

An Introduction to 3D User Interface for Operating System

Vaishali Jadhav

Professor, Department of Computer Engineering, Babasaheb Gawde Institute of Technology, Mumbai, Maharashtra, India

Abstract: *There has been considerable advancements, growth and development in the 3D displays/screens/monitors or 3D capable displays/screens/monitors. This emergence has caused a rise to many 3D - movies, games and virtual environments in very myriad quantity. Now these systems also open a large range of possibilities in the way an individual interacts with the computer. The primary operating systems of the traditional personal computer does not make use of the 3D capabilities of the display devices. This paper talks about how the traditional personal computers interacting methods can be revised and a 3D User Interface environment can be created for the user to interact and make the optimum use of the computer. The denouement of this innovation can be used in the future in various fields like medical, education and also for entertainment purpose.*

Keywords: Anaglyph 3D, Stereoscopic Images, Augmented Reality, Gesture Recognition, Shallow-depth 3D.

1. Introduction

To provide an intuitive natural interface for users to interact with computers, through gestures and operating directly on perceived 3D objects on the 3D display to drive the operating system. A common interface must be provided for platform-independency and hardware abstraction. Users must be able to use the 3D UI interface on different OS's using technology they prefer to implement the 3D output and gesture recognition. A new manner of interacting with windows and controls must be introduced, suited particularly to 3D user interface, such as more natural gestures for actions such as minimizing, restoring and closing and resizing windows. The system will be primarily developed and tested on a Linux based operating system, as it is open source technology as easy to patch. However, the platform-dependent modules and the common API's will be separate such that porting to a different operating system is easy and convenient.

The system should be portable to various platforms with appropriate 3D display drivers and sensor drivers. The API to interface with the 3D UI system should be common on all the platforms. The system should be independent on the 3D display technology or the sensors used. Therefore to enrich the interface to the user this paper presents a shallow-depth 3D environment with a limited depth. Within the shallow-depth 3D environment several common interaction methods need to be reconsidered.

2. Current/Existing System

In an anaglyph, the two images are superimposed in an additive light setting through two filters, one red and one cyan. In a subtractive light setting, the two images are printed in the same complementary colors on white paper. Glasses with colored filters in each eye separate the appropriate images by canceling the filter color out and rendering the complementary color black. A compensating technique, commonly known as Ana-chrome, uses a slightly more transparent cyan filter in the patented glasses associated with



Figure 1: Nvidia 3D glasses used for Anaglyph 3D

the technique. Process reconfigures the typical anaglyph image to have less parallax. An alternative to the usual red and cyan filter system of anaglyph is Color-Code 3-D, a patented anaglyph system which was invented in order to present an anaglyph image in conjunction with the NTSC television standard, in which the red channel is often compromised. Color-Code uses the complementary colors of yellow and dark blue on-screen, and the colors of the glasses' lenses are amber and dark blue.

- Red Color: for the left eye
- Cyan Color: for right eye

To present a stereoscopic picture, two images are projected superimposed onto the same screen through different polarizing filters. The viewer wears eyeglasses which also contains a pair of polarizing filters oriented differently (clockwise/counterclockwise with circular polarization or at 90 degree angles, usually 45 and 135 degrees, with linear polarization). As each filter passes only that light which is similarly polarized and blocks the light polarized differently, each eye sees a different image. This is used to produce a three-dimensional effect by projecting the same scene into both eyes, but depicted from slightly different perspectives. Additionally, since both lenses have the same color, people

with one dominant eye (amblyopia), where one eye is used more, are able to see the 3D effect, previously negated by the separation of the two colors. Circular polarization has an advantage over linear polarization, in that the viewer does not need to have their head upright and aligned with the screen for the polarization to work properly. With linear polarization, turning the glasses sideways causes the filters to go out of alignment with the screen filters causing the image to fade and for each eye to see the opposite frame more easily. For circular polarization, the polarizing effect works regardless of how the viewer's head is aligned with the screen such as tilted sideways, or even upside down. The left eye will still only see the image intended for it, and vice versa, without fading or crosstalk. All types of polarization will result in a darkening of the displayed image and poorer contrast compared to non-3D images. Light from lamps is normally emitted as a random collection of polarizations, while a polarization filter only passes a fraction of the light. As a result the screen image is darker. This darkening can be compensated by increasing the brightness of the projector light source. If the initial polarization filter is inserted between the lamp and the image generation element, the light intensity striking the image element is not any higher than normal without the polarizing filter, and overall image contrast transmitted to the screen is not affected.

An eclipse method, a shutter blocks light from each appropriate eye when the converse eye's image is projected on the screen. The display alternates between left and right images, and opens and closes the shutters in the glasses or viewer in synchronization with the images on the screen. This was the basis of the Televue system which was used briefly in 1922. A variation on the eclipse method is used in LCD shutter glasses. Glasses containing liquid crystal that will let light through in synchronization with the images on the cinema, television or computer screen, using the concept of alternate-frame sequencing. This is the method used by nVidia, XpanD 3D, and earlier IMAX systems. A drawback of this method is the need for each person viewing to wear expensive, electronic glasses that must be synchronized with the display system using a wireless signal or attached wire. The shutter-glasses are heavier than most polarized glasses, though lighter models are no heavier than some sunglasses or deluxe polarized glasses. However these systems do not require a silver screen for projected images. Liquid crystal light valves work by rotating light between two polarizing filters. Due to these internal polarizers, LCD shutter-glasses darken the display image of any LCD, plasma, or projector image source, which has the result that images appear dimmer and contrast is lower than for normal non-3D viewing. This is not necessarily a usage problem; for some types of displays which are already very bright with poor grayish black levels, LCD shutter glasses may actually improve the image quality.

3. Proposed System

The system helps to take advantage of the 3D display technology to display the operating system in 3D. It keeps track of the depths of windows to display. This tracking includes lower depth for recently active windows and higher

for least recently activated windows and depths to window controls. This system helps to interact with the 3D objects by touching the objects in 3D space and no touchscreen required. It will detect 3D movement and recognition of hands and gestures and also handle common window tasks such as Minimize, Restore, and closing using gestures. This system will prepare a platform for 3D entertainment making possible better interactions for 3D Games with 3D inputs, 3D Movies, 3D Movies with user interaction.

The system make you feel that you are the only viewing Operating System in 3D space but also interacting with it in 3D space.

You must have seen science fiction movies, such as Minority Report, Iron Man, Star Wars and yada yada. The futuristic computers and the technology you see in such movies, have one thing in common, when it comes to the user interface. Everything is 3D, and windows are objects – that you can hold, move around and throw around. The goal of this paper is to make this form of HCI a reality, and it really isn't all that hard to do so. For this, a specialized hardware – known as a depth sensor (for example, Kinect) is used. A PrimeSense Carmine 1.09 will be used.

4. Computation of 3D Technology

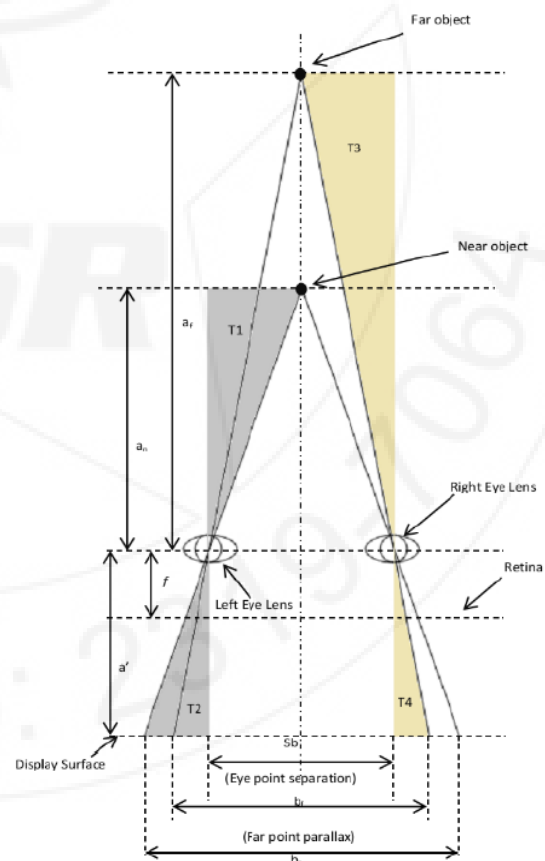


Figure 2: Mathematical relationship of stereo base (S_b) to the stereoscopic eye points, stereoscopic object space and the 3D stereo display surface.

The integral of viewing 3D content on a display screen or display system is the appropriate offset of two images in each

stereo pair. Therefore in order to generate these 3D stereo pairs with adequate parallax or the effect whereby the position or direction of an object appears to differ when viewed from different positions, e.g. through the viewfinder and the lens of a camera. To generate these parallax a representative sense of depth and the viewpoints of each images to be positioned in such a way that they will ultimately be at an appropriate distance from the projection plane, the assumed viewing distance. The distance far from the objects and the near objects will be computed automatically for each scene and will be adjusted to insure that there are no drastic changes in scene geometry from once scene to the next. Nothing is accepted in this world unless and until you prove it mathematically, therefore the mathematical equation of this computation is,

$$b_n = \left[\frac{a' x Sb}{a_n} \right] + Sb$$

$$b_f = \left[\frac{a' x Sb}{a_f} \right] + Sb$$

$$\frac{(b_f - b_n)}{2} = 2 * a \tan \left[\frac{Sb / 2}{a_n - a'} \right]$$

$$Sb = d \left[\frac{a_f x a_n}{a_f - a_n} \right] \frac{l}{a'}$$

Where:

T1 and T2 are similar triangles

T3 and T4 are similar triangles

f = ~22.22 mm (lens to retina distance)

a' = viewing distance based on screen size

a_n = a' for zero parallax

$\frac{(b_f - b_n)}{2}$ = left/right eye image offset

a_f Derived from current source content

5. Modus-Operandi of Gesture Recognition

Gesture recognition is a topic in computer science and language technology with the goal of interpreting human gestures via mathematical algorithms. Gestures can originate from any bodily motion or state but commonly originate from the face or hand. Current focuses in the field include emotion recognition from the face and hand gesture recognition. Many approaches have been made using cameras and computer vision algorithms to interpret sign language. However, the identification and recognition of posture, gait, proxemics, and human behaviors is also the subject of gesture recognition techniques. Gesture recognition can be seen as a way for computers to begin to understand human body language, thus building a richer bridge between machines and humans than primitive text user interfaces or

even GUIs (graphical user interfaces), which still limit the majority of input to keyboard and mouse. Gesture recognition enables humans to communicate with the machine (HMI) and interact naturally without any mechanical devices. Using the concept of gesture recognition, it is possible to point a finger at the computer screen so that the cursor will move accordingly. This could potentially make conventional input devices such as mouse, keyboards and even touch-screens redundant. Gesture recognition can be conducted with techniques from computer vision and image processing.

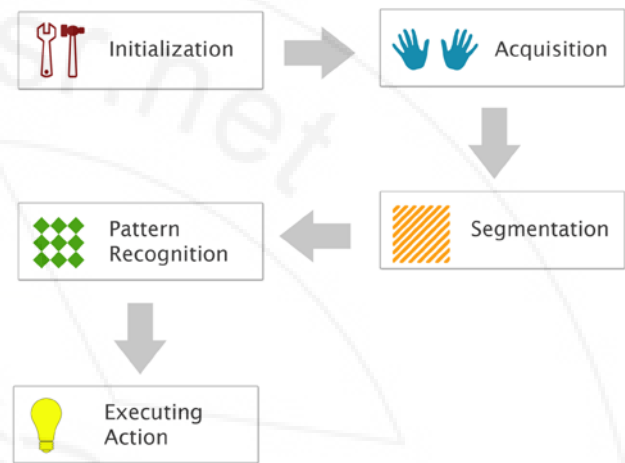


Figure 3: Process of Gesture Recognition

Gesture recognition system is based on two important parameters that is user hand posture and this posture location. One of the cheapest method of implementing this would be using a simple USB web cam. But the computer lacks the ability of a human eye and hence much more lucid algorithm need to be designed using OpenCV to make the computer see and differentiate between the hand and the other objects in the images and also to detect the various fingers/end points of the hand.

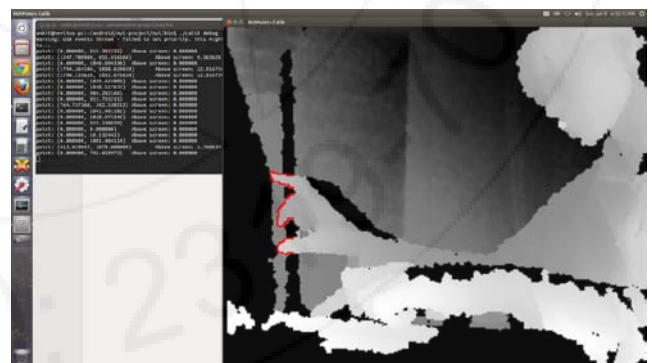


Figure 4: Gesture detection by 3D depth sensor

The modus-operandi or the process of Gesture Recognition is illustrated in the figure-3 above. The following processes compose and explain the working architecture of this mechanism.

1) **Initialization:** This is the first and the basic stage wherein the postures are stored in the virtual memory which is created in a start-up step. In order to configure this memory, different ways are proposed.

- 2) **Acquisition:** This is an image capturing step done by the camera.
- 3) **Segmentation:** When the camera captures the image it is necessary to disintegrate the captured image by separating each frame and processing it separately and making it ready for analysis.
- 4) **Patter Recognition:** Once the user's hand has been segmented, its posture is computed with those stored in the system's memory. Sometimes an innovative approach would be using Hausdroff matching algorithm approach to speed-up the computation and get the result in less amount of time.
- 5) **Executing Action:** Thus, the system carries out the corresponding action according to the recognized hand posture.

6. Conclusion

The implementation of the proposed system can not only give 3D views but also gives a new method to interact with the personal computer in a 3D environment, more precisely it can be called a 3D interaction or a 3D user interface to use the personal computer. Generally most algorithms usually ignore the problem of barricade due to the associated computational complexity. But however, recently approximate algorithms for similar problems in computer graphics have been proposed. Finally, many possibilities have been touched and many areas are still and far from the sight of the research which will lead to further advancements of this technique in the coming years.

References

- [1] Mark Hancock, Sheelagh Carpendale, "Shallow-Depth 3D Interaction: Design and Evaluation of One-, Two- and Three-Touch Techniques," ACM Press, 2007.
- [2] Burdea, G., "Force and Touch Feedback for Virtual Reality," Wiley Interscience, 1996.
- [3] Elena Sanchez-Nielsen, Lius Anton-Canalis, Mario Hernandez-Tejera, "Hand Gesture Recognition for Human-Machine Interaction," Journal of WSCG, 2003.
- [4] Barry Bitters, "Advances in Desktop 3D-Stereoscopic Visualizations of Geospatial Data," ISPRS Technical Commission IV & AutoCarto, 2000.
- [5] Watsen K., Darken R., "A Handheld Computer as an Interaction Device to a Virtual Environment," Third Immersive Projection Technology, 1999.
- [6] Sushmita Mirta, Tinku Acharya, "Gesture Recognition: A Survey," IEEE, 2007.
- [7] Liang J., Green M., "JDCAD: A Highly Interactive 3D Modelling System," Computer & Graphics, 1994.

Author Profile



Prof. Vaishali Jadhav, is a lecturer in Maratha Mandir's Babasaheb Gawde Institute of Technology, Mumbai. She completed her under-graduation in Information Technology from Government College of Engineering, Karad. Her fields of interest include Artificial Intelligence, Machine Learning, Algorithms and she is passionate about Augmented Reality, Computer Vision and Human Computer Interface (HCI).