A system for panoramic navigation inside a 3D environment

Polceanu Mihai
"Ovidius" University, Faculty of Mathematics and Computer Science
Bd. Mamaia nr.124
900527, Constanta, Romania
polceanum@gmail.com

Popovici Alexandru
"Mircea cel Batran" National College
Bd. Stefan cel Mare nr.6
900726, Constanta, Romania
popovici.alexandru@gmail.com

Popovici Dorin-Mircea
"Ovidius" University, Faculty of Mathematics and Computer Science
Bd. Mamaia nr.124
900527, Constanta, Romania
dmpopovici@gmail.com

ABSTRACT

This paper presents a physical user interface intended to help the user (or multiple simultaneous users) to achieve an intuitive movement inside a 3D environment without using common interaction devices such as mouse or keyboard, while stressing the aspect of reducing financial investments. After exposing an analysis of current solutions and implementations of related topics, we argument our implementation and give detailed aspects of hardware and software architecture of the system, as well as a comprehensive efficiency study and explore the use cases with people with motor impairment. As future work, we intend to extend the usability of the system and release it under the GNU General Public License (GPL) for free use and further development by other parties.

Keywords

Virtual Reality, 3D navigation, user interface for physically disabled individuals.

1. INTRODUCTION

Regardless of the quality of simulated 3D worlds, people are still conscious of the barrier between them and what they see; this is because they only benefit from a keyboard, mouse or other common input devices. This paper presents a system through which we try to whittle this barrier and give users a natural interaction tool which they can intuitively use to navigate at will with natural body movements. The concept was also designed to be easily configurable "at home" and to be a low-cost solution. Structure, efficiency and the possible uses as an enhancement for physically disabled individuals are explored in this paper.

2. MOTIVATION

The system discussed in this paper is an improvement of the IIUBAR setup (acronym for Interactive Informative Unit Based on Augmented Reality) described in [Pop08]. It refines the hardware setup used by its previous version, and uses Virtual Reality instead of Augmented Reality; this makes its uses slightly different: user immersion instead of fixed-point informational unit.

The goal of this paper is to extend the possible use cases of this system architecture and to exhibit its advantages and drawbacks compared to specialized hardware. This system has been developed with respect to product quality and reduced financial investments.

3. RELATED WORK

Throughout the history of user interfaces there have been many metaphors for addressing visualization and navigation inside three dimensional virtual worlds. Because of the impossibility of fully recreating a three dimensional space on a two dimensional display system, the ideas that were developed in this domain can be split into two main categories (Figure 1): fixed display metaphors with interaction devices such as mouse, keyboard, space mouse, etc, and mobile spatially aware systems.

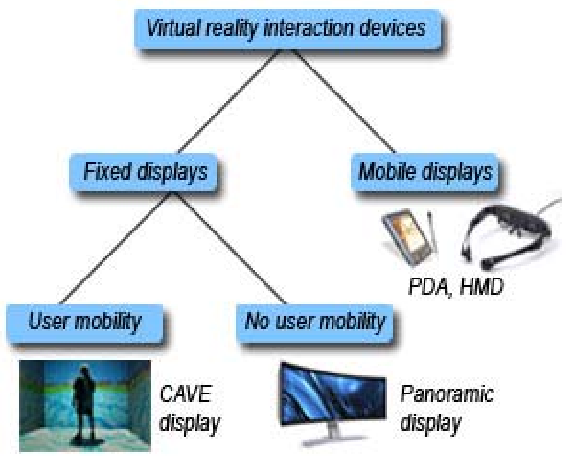

Figure 1. Interaction devices categories

The oldest member of the fixed displays is represented by the classic desktop environment on which 3D applications can run and the user can "move" inside them by pressing keys or by dragging the mouse. Display sizes and image quality have grown in direct dependence with technology, and today a wide variety is available for purchase; this technological advancement came to enhance the experience of 3D navigation by enlarging the user's field of view and therefore contributing to a higher level of immersion inside the virtual environment. Famous examples of this technique include connecting multiple screens to form a panoramic display around the user, a similar metaphor that uses projections which materialized as the CAVE system [Cru93], wide panoramic screens [Bau05], spherical display systems, and other means of extending the field of view as much as possible. Unfortunately, as marvelous as they prove to be, extra technology comes with extra costs, and may not be easily accessible for the usual user and even for educational institutions because of the lack of funds. The category of fixed display devices can be further split into two scenarios depending on the user's mobility and the lack of it; the CAVE system allows users to move freely in a designated space and reacts to his/her movements. The latter, although it allows the implication of multiple users, cannot give each user a personalized viewpoint but only a group-oriented interaction. So far we've identified some key advantages of this category: extended field of view, multiple users and high level of immersion. The inversion of the fixed display concept takes us to explore devices that use their own spatial position to transmit visual data to the user. These devices range from personal digital assistants (PDA) to head mounted displays (HMD) which come in a wide variety of designs. This category also consists of hardware that is specialized for performing precise tasks which, through the prism of virtual reality, consist of binding the navigation to the user's view point; this way the user can specify the desired focus in the environment either by pointing the device or using a pen also known as the peephole display concept [Yee03], or in the case of the HMD by tilting the head in direction. An example from this category is represented by the use of a palm computer for interacting with a virtual environment [Pig08]. This feature gives the great advantage of user mobility, being only constrained by the physical space available to move into. There are several drawbacks to these methods: hand-held devices can only display a small portion of the visualized virtual world and the HMD type devices should allow the user to be aware of the surrounding real environment to prevent accidental collisions while moving around. The latter is achievable through augmented reality but another impediment arises due to the low video resolution relative to natural sight; this can be overcome by using see-through lens technology, but again we stumble into cost issues. From this category, we can derive two new advantages: mobility and user-bound viewpoint.

After this analysis we propose the following question: is it possible to achieve similar performances with relatively basic cost-wise accessible hardware? One of the most used input devices used today is video. In 2004, Microsoft reckoned more than 18.5 million webcam users only with instant messaging applications [Web09a], and the number is ever growing. Webcams have become an accessible and necessary possession for internet users, and they can be used with a wide range of applications and operating systems. Using webcams as input devices is cost-efficient, but require a greater effort to create software capable of interpreting the input data; fortunately, open-source frameworks are freely available which do most of the work. Navigating inside a three dimensional environment requires linear view-point movement and the ability to rotate. In this paper we propose a model through which we try to absorb the mentioned advantages using low-cost equipment. Our solution consists of a fixed but rotatable display that interprets the user's turning movements and level of approach for navigation. To navigate inside a three dimensional environment the basic requirements are speed and rotation, which can be achieved by determining the angle of rotation relative to a point of reference, and using the distance to the face of the user as directional speed input. In the following sections we describe the architecture of this system, its use cases and efficiency evaluation results.

4. SYSTEM ARCHITECTURE AND USE CASES

Considering the available technology previously analyzed, we decided that the best method of reducing the cost of the system is to recreate the functionalities of the specialized hardware through the means of software. This way the equipment requirements can be reduced, but we must emphasize the used software solution. Another aspect of the system is that any user with minimum knowledge on software installation and basic experience with material carving (for the special support table) can create a replica of the system at home without significant investments; we like to believe this can be a motivational factor through the satisfaction of building it. In the following subsections we present the aspects of the hardware and software used, followed by the exhibition of an official use of the system within an ongoing project, and finally we describe the multiuser support and the possibility of creating networks with multiple implementations of the system.

Hardware architecture

The system is composed of two simple webcams connected to a laptop placed on a rotatable support like illustrated in Figure 2. The purpose of the tripod table is to allow the hidden camera to look down on a cardboard marker which is needed by the software to extract rotation coordinates. When in use, the top camera is always directed toward the user(s). Each component is adjusted to fit the others, thus making the system stable for user interaction.

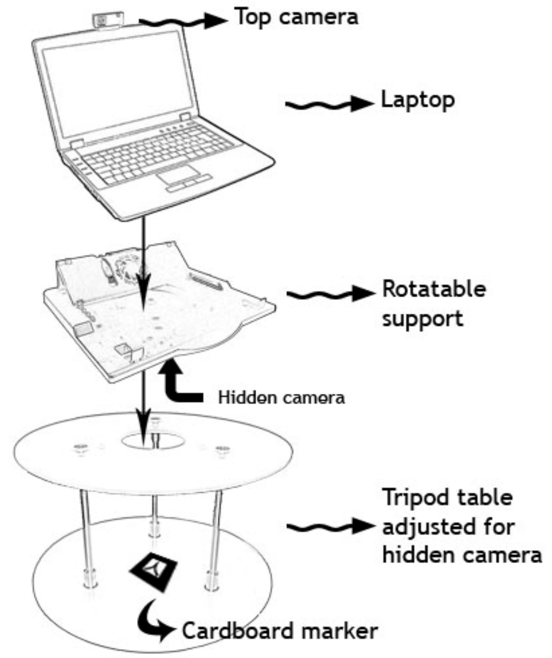

Figure 2. System schematic

Our implementation uses a laptop, a rotating laptop stand, two low-cost webcams and a hand-made tripod table, although any similar hardware can be properly used for which drivers, if required, are compatible with the utilized operating system.

Software

The application is developed in C++ and runs on UNIX (our implementation uses Ubuntu [Web09b]). It combines three open-source frameworks as follows:

- AReVi (Atelier de Realite Virtuelle) [Web09c] for virtual environment representation. AReVi is a powerful agent-based library that provides services for multi agent systems and 3D graphics.
- OpenCV (Open Computer Vision library) [Web09d] for face detection support. Visual algorithms play an important role in deciding how far away a user is from a camera. To advance or retreat in the virtual environment we chose to use the distance of the user relative to the top camera. Achieving this effect resides in detecting the user's face; the difference in face size from different positions give away the distance, i.e. if the image of the face appears larger implies that the user is closer to the camera and vice-versa. A basic threshold based noise reduction algorithm is used to prevent the navigation speed from trebling.
- ARToolkit (Augmented Reality Toolkit) [Web09e] for viewpoint orientation. Similarly to OpenCV, ARToolkit uses image processing algorithms to extract position and rotation information from a physical cardboard marker and returns a rotation matrix. For rotation in the horizontal plane, we only need one rotation angle (around the z axis); we can calculate this angle with Equation 1, where R = Rx*Ry*Rz (a 3 by 3 matrix), and Rx, Ry and Rz are defined in [Fol93]. This way, the resulting angle is applied to the viewpoint inside the 3D environment and the effect of rotation is achieved.

$$|\gamma| = acos(\frac{R(1,1)}{cos(asin(R(1,3)))})$$

Equation 1. Rotation angle calculation

The main application works by reading data from OpenCV and ARToolkit through a local shared memory mechanism (Figure 3); this concept is not new, but makes individual builds independent from each other, and enhances simplicity of the code and extensibility of the software.

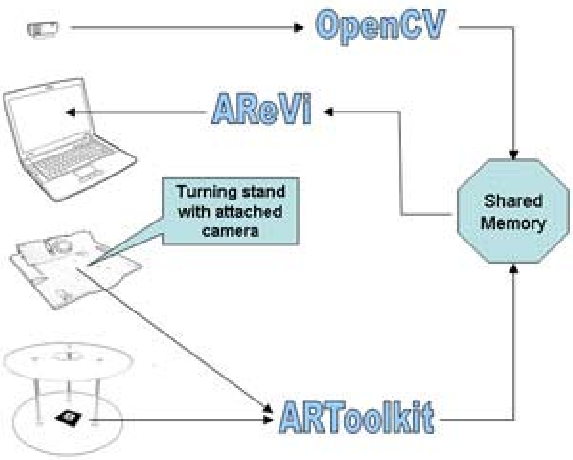

Figure 3. Communication between components

Demonstrating the system's features

For testing the system we developed two applications: a game entitled "Crystal Island" in which players enroll to solve quests by collecting crystals of different colors and bringing them to the totems which required them, and an interactive lesson about the solar system entitled "UFO Driver" in which the user controls a flying saucer and navigates through our solar system to discover the planets, our sun, and the asteroid belt (Figure 4).

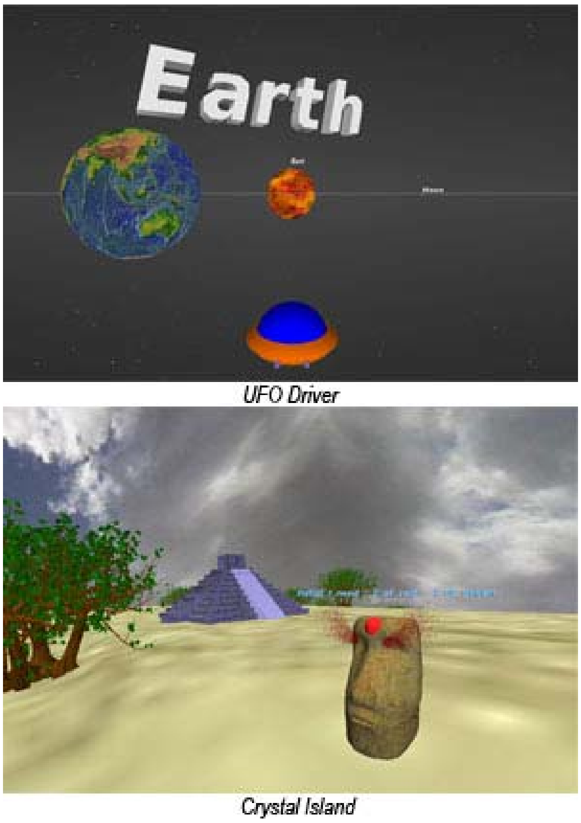

Figure 4. Applications of the system

This lesson about the solar system was created with respect to the real dimensions of the planets, to give scholars a feel of the great distances between celestial bodies and to help them grasp this information in an entertaining and intuitive way.

Virtual tour of archeological sites

Today, a lot of emphasis is put on virtually reconstructing lost cultures from different times in the history of mankind and even before. History lessons have evolved into interactive game-like experiences in which users can explore 3D replicas of ancient artifacts, buildings and even people. The immersion of the user in these environments can give visual, auditory or haptic feedback which helps to better grasp the details that were specific to a certain time in the past.

In this sense, apart from the games that we developed to evaluate the system, we also integrated the system in the TOMIS project which aims to virtually reconstruct the ancient Roman Edifice with Mosaic from Constanta, Romania, through designing, implementing, experimenting and demonstrating an interactive and collaborative multi-sensorial system based on VR/AR technologies. Although not yet complete, the reconstruction of the site has been integrated with the system and allows users to walk through the edifice like it was between the years 46 AD and 610 AD (Figure 5). The starting point of the virtual tour corresponds to the system's location so users can grasp the feel of orientation, and presence in the Tomis colony during the Roman period. The reconstruction has a hypothetical approach as the archeological information from the colony is not entirely complete. This enhancement to the project aims at promoting culture and tourism in the region.

Figure 5. Using the system in the Roman Edifice with Mosaic from Constanta, Romania

Multiuser interactivity

Having more than one system can be used to create a network of interactive "3D browsers" through which users can compete in games or explore virtual sites. The networks can be either local, using a wireless router for increased mobility, or distributed over the internet, or both local and wide area networks connected through a server located on one of the machines.

Along with the network possibility, the system supports multiuser on the same machine (Figure 6). In this situation, the control of the unit or avatar is distributed to each user participating in the interaction. To allow other people to observe without disturbing the users, adjustments have been made to this feature so that people who are more than approximately 1.5 meters away from the system (the limit of physically maneuvering the device) cannot influence the acceleration.

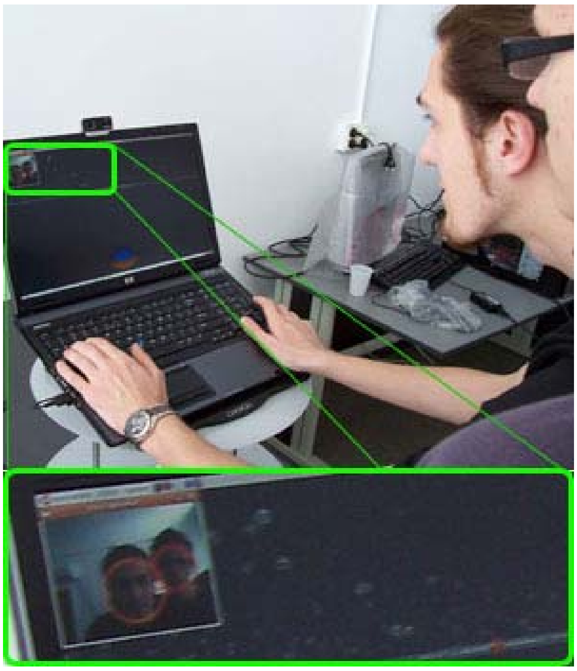

Figure 6. Illustration of multi-user support: average of the face positions is computed (magnified at bottom of picture)

When more users engage in interaction with the system, an average of the users' positions is made and requires them to work in a team to achieve the desired result. Hence to move forward in the 3D environment all users must lean forward and if one of them leans backwards the average acceleration would decrease and cause inefficient movement.

5. EFFICIENCY STUDY

With the occasion of the "Laval Virtual" contest that took place in Laval, France in 2009 [Web10] where we participated with the system under the name of "Navoramique" (which stands for Navigation Panoramique), we took the opportunity to test its efficiency with the help of the people who tried it (Figure 7). In this section we discuss the results of the survey, and some of the suggestions received from our users.

Figure 7. An user testing the system at Laval Virtual

A survey was prepared which contained six questions about the system, as follows:

- Q1: "I think Navoramique is intuitive and easy to use."
- Q2: "Using Navoramique is more appealing than using a keyboard and a mouse to navigate."
- Q3: "I easily learnt how to control my movements with Navoramique."
- Q4: "I think 'UFO Driver' is an interesting lesson about the Solar System."
- Q5: "I found 'Crystal Island' to be an attractive quest game."
- Q6: "I would like to have a version of this system at home."

The possible answer choices were: "Totally disagree", "Disagree", "Neutral", "Agree" and "Totally agree". We made the surveys available both on paper and online within an application, but most users preferred the paper forms because they were faster to fill in, and more persons could submit them simultaneously (we only had one computer available for this task). Figure 8 shows the comparative results for this survey on a number of 112 users of all ages, and different nationalities.

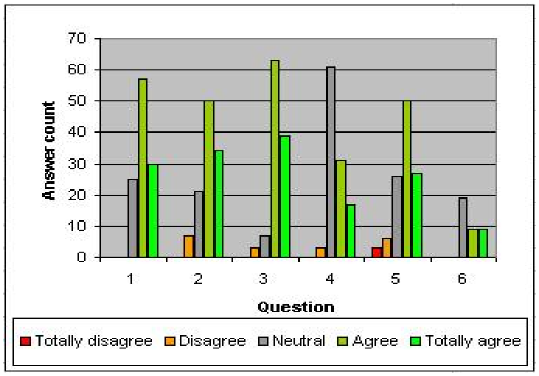

Figure 8. Answers for the evaluation survey

From the results we learnt that an average of 81% of the users gave a positive answer for the first three form items while 18 users expressed the wish for their own version of the system. As not all users played the two available games, the positive answer average was 89.91% of the people who did try the games (Q4, Q5). There were 59.5% people who tried the games from the total who submitted a survey.

As an additional note, in the case of "Crystal Island", the player's performance was measured by keeping the score; the score was calculated with respect to the time that the user needed to complete the tasks of the game. The users also had visual feedback of their current score and about their position in the top scores and after completing the game. The observed results were that while some users managed to achieve a top score from the first try, the others learnt quickly and were able to improve their scores after playing several times.

Some of the users also submitted comments about the system which helped us identify some weak and strong points.

One user pointed out that the table should be adjustable for each person's height. Unfortunately we underestimated the possibility of the difference in height; most tall users had problems with the system because the laptop monitor did not permit a very wide angle of inclination, and therefore they could not adopt a comfortable position while testing. The best results appeared when the line between the user's face and the top camera was close to the horizontal.

Another drawback would be the energy necessary to physically rotate around the table. We also encountered issues with the lighting in the room which caused problems with the face detection when the user stood between the camera and the light source.

The idea was better accepted by children who enjoyed the games and the fact that they had to move around to navigate, and by people who were not comfortable with navigating with the mouse and keyboard. One user mentioned that "as a learning tool it would be enhanced if students have to search for information"; this underlines the factor of motivation in learning.

To conclude the statistics, the system can benefit from small comfort-related improvements, and can serve as an efficient interactive learning tool for primary school pupils and for students. It can also be used as a navigation tool, complementary to the standard input devices.

6. AIDING PERSONS WITH MOTOR DISABILITY

While at the Laval exposition, we were most moved when two persons in wheelchairs asked us if they can try the system (Figure 9). We discussed the possibility of adapting the mechanism to minimize the effort needed for navigation, and one of them suggested that the facial recognition could be also used for turning, so they can only use the head movements. Another solution which we discussed was to allow left-right navigation without having to make a whole turn, but only to slightly rotate the stand; this would give similar feedback effect and would be a lot more convenient in this case. As noted in [Hol06], the weaknesses of one modality are offset by the strengths of another, and by modality the means of interaction is implied. Slight changes to the system like the ones previously mentioned can substitute for the impairment of lower body movements.

Figure 9. Users with motor disability

Although we had not foreseen this use case, we were deeply moved by the fact that their impression was a positive one, and we hope we will collaborate with the asylum in Laval to share the technology.

7. RESULTS

The main aspect which we tried to demonstrate is that using innovative ideas together with common and accessible hardware and software resources, one can achieve efficient low-cost solutions to enhance human-computer interaction either for gaming, learning or aiding persons with disabilities. The presented system has been used for educational, entertainment and aiding purposes. It also represents a method for museums to exhibit a new, modern point of view to the visitors.

8. ACKNOWLEDGMENTS

The present system has been implemented in the Laboratory of Virtual and Augmented Reality Research (http://www.univ-ovidius.ro/cerva/) within the Faculty of Mathematics and Informatics of the OVIDIUS University in Constanta, Romania.

This work is supported by the TOMIS project, no: 11-041/2007, by the National Centre of Programs Management, PNCDI-2 - Partnerships program.

Thanks to everyone who submitted the system evaluation surveys and shared their impressions and suggestions with us.

9. REFERENCES

[Bau05] Baudisch, P., Tan, D., Steedly, D., Rudolph, E., Uyttendaele, M., Pal, C., Szeliski, R., Panoramic viewfinder: providing a real-time preview to help users avoid flaws in panoramic pictures, Proceedings of the 19th conference of the computer-human interaction special interest group (CHISIG) of Australia on Computer-human interaction: citizens online: considerations for today and the future, Canberra, Australia, November 21-25, ISBN 1-59593-222-4, 2005.

[Cru93] Cruz-Neira, C., Sandin, D.J., DeFanti, T.A., Surround-screen projection-based virtual reality: the design and implementation of the CAVE, Proceedings of the 20th annual conference on Computer graphics and interactive techniques, ISBN 0-89791-601-8, 1993.

[Fol93] J.Foley, A.van Dam, S.Feiner, J.Hughes. Computer Graphics: Principles and Practice. Addison-Wesley Co., ISBN 0-201-12110-7, 1993.

[Hol06] Holzinger, A., Nischelwitzer, A.K., People with Motor and Mobility Impairment: Innovative Multimodal Interfaces to Wheelchairs, Computers Helping People with Special Needs, Springer Berlin / Heidelberg, ISBN 978-3-540-36020-9,2006.

[Pig08] Pignatel, A., Farella, E., Brevi, F., Benini, L., Gaiani, M., On the use of a palm computer for design review interaction in a virtual room, The 16th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG 2008), Full Paper Proceedings, ISBN 978-80-86943-16-9, 2008

[Pop08] Popovici, D.M., Polceanu, M., Interactive Informative Unit Based on Augmented Reality Technology, In Proceedings of The 3rd International Conference on Virtual learning (ICVL2008), Bucharest Univ. Press, ISSN 1844 - 8933, 2008.

[Web09a] Website of BBC (article from 2004): http://news.bbc.co.uk/2/hi/technology/3833831.stm (checked Oct 2009).

[Web09b] Website of Ubuntu operating system: http://www.ubuntu.com/ (checked Oct 2009).

[Web09c] Website of AReVi framework: http://svn.cerv.fr/trac/AReVi (checked Oct 2009).

[Web09d] Website for OpenCV library containing links to documentation and source: http://opencv.willowgarage.com/wiki/ (checked Oct 2009).

[Web09e] Website of ARToolkit library: http://www.hitl.washington.edu/artoolkit/ (checked Oct 2009).

[Web10] Official website of the Laval Virtual event: http://www.laval-virtual.org (checked Jan 2010).

[Yee03] Yee, K.-P, Peephole displays: Pen interaction on spatially aware handheld computers, in Proceedings of CHI 2003, ACM Press, 2003.