A Point-and-Shoot Technique for Immersive 3D Virtual Environments

Figure 1: Users using the proposed point-and-shoot interaction technique. From left to right: a user wearing the virtual reality devices used by the technique; a close-up on another user performing the tasks; placing the 3D tracker on user's arm; a screenshot of the testbed application where the users performed the tasks.

Marcos Slomp   Andre Spritzer   Rafael Torchelsen

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

{slomp, spritzer, rptorchelsen} @ inf.ufrgs.br

Abstract

Modern computer games often explore graphics quality while leaving interaction and gameplay behind. Arcade machines may improve gamer experience by using special devices (instead of traditional joysticks) to interact with the game. We propose a technique for 3D "point-and-click" interaction that separates both camera and pointer control. It also explores audio and visual collision detection feedback to improve user spatial sense. The experiments were done within a virtual 3D environment, similar to that used in first person shooter games (FPS), a genre popular and well known by both the entertainment industry and gamers. It can also be used in applications other than games, to perform tasks such as "drag-and-dropping" objects, drawing and selecting in general. The results were compared with the traditional FPS interaction technique (mouse-driven camera with point-and-click interaction).

1. Introduction

A common problem with 3D interaction is finding a way to identify and interact with multiple objects at the same time [1]. In such applications, the user is often immersed in a virtual environment and must perform tasks while exploring. The point-and-shoot task exemplifies such a situation, in which the user must travel within the virtual world to seek and destroy targets that may be vulnerable to some specific action (such as different weapons). First-Person-Shooter games (FPS) are a traditional type of point-and-shoot application with these characteristics, being one of the most popular game genres, which includes titles such as the well known Quake and Unreal franchises.

One of the reasons that make FPS game so popular is the degree of immersion that the player experiences. Traditionally a camera is placed in the position of the eyes of the avatar to represent the point of view of a person in the game's world. The player can control the direction where the avatar is heading independently from the direction of view, which can also be controlled. However he cannot move the aim independently from the camera. Most FPS games mimic the behavior of a soldier that has a gun in his hands. For instance, if the weapon were a bazooka, the player would not be able to control the aiming independently from the camera. However, if it were a pistol, something like that is naturally not always true. This traditional gameplay also leads to another problem: when the player wants to move in a direction while shooting in another, he would not be able to see where he is going.

This work presents a technique where the camera control and target aiming are completely independent from each other and weapon switching can be performed quickly and easily. All of this is made intuitive the user by the mapping of his natural real-word gestures to the actions of his avatar. A head-mounted display is used to dictate the direction of view, a pinch-glove for both shooting and weapon selecting and a 3D tracker for aiming. The shooting metaphor was chosen for the testbed application due to the popularity of FPS games and the intuitive translation of real-world gestures to this type of virtual experience.

The technique works well for FPS applications and it can also be easily adapted for uses other than games, such as "drag-and-dropping" objects, drawing and selecting in general.

Empirical results were presented to support the use of the technique as a means of improving user immersion and performance on switching weapons and avoiding "false hits", as described in more details in section 3.

Since the testbed application is very similar to a shooter game, this document starts with some background information and an overview of works describing related techniques. Next, the featured technique is explained in details, including its basic idea, hardware, software, experiment and data collection. This is followed by the results obtained and the problems encountered and pointed by the subjects of the experiment. Finally, conclusions are summarized and a set of improvements, future works and research opportunities are listed.

2. Background

Despite the lack of publications from the game industry and developers, it is interesting to examine how shooter games emerged and evolved. After this briefing, some relevant related works will be discussed.

Arcade games were a very popular entertainment choice during the 1980's. Missile Command, a game from Atari, was launch in mid 1980 to become on of the greatest classic video games ever. The gameplay was very simple: move a cursor on the screen and press a button to shoot missiles at enemy missiles in order to keep them from hitting cities that were located below the screen. All interaction was done using a trackball and three buttons correspondent to three onscreen cannons that could be used by the player to accomplish his in-game mission. Within a level, the cursor moves always at the same speed, which is increased as the player advances in the game. The view was static, which means that the player was not free to explore the world. Years later, the same game was ported to Atari 2600, which introduced an axis-based joystick with a single button, but its gameplay barely changed.

Another relevant game appeared in 1984, when Duck Hunt was launched by Nintendo in Japan for its Nintendo Entertainment System (NES). The same idea from Missile Command was present here, but the player was able to interact with a more realistic metaphor: a pistol, which was used by pointing it at the screen and pressing its trigger. The screen was also static. Variants of the same game allows the player to interact using the default joystick from NES, which consists of a digital pad with four directions (D-Pad).

Many other games emerged from Missile Command and Duck Hunt, but, in 1992, Konami launched for the arcade the game Lethal Enforces, which uses the pistol metaphor and also allows limited weapon switching and camera movement. During gameplay, an icon of a weapon appears on the screen. If this icon is shot at, the current weapon becomes the one it corresponds to and During gameplay, an icon of a weapon appears on the screen and if the user shoots at it, the current weapon becomes the one represented by the icon and it will be active until it runs out of bullets, in which case it will switch back to the initial gun. The camera moves horizontally (scrolling) after the player eliminates all enemies in an area (screen). VirtuaCop, an arcade game by SEGA released in 1994, allowed the player to move in a 3D virtual environment, but the camera still moved without any interaction from the player. From this day on, no relevant modifications were made on the gameplay of this type of shooting game.

Two years later from VirtuaCop a new game style was born: Wolfenstein 3D was launched by id Software for the PC (DOS). The player was able to freely move within the virtual environments of the game, moving forward and backward and looking right or left. Looking up or down was not possible. A virtual gun was placed (drawn) on the bottom of the screen, centered, thus locking the aiming to the camera movement. Basically, all interaction was done with the traditional keyboard, including weapon switching. It was the beginning of the First-Person Shooter era, and many games derived from Wolfenstein 3D, including the well known Doom and Quake franchises, also from id Software.

The most significant advance in gameplay from Wolfenstein 3D to modern FPS gaming is the possibility of looking freely by using the mouse to control the camera. But again, aim and camera control are bound together. Some games, such as Freelancer, from Microsoft, allowed aiming to be done independently from the camera, but since navigation and aiming were done using mouse, they tended to mutually interfere.

There were also FPS games that used virtual reality devices (head-mounted displays) to improve the gamer experience. In these cases, looking and aiming were again bound together and shooting and weapon switching was accomplished in non-intuitive ways, using special joysticks which were held by the player.

3. Method

Within this section, all implementation details and issues will be discussed. First, an overview of the basic concept of the proposed technique is provided. The hardware available for the implementation of the concept is then presented, followed by the software used and implemented. The experiment is then discussed in details, which include the hypothesis, dependent and independent variables, and what data was collected from the testbed application and from the questionnaire applied to each participant.

3.1 Concept

As previously mentioned, the presented technique works very well for point-and-shoot applications. Since its goal is to skirt problems and limitations found in existing interaction techniques, the following items should be considered:

- user should be able to look freely;
- aiming independent from camera control;
- ability to switch between weapons;
- shooting.

It is important to perform all these actions naturally, i.e., the devices employed should allow intuitive and simple ways of mapping real world intentions and gestures to the virtual environment actions.

To implement the camera movements, a head-mounted display with rotation tracking is used. For aiming, a single 3D tracker with only rotational degrees of freedom is sufficient. Shooting and weapon switching are performed using a data glove with contact sensors on each fingertip (pinch glove).

The user wears the head-mounted display an all of his movements are naturally mapped to the testbed application, which properly rotates the camera in accord with the user's real world movements. The user also wears a data glove on his dominant hand. When a contact between the thumb and any other fingers occurs, it means that the user wants to shoot and that the weapon of choice is the one correspondent to the finger that touched the thumb. The tracker is placed on the user's arm, since placing it on the data glove is not a viable option because hand movements may change its rotational state so that it points to a different place from where the user wants to aim at, which is not desirable (Figure 2).

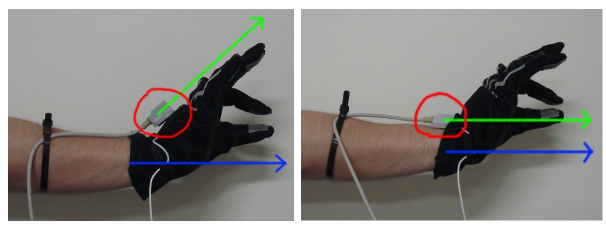

Figure 2: Tracker positioned on user's hand and arm respectively (red circles). Blue arrows represent where the user intends to aim and green arrows represent the tracker's effective aim. It is preferable to place the tracker on the arm, since placing it on the glove may result in a discrepancy between where the user wants to shoot and where is actually shot at.

Note that the user is not able to move freely (translation) within the environment, but as the main objective of the technique is pointing-and-shooting, this limitation also helps him to keep in mind his goal, avoiding distractions such as exploring the environment excessively and unnecessarily.

3.2 Hardware

The head-mounted display used was IISVR's VFX3D. The 3D tracker was Ascention's Flock of Birds and only its rotational features were used. The data gloves used were Fakespace's Pinch Gloves. Figure 3 shows the virtual reality devices used.

The computer used was a Pentium 4 1.4GHz with 512MB RAM, running Windows XP Professional. The video card was a GeForce Quadro FX2000.

This was the equipment chosen due to its availability at the laboratory where the technique was developed. As will be further mentioned, additional trackers and a more accurate head-mounted display may contribute positively to the success of the technique.

Figure 3: Devices used. Top-left: VFX3D (head-mounted display). Top-right: Pinch Gloves (data glove). Bottom: Flock of Birds (3D tracker).

3.3 Software

In terms of software, OpenGL was used for 3D rendering. The audio was implemented using OpenAL. A simple collision detection system was also implemented, dealing with boxes, spheres and planes. Since boxes later proved not to be helpful to the testbed application, they were not used. An exporter for the 3DS file format was implemented in order to import models created within 3D Studio MAX. Particles and textures were made using GIMP and Microsoft Paint.

The virtual environment thus provides much feedback to the user: particles, fog, radiosity-based illumination (using light maps), textured models, 3D sound effects, music and collision response.

There are many virtual reality and 3D interaction applications, but few explore visually complex scenes. This is particularly important to this experiment since it can be essential for a higher immersion in the virtual world, which is one of the requirements of the experiment. The idea is to keep a simple and clean scene, without much details, but that still is visually pleasant to the user.

For device programming and communication, support SDK from manufacturers was used. Unfortunately, no helpful SDK or driver was found for the Pinch Gloves, so a driver for it had to be developed from scratch, based on code fragments found from a few sources on the Internet.

More details about the design of the virtual environment and the user tasks are described in the next subsection.

3.4 Experiment

The testbed application consists of a simple first-person shooter game, that was codenamed *GhostHunt*. The user must seek for ghosts (targets) and destroy them using his weapons. When all ghosts of an area are destroyed, the camera moves (translates) to another area, where a new gate appears and more ghosts emanate from it.

Four types of weapons were made available: yellow, green, blue and pink. Ghosts appear in these same four colors. Every time a ghost gets hit by a shot of its color ("true-hit"), its color randomly changes to one of the other weapon colors. After four hits, the ghost is destroyed. If a shot from a different color hits a ghost, the "false-hit" counter is incremented. Refer to Figure 4 to some screenshots of the testbed application.

The user task is quite simple: clear all ghosts in the area using his weapons and finally face a "boss". Ghost size and initial color are chosen randomly. They move using route points within an area (near its spawn gate). When a route point is reached, it randomly chooses another one and selects a new linear and angular velocity. The boss acts just like the ordinary ghosts, but instead of being completely destroyed after four hits, each time it gets a true-hit, it is destroyed but spawns two "mini-bosses", which are identical to the boss, but smaller in size. The same applies to these mini-bosses and this goes on until no additional spawns are permitted.

Audio feedback is provided, using 3D sound effects to help the user locate the ghosts that are outside his field of view. Music was added to help cutting off the user from the real world.

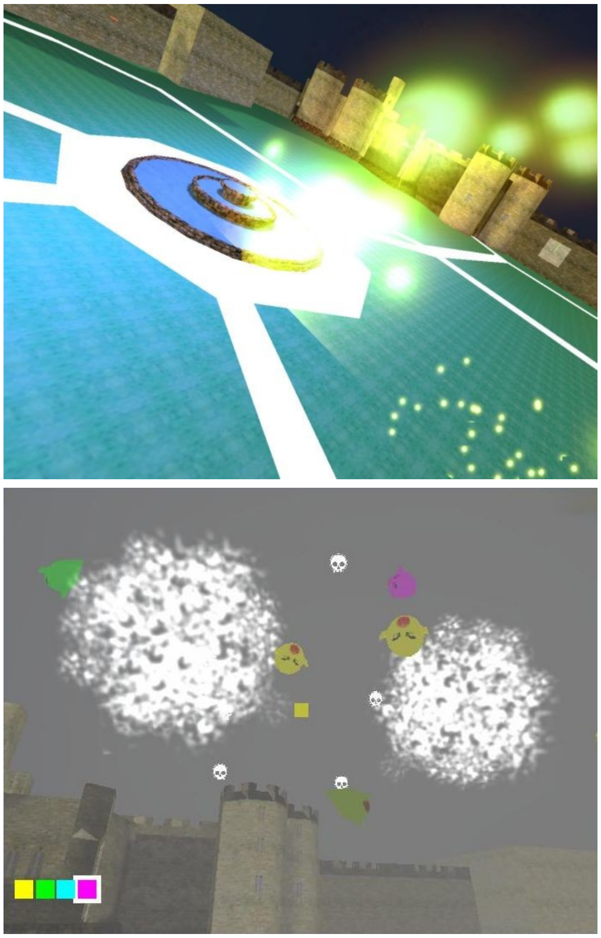

Figure 4: Screenshots from the testbed application. On top, an overview of the virtual environment used, with some particle effects and without fog. On the bottom, a screenshot of the in-game experience (fog is enabled): the yellow square near the center of the picture represents the aim of the user; the squares on the bottom of the picture represent the weapons (white border marks the selected weapon); ghosts appears in different colors, spawning from gates (white clouds); white skulls are particles emitted when a true-hit occurs.

The user had to clear the level twice: one time using the traditional interaction technique of first-person shooters, using the keyboard (for weapon switching) and mouse (looking, aiming and shooting), and another using the devices and the presently proposed interaction technique. Before the experiment, the user was given the opportunity to practice in both modes. A simple launcher program was developed to collect user information (such as identification and dominant hand) and then create configuration files that were used as input by the *GhostHunt* application. The launcher also randomly sorted the order of the practice sections as well as the order of the experiment sections. Figure 5 shows a user interacting with the environment using the technique described in this document.

Figure 5: User playing GhostHunt with the proposed technique: he is wearing a head-mounted display, a pinch glove and a 3D tracker on his arm.

3.5 Hypothesis

When comparing the new interaction technique with the traditional first-person shooter, the following items are supposed to be verified:

- it will be easier to seek enemies;
- it will be easier to aim and hit enemies;
- weapon switching will be easier and faster;
- performance will be improved with independent aiming and looking.

Only one independent variable is tested: performing the task using the traditional interaction technique and the new one, in which camera rotation and aiming are independent from each other. For the dependent variables, the following items were considered:

- time to complete the task;
- "true-hit" ratio;
- "false-hit" ratio;
- immersion level (subjective);
- fun level (subjective).

After the tests, when analyzing the resulting logs and the feedback given by the users, many interesting facts were noticed and are described in more details in sections 4, 5 and 6.

3.6 Data Collection

The log file resulting from the experiment contains the following information:

- clear time: time spent to complete the task;
- shots: number of shots fired;
- false hits: shots that collided with a target but are not in the same color as it;
- hits: shots that effectively collided with targets. Since the number of targets is always the same in all executions of the testbed application, this data is also always the same and is thus irrelevant.

The launcher application also asks the user what is his dominant hand before calling the practices and tests.

All participants were instructed to answer a questionnaire, with pre and post experiment questions. It was composed only from simple multiple-choice questions, helping the user to be as direct and clear as possible on his answers.

At the pre-test, the following was asked to the users:

- name
- age
- sex
- level of education
- video-game experience
- first-person shooter games experience
- virtual reality devices experience

where the experience was measured with four levels: 1-None; 2-Poor; 3-Medium; 4-High.

For the post-test, the following questions were asked to the subjects, comparing the traditional FPS experience with the new technique:

- Which one was more fun?
- Which one was more tiring?
- In which one was it easier to seek enemies?
- In which one was it easier to aim and hit?
- In which one was it easier to switch weapons?
- Which one was more immersive?
- In which one the sound was more helpful?
- Was independent looking and aiming helpful?

A subjective and optional question was provided at the end of the questionnaire, in which the subjects could write their comments about the experiment. The information collected from these notes proved to be very valuable to confirm some already expected behaviors and problems. This described in more details in Section 5.

4. Results

To validate the hypotheses, the questionnaires and logs from the tests were analyzed, resulting in the following:

1. **It will be easier to seek enemies.** The opinion of the test subjects confirms the hypothesis. Figure 6 shows that 60% of the testers prefer the head-mounted device to search for the ghosts;
2. **It will be easier to aim and hit enemies.** The opinion of the testers and the logs refutes the hypothesis. From the questionnaire we can see that 85% prefer the mouse to aim and from the logs we can see that the hit-ratio (Figure 7) is better in the traditional simulation;
3. **Weapon switching will be easier and faster.** The questionnaire confirm the hypotheses with 85% of the testers preferring the Pinch Glove to switch amongst weapons;
4. **Performance will be improved with independent aiming and looking.** The hypothesis can not be confirmed because the opinion of the testers was divided, with 50% preferring the new method and 50%, the traditional. The time to complete the task suggests that the hypothesis is not valid. Figure 7 shows that the time taken to complete the tasks is increased by 10% when the devices are used.

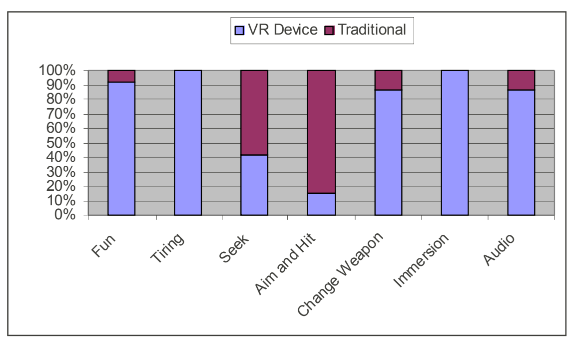

Figure 6: Results of the questionnaire.

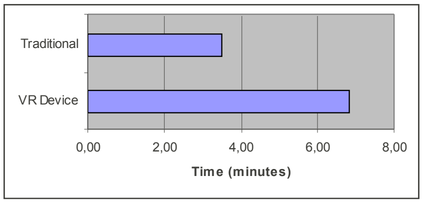

Figure 7: Time taken to complete the task.

The increase of time to complete the task is linked to the fact that about 10% of the testers had never used virtual reality devices in the past. Figure 8 illustrate the experience of the testers with games, FPS games and virtual reality devices.

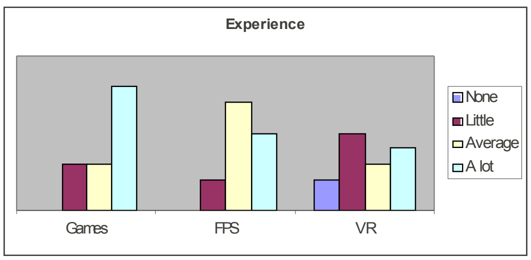

Figure 8: Past experience of the testers.

As to the amount of shots fired and the amount of true hits, Figure 9 shows that the traditional simulation generates 16% more true hits. This data is linked to the fact that most of the users had a lot of experience with FPS and little with virtual reality devices.

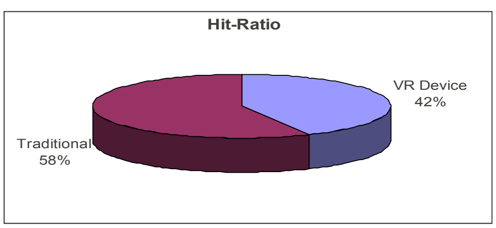

Figure 9: Hit Ratio.

An interesting data collected in the logs and illustrated in Figure 10 is that there were 32% more false hits in the traditional simulation. It indicates that the selection of weapons using the virtual reality devices is more efficient.

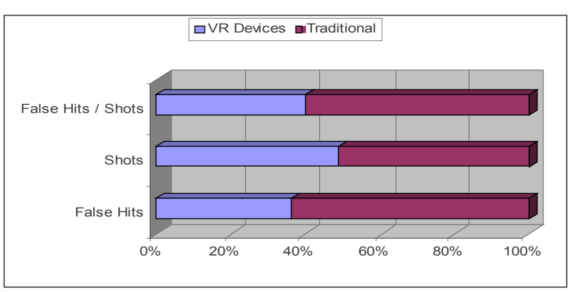

Figure 10: Shots and false hits data collected from the logs.

Another data that indicates that the presented technique is promising is that 100% of the testers affirmed that it is more immersive and 90% indicated it is more fun.

Moreover, the opinion of the testers as to the audio is that it is more helpful in the presented technique. In their opinion, it makes the process of seeking the ghosts with the head-mounted display easier.

It is an unanimous opinion, though, that the intense use of the head-mounted display and the arms make the technique physically tiring. An interesting result was that the fatigue was only noticed by the testers after the end of the test – while they were playing, they didn't feel tired. This is a probable result of the test being entertaining, which is one of the goals of the technique.

5. Feedback and Problems

During the development of the testbed application, several issues were noticed: difficulties of device calibration, problems when converting rotational data between the device driver and the application, rendering fill rate problems when a large amount of particles appears at the same time on the screen, the limited resolution of the available head-mounted display and electromagnetic interference on the 3D tracker. A lot of time was spent adjusting parameters and making conversions and data wrapping to try to minimize some of them. More time was spent on this due the lack of software for the Pinch Gloves, for which a new driver had to be developed. The authors are not any experts in virtual reality, and the lack of experience with devices and interaction techniques was supposed to be surpassed in a very short time (about a month of development time).

The participants feedback notes have proved themselves to be a really valuable source of information. They helped confirm many speculations that were raised during the development, when the problems previously noticed and described above were faced. Below there is a list of the most relevant user feedback that was given after the experiments, followed by the authors' comments about it:

- **The aim was difficult to see when using the head-mounted display**. Nothing much can be done about this, since the available head-mounted display has a very poor resolution, field of view and color quality.

• **The aim color could be changed according to the selected weapon**. Very easy to implement, despite not being common in first-person shooter games.

• **The aim's style could be changed**. A simple yellow square was used for the aim, but it can be replaced by any other texture.

• **Lack of practice and contact with virtual reality devices**. Since virtual reality is not very popular amongst people in general, there is not much that can be done about this, but more training sections may improve user performance.

• **Shots are too slow**. The speed can be increased, but it is not very clear if this was a real problem, since just one subject pointed this out.

• **Slowdowns while playing**. It happens when many particles appears on the screen at the same time, causing a fill rate problem. One option is to limit the number of simultaneous particles, since many of the users tended to use their weapons like a machine gun.

• **The Pinch Gloves failed to respond sometimes**. It was noticed before that the Pinch Gloves do not fit very well in all hand sizes, which can cause difficulties in detecting the contact between the fingers – mainly when trying to touch the thumb with the little finger.

• **Aiming problems when the hand is too far away from the body**. The data retrieved from the tracker presents an increasing and significant error as the tracker is moved away from its magnetic base. This error is accentuated by the presence of nearby metal objects, which may cause a magnetic interference.

With more appropriate devices and more calibration and training, the authors believe that it is possible to highly increase user performance and validate the second hypothesis. Despite all these problems, all users preferred the new technique, considering it more fun and immersive. Another interesting fact is that the weight of the head-mounted display was barely noticed when doing the tests of the new technique, and the authors believe that the immersion level provided by the devices and the feedback given by the testbed application helped the user to ignore its weight.

6. Conclusions

Since it has been some time that first-person-shooters use the same type of interaction, the proposed technique comes to revigorate the genre. To that goal the proposed point-and-shooting technique presents positive results and feedback. The results showed that weapon switching was more entertaining and efficient with the proposed technique. It was also noticed that users tended to perform less false-hits using the new method. It received mostly positive feedback from testers, especially in regards to the entertainment level and the mapping of real-life movements. Such mapping can be used not only in FPS games, but also in any application that requires fast switching of tools, selecting, point-and-clicking, etc..

From the results obtained with the experiment, it can be said that the presented technique allows for a more immersive and entertaining interaction with a 3D environment, to a point that the fatigue felt by the users due to the head-mounted display and the arm movements necessary to control the aiming passed unnoticed by the subjects until after the tests were over.

7. Future Work

With the feedback given by the participants and due to problems encountered during the development stage, a series of improvements can be made to improve user performance and then, possibly, validate the second hypothesis:

• improve the graphics engine to avoid slowdowns;

• include extra scenes;

• training some users for some time using the new technique and confront them with users using the traditional interaction;

• use a rotatory chair to allow 360 degrees for yaw;

• develop tasks that need the use of both hands;

• add haptic and wind feedback;

• study the possibility of adding translational degrees of freedom to allow the user to move and interact with the virtual environment;

• add offensive actions for ghosts to study user's reactions;

• command the experiments in a room isolated from any metal that can interfere with the 3D tracker;

• use more than one tracker to estimate the user's aim, using translational degrees of freedom to improve tracking stability and effectiveness;

• add tasks with stationary targets and better calibrate gameplay parameters such as shot speed, the ghosts' artificial intelligence, and colors that can be even more easily identified in poor head-mounted displays;

• develop a “bracelet” to comfortably integrate the tracker and pinch gloves, giving the user more accuracy and steadiness; it would also reduce the amount of wires scattered around the room;

• add haptic and wind feedback [2].

References

[1] Lucas, J.F., Bowman, D.A., Chen, J., and Wingrave, C.A., "Design and Evaluation of 3D Multiple Object Selection Techniques," Proc. of the ACM I3D, March 2005.

[2] Deligiannidis, L. and Jacob, R.J.K. 2006. "The VR Scooter: Wind and Tactile Feedback Improve User Performance". IEEE Symposium on 3D User Interfaces 2006 (3DUI 2006). March 2006. pages: 143-150.