most efficient for that task, but further work is required to assess the effect of training on the user performance with such input devices.

Fig. 1. On the left, the Microsoft XBox 360 Gamepad. On the right, the Nintendo Wii Remote, front and back. Labels a, b, c, d, e and f, are superimposed to the picture for explanation purposes.

III. THE DEVICES

To choose which consumer devices to consider for our testing purposes, we selected them by evaluating how easy to obtain they were, how expensive and finally, how easy to set them up with a desktop PC they were. Among gamepads especially built for the PC, for example, there are several models available, each of them having a particular button layout. The one we chose is the Microsoft XBox 360 gamepad, because of its wide availability, great ergonomic comfort and the easiness of plugging it to a desktop PC. The other controller examined, the Nintendo Wii Remote, was obviously not specifically developed for use with a PC and requires a bit more effort to set it up. In the next section, the details of the input mappings are explained.

A. *The Microsoft X-Box 360 Controller*

The XBox 360 controller is the standard gamepad supplied with the Microsoft gaming console. A special PC version exists but, for the purposes of our tests, we used a XBox 360 controller connected to the PC via the USB wireless receiver. This controller has two analog sticks, which are mapped to the translation (a in Fig. 1) and rotation (b) actions. These analog sticks can only express a direction of movement on two axes, while a further one is obviously required. To compensate for this shortcoming, the button *A* (c) switches the currently used Y axis on the analog sticks to toggle between the real world's Y and the Z axes. The other buttons are not used.

B. *The Nintendo Wii Remote*

The Wiimote can be connected to a desktop PC through a Bluetooth interface. Internally, it uses a 3 axes linear accelerometer to sense any change in orientation. It also has an optical sensor capable of tracking four different infrared emitters. These infrared LEDs are usually placed on the Nintendo sensor bar (five at each end): the Wii CPU is then able to calculate the distance between the Wiimote and the sensor bar through triangulation. Applications can make use of this information to calculate position and orientation of the Wiimote. Unfortunately a Wiimote connected to a PC cannot make use of this feature due to their unavailability as a mainstream product. As mentioned, many enthusiasts recurred to building their own ones. We did not take this route because we wanted to evaluate the effectiveness of off the shelf technology which required as less effort as possible to get it working. A stand alone Wiimote can only be used to express 2DOF movement direction. Therefore the button *B* (d in Fig. 1) is used (by holding it), as in the XBox controller case, to switch between axes, while button *1* (e) is used to enter into translation mode and button *2* (f) is used to enter rotation mode.

IV. THE TEST ENVIRONMENT

In order to carry out our experimental study, we created an application that allows users to perform tasks by interacting with a 3D environment. For each task, the application records the time to complete it. All tasks have a time limit of 60 seconds, after which the results are recorded anyway, but the task is not considered complete.

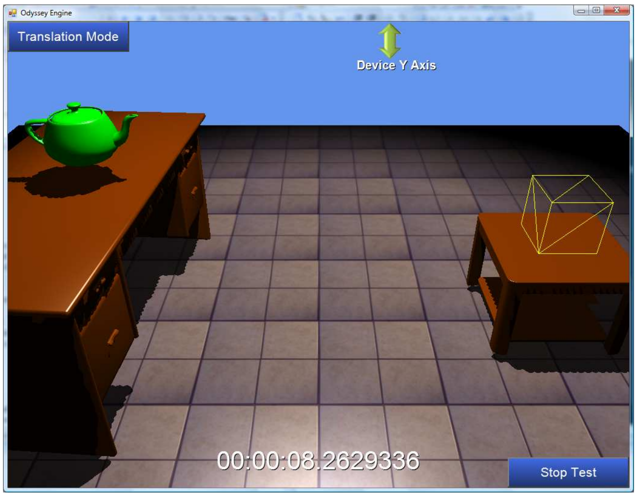

Fig. 2. The “translation” task

Fig. 2 shows the 3D environment for the first task, called “translation” task since participants are required to move the classic "Utah teapot" model from the left table to the right table. Since the controllers can only express a two-dimensional direction, the destination location is placed so to require users to switch axis if they want to position it correctly. A light source projecting shadows is placed straight above the scene, rendered using a classic shadow mapping algorithm. In this way, shadows helps users to better judge the object depth in the virtual scene. A bounding box placed on the destination location also helps users understand where exactly to place the object. The system also checks for collisions by forbidding to move the teapot through the tables or the floor, for example. The task is completed when the object is in the destination location or the time runs out.

The second task is called “rotation” (Fig. 3) since the user has to rotate the teapot to match the orientation of the tip of the teapot with the orientation indicated by a 3D arrow model, shown on the right of Fig. 3 by rotating the teapot so that it matches the arrow’s orientation shown on the right of Fig.3. The last task is called “path” task since the user has to move the teapot through a series of hoops placed in the VE in the correct order (Fig. 4).

This application was developed using the .NET open source port of the DirectX, SlimDX [13]. Input device operativity was provided by the XInput API, for the XBox gamepad and the Wii Library [14], for the Wiimote.

According to a user-centered approach, the application has been developed with an iterative process, in which prototypes have been evaluated with users. We involved 6 computer science students in such evaluations, which consisted in informal observation of a user at a time performing the described three tasks, followed by an interview.

The feedback received from some users made us introduce some textual labels in the interface that helped to report some conditions about the state of the device they were currently using. In fact, the considered input devices cannot express a 3D direction; so as previously mentioned, the Y-axis must be sometimes switched between up/down movement on the world’s Y-Axis and forward/backward movement on the world’s Z-Axis. This, however, caused some users to forget in which state the input device's Y axis was set to. So to avoid or minimize this effect, an icon was placed to the top of the screen that reported whether the input device's Y axis was mapped to the 3D's Y or Z axes. For example, in Fig. 2, the icon for up/down movement is displayed at the top right corner of the screen, while in Fig. 4, the forward/backward movement is displayed instead. The same also applies for rotation on the world’s Y or Z axes (Fig. 3).

Another issue regarded the other possible state change, from the *”translation mode"* to the *“rotation mode"*. To this end, a text icon reminding users in which state they are in was placed at the top right corner of the screen (see Fig. 2-4). Finally, some users asked about a more easily identifiable visual indicator for the rotation task. Previously, only a small status message reported them the distance from the correct orientation. This was removed in favor of a bigger icon in which the distance is written on top of a red background (see Fig. 3) until it falls within a certain limit; when this happens the task is considered passed and the background becomes green. These changes resulted into a new version of the application, which is the one used for study and is shown in Fig. 2-4.

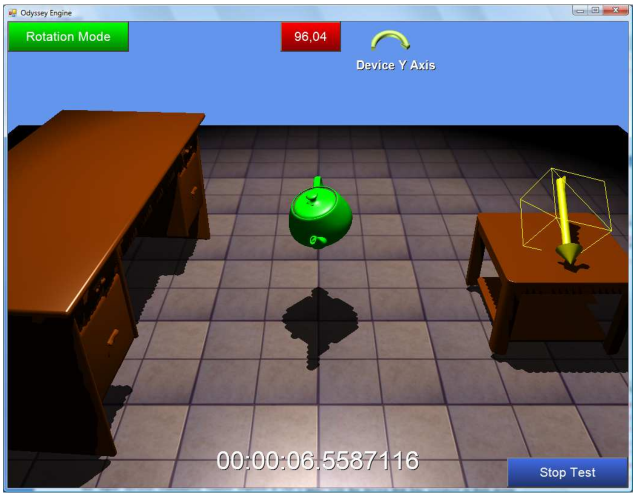

Fig. 3. The “rotation” task

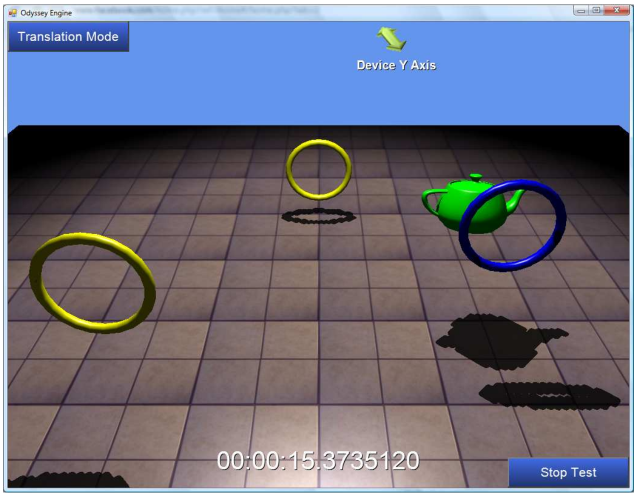

Fig. 4. The "path" task

V. THE FIRST PILOT STUDY

A group of 6 Computer Science students (3M, 3F) participated to the experiment. Participants were divided in three groups. Each group had to perform three tasks (translation, rotation and path) using only one of the three input devices. None of the participants assigned to the groups interacting with either the Wiimote or the XBox Gamepad had any prior experience with these devices. They had none or very little experience with VEs.

Two HCI researchers conducted the experiment. One of them explained to each user the purpose of the test and the control mappings of the input device they were going to use. Before performing each task, the user looked a brief example movie of what they had to do to pass it. The second researcher observed participants behavior during the interaction and took notes. When the tests were over the researcher had a brief discussion with each participant to collect their impressions and feedbacks about difficulties they encountered during the test.

We observed that users having little experience with VEs had problems with judging depth distances correctly. For example most of them did not understand that, in the third task, the hoops were placed at different distances; when they had to move the teapot through the hoops, they only tried aligning the teapot to the two-dimensional projection of the hoops, not realizing that they were at different distances (despite the sample movie explicitly showing the correct way to complete the task). The shadows projected by the teapot and hoops did not help users. Thus, we updated the prototype of the test application so to help participants to correctly recognize distances. In the previous version of the test environment, the ground platform was rendered with an uniform blue color. In the new version, the ground platform is rendered with a texture showing a common tiled floor. In this way, users could be able to better judge the depth of objects placed in the virtual scene, thanks to their projected shadow on the floor.

Another important issue emerged in the pilot study: the users seemed to not be paying much attention to the explanation of the input control mappings given by the researcher. In fact we observed they routinely fumbled with controls when the teapot did not respond as they expected to. Overall though, by the third task users had understood how to use the device and in fact they got better results.

VI. THE EXPERIMENT

In the final experiment, a group of 10 teenagers (5M, 5F) was involved. The experimental design was changed, as described in the following.

A. *Procedure*

Differently from the previous experimental design, we decided in favor of a within-subjects design, letting each participant perform each task with all three devices. For each device the participants performed the translation task as training, because in the pilot study we noticed that only explaining the input control mappings was not enough to have a successful interaction. The time for executing the rotation task was not recorded.

Devices (Mouse & Keyboard, Gamepad and Wiimote) and tasks orders ("rotation" and "path") were counterbalanced. As in the pilot study, an example movie was shown before each task. The participants were required to rate their experience with the device they were using at the end of each task.

The experiment ends by filling a questionnaire about user skills with computer games or VEs in general, how frequently they have previously used each device, etc. Some questions regarded the user experience with the device employed.

Fig. 5. Participant interacting with the Wiimote

B. *Results*

Concerning the quantitative analysis, the task execution time of completion and number of errors were considered. The average time to perform both tasks (rotation and path) shows that the Wiimote was decidedly slower (29,09 s) than the Xbox 360 Gamepad (21,55 s) and the Mouse & Keyboard (20,40 s) (see Fig. 6).

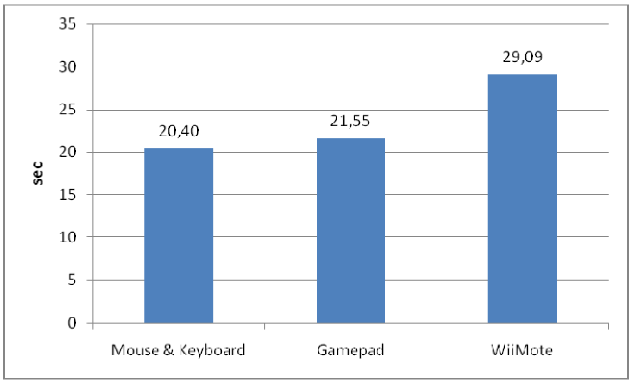

Fig. 6. Mean times for the three devices averaged on all three tasks

Even analyzing each task separately, the Wiimote is slower than the other two devices (Fig. 7). Only in the rotation task, the Wiimote appears to be faster; but looking the error graph, this task was not completed in 80% of the total attempts, compared to 20% for the Mouse and Keyboard and 10% for the Gamepad. In the "path" task, the error rate is still very high (40%) with the Wiimote, while participants did not make any error with the two other devices (Fig. 8).

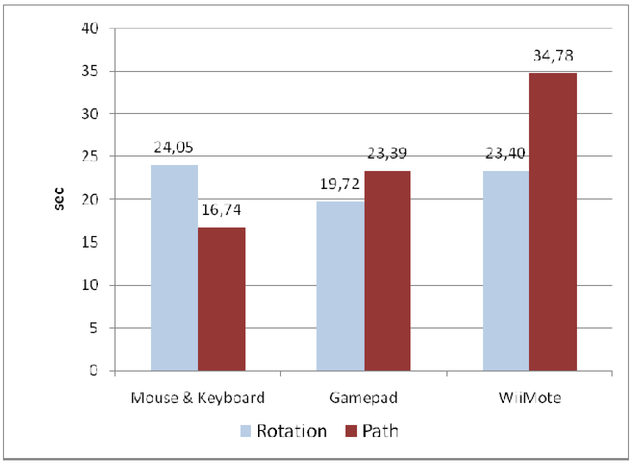

Fig. 7. Mean execution times for each device. Left column refers to the “rotation” task, right to the “path” task

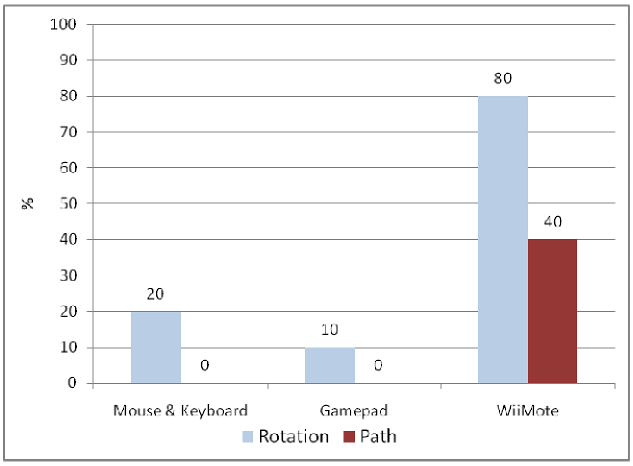

Fig. 8. Percentages of tasks not succesfully completed in the given time. Left column refers to the “rotation” task, right to the “path” task.

The analysis of the questionnaires demonstrate that 60% of participants said that the device they preferred using during the experiment is the Gamepad, while 90% of them stated that the one they disliked the most was the Wiimote (Fig. 9).

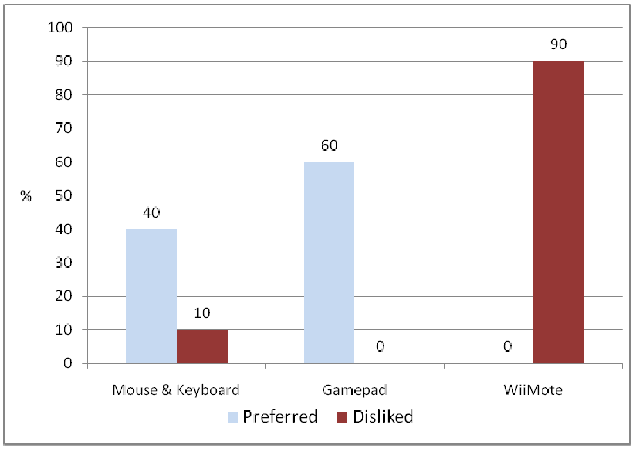

Fig. 9. Percentages of preference for each device. Left column represents their appreciation rating, the right column their dissatisfaction rating .

Participants were asked to express their opinion about the difficulty they encountered in performing the two tasks: rotation and path, with the different devices. This rating was on a scale from 1 (easiest) to 7 (hardest). According to the participants, it was harder to complete the tasks with the Wiimote (mean = 5,3) while Mouse and Keyboard (mean = 2,95) and Gamepad (mean = 2,85) received similar ratings (Fig. 10).

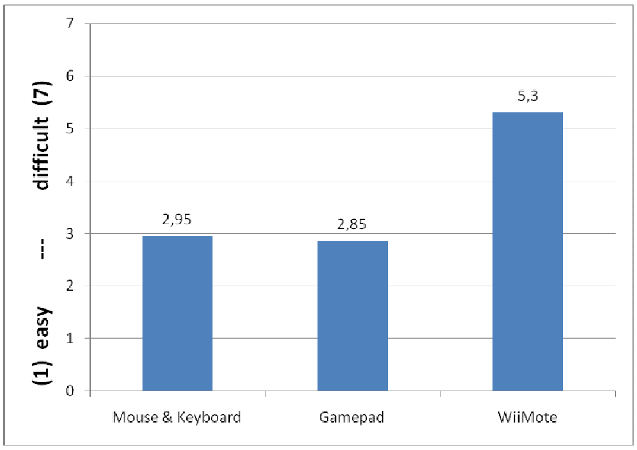

Fig. 10. Difficulty ratings for each device, expressed on a scale from 1 (easiest) to 7 (hardest).

VII. DISCUSSION

From what we were able to observe, the presence of a tiled floor virtually eliminated the problem concerning the judgment of distances that users experienced in the pilot study. In fact the tiled floor, together with the projected shadows, helped users to immediately recognize and identify those objects who were closer or farther than the teapot, enabling them to react accordingly. Interestingly enough, the Gamepad appears to be preferred over the newer Wiimote. This is probably due to the fact that the Gamepad, by employing the use of analog joysticks, is somewhat more precise than the Wiimote. The accelerometers are very sensible and report even small variations (although we did filter out very small values) so a certain degree of training is necessary to be able to use the Wiimote. Another important aspect to consider is that the Wiimote is newer as a “popular” device, although the technology it uses is not new in the entertainment world, as it has been used years before, albeit with limited success. The gamepad, available since a longer time, is more “familiar”: even if someone has never actually used it, it is highly more probable that s/he has at least seen somebody else using it. Furthermore, the presence of two conventional analog sticks conveys more easily the idea of how it should be operated. The Wiimote, instead, requires some explanation before understanding how to use it, because it is the first example of its genre. In the future though it can be assumed to gain a better degree of familiarity, if the Nintendo Wii platform will continue to be developed on.

According to the results obtained thus far, the Gamepad seems to be the most appropriate device to use in such settings where the use of mouse and keyboard is unpractical: it can be used while standing, it is easy to learn and the presence of two analog stick, several buttons offer a great variety of input possibilities. As of now, though, the majority of virtual environments using a Gamepad as an input device allow users to just navigate and look around and offer a severely limited degree of interactivity.

VIII. CONCLUSIONS

This paper presents results obtained from a user study in which two low cost input devices were tested and compared to performances of classic mouse and keyboard, namely: the Microsoft Xbox 360 gamepad and the Nintendo Wii remote. Aim of the study is to investigate the feasibility of using such low cost devices as valid alternatives, especially for settings where the use of the mouse and keyboard is less convenient. The study findings indicate that mouse and keyboard are still the most efficient input devices as far as task performance is concerned. The gamepad is preferred by users, possibly because it is already popular in videogames. Thus, it would be worth carrying out further studies in order to evaluate if longer training with the new devices would have an impact on the overall performance. Future works will include the possibility of evaluating the impact of such devices when applied to "real" VEs with graphically realistic settings, which could perhaps mitigate the problems we encountered with the cognitive aspect of correctly interpreting distances in the virtual world. In fact, when performing the tests, we did not foresee those problems to arise with such relevance. The obvious reason is that users who do not deal with 3D VEs on a daily basis do not share the same level of expertise of researchers and practitioners. Some details that are easily identifiable by domain experts are instead hard to spot for casual users. Special care must then be placed in making sure that these issues do not influence the outcome of the test.

IX. REFERENCES

[1] H. Patel, O. Stefani, S. Sharples, H. Hoffmann, I. Karaseitanidis and A. Amditis, "Human centred design of 3-d interaction devices to control virtual environments", *International Journal of Human-Computer Studies*, vol 64, Mar. 2006, pp. 207 – 220.

[2] C. Hand, "A survey of 3d interaction techniques", *Computer Graphics Forum*, vol. 16, Dec. 1997, pp. 269–281.

[3] M. Mine, "Virtual environment interaction techniques", *University of North Carolina Computer Science Technical Report TR95-018*, 1995.

[4] I. Poupyrev, M. Billinghurst, S. Weghorst and T. Ichikawa, "The go-go interaction technique: non-linear mapping for direct manipulation in VR, in *Proc. of the 9th annual ACM symposium on User interface software and technology*, Seattle, USA, 1996, pp. 79-80.

[5] D. A. Bowman and L. F. Hodges, "An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments", in *Proc. of the 1997 symposium on Interactive 3D graphics*, Providence, USA, 1997, pp 35-ff.

[6] D. A. Bowman, D. B. Johnson and F. Hodgesl, "Testbed evaluation of virtual environment interaction techniques", *Presence: Teleoperators & Virtual Environments*, vol. 10, Feb. 2001, pp 75–95.

[7] K. Hinckley, J. Tullio, R. Pausch, D. Proffitt and N. Kassell, "Usability analysis of 3d rotation techniques", in *UIST '97: Proc. of the 10th annual ACM symposium on User interface software and technology*, Banff, Canada, 1997, pp. 1–10.

[8] R. P. Darken and R. Durost, "Mixed-dimension interaction in virtual environments", in *VRST '05: Proc. of the ACM symposium on Virtual reality software and technology*, Monterey, USA, 2005, pp. 38–45.

[9] L. Gallo and I. Marra, "3d interaction with volumetric medical data: experiencing the wiimote", in Proc. of the 1st international conference on Ambient media and systems, Quebec, Canada, 2008, pp. 1-6.

[10] J. C. Lee, Hacking the Nintendo Wii Remote, *IEEE Pervasive Computing*, vol. 7, July 2008, pp. 39–45.

[11] A. Henrysson, M. Billinghurst and M. Ollila, "Virtual object manipulation using a mobile phone", in *Proc. of the 2005 international conference on Augmented tele-existence*, Christchurch, New Zealand, pp. 164–171.

[12] P. Isokoski and B. Martin, "Performance of input devices in fps target acquisition", in *Proc. of the international conference on Advances in computer entertainment technology*, Salzburg, Austria, 2007, pp. 240–241.

[13] SlimDX, http://slimdx.mdxinfo.com, last retrieved on January the 12th, 2009.

[14] Wii Library, http://www.codeplex.com/WiimoteLib, last retrieved on January the 12th, 2009.