Using Head and Finger Tracking with Wiimote For Google Earth Control

Gürkan Vural
Department of Computer Engineering
Middle East Technical University
gurkan.vural@gmail.com

Gökhan Tekkaya
Department of Computer Engineering
Middle East Technical University
tekkaya@gmail.com

Can Eroğul
Department of Computer Engineering
Middle East Technical University
erogul@ceng.metu.edu.tr

***Abstract*—Using an LED array and some reflective tape, you can use the infrared camera in the Wii remote to track objects, like your fingers, in 2D space. This lets you interact with your computer simply by waving your hands in the air similar to the interaction seen in the movie ”Minority Report”. The Wiimote can track up to 4 points simultaneously. A demo video for this method can be seen in [1].**

Using the infrared camera in the Wii remote and a head mounted sensor bar (two IR LEDs), you can accurately track the location of your head and render view dependent images on the screen. This effectively transforms your display into a portal to a virtual environment. The display properly reacts to head and body movement as if it were a real window creating a realistic illusion of depth and space. A demo video for this method can be seen in [2].

In this study, we intend to combine these two methods to use the Google Earth application without mouse and keyboard. We can scroll, zoom in and zoom out by using our fingers, and the monitor will display the position of the world according to the position of our head, so we will have a small world simulation flying in front of us, which we can interact with our fingers with a futuristic way of human computer interaction.

I. Introduction

Recent enhancements in technology enables us to implement cost-effective virtual reality systems with COTS (common-off-the-shelf) components. In this work we tried to transform GoogleEarth into a virtual reality world. In order to achieve that, we are going to use the relatively cheaper Wii remote and a head mounted sensor bar (two IR LEDs) to track the location of the head and render view point images on the screen. This effectively transforms the display screen into a portal to a virtual environment. The display properly reacts to head and body movement as if it were a real window creating a realistic illusion of depth and space. The concept is explained in detail by video on the web [2]. Wii remote is also used widely in the world, which makes this project's output a more widely usable product.

To achieve better VR interaction, fingers of the user is also are tracked to allow the user to use his fingers in the air to control the GIS system (Google Earth in particular). This kind of interaction is seen in the movie Minority Report. The concept is explained in detail by the video on the web[1]. Although it is a fun and futuristic way of interaction, using your fingers in the air is unfortunately a tiring way of interaction to use in production environment. A wii remote, an infra red LED array, and some reflective tape will be used, so again we are building a VR system as cheap as possible integrated with open source software.

Today's desktop computers have enough computation power to provide real-time feedback to the users. We will try to make a good use of the CPU in order to get good performance.

This project can also be implemented by an IR camera instead of Nintendo Wii Remote. The image output of the IR camera can be captured by the computer and using a hot spot detection algorithm, we can easily find the positions of the fingers. The downside of this method is standard cameras captures 25 or 30 frames per second, yet Nintendo Wii has a 100 Hz output rate, which makes it far more precise than cameras. In addition to its high frame rate, since the detection is performed on-board the wii-remote, the CPU on the computer is idle and available for our usage. The downside of using wii-remote is that wii-remote can only send four points through blue-tooth connection to the computer, which is enough for us right now. But if we wanted to track 10 fingers it wouldn't be possible with wii-remote. For projects with such requirements, an IR camera would be a better solution.

We made some experiments with the IR camera, and the wii remote. We experimented with the sensor bar of the Nintendo Wii and an IR led array in the IR camera. After capturing video and viewing them on the television, we saw that the Nintendo Wii's sensor bar are far more powerful than a standard IR camera's IR leds. Initial tests performed with the reflective tapes reflected enough infra red light, which made us optimistic that wii remote would produce good results in tracking head and fingers' positions. Unfortunately wii-remote has a very short range when we use the led array of the IR camera. Wii remote only detects really bright light, so instead of using the IR camera's IR led array, we decided to use another piece of the Nintendo Wii, the sensor bar (with IR LEDs in it), which is sold with the Wii itself.

The rest of this work is organized as follows: Section 2 gives information about related projects, Section 3 provides the details of the proposed system and Section 4 presents the results of the experiments. Section 5 discusses about possible future work and finally Section 6 summarizes and concludes the study.

II. Related Work

A. Visualizing the Future of Virtual Reality[3]

Virtual reality systems have been used for more than 20 years in science fiction movies, immersive video games and television shows. However the primary purpose of VR systems is to aid engineers and scientists in visualizing the systems which they are analyzing or designing. Viewing objects in three dimensions and using intelligent control mechanisms, enable users to fulfill tasks in an efficient and an accurate way.

As a broad range of VR systems available in the market, from very complex and expensive systems such as The Cave Automatic Virtual Environment (CAVE with extremely high price tags up to US $10 million), to new generation of very powerful VR systems such as GeoWall (www.geowall.org), VizTek's P1 Virtual Wall (www.viz-tek.com), and the VisBox (www.visbox.com) which provide a substantial fraction of the CAVE's performance, but at less than 1 percent of the price.

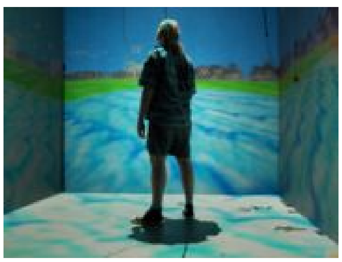

Fig. 1. A snapshot of the CAVE

Linux, as a synonym of software development, and these new VR systems have naturally adopted an open philosophy as well. This attitude toward software development has led to the distribution and continuous improvement of VR software libraries, among the most important of which are FreeVR (www.freevr.org), VR Juggler (www.vrjuggler.org), and OpenGL (www.opengl.org). The use of such standardized libraries lets the open-source community take advantage of cross-platform code transfer, code reuse, and cooperative development on different systems.

VisBox: This computer takes advantage of well-known computer vision and optical-sensing algorithms to determine the location and orientation of the user's head in the 3D space in front of the screen. It then uses this information to adjust the projected images, accordingly. This lets the user walk into a projection or look under a device simply by moving around in front of the screen. It is very difficult to explain the power of head tracking and how much more immersive and real images seem with it. However, we recommend that you not invest in a system without head tracking until you have tried one with it. Visbox and it's important parts are represented in Figures 2 3 and 4. As you can see in the Figure 3 retro reflective spheres are used on each side of the glasses, so that where ever the camera is the IR light coming from the IR LEDs will reflect to the camera. The powerful IR-LED arrays (seen in Figure 4) ensure good range for usage. Visbox can be seen in action in Figure 5

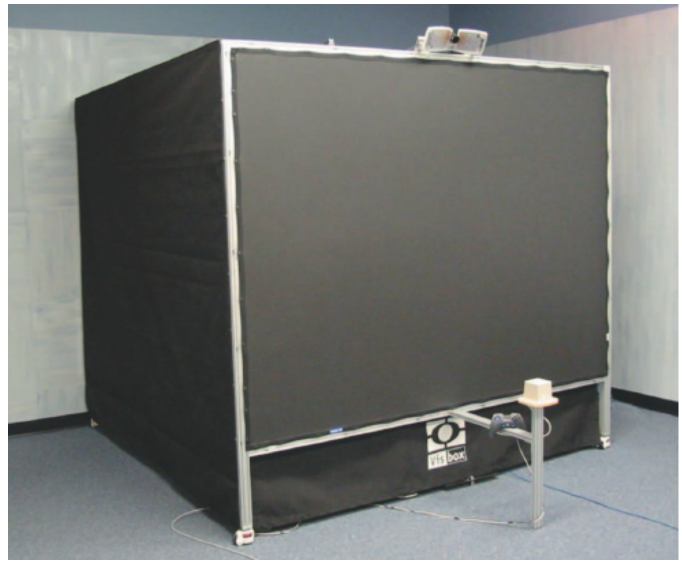

Fig. 2. VisBox virtual reality system. The unit ts into an eight-foot cube.

Fig. 3. VisBox's glasses with retro reflective spheres for head tracking

B. Swordplay: Innovating Game Development Through VR[4]

Video gaming technology revolved over the past three decades from simple 2D no-sound gray-scale games to 3D 5.1 surround systems 16 million color games. Video games have been an important driving force for the advancement of real-time computer graphics, sound, and, more recently, artificial intelligence. For example, with each new generation of console game hardware, the graphics and sound provide greater realism and the AI is more formidable. Unfortunately, until recently the same arguments cannot be made when it comes to the advancement of user interaction and game play. The best example for new generation of input devices is Nintendo's Wii console controller (the Wiimote).

C. TAVERNS: Visualization and Manipulation of GIS data in 3D Large Screen Immersive Environments[5]

Currently there are many software packages that allow analysis of Geographic Information Systems (GIS) data. However, these software packages limit user interaction and exploration

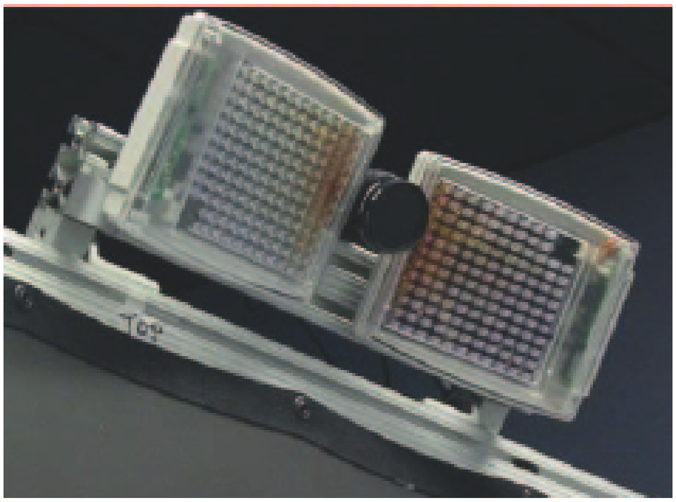

Fig. 4. VisBox's camera and infrared-LED array for head tracking.

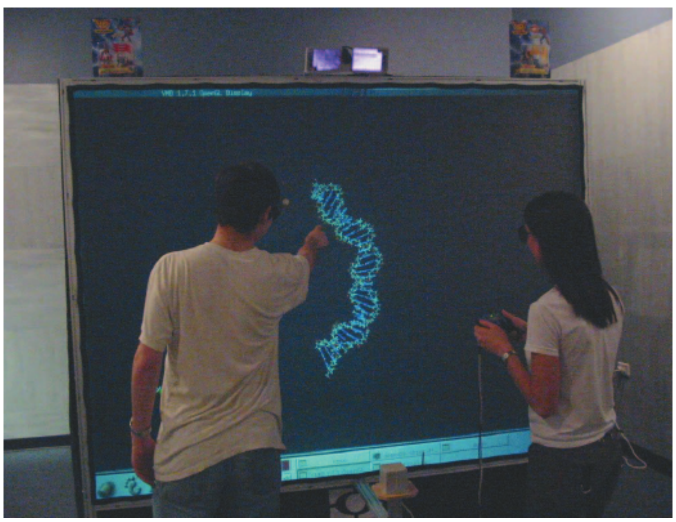

Fig. 5. The VisBox in action. Two Valparaiso University students view a DNA molecule using the VisBox.

of the GIS data to 2-dimensional (2D) interface. Virtual Reality (VR) technology provides a great way for manipulating and visualizing 3-dimensional (3D) data in three dimensions.

The creation of realistic training environments is crucial to enable the military to train its personnel in computer-simulated environments. The fields of computer graphics and visualization are the crucial components in the creation of realistic training environments. Although advanced visualization techniques help to establish realism, a realistic recreation of real world geographic data is also necessary to achieve a fully immersive and realistic simulation environment.

A GIS is a computer system used for capturing, analyzing, and displaying geographically referenced information. In short, GIS revolves around the creation and management of spatial data identified according to location. It is a computer system that allows users to integrate, store, edit, analyze and display geographic data. Some of the practical applications of

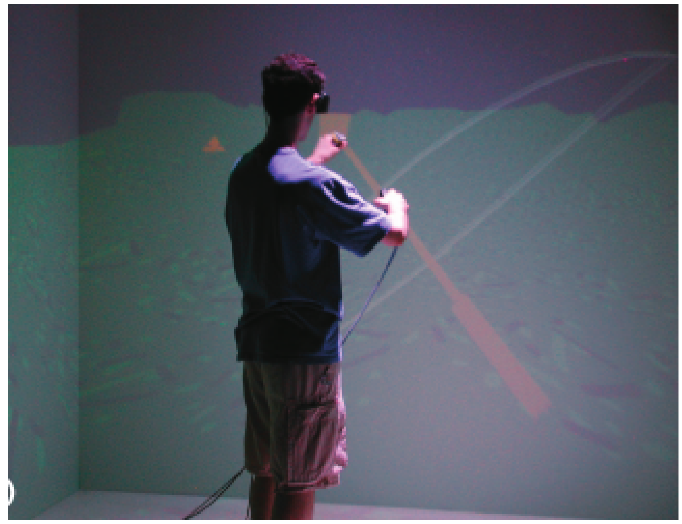

Fig. 6. The game Swordplay: A player prepares to fire an arrow with a bow.

GIS are scientific investigations and military training.

VR provides a medium composed of immersive interactive computer simulations that provide real-time feedback to the users. VR is a technology that can provide sophisticated real time 3D user interface for users to interact with 3D applications. Therefore, VR technology is a good candidate as a user interface for interaction with 3D GIS data.

III. Tracking Based GoogleEarth Control

There are two options in our system for controlling the GoogleEarth application (using head tracking or finger tracking). In both of the setups some additional items are required besides the wiimote. First of all what we need is a IR light source. For this we have many options, we can use the original sensor bar of Nintendo Wii or build an IR LED array as in [1], or even a couple of candles will be enough. And for reflecting light we can use reflective tape or any other reflective material. In addition to GoogleEarthAirlines library for embedding GoogleEarth in a .NET application as an activex control we use the wiimotelib provided by [6]. This library provides positions of the four points tracked by the wiimote and all possible output values.

As as result we want to be able to use Google Earth witouth touching the keyboard or mouse. We want to be able to turn the world, go south, west,... and zoom in or zoom out. So the image in front of you goes east when you move your head to the right. It moves south when you move down. Also you can zoom in or out by getting close to the screen or going far away from the screen.

Normally Nintendo Wii has a range of approximately 3 meters, but we need to capture the reflections of the fingers because we do not have IR leds on our fingers, so our range gets divided by 2, which is a close range, but generally, that is enough for using a computer. Range can be increased by using better IR leds, or using IR camera's for better detection, rather than using wii remote's detection.

Algorithm 1 Head tracking control for Wiimote.

Require: Height screen is known.
Require: Standard wiimote with $45\deg$ field of view is used.

repeat
    $p1 \leftarrow$first detected point by wiimote
    $p2 \leftarrow$second detected point by wiimote
    $dx \leftarrow p1.x - p2.x$
    $dy \leftarrow p1.y - p2.y$
    $avgx \leftarrow (p1.x + p2.x)/2$
    $avgy \leftarrow (p1.y + p2.y)/2$
    $pointDist \leftarrow \sqrt{dx^2 + dy^2}$
    $screenX \leftarrow (avgx - 512.0)/512.0$
    $screenY \leftarrow (avgy - 384.0)/384.0$ {$screenX$ and $screenY$ normalized screen coordinates with respect to GoogleEarth's coordinate system whose origin is window center .}
    $angle = \frac{\pi/4}{1024} * \frac{pointDist}{2}$
    $headDist = \frac{IRdotDistanceInMM/2}{\tan angle * screenHeightinMM}$;
    $worldCoord \leftarrow GE . FromScreenCoords ToWorldCoords()$ {GE stands for Google Earth}
    $camParam \leftarrow GE.CamParams()$
    $camParam.range* = headDist$
    $GE.CamParam \leftarrow (worldCrd.Lat, worldCrd .Long, camParams.range)$ {GE Camera Parameters are assigned by Latitude Longitude and range which are computed above.}
until program termination

In order to integrate the tracking software module we use a web plug-in[7] for GoogleEarth. The plug-in wraps the most of the functionality provided by GoogleEarth user interface, provides interface function in java script. With this plug-in we can easily embed the activex component wrapping the GoogleEarth and interact with the application in a smooth and reliable way. We also consider using GoogleMaps for spatial data display, yet since GoogleEarth can work without internet connection we preferred to use GoogleEarth. If internet connection is not a problem then using GoogleMaps is better to use, since its API is officially provided by Google and you do not need any 3rd party plug-in to interact with the application.

Control with finger tracking is very much similar to the control with head tracking. The computation of normalized screen coordinates for the found point by the wiimote is exactly the same as in 1. If only one IR dot is acquired by the sensor then a transformation in xy plane is performed. However if two dots are detected by the sensor then either zoom in or zoom out operation is performed according to the previous position of the detected dots. If the previous distance between the dots is larger than the current distance, a zoom out is performed, otherwise zoom in operation is performed.

A. Hardware

Several different hardware pieces are needed for this project. Using better material improves range and quality of tracking. We need the IR camera to detect the position of the head. There are 2 types of camera that we can use:

1) Nintendo Wii-mote (Figure 7)
2) A Regular IR Camera (Figure 8)

Fig. 7. Nintendo Wii-mote

Fig. 8. Infrared Camera

As mentioned above, wii-mote has 100 Hz input with 1024*768 resolution, which is really good. But the downside of it, it only detects 4 point, and we cannot change the code of point detection. Unfortunately, wii detects only strong infrared sources, because they do not want any other IR source to interfere. This is the main problem of our implementation, we are limited with the sources that wiimote detects.

On the other hand, a regular IR camera can be used. We also bought one of them, and made some experiments, but we didn't re-write a point detection algorithm. Also we would need to integrate the video input of the camera, with the detection algorithm, but with wiimote, we just connected wiimote over blue tooth, and fetched the coordinates of the detected points via wiimote library. This was the easier solution, with a higher frequency and resolution.

Then we need IR sources so the camera can detect the reflections of head and fingers. Here is a list of IR Sources:

1) Nintendo Wii Sensor Bar (Figure 9)
2) Color Infrared Camera (Figure 8)
3) IR LED (Figure 10 & 11)

Fig. 9. Nintendo Wii Sensor Bar

Fig. 10. Little IR LEDs

Fig. 11. The IR LED array that Johnny Lee used in his experiments [1]

Also some of these sources are portable, so we can put them on our head, or hold them like pens, and write or draw in the air, so the IR camera can detect them, we can use the light sources like our fingers. Or we can use them just as IR light sources, and use some reflective material on our heads and fingers, so the camera will detect the reflected light. As you can see in Figure 13, Johnny Lee also used this method on his fingers[1].

Here is a list of reflective material:

1) Reflective Tape (Figure 12)
2) Aluminum Foil.
3) Metal (For example a ring)
4) Your actual fingers

Fig. 12. Reflective tape

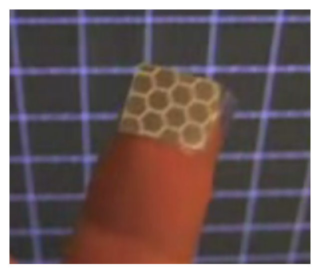

Fig. 13. Reflective tape placed on Johnny Lee's finger [1]

If you use your actual fingers, it will reflect the IR light with a range of approximately 10 cm. We use the other materials to increase the range, so that we don't have to be sitting next to the source and camera. But even with the best of these reflectives the range is halved, because the light comes from the source, gets reflected, and reaches the camera. So we are looking for ways to put the IR sources on the head or fingers, so we will have a longer range. IR LEDs(Figure 10) are practical for this purpose. You can attach them where ever you like. You can have glasses that have attached IR LEDs, as you can see in Figure 14. Unfortunately we were not able to find these IR LEDs in Turkey.

A cheaper solution would be to mount Nintendo Wii sensor bar on a cap like in Figure 15. It works good but it is not practical, because you will either have a cord connecting your head to Actual Wii, or you need to buy a wireless sensor bar, which is not the default sensor bar, so this means extra cost for you. Also this item is not a common item. We couldn't find one in the stores in Turkey. We made our experiments by holding the sensor bar under our head, and it wasn't a big problem for testing, but no one would like to use it to play games or use Google Earth, so this method is only usable for testing.

Fig. 14. Glasses with IR LEDs on it.[1]

Fig. 15. Nintendo Wii Sensor bar, located on a cap by Johnny Lee. [1]

IV. Experimental Evaluation

The users which had previous experience with Google Earth adapted faster than the ones which had no experience. Most of the users wanted smoother movement and zooming abilities. These abilities are also lacking in Google Earth. Also a switching mechanism between the tracking methods was mentioned, but it is not an important problem.

Google Earth is a slow application to test our method. We have an input of 100 Hz, but google earth cannot respond with that speed, and we don't have any chance to change anything about Google Earth. Choosing an open source project for head and finger tracking would be a wiser choice. So the overall experience was not an immersive experience, but what ever the interface is Google Earth is not an immersive program, so in order to test our project, combining the project with a faster and more immersive program would give better results.

V. Future Work

The project is fully completed, but it can be improved in several ways. Although tracking only the head or only fingers is enough to implement all the necessary user interfaces, head and fingers can be tracked at the same time. Tracking both of them at the same time will cause new problems. Using two wii-motes will help solving these problems. Also using two wii-motes can enhance the tracking quality in tracking only head or fingers.

Using a better IR source will enhance the wii-mote input drastically. We were not able to build an IR source as good as Johnny Lee demonstrated in the video[2]. We couldn't find the necessary parts in the stores in Turkey. The IR source he used was also hand made, which was specifically designed for this kind of usage.

Another enhancement will be smoother movement in Google Earth, which was asked by the test users. This can be done by optimizing the parameters in our algorithm. Test users also demanded to make a switch between the tracking methods.

Best improvement will be to integrate the user interface with a faster and more immersive program to measure the immersiveness of the users more precisely.

Another version of our system can be implemented with a regular IR camera, with our own point detection algorithm, so that we would have more control over the system, and we wouldn't be dependent on the wiimote. Most of the cameras have output of 25 or 30 frames per second, but we will have the ability to detect the points produced by low quality products, so we will have a cheaper system.

VI. Conclusion

This project is a re-implementation and combination of the methods developed by Johnny Chung Lee. Although the implementation was successful, we saw that the hardware quality affects the quality of the outcome, and it is not easy to find good IR equipment in our country. With better hardware items, we can improve the range of our system. An IR LED array like Johnny Lee used (in Figure 11) would definitely increase our range on finger tracking, and glasses with IR Leds would also eliminate most of the errors we are struggling with in head tracking. The IR LED array could also be used for head tracking.

The system has the potential for an immersive ambiance, but the program we used for integration is not an immersive application. Google Earth is not immersive even when you use it with keyboard and mouse, on a fast computer, because, it needs to communicate with google servers, and if you want to observe different places, you have to wait for the data to be streamed from the servers. Although Google Earth can run without internet connection, but even then it is not a fast responding program. Our system has an upper limit of 100 Hz, which is the output limit of Nintendo Wiimote, but we cannot test this upper limit, because Google Earth cannot respond even near these speed limits.

In the end, we couldn't find the immersiveness sense that we were hoping for, but the user interface is completed as expected. It works like Google Earth. The parameters can be improved for better user satisfaction.

REFERENCES

[1] J. C. Lee, "Tracking fingers with the wii remote," 2007. [Online]. Available: http://www.youtube.com/watch?v=0awjPUkBXOU

[2] ——, "Head tracking for desktop vr displays using the wiiremote," 2007. [Online]. Available: http://www.youtube.com/watch?v=Jd3-eiid-Uw

[3] D. Tougaw and J. Will., "Visualizing the future of virtual reality," *Computing in Science and Engineering*, vol. 05, no. 4, pp. 8–11, 2003.

[4] M. Katzourin, D. Ignatoff, L. Quirk, J. J. L. Jr., and O. C. Jenkins, "Swordplay: Innovating game development through vr," *IEEE Comput. Graph. Appl.*, vol. 26, no. 6, pp. 15–19, 2006.

[5] S. Su, W. Sherman, F. C. H. Jr., and M. Dye, "Taverns: Visualization and manipulation of gis data in 3d large screen immersive environments," in *ICAT Workshops*, 2006, pp. 656–661.

[6] "Wiimotelib: A managed library for using the wii remote from .net." [Online]. Available: http://code.google.com/p/giimote/wiki/WiimoteLib

[7] "Google earth airlines - web plugin for google earth." [Online]. Available: http://www.googleearthairlines.com