Table of Contents

Chapter 1

Introduction: Virtual Reality, Problems, and Solution Models

The concept of virtual reality has been explored for a long time and this is a field of research which involves many aspects. Virtual reality encompasses everything we know about the world, trying to create a virtual copy of reality. A main purpose among researchers in this field is to find new and more natural ways of interacting and displaying virtual environments.

This thesis contributes to the field of virtual reality, by proposing a system for interaction on a workbench display system. Furthermore a comparison of different stereo visualization techniques (passive anaglyph and active stereo) is proposed.

This thesis content is outlined in the following chapters.

1. **Introduction: Virtual Reality, Problems, and Solution Models:** This chapter provides an overview of the term and technology of virtual reality. The main problem of virtual reality tracking and stereoscopic

visualization is discussed. The approach proposed in this thesis is then introduced, which represents a solution to some of the problems.

2. **Tracking Systems and 3D Visualization:** In this chapter different techniques for tracking position and for 3D stereo visualization are presented. The problems inherent to different techniques are also discussed.

3. **Workbench Developments and Interaction Methods:** This chapter provides a brief overview of selected techniques and developments in the fields of workbench display systems and tracking (state of the art). Each technique and development is followed by a comparison to the proposed system.

4. **The Proposed Investigation:** This chapter presents the core idea and argumentation. The outlines of the experimentation and the research development plan are also introduced.

5. **Experimentation and Discussion:** In this chapter the development of the workbench and the proposed tracking system are presented. The systems are then tested and discussed.

6. **Conclusion and Future Work:** This chapter summarizes the main characteristics of the proposed system development to the field. Furthermore suggestions for future research are discussed.

1.1 Virtual Reality Systems

Although the basic concept of virtual reality (VR) is quite common and it has been known for some time, it is important to understand the definition of virtual reality. First of all, let us define what comes closest to us, i.e. reality. Looking at a dictionary, *reality* points to all things that are actual and in existence, giving a presence of being. The word virtual is defined by not being real or physical. These definitions can be summed into the following definition of virtual reality:

> *"Immersion of one or more individuals in a virtual environment, with the aim of achieving the illusion that they [the user(s)] are in a place, time, or situation different from their actual real-world location and/or time"*[17]

But virtual reality has also taken other definitions in present time, relating more to the combination of computers and interface devices (glasses, gloves, tracker, etc.) that give the illusion of being in the virtual world. Consider also the notion of immersion, or the illusion of being in the virtual world, is not always a prerequisite for virtual reality. In fact, VR also covers fields that do not require full immersion, but instead provide a view of objects without those actually exist. The above concept includes the research field which this thesis will explore.

Virtual reality was born when computers where first introduced. The US American army research facilities, who had been doing crude simulations of combat fields with models, adopted the computer. The computer allowed for faster processing of data which gave light to the early stages of virtual simulations. Researches realized the unique potential of being able to test and experiment with models for airflow, etc. Today the technology is being used in many different fields (concept visualization, data mining and simulation) from research to software development (Moeslund el at. 2000:1) [16]. Virtual reality has evolved, and with it displays systems. From large and expensive

panoramic and interactive setups (e.g. 6-side CAVE, Panorama, Powerwalls, etc.) and wearable Head Mounted Display systems, to more affordable and modular solutions.
An example of the latter is represented by previous work of the author of this thesis. In particular the semester 9 project [13], and a publication at the Eurographics IC conference last February [8].

We can think about dividing VR related display systems into the following categories:

- **Non-immersive virtual reality system:** In a non- immersive virtual reality system the border between the display surface and the virtual world is clearly visible. The virtual world is regarded as by a window, therefore these systems are also called Window on World (WoW) or Desktop VR. Many of these systems use a conventional computer monitor for the visual feedback, but also workbench type display systems fall under this category. A non-immersive display system is in many cases also preferred when one or more users view and collaborate around virtual objects.

- **Immersive Virtual reality System:** In an immersive virtual reality system arrange the user is under the impression of being in an artificial environment. In order to keep the impression realistic, the user must be integrated as much as possible into the virtual world. Many immersive virtual reality systems may use the so-called Head Mounted Display (HMD) and the CAVE [2](see also Chapter 2.2.2).

- **Augmented Reality System:** In an augmented reality system the user sees both the real and the virtual world. Users normally wear a HMD with see-through glasses (see also Chapter 2.2.2), allowing the user to see a real world view with superimposed virtual objects or text. This technology is for example used by military pilots, who have additional information displayed on their helmets visor, but is now also showing up in the auto industry where information such as speed and direction, is displayed on to the windscreen.

When working with virtual environments (VE) it is important to know the position of the viewer in order to provide the correct perspective view of the environment. Although not essential for all applications and display systems, tracking of viewer position is adopted by systems such as the CAVE, Workbench and HMD. Tracking can be done optically or magnetic (see Chapter 2.1), the latter being the most applied. A tracker system provides and outputs the position and orientation of the object being tracked, e.g. the human head and hand.

Using a virtual reality system usually also gives the sensation of depth "traversing" the display surface. The perception of depth is possible by using special visualization systems which typically require the use of glasses made to provide users with the correct view for each eye. There are different 3D stereo visualization approaches, which can be categorized as *passive*, *active* or a*utostereoscopic* (see Chapter 2.2.1).

1.2 The Workbench System

The motivation for working with virtual reality and interaction comes from the author's great interest in the field and an interest improving performance of current VR systems. Furthermore, a workbench environment where various tasks can be performed and users can collaborate around virtual objects displayed on the surface, is great interest in many fields.

The research in visualization of virtual environments has given a wide variety of displays on which viewing 3D content. Those displays give different levels of immersion. The CAVE (Cruz-Neira el at. 1993) [2] is supposed to give best immersion performance. Although the CAVE encompasses some of the key point off the ultimate medium for virtual reality, it is however dependent on the most advanced computers and displays available. The CAVE as well as other such multi-projection displays are universal and try to accommodate multiple usages and applications (Krüger el at.) [14]. In addition to this users are constrained to using wired tracking devices to interact with the virtual environment.

The developers of the Responsive workbench (Krüger and Fröhlich 1994) [15] at GMD (German National Research Institute for Information Technology) and later in collaboration with TAN Projektionstechnologie, approached the design of the display by looking at it from the users point of view. So instead of designing a universal interface, the interface would be designed with distinct tasks in mind.

Through extensive user task analysis for different tasks performed by physicians, architects and automotive engineers, the developers found that most of their tasks where performed within a workbench scenario. Furthermore, the analysis concluded that physicians, architects and automotive engineers all had different areas of focus concerning which tasks that they wished to be able to perform on a workbench (Krüger el at.1995) [14]. This realization led them to conclude that most co-operative tasks rely on a workbench.

The analysis drove them to develop the “Responsive Workbench” as a test bed for different use case scenarios.

A workbench consists of a display surface where stereoscopic images are projected onto, allowing the user to see virtual objects standing on top of the workbench surface [Fig. 1.1]. The visualized stereoscopic images are viewed with active stereo technology based on shutter glasses. When collaborating around a task a “guide” operates the table while several observers can follow the event with additional shutter glasses.

A magnetic tracking system provides the correct perspective view, to the “guide” together with the position of an interaction device.

However if more than one user is viewing the workbench, it will only be the "guide" that receives the correct view, while for second users the view can "nearly" be correct if he/she is placed close to the guide, otherwise it is distorted. The above represents a typical problem for Workbench display systems. Tracking position of users and devices requires the users to wear tethered sensors. This is invasive and leads to un-natural interaction.

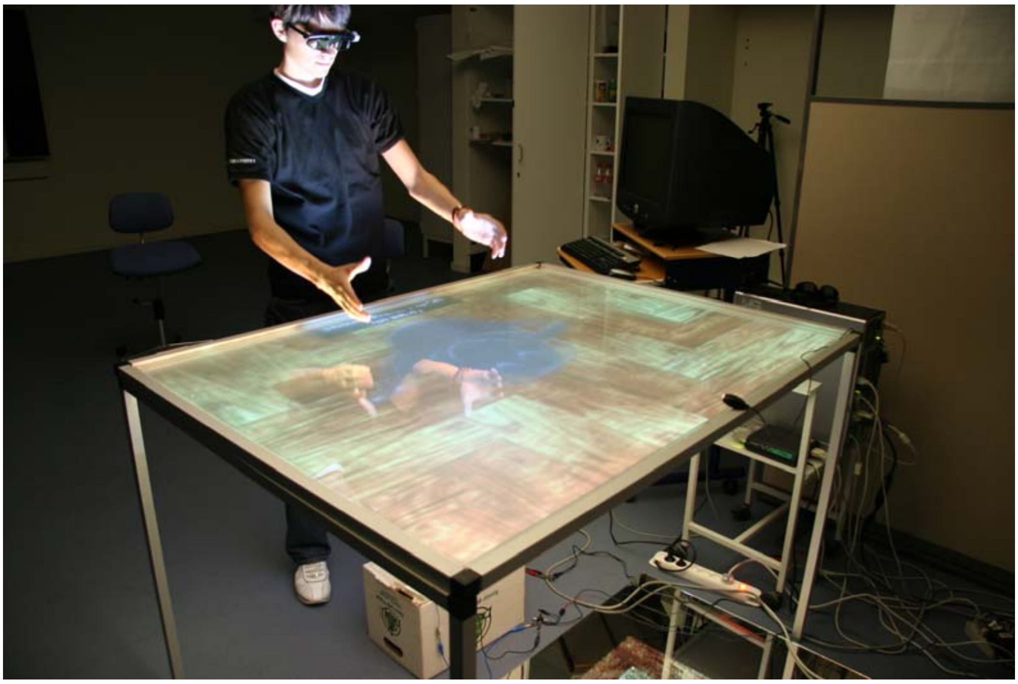


Figure 1.1 The Developed Workbench

1. 3 Problems

The investigation into virtual reality and has shown, that most virtual reality systems use tracking and 3D stereo visualization techniques. Although the technology is close to maturity, there still exist problems that need to be addressed. In particular in several cases the users have to sacrifice their experience due to invasive sensors and approximated stereoscopic visualization.

1. 3.1 Tracking

A tracking system must be used to obtain the users position. Most common is magnetic tracking, which can be very precise.

The magnetic tracking system produces a magnetic field to acquire the position of the sensors carried by the user. Unfortunately, the magnetic field can be influenced by metals in the nearby surroundings, resulting in an imprecise tracked position. Furthermore, the magnetic trackers can be invasive i.e. the user has to carry a wired sensor on his hand, expensive depending on the number of sensors and the range of the emitter.

1.3.2 3D Stereo

Most 3D stereo systems rely on the user wearing special glasses to perceive depth in an image. These systems can either be passive, active or autostereoscopic, but they all have some limitations. A common problem with all 3D stereo systems is crosstalk. The perception of crosstalk is known as ghosting. The effect of crosstalk, which is different from system to system, is caused when the image from one eye bleeds to the image of the other eye. This causes the user to see a dimmer version of one of the images superimposed on the other. In this context it is also worth emphasizing that ghosting can diminish the 3D effect, and lead to nausea.

Wearing special glasses to perceive depth can also be invasive. Furthermore, a prolonged viewing through glasses can result in eyestrain, (caused by the eyes not being able to adapt to stereoscopic imagery). Another factor leading to eyestrain is image misalignment.

Nevertheless, when choosing a 3D stereo system it is the cost of it, which directly bounds the technology adopted.

1.4 Proposed Investigation

The investigation into some of the known problems around tracking in virtual environments has led the author to propose a new system for tracking hand position in 3D environments. In addition, the analysis of different visualization technologies has led the author to consider a study of the proposed tracker with different approaches to stereo visualization.

1.4.1 New Concept Tracking

This thesis presents a method for tracking hand position on a workbench, based on computer vision techniques. A main advantage of a computer vision based tracking (optical tracking) is that the user does not have to carry a sensor in his/her hand, which gives the user more intuitive and natural interaction capabilities (avoiding the use of invasive wired sensors). The innovative aspect of the proposed system is a workbench where there is no need for wearing data gloves. The author of this thesis has not found any literature work where such a method is proposed.

The proposed workbench is a table sized display which surface consists of a ground glass plate, and where images are back projected. It is proposed an optical tracking system capable of:

1) detecting the reflection of one hand of the user captured by a camera.
2) tracking the hand while it is moving.

The proposed system is based on the use of infrared light. In particular, thanks to infrared light sources illuminating the display surface and to an infrared filter set in front of the camera lens the camera system can isolate the users hand from the rest of the image in the camera image plane.

1.4.2 Comparison of different 3D stereo techniques

Another objective of this thesis is to analyze the effect of using two different stereo visualization techniques on the proposed workbench system. In particular, the use of passive anaglyph and active stereo is proposed. The proposed analysis should possibly conclude about which of the methods is most effective for the workbench.

The two proposed approaches to stereo visualization will therefore be applied to our workbench system. The resulting characteristics will be assessed through experimentation.

The comparison of the passive anaglyph opposed to active shutter glasses, will also define the outlines for a more comprehensive study of the different systems as a part of further studies in future.

Please note that the different stereoscopic techniques proposed in the investigation have different level of affordability, which makes it relevant to assess system performance.

Chapter 2

Tracking Systems and 3D Stereo Visualization

This chapter presents different technologies for tracking user position in virtual environments and then it presents different technologies for 3D stereo visualization. In particular, the first section describes two among many most popular tracking methods. The second section describes how humans perceive depth and which kinds of systems have been proposed for replicating such experience when observing virtual worlds. At the end of each section there is a short summary of key points.

2.1 Tracking systems

Tracking is a broad term, but in our context it has to do with the tracking of motion. Motion tracking can be defined as the acquisition of position and orientation values (coordinates) of any moving object relative to a stationary reference. Among the two most common technologies, the magnetic and optical tracking.

2.1.1 Magnetic tracking

Magnetic tracking is the most common technology used for interfacing human (real objects) with a virtual environment. This system can capture 3D position (X, Y & Z coordinates) 3D orientation (Y, P, R coordinates corresponding to yaw, pitch, roll), of any real world objects. Real world objects that have 6 degrees of freedom (6-DOF) or less e.g. the head and hand can be tracked. Capturing all 6-DOF is central for realistic interaction with any virtual environment.

A typical magnetic tracker contains the following components:

1. a *transmitter*, which is usually fixed on the workspace ceiling or in the nearby surroundings.
2. one or more *sensors* (cabled to the interface device).
3. an *interface* device (often called the *filter*).
4. a *computer*.

In a 6–DOF magnetic tracking system, the transmitter consists of three coils on orthogonal X, Y, Z axis. An electrical field (either AC or DC) passes through each coil and creates a magnetic field in a desired direction. The sensor consists of a similar set of three coils, but they are passive. The transmitted field creates a current in the passive coil; the strength of the received electricity is directly proportional to the distance between the transmitter and sensor. Orientation coordinates are also acquired based on the strength of the received current. In particular, when the sensor is pointed away from the magnetic field its signal becomes weaker. Through mathematical calculations the different coordinates are distinguished from each other. In this process jittery readings are also filtrated away in the interface device [18].

The key benefit of using magnetic tracking is the precision of the resulting measure. Furthermore, the electromagnetic field can travel through humans, giving freedom of movement. Magnetic tracking systems are precise (up to 5 meters depending on the conditions) and low cost (typically 2500 euro for a 1 senor system).

The magnetic tracking method has also drawbacks which in some cases can influence the precision of the system leading to non-valid data acquisition, as well as causing discomfort to the user.

Among main problems are as follows:

- *Interference* with surrounding objects when these are composed by metals.
- *Invasiveness* of the tracker system i.e. the user has to carry a wired sensor on his hand, head, etc.
- *Latency*, i.e. delay of the object related to user movements.

When using a magnetic tracking system, the magnetic field produced can be distorted, due for example to metals close to the transmitter and sensors, resulting in a wrong measurement xxx to the system. In most workspace locations there is metal nearby to the system, centered in objects like chairs, tables and walls. The electro-magnetic field will be influenced by ferromagnetic substances (metals attracted by magnets; like iron and steel). The field distortion results in inaccurate values for position and orientation.
In conjunction the with field distortion the presence of nearby metal can also produce jitter. This can be seen when using head tracking and the view is “jumping” around resulting in nausea.

Besides this, jitter is often observed, when reaching the boundary of the electro-magnetic range. Although the filter compensates for jitter, it is impossible to get rid of at the edge of the magnetic field.

As previously mentioned, a magnetic tracker, acquires in most cases the object position from sensors cabled on the body. The cables can limit user interaction with the virtual environment.

Recent research in new wireless technologies has led to wireless sensors for magnetic trackers. Although rather new, the implication of having wireless sensors for acquisition of position and orientation data widens the use of the technology.

Earlier models of the magnetic tracking technology also had a problem of latency. The latency was caused by the computational heaviness of calculating and generating position and orientation. In addition filtering of the signals had to be performed to get rid of the jitter, and the different strengths of signal had to be equalized to get steady values. Furthermore the number of measurements per second on each sensor (sample rate), is only 60 Hz so some people might perceive latency, because the human sensor organs are quick enough to notice the gaps between each sample.

Nevertheless, newer models have been produced which are capable of over double the sampling rate at 120 Hz and even up to 240 Hz per sensor.

2.1.2 Optical tracking

Optical tracking is based on a vision system which captures object movement and provides the objects three dimensional coordinates based on movement capture movements analysis.

A typical optical tracking system contains the following components:

1. two or more *cameras*
2. *targets* (e.g. markers or silhouettes)
3. a *computer*

The method for acquiring object coordinates can either be marker- or markerless based.

Marker-based optical tracking is when users or objects are equipped with markers that often are covered with retroreflective material. This material lights up much like a traffic sign in the dark or the reflectors on a bicycle. Cameras are calibrated to look for the characteristics of these markers. The data from the cameras are processed on a computer, which outputs the coordinates of the marker, i.e. the position of the user or object.

In markerless based optical tracking the cameras try to determine the position without the use of markers. There exist different approaches for this method, among them silhouette and anatomy based systems. It follows a short description of examples of the two approaches:

a) Tracking silhouettes with help from infrared filters. With this method a user moves in front of a canvas. The canvas is rear projected with some lights, which radiate infrared light towards a camera. Most cameras are able to see infrared light. The user standing in front of the canvas blocks the infrared light, thereby creating a silhouette. Mounting an infrared filter in front of the camera creates a better silhouette as all visible light is removed, producing a more distinct edge around it. The silhouettes position can then be calculated by image processing techniques.

b) Optical tracking based on anatomy is a combination of data about the human body and a cameras image. Information about the body and the structural hierarchy of its parts are defined by a mathematical model, which is able to recognize the individual parts. The user stands in front of a uniformed colored background, where after his/her silhouette is extracted. The position of the parts can then be calculated allowing for a three dimensional representation of the data.

The following table outlines the advantages and disadvantages of different optical tracking approaches.

	Advantages	Disadvantages
Marker based Tracking	• above average accuracy • user is not impaired by cables. • can register fast movements	• line of sight can be obscured • illumination of the scene • need frequent re-calibration
Markerless Silhouette based Tracking	• no preparation of user • real-time analysis • only one camera needed	• limited camera angle • overlap w/ several users • only 2D movement is analyzed
Markerless Anatomy based Tracking	• only one camera needed • no preparation of user	• limited camera angle • cloths not same color as background • restart if users have different skin color • 3D reconstruction problem /w several people • Less exact than marker-based tracking

The table illustrates that there are a lot of key benefits with the use of optical tracking. One thing worth noting is that there is no preparation time for the user, allowing ease of use and not inhibiting user movement. Another thing worth noting is the real-time analysis enabled by the markerless silhouette approach; this could imply low latency.

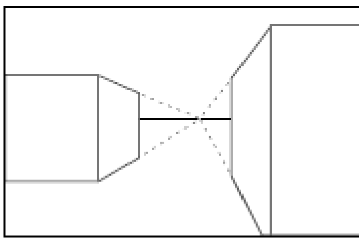
2.1.3 Summary

An analysis of the above described methods for tracking in virtual environments leads to the conclusion that the magnetic tracking is to be preferred system for acquiring three dimensional position and orientation coordinates of a moving object. This is probably due to the fact that precision is vital for experiencing immersion: In fact, if the user does not see what he/she is expecting then the “illusion” of being there disappears. The analysis also leads to the conclusion that for some virtual reality systems an optical based tracking solution would be more suitable, due to the fact that the user can start moving and interacting without the need for re-calibration and preparation. Furthermore the use of an infrared filter will reduce the illumination problem, which is a disadvantage of optical based solutions.

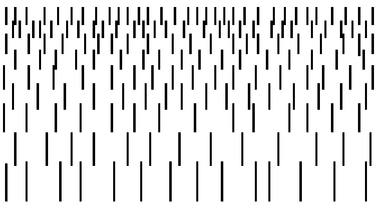
2.2. 3D stereo visualization

To understand 3D visualization, we need to known the term *stereopsis*. Stereopsis is the ability we humans have to perceive depth in the world. We do it by fusing the images in each eye to a new image which is "greater" than the sum of the two. This means that we are able to perceive depth of what we see.
Before going further we need to known basics of human perception. The fundamental cues that help us perceive distance or depth are called monocular depth cues.

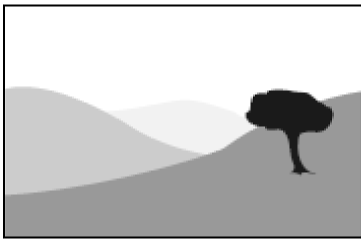
The following points outline the main cues for monocular viewing.



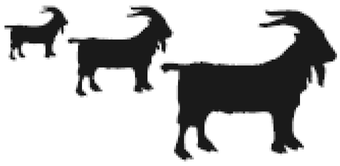
Perspective is the most important cue when combined with a 3D stereo approach. If this cue is exaggerated, or if the vanishing point can be seen, the image's depth will is improved greatly.



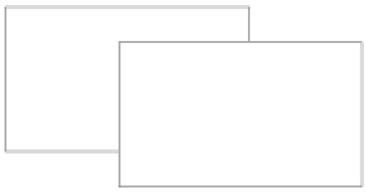
Texture relates to the fact that texture of e.g. grass will be more noticeable when close to the viewer and more faded on a far away distant hill.



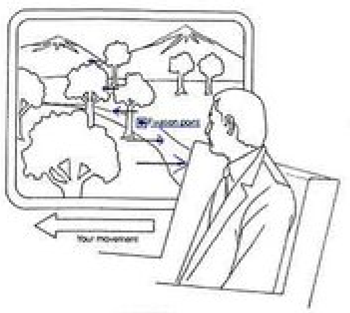
Contrast decreases with distance (haze is reducing visibility further away).



Size of familiar objects is reduced at further distance and through experience we know that object that are close are bigger than the objects further away.



Occlusion of objects close to the viewer cut into the shapes of more distant objects.



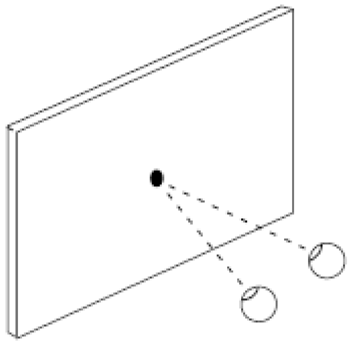
Motion parallax is observed with movement i.e. objects seem to move faster close to the viewer and slower at a distance. Furthermore objects closer than fixation appear to move in the direction opposite to your direction of movement. Objects further appear to move in the same direction as you are moving.

Although the monocular depth cues do not give stereoscopic depth, they are however very important to enhance the effect and can deliver a convincing three dimensional image. Especially the perspective and motion parallax cues enhance the depth, and deliver a better stereoscopic depth [8] [10]. The monocular depth cues are indirect cues that enhance our depth perception. Stereopsis is the direct method, which yields "true" depth information. The reason we perceive depth is because of retinal disparity. *Disparity* is the distance between our eyes that gives each eye a slightly different perspective view of the world. The different views are then fused in our mind to give the sensation of depth.

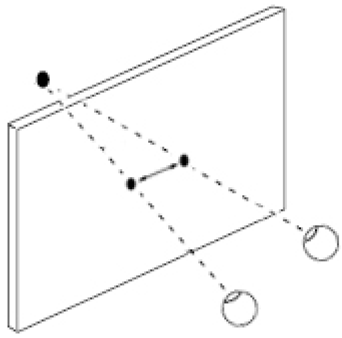
Disparity is close coupled with the term *parallax*, which is what a viewer sees on a display system, where as disparity is what is seen on the retina.

The parallax on the display produces retinal disparity, and the disparity procedures stereopsis.

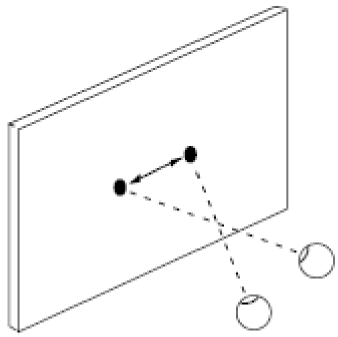
There exist different types of parallax, which can be experienced on a display system.



Zero parallax occurs when the eyes are converging on the same point on the display screen.



Positive parallax is when the eyes are converging on a point behind the screen.



Negative parallax occurs when the eyes are converging on a point in front of the screen. It is negative parallax that gives the effect of objects “popping” out of the screen.

2.2.1 Stereoscopic Approaches

The definition of a stereoscopic approach is in how the system separates the different perspective views. In particular the hardware which is being used, and the effectiveness of it.

A common problem with almost all the stereoscopic approaches is *crosstalk*. Crosstalk is when one eye sees the perspective view of the other eye. This happens because the hardware is not "good enough" to exclude the views from each other. The perception of crosstalk is known as *ghosting*. Ghosting causes the user to see a dimmer version of one eyes image superimposed onto the other. The perception of ghosting is different from system to system, but parallax and image contrast intensifies the perception of it. The more parallax the worse it gets. In this context it is also worth emphasizing that ghosting can ruin the 3D effect, and lead to nausea and sickness. [10]

Passive Stereo

Passive stereo systems multiplex images in space to achieve depth. The left and right eye images are overlapped to create the desired separation. There exist three main approaches of passive stereo: *Anaglyph*, *Polarized* and *Separated Displays.*

Anaglyph Stereo

Anaglyph uses color filters to separate left and right eye images. The images are multiplexed in space and special glasses with color filters allow each image for "passing through" the correct filter. Figure 2.1 describes the approach.

The advantage of using anaglyph is that it is cheap and simple, and does not require expensive hardware to generate the images. The disadvantage is that this method does not display true color and in some cases ghosting can occur, in cases when the color of the filters in the glasses does not precisely correspond with the colors of the projected image. Great results can be experienced with grayscale images.

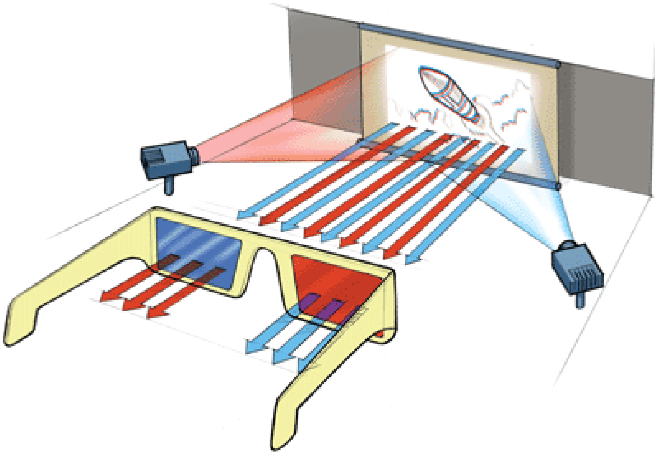


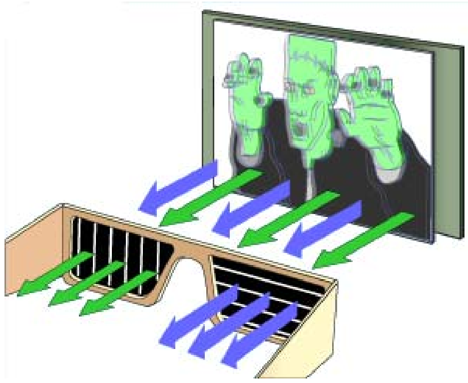
Figure 2.1 The red and blue lenses filter the two projected images allowing only one image to enter each eye. (www.howstuffworks.com)

Polarized Stereo

Much like anaglyph stereo, this method uses filters to separate left and right eye images. These filters are not colored, but polarized. Polarized filters affect the light waves, so they only travel in one plane. The filters are built into the glasses and mounted in front of two synchronized projectors. They project two views onto a screen, each with a different polarization. The glasses allow only one of the images to reach each eye. This is because the glasses have corresponding filters with the different polarization. Figure 2.2b shows how the different polarization reaches the correct eye. This system can yield true colors opposed to anaglyph. A necessity for this system to work is a special screen called a silverscreen. A silverscreen keeps the polarization of the incoming light from the projectors, as the two images are kept separate by the different polarization. If the screen randomizes the polarization while reflecting it, the glasses cannot properly filter out the inappropriate image from each eye, hereby ghosting can occur.

As mentioned the hardware requirements call for two projectors that in most cases have to be mounted on a special rig that can be aligned and hold filter brackets for the projectors (Fig. 2.2a). The alignment of the projector images has to be quite accurate to give the correct perception, but also giving the least amount of ghosting.

(a) Two synchronized projectors with filters (Aalborg University Copenhagen)



(b) Polarization glasses (www.howstuffworks.com)

Figure 2.2 Polarized system

Although the wrong alignment can lead to ghosting, the choice of the screen material can also generate this effect. As mentioned the type of screen needed for this system is called a silverscreen or non-depolarized screen, which preserves the polarization of the two projected light beams. So to summarize the polarized system has the following characteristics to be aware of:

- Hardware (expensive)
- Screen Material (expensive)
- Alignment and calibration

Separated Displays stereo

Unlike the two previous methods the images are not multiplexed but sent to separate displays. No crosstalk is obtained by setting displays close in front of each eye. A common system that utilizes this approach is the Head Mounted Display (HMD).

Active Stereo

In relation to passive systems, the active stereo system multiplexes images over time. This system mostly uses shutterglasses to reach the desired separation effect. Figure 2.3 shows the glasses used at Aalborg University Copenhagen. Shutterglasses are synchronized to the displays refresh rate and alternates between the left and right eye images in rapid succession, [5]. Each eye is shut electronically: hence the term shutterglasses. The perceived image has true color and in a great deal of cases can be free of ghosting with the introduction of new screens and DLP projector technology.

Figure 2.3 E-Dimensional shutter glasses

When using active shutter glasses the synchronization can be lost and has to be reset before it can work again.

Autostereoscopic stereo

Autostereoscopic stereo is an approach where no glasses are needed to perceive depth. Unlike the other methods, one could say that the screen is actually wearing the glasses. A lenticular lens plastic sheet is placed in front of the display. The sheet has very narrow vertical cylindrical lenslets spaced to correspond to the columns of an interlaced stereo pair. The stereo pair (*i.e.*, images corresponding to left and right eye viewpoints) is interlaced into alternate columns in a two-dimensional image. In this way, the correct images of the stereo pair are directed to its corresponding eye and here by generating a three-dimensional image.

The narrow vertical lenslets only allow for a small amount of user to see the image correctly, as the viewing angle of the directed stereo pair is very small.

2.2.2 Display systems

Coupling different stereoscopic approaches to a display system provides a mean of visualization and immersion. Depending on the degree of immersion and interactivity, different systems are available. Based on Livatino & Privetera (2006) [9] the main attributes which characterize a display system are:

- **Display Size –** virtual reality displays range from big auditorium sized to tiny HMD sized displays
- **Display Surface –** there exists different kinds of surfaces like a flat, curved, round, cubic, wall etc.
- **Projection Modality –** refers to displays that a computer can output to e.g. CRT or LCD monitors and front and rear projected screens
- **Image Quality –** resolution, contrast, brightness and color contribute to the image quality

The attributes mentioned above define main display system characteristics that can be found within virtual reality. At Aalborg University VR Media Lab and Aalborg University Copenhagen we feel lucky to have the state of the art facilities, where users can experience wide range of virtual reality experiences from HMD, to 3D-panorama and CAVE systems. Unfortunately, these facilities do not include a workbench type display.

CRT or LCD display

The CRT display is the most common display and therefore is the first display that people get in touch with, when viewing 3D. The above mentioned 3D stereo approaches work on such a display except for the polarized stereo system. The display has to run at refresh rates from 85-120Hz in order to avoid the active stereo approach from flickering. LCD display technology is not capable of such high refresh rates yet, so only the anaglyph stereo approach is possible on LCD displays. The benefits of CRT displays are that they are widely available and can yield high image quality. Furthermore, CRT displays can be used with almost all the stereo approaches.

Panorama

The Panorama at Aalborg University's VR Media Lab accommodates up to 28 persons placed in front of a large cylindrical screen (width 7.1m, height 3.5m arched 160 degrees) (Fig. 2.4). The display consists of 3 segments, where each segment is projected by one projector. The benefit of having a big cylindrical screen is that the viewers are given a convincing spatial presentation of the virtual world. The system can display both in 2D and in 3D using active stereo glasses. The main applications for the panorama are presenting and simulation.

Figure 2.4 Panorama at Aalborg University VR Media Lab

CAVE

The first CAVE (Cave Automated Virtual Environment) was developed by Carolina Cruz-Neira, Tom DeFanti and Dan Sandin, at the University of Illinois at Chicago in 1992 [2]. Its' main purpose was scientific visualization to let computational scientist discover new things faster. It consisted of 3 rear projected walls and 1 floor projected from above. Many have since copied the idea and some have added more walls, like the CAVE at Aalborg University VR Media Lab.

The CAVE at VR Media Lab is a display system that is constructed like a box (width 2.5m, height 2.5m depth 2.5m). Figure 2.5 shows the system. Projected images are projected on each side of the box and with the active stereo glasses;

it presents a fully immersive 3D experience. The viewer's head movements are magnetically tracked giving the possibility to look under objects and giving the sensation of being there, surrounded by a virtual world. The interaction is done with different peripheral devices that allows for 6 DOF. Having multiple passive viewers' poses a problem. In fact, only one viewer gets the full benefits. Passive viewer's makes them observers rather than being able to participate actively. All the observers perceive the virtual world differently.

Figure 2.5 CAVE at Aalborg University VR Media Lab

Head-Mounted Display (HMD)

This form of display was invented by Ivan Sutherland in the 1970s. The HMD is carried on the head and uses the separated display stereo approach (Fig. 2.6a.). Unlike the other systems, HMD does not rely on any filtering to give depth perception. Combining the HMD with a head tracking device, the user can freely look around in the virtual environment. Some head-mounted displays also have the ability to view a see-through image superimposed upon a real world view, creating what is called augmented reality (Fig. 2.6b). A common factor is that they are bulky, expensive and have wires limiting movement. However new technology development has reduced both the cost and size. Moreover the image quality has improved due to higher display resolutions.

(a) Normal HMD (b) See-through HMD

Figure 2.6

Workbench

Workbench type displays were developed to accommodate the wishes of physicians, architects and automotive engineers. They lacked a system for close collaborative work. The following are three examples of such display systems.

- The Responsive Workbench (RWB) is a cooperative virtual environment where users can share the same space and information for specific tasks (Fig. 2.7a). In contrast the virtual environment CAVE, the RWB has a universal interface which is intended for many different users and usages. The RWB display resembles a table giving the impression of virtual objects to standing on the tabletop.
- The ImmersaDesk is a screen tilted at a 45 degree angle (Fig. 2.7b). The tilted screen allows the user to look forward and down at the same time. The reason for this is to resemble the CAVE has the same ability to look forward and down and this contributes to the virtual reality experience.
- The Holobench is an L-shaped 3D projection display with two orthogonal projection surfaces (Fig. 2.7.c.). The benefit of having an L-shaped display is the ability to interact with virtual object with the obstruction of the screen. Only in a CAVE or tabletop display this is possible. The system also uses 2 graphic pipelines whereas the other displays only use 1.

(a) The responsive workbench

(b) The ImmersaDesk

(c) The Holobench

Figure 2.7

2.2.3 Summary

In summary we can state that no 3D stereo visualization approach is free from problems.

	Advantages	Disadvantages
Passive Stereo	• Cheap • simple processing • anaglyph can be viewed on any display • polarized and separated displays systems can yield true color	• anaglyph can not yield true color • ghosting can occur while using anaglyph • polarized system is expensive • polarized system needs calibration and alignment
Active Stereo	• true color representation • ghosting is minimal	• requires expensive equipment • synchronization has to be reset if signal is lost • flicker at low refresh rates
Autostereoscopic Stereo	• no glasses needed • true color representation	• limited viewing angle • limited amount of viewers

The table outlines the advantages and disadvantages of the different stereoscopic approaches. Ideally Autostereoscopic Stereo would be preferred, but the technology is still new and the limited field of view for viewers, makes it unusable for presentation. However the author of this thesis believes that the active system is the next best solution as it can produce true colors with the smallest amount of ghosting and calibration. Surprisingly Anaglyph Stereo has a lot of advantages if one looks besides the color bleeding of the filters. It provides users with a cheap and ready solution which can be used anywhere on any display.

The different display systems can be coupled with the different stereoscopic approaches, depending on the amount of immersion and interaction wished for use. For full immersion users should choose the CAVE or HMD, which surrounds the user with a virtual environment. Presentations for large audiences would benefit from using a passive system, as glasses are cheap and the hardware less expensive. Collaborative and manipulative work for rapid prototyping could be done on a type display that allows this, e.g. Workbench.

Chapter 3

Workbench Developments and Interaction Methods

This chapter provides a brief overview of selected techniques and developments in the fields of workbench display systems and tracking. The overview is not to be considered as a exhaustive analysis of the research fields. Rather, it is in order to provide, the reader with a description of different creative solutions for workbenches and tracking.
The proposed system does have similarities with some of literature. In fact, the works represented in this chapter have inspired the author of this thesis to the proposed method.

In the following pages, each literature is described through a short overview which points out the technique or development proposed as well as main approach characteristics. At the end of each overview, a comparison of the presented work with to the solution proposed in this thesis is provided.

3.1 Krüger, Bohn, Fröhlich, Schüth, Strauss and Wesche, [14] [15]

The proposed concept is an alternative virtual reality display for collaborative environments and applications, called "The Responsive Workbench".

This virtual environment display is designed to support users working on desks, workbenches, and tables, such as physicians, architects and automotive engineers, with an adequate human-computer interface. Virtual objects are located on a real "workbench". The objects, displayed as computer generated stereoscopic images, are projected onto the surface of a table. The users operate within a non-immersive virtual environment. A "guide" uses the virtual environment while several observers also can watch events by using shutterglasses. Depending on the application, various input and output modules can be integrated, such as motion, gesture and speech recognition systems.

Before implementing the "Responsive Workbench" the authors performed a user task analysis to see which tasks that where carried out by physicians, architects and automotive engineers. From the result of the analysis the authors came to the conclusion that their cooperative tasks relied on a workbench scenario rather than on any other virtual reality display.

No real problems are described, but the only thing to take note of, is the distortion experienced by observers whose head position is not tracked. The reason for this is that the projection follows the "guide" tracked head position.

The main characteristics of the developed VR display can be summarized as:

- Tabletop display, which gives a highly *natural setting* to share experiences.
- The support for *collaborative* and *manipulative* tasks which benefits the above mentioned type of users.
- The use *stereoscopic approache*s to give users a perception of virtual objects standing on display surface.

Comparison

The proposed display system fills in a gap between immersive rooms (such as the CAVE [2]) and the HMD. Both displays are not especially user orientated, based on the fact that they focus more on the ability to be immersed rather than the reason for it. “The Responsive Workbench” is purpose built, as it can facilitate specific user needs, e.g. a physician has no interest in standing inside a virtual body as it is not a natural part of their work process, but he/she would rather see a virtual body from a distance on a table.

In the wake of the responsive workbench, other workbench like displays appeared. The ImmersaBench [11] developed at EVL (Electronic Visualization Laboratory at University of Illinois) and the Barco Tan Holobench, extend the use and appeal of such a display.

The display system proposed in this thesis is similar to the one proposed by the presented work. However the size and hardware requirements are different, due to the simplification of this thesis’s design. Both systems us active shutterglasses to display three-dimensional images, but this thesis will, also use passive anaglyph approach and will compare performance of the two systems.

3.2 McDowall and Bolas, [4]

This article introduces two new technology developments in the field of virtual reality displays.

The first new technology, called Duo, allows two simultaneous and individually head tracked 3D stereo pairs to be displayed on a single display surface. This gives the two viewers the possibility to share a workspace by pointing at a model and sharing virtual objects to manipulate. The motion of either viewer does not affect the view of the other. This technique is also described in another paper by Agrawala et al. (1997) [1]. The authors developed custom shutterglasses, which use the time sequential technique. The perspective views of e.g. user 1 (U1) and user 2 (U2) are displayed in rapid succession in the following order; U1 Left Right, U2 Left Right, U1 Left (U2) Left Right etc.. This produces flicker because of the image display sequence where each eye only sees 25 % of the time; otherwise it is opaque. Tests however showed an enhanced interaction with virtual objects in a virtual environment.

The second new development in is the ability to adjust the angle of the projection surface to provide a natural workspace depending on the application and associated virtual models. Adjusting the angle of the workbench projection surface between vertical and 20 degrees was preferred. The reason for this being that a steeper angle would interfere with the interaction, because the users hand could touch the projection surface.

The main characteristics of the new development technologies:

- Two simultaneous and individually head tracked 3D stereo images on one display surface, through modification of the time sequential imaging sequence.
- Display surface adjustment, to facilitate different applications.

Comparison

The display system proposed in thesis is inspired by this article. In particular, the implementation and design of an adjustable screen. This type of screen will make the system more versatile and it will widen the range of applications. The adjustable screen is also inspired by the ImmersaDesk [11], allowing the user to look forward and down at the same time, (i.e. like in the CAVE [2]).

Although the two user stereo is a welcome new development, it will not be implemented as it is not in the scope of this thesis.

3.3 Rekimoto and Matsushita, [6]

The proposed system called *HoloWall* is a wall sized computer display that allows users to interact without pointing devices. Users can instead use there fingers, hands, body and physical objects to interact. The concept is based on displaying projections on surfaces such as walls, floors and tabletops. These computer surface displays require a new method for pointing. The authors point out that to achieve this, the cameras play an important role, as users do not carry pointing devices. The display surface is rear projected with a video projector and e.g. the users hand is tracked by a camera with an IR filter (opaque filter that blocks out light below 840 nm). Two IR light arrays then emitted onto the back of the display surface and the camera picks up any reflected light. The reflection is processed with the help of simple image processing techniques to determine the position. By adjusting the threshold values *hovering* object from 0-30 cm from the display surface, can be recognized.

The proposed method has the following main characteristics:

- *Non-invasive* interaction, as it is based on markerless optical tracking.
- Infrared tracking, to track object reflection.
- Simple image processing techniques to acquire position coordinates.

Comparison

The proposed system has the advantage that users do not require any data glove or pointing devices to interact with the display surface. Furthermore the use of IR light for tracking eliminates the problem of poor and low illumination. Moreover the light from the video projector is eliminated with the IR filter.

The system proposed in this thesis, will also use IR light emitters to light up the display surface from behind, but the system will have twice as many emitters, because the large surface, (more light is then needed). In addition this thesis will use the image threshold values and blob size to determine the height of the tracked hand above the surface.

3.4 Other related works

There are other related works, mentioned below which all contributed to the design of the workbench display and interaction methods proposed in this thesis.

As mentioned earlier the system proposed in this thesis will use infrared light for tracking the hand position. Infrared light was also proposed by Moeslund et. al. (2000) [16]. Where the system was optical based and it used four infrared cameras to detect where users where pointing in the 3D environment. By doing it this way, the authors wanted to eliminate the use of the Wanda (a pointing device for immersive environments based on magnetic tracking), so that users would have a more intuitive way of pointing, similar to what they would in the real world. The system relies on retroreflective markers to calculate the position of the user and where he/she is pointing. In addition, the system proves that the use of an infrared filter (opaque filter) can eliminate nearly all the light projected on the CAVE walls, thereby making it very effective.

Infrared light is also used by J. Han on his multi touch system [3]. Compared to the approach of Rekimoto and Matsushita, Han uses an optical phenomenon called *frustrated total internal reflection (FTIR)*, where light is refracted (bent) inside a medium, e.g. glass and scatters to the surface of the medium where it is touched, thus lighting up the touching object. This approach yields a very high spatial and temporal resolution, but it does not allow for any tracking of *hovering* objects above the display surface. The reason for this is that the light is "trapped" inside the display and is not passing through, as on the HoloWall. Initially, the described technique was intended to be applied in this thesis but, this was not possible, because the hovering is not applicable and the display required advanced knowledge of optics and FTIR.

Interaction in some virtual environments requires precise knowledge of virtual objects movements. This is a difficult task being the user is unaware of the force he/she is applying to an object. The "virtual spring manipulator" is a solution to this problem [7]. This system uses virtual springs as a way of giving *visual force feedback* to the user. Springs are bent and compressed depending on the direction and amount of force applied. Although the system was developed for steering particles in molecular dynamics, it can also be applied for pushing and pulling in other applications. The system of this thesis will also implement physics in to the virtual world as a way to move and apply force to an object.

3.5 Summary and Analysis

Summary

The described approaches to workbench development and interaction methods all provide solutions to their problem at hand. The entire workbench development is based on the original work by Krüger et al. [14] [15], where a virtual reality workbench display is proposed for collaborative work, based on test and analysis of different users. The concept of a workbench environment has been further developed by Agrawala et. al. (1997) [1], McDowall & Bolas (1997) [4], and Czernuszenko et. al. (1997) [11], who have developed new technologies and modified the original design to accommodate a wider application. These new developments include a multi user 3D stereo system and an adjustable screen.

The approach by Rekimoto and Matsushita (1997) [6] to bring interactivity to users on walls makes use of a basic camera setup to gain a substantial and high level of interaction. In particular the use of infrared emitters to light up the display surface is worth noting. Similar approaches by J. Han (2005) [3] and Moeslund et. al. (2000) [16] rely on the use of infrared emitters and opaque filters to track positions of hands and fingers. In continuation a solution for visual force feedback is presented by Koutek et. al. (2002) [7]. The solution use “springs” as a way to determine forces applied to an object.

Analysis

All the presented papers provide solutions for tracking the position of head, hand, body and physical objects. The systems described rely on retro-reflective markers for tracking which is also a very common approach as pointed out in Chapter 2. Although an optical approach has many advantages in relation to a magnetic approach, the user is forced to wear markers either on his/her body or on objects held in the hand.

When look at the responsive workbench and other virtual reality systems, they all use active shutterglasses, but if we look at what we learned in Chapter 2, using passive anaglyph should also yield a good stereoscopic depth perception.

A stereoscopic workbench system based on using infrared light emitters for tracking hand position can not be found in the literature. In particular, it is hard to find literature on tracking markerless reflected light in virtual reality system. In addition, little can be found on the literature concerning the comparison of different stereoscopic approaches. In fact, the literature mainly focuses on specific approaches.

This thesis consequently aims to propose:

1. an approach to interacting in 3D environments on a workbench, based on markerless optical tracking.

2. the development of a workbench display supporting tracking based on infrared light reflection.

3. a comparison of passive anaglyph and active shutterglasses on a workbench display system.

Chapter 4

The Proposed Investigation

This chapter proposes a method for tracking hand position in 3D on a workbench display system. In the chapter 2 we have learned that the current technology for tracking 3D position has some limitations, e.g. the invasive sensors. Furthermore the literature proposes new technology developments that that is taken into consideration in the proposed system.

4.1 Core Idea and Argumentation

This thesis proposes the investigation of a new system for tracking position in 3D environments. In particular the proposed solution will have its main focus on the building of a workbench display system and the development of a markerless vision based (optical) interaction method for tracking hand position. This thesis will also propose a study of the new system for tracking with different approaches to stereo visualization. In particular the interaction will be performed on a workbench running different 3D stereo approaches, developed for this purpose. The two approaches for investigating stereo visualization will be: Passive Anaglyph and Active Shutter glasses.

The motivation for working with virtual reality and interaction comes from a great interest in the field and an interest in creating new approaches to a problem. A workbench environment where various tasks can be performed around a virtual object displayed on the surface, provides an ideal test bed for the proposed tracking system.

A workbench is the preferred display because of the way users can collaborate and manipulate virtual objects "standing" on the display surface. As discussed in chapter 2, virtual reality needs tracking and it needs them better. The proposed tracking system would have:

- the advantage of being non-invasive upon the user and giving a more natural method of interacting.
- potential for new application like games and enhanced collaborative manipulation.
- some uncertainty about whether it is able to work probably, but it is believed that the simplicity of the system would lean towards the opposite notion.

4.2 Investigation Design

The Infrared Vision-based Tracker

The tracking setup can be done in several ways; either a camera is placed above the workbench display surface, below it or it can be placed besides the projector looking into the workbench mirror. The latter has been chosen, as a camera placed above will be obscured by the user and a camera below can not view the entire display surface.

A method for tracking on the display surface is proposed based on, an investigation into different interaction and touch screen techniques (see chapter 3). The investigation led to an infrared solution. The light from the video projector is hard to track and constantly varying. Using infrared light, the light conditions will be more stable and using a visual opaque filter (infrared filter) in front of the camera, infrared light can be separated from the video projector light.

Simple image processing techniques can be performed to separate the hand from the rest of the image. Further more, adjusting the image threshold, the system is capable of detecting hovering above the display surface.

3D Stereo Approaches

At Aalborg University we are privileged by having the ability to use different stereo approaches and display systems. On the developed workbench, both passive anaglyph and active stereo approaches will be used and a test will compare performance.

Objects displayed on the surface will be viewed with negative parallax giving the perception of them standing on the workbench surface *within the user space*. In combination with displaying the correct parallax, the virtual camera view has to be consistent with what the users perceives.

The Workbench Display System

Research into different workbench displays and techniques has been performed to determine the approach for the proposed workbench.

As mentioned in chapter 3 workbench display systems are built for interacting with 3D objects. This property makes it ideal for the proposed tracking system, but also sets the design boundary of what such a system should be capable of. The design should also be able to display stereo imagery with passive anaglyph and active stereo approaches.

The workbench display surface is made of ground glass, creating a diffuse surface on which images can be projected. In addition glass is very stiff and easy to clean. Originally a wooden frame was preferred, giving a big and heavy construction. However previous work by the author of this thesis used aluminum profiles for building a display frame [13] these are light, strong and easy to assemble. Implementing an adjustable screen will allow a wide range of applications.

4.3 Experimentation

The proposed tracking system is more a proof of concept, rather than being designed for a specific application, therefore the experimental tasks will be simple. The tasks will consist of moving virtual objects, by pushing them around. Users will push cubes from a position A to a position B, using the proposed tracking system and the magnetic tracking system. The comparison of the two systems will determine the optical tracking systems performance opposed to the magnetic trackers.

In connection with the tracking experiments, an analysis of the effect of using two different stereo visualization techniques; on the proposed workbench system should possibly conclude about which of the methods is most effective for this type of display.

4.4 Research Development Plan

The development of the Workbench and tracking system will contains the following tasks:

Workbench development: Building the workbench should be based on the workbenches in the literature. The development can be divided into two parts: a pilot study and full scale development. The pilot study should be performed to see if it is possible to track with infrared light and to try out different materials. The lessons learned by the pilot study should be transferred to the full scale workbench.

Image acquisition: Different cameras will be tested for performance. The camera tests should determine whether a cheap web camera can be used instead of an expensive firewire camera. Furthermore an infrared filter should be placed in front of the camera to eliminate all visible light. The pilot study will also conclude if infrared light can be used on the proposed workbench.

Position tracking and mapping: During the pilot study, the simple image processing should be developed and should easily be transferred to the full scale workbench. The position of the hand should then be mapped to a object in the virtual world.

Application: As mentioned the proposed tracker is more a proof of concept, so therefore a specific application is not vital. A test environment should be made to see the tracking is working.

Chapter 5

Experimentation and Discussion

This chapter describes the development of the proposed tracking system and workbench displays system. The experiments are then defined, where after the results of the experimentation are pointed out and discussed.

- *Included on the CD are images of the Workbench and movie clips of the tracking*

5.1 Hardware Setup

Before starting to build and experiment a pilot study was performed to determine if it was possible to track the hand with help from infrared light.

5.1.1 Pilot Study

The setup of the pilot study system contained a little display, built of a matte acrylic sheet (20 x 20 cm) mounted on an aluminum profile. The plate was rear projected and a web camera with an infrared filter was also placed behind it. The little display was then rear lit with infrared light. Using simple image processing techniques, e.g. background subtraction and threshold, the hand was tracked. Although the pilot study yielded positive results, a full scale development might not give same results.

5.1.2 Full Scale Development

Now that we know that tracking can be done on a small scale, we need to develop a full scale system, to see if it is also possible on a large display surface also running 3D stereo visualization

The full scale system will consist of the following components:

- workbench with mirror
- 3D stereo projector
- camera
- magnetic tracking with 1 sensor
- two IR arrays
- computer

The Workbench

The workbench display is built on the background of knowledge gained through research (see chapter 3) and prior experience.

In comparison to the pilot study display the proposed workbench will have a glass display surface, because an acrylic sheet of this size would not be stiff enough. In addition the display surface is ground glass, giving a diffuse surface on which to project on to.

Most workbench displays have wider aspect ratio screens (similar to 16:9) running non-uniform resolutions. The proposed workbench has a 4:3 aspect ratio screen running native resolutions, which is standard on most computer screens today. Inspired by the responsive workbench [15] and the ImmersaDesks [11], the screen is hinged at one end allowing different vertical viewing angles of the screen. The tilting of the screen gives more possibilities for different usages and applications.

In order to project images on to the display surface a mirror is used to direct the light cone from the projector. The entry angle of the light cone is equal to the exit angle and because this is not at a 90^{o} angle, keystone on the projector must be used.

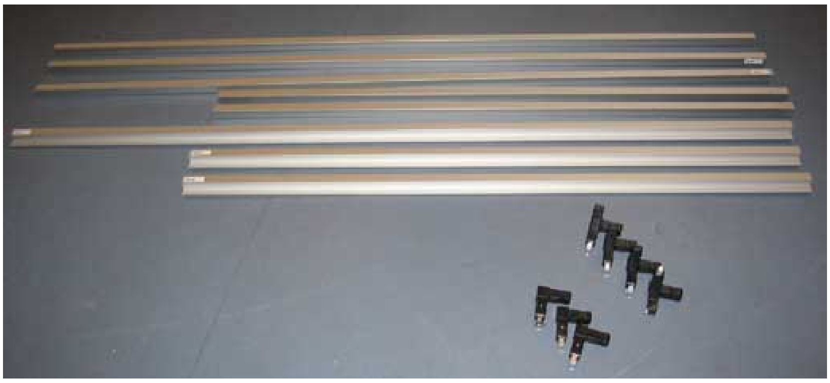
The workbench frame is made of aluminum profiles and reinforced plastic joints, giving a light and strong frame, but also easy assembled. The adjustment and assembly of the profiles took about an hour, in assuming that a wooden construction would be heavier and require more materials.

Materials where bought in the Danish DIY (Do It Yourself) store Silvan, and the glass was bought at a glazier, in a custom made size.

The overall dimensions of the workbench are:

Width:	133 cm / 52,36 in
Depth:	100 cm / 39,37 in
Height:	90 cm / 35,43 in

The following shows the step by step assembly of the workbench:



Step 1: The materials needed for the frame

Step 2: The 4:3 frame for the glass screen surface is assembled.

Step 3: Here the frame for supporting the glass screen frame is assembled.

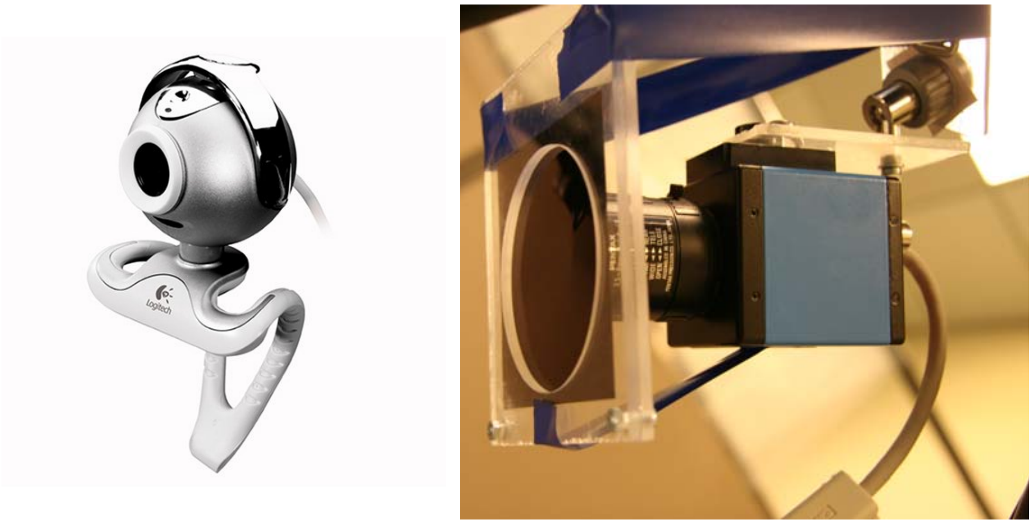
Step 4: Here the final frame for the screen and the supporting frame are combined.

The Camera

The image acquisition in the pilot study was with a Logitech QuickCam Zoom web camera (Fig. 5.1a), but due to a high level of noise a high quality camera (Fig. 5.1b) was used for the proposed system.

All CCD based cameras can detect infrared light. An easy example is to take you TV remote and then press a button. What you will see is that every time you press a button, the infrared light at the tip of the remote lights up. But the problem with the camera is that is also sees visible light, which we do not want. We only want to see the infrared light, and by placing an infrared filter that blocks out 99% of the visible light in front of the camera, we only see the infrared light from the two IR arrays.

This solution works very well, but through research it was discovered that all cameras have an IR cut filter built in. The IR cut filter, blocks some of the infrared light. So by removing this filter which is a thin film placed on a lens, an increase in sensitivity could be obtained.



(a) Logitech QuickCam Zoom

(b) ImagingSource Firewire Monochrome Camera DMK 21F04 in the system setup

Figure 5.1 Cameras Tested

The camera is placed below the projector pointing into the mirror, so it can see the entire display surface.

Additional hardware

Besides the above mentioned hardware, the system also consists of a 3D stereo projector, magnetic tracker and computer. The 3D stereo projector is capable of running refresh rates up to 120Hz, so it can be used with active stereo shutter glasses. To get the correct perspective view to the user, a magnetic tracker is used to track the head position. The tracker used in this system is unstable even at short range, due to an unoriginal power supply. Last a powerful computer is running the developed application and projection.

5.2 Position Tracking

To be able to interact in 3D environments requires minimum 3DOF (X, Y, Z). The X and Z coordinates are easy to acquire, by the position of the hand across the display surface. Calculating the Y coordinate poses a challenge. The proposed system will use the area of hand and convert it to height positions. In theory, the closer the hand is to the display surface the bigger an area is occupies. Tracking of the hand is done in *EyesWeb*, a program specially designed for analysis and processing of movement, midi, audio, and music signals. Eyesweb uses graphical programming, where functions are represented by small building blocks that can be dragged into a workspace, creating patches.

Figure 5.2 shows the patch for the proposed optical tracker.

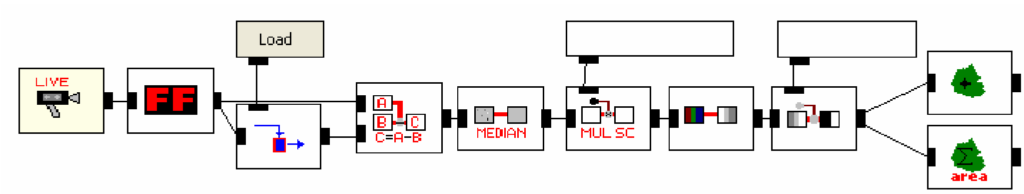


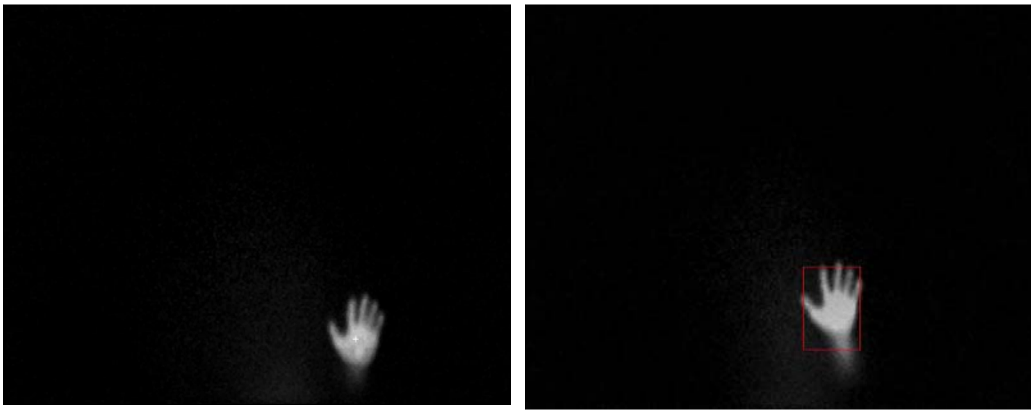
Figure 5.2 EyesWeb Tracking Patch

Here are the blocks explained, starting from the left:

Block Nr.	Function
1	Frame grabber that retrieves the camera images continuously
2	Keystone
3	Buffer that stores the first image on patch start
4	Subtraction of two inputs that outputs a background subtracted image
5	3 x 3 Median filter to remove any noise remaining
6	Multiplies the incoming images with a value between 0-255 to give a clearer image
7	Color to grayscale

8	Thresholds the image
9	These two blocks retrieve the center of gravity and area of the tracked object

In Figure 5.3 the hand is clearly separated from the background image providing a good feature to track. The center of gravity coordinates and the area of tracked hand is then saved to a text file for further processing.



(a) Center of Gravity (b) Bounding Box (Area)

Figure 5.3 Tracked Hand

5.3 Application

The 3D visualization was done in *Virtools*, which is used for rapid prototyping and has built in VR functions. Like EyesWeb, Virtools uses graphical programming, where functions are represented by building blocks that can be dragged into a schematic or object.

The developed application has no purpose other than as a test bed for the 3D stereo visualization and proposed tracker.

The basic setup of the virtual environment consists of a projection plane which is placed at zero parallax corresponding to the display surface. The projection plane is used as a reference by a tracked camera, giving the correct perspective view to the user. Furthermore 3D objects which are placed above the projection plane are perceived as "standing" on the display surface (Fig. 5.4).

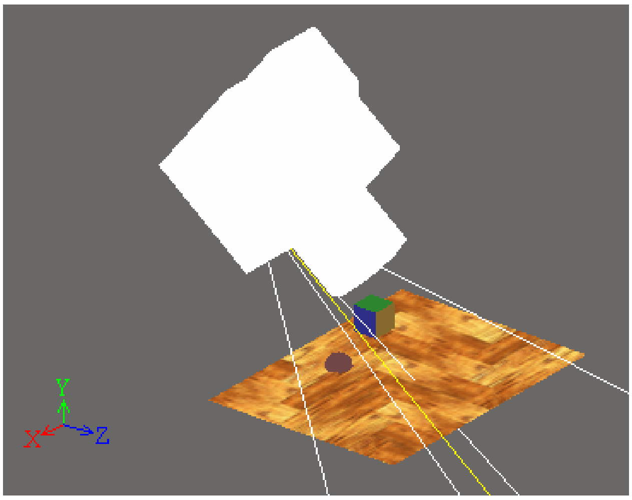


Figure 5.4 The virtual environment setup

In figure 5.4, the top of a little sphere can be seen representing the users' hand. In Virtools the data stored in the EyesWeb text files is loaded into an array and mapped as the spheres position coordinates. All objects in the scene are then physicalized, which gives them mass and they are able to collide with each other.

5.4 Procedure

A first phase experimentation was conducted to test the 3D stereo approaches and calibration of the workbench. Few VR-experienced users where asked to try the active stereo approach on the workbench and describe their perception of the presented scene when looking around. A few of the users also tried the passive anaglyph approach for comparison.

The author of this thesis has additionally tested the proposed tracking system with different stereo approaches. Due to the time frame of this thesis, solving problems with sending data between EyesWeb and Virtools has prevented an intensive user test.

5.5 Results and Discussions

In first phase experiments the objective was for users to test the different 3D stereo approaches and the calibration of the workbench. Consequently, users experienced great depth, when viewing the virtual object. In particular, the strong perception of objects "standing" on the display surface was noted. Several users observed flicker when wearing the active shutterglasses, while viewing the workbench display.

A mentioned a few users also tried using passive anaglyph glasses, resulting in the same great depth perception. Besides not yielding true color and producing a little amount of ghosting, users pointed out that the red filter was irritating the eye.

In the first phase experiment a slight calibration problem was found. When users moved there head the object was slightly distorted. In fact, some of the distortion was produced by the magnetic tracking.

The author of this thesis additionally tested the proposed tracking system with the different 3D stereo approaches. As experienced by the users of the first phase experiment overall depth perception was strong. The proposed tracking system worked as anticipated in the X, Z direction, but the height calculated from the area of the tracked hand was not working as well. Pushing a virtual box from one place to another was effective, both in passive anaglyph and active stereo. The only time the 3D effect is ruined, is when the user's hand intersects the 3D object.

The comparison of the two stereo approaches gave the expected results as also pointed out in chapter 2.2.3 that both systems are equally good for viewing 3D, but have some downsides as well. However the active stereo approach did give the best, due to the true color image.

In summary, the results from the first phase user test and the test with the proposed tracker show that the different 3D stereo approaches work well on the developed workbench. Though the proposed tracker works well in the horizontal plane, further development is needed to get a precise height value.

Chapter 6

Conclusion and Future Work

6.1 Concluding Remarks

This thesis has proposed a new system for tracking hand position in 3D on a workbench display. The workbench consists of a table sized display with a surface of a ground glass plate on which 3D stereo imagery is projected.

Furthermore, the developed tracking system can detect the reflection of a hand moving above the display surface. The underlying technology is infrared light sources illuminating the display surface and an infrared filter placed in front of the camera lens, thereby eliminating visual light.

The importance of the technology developed in this thesis is underlined by the fact that no such technology exists. However, the comparative analyses showed that elements from the technologies on optical tracking and virtual reality displays are applicable parts of the proposed system. The infrared markerless tracking method, which is a part of the optical tracking technology, was applied in order to track the hand. Further, the general concept of workbench type displays was also applied.

The comparison between 3D stereo approaches passive anaglyph and active shutter glasses was performed in order to assess which had the best

performance on a workbench display. This comparison showed that none of the two were superior in terms of depth perception. However it should still be kept in mind that active shutter glasses yield true color compared to the anaglyph approach. On the other hand anaglyph is a more affordable technology and can be used on any display system.
The tests of the proposed tracking system showed that some interaction was possible. In addition when comparing both 3D stereo approaches with the proposed tracker, none of them stood out more than the other.

Conclusively, the proposed tracking system of this thesis enables interaction in 3D on a workbench display system. Furthermore the developed workbench is capable of displaying passive anaglyph and active stereo imagery.

6.2 Future research

Some aspects of the proposed system still represent open issues that need more future development. Among them: acquiring the height of the hand above the display surface and a more comprehensive comparison of different stereo approaches.

- **Height above the display:** The experimentation showed that the height estimation from the area of the tracked hand was very imprecise. Future work should explore new methods for acquiring height values. This could be done by measuring the grayscale value of the tracked point. The reflected light of the users hand is perceived as white near the display surface and fades to black when rising above the displays.

- **Stereo Comparison:** Although user where generally pleased with the 3D stereo approaches, this thesis lacks a more comprehensive study of the differences perceived by viewers. Specific questionnaires should be handed out after each test to get a combined overview of the differences between the systems.

Future work could also look into the tracking of multiple hands and gesture recognition to give the user more tools to work with when interacting with a virtual environment.

References

[1] Agrawala, M., A. C. Beers, B. Fröhlich, P. Hanrahan, I. McDowall, and M. Bolas. The Two-User Responsive Workbench: Support for Collaboration Through Individual Views of a Shared Space. In *Proceedings of ACM SIGGRAPH'97,* pp. 327–332, 1997

[2] Carolina Cruz-Neira, Daniel J. Sandin, and Thomas A. DeFanti. Surround-screen projection-based virtual reality: The design and implementation of the cave. In James T. Kajiya, editor, *Computer Graphics (SIGGRAPH '93 Proceedings),* volume 27, pages 135–142, August 1993.

[3] Han, J. Y. Low-Cost Multi-Touch Sensing through Frustrated Total Internal Reflection. In Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology, 2005.

[4] I.E. McDowall and M.T. Bolas. New Development for Virtual Model Displays. *Computer Graphics*, 32 (2):32-34, May 1997.

[5] John A. Roese , Lawrence E. McCleary, Stereoscopic computer graphics for simulation and modeling, *Proceedings of the 6th annual conference on Computer graphics and interactive techniques,* p.41-47, August 08-10, 1979, Chicago, Illinois, United States

[6] J. Rekimoto and N. Matsushita, "Perceptual Surfaces: Towards a Human and Object Sensitive Interactive Display," *Workshop on Perceptual User Interfaces* (PUI-97), Banff, Alberta, Canada, Microsoft (October 1997), pp. 3032.

[7] Koutek M., van Hees J., Post F. H. and Bakker A. F. Virtual Spring Manipulators for Particle Steering in Molecular Dynamics on the Responsive Workbench. *In Eighth Eurographics Workshop on Virtual Environments, pp. 53-62, Eurographics Association, 2002.*

[8] Livatino, S, Agerbech, V, Johansen, B, Johansen, A. Designing a Virtual Reality Game for the CAVE *Proc. 4th Eurographics IC*, Catania, Italy, 2006.

[9] Livatino, S, Privitera F. 3D Stereo Visualization for Mobile Robot Tele-guide, *will be published in 2006*

[10] L.Lipton, Stereographics, Developers Handbook, *Stereographics Corporation, 1997*

[11] Marek Czernuszenko, Dave Pape, Dan Sandin, Tom DeFanti, Greg Dawe, Maxine Brown: *The ImmersaDesk and Infinity Wall projection-based virtual reality displays*. Computer Graphics, 31(2), 1997, pp. 46-49.

[12] P. S. Iversen, M. Jakobsen, J. B. Johnsen, N. Larsen, T. B. Olesen and M. Thomsen. CAPE: Cave Pointing Environment. 2000, *Department of Computer Science, Aalborg University DAT3.*

[13] V. Agerbech, A. Johansen, B. Johansen, *The Jedi Training Remote*. Tech. rep., Aalborg University Copenhagen, Denmark, 2005.

[14] Wolfgang Krüger, Christina-A. Bohn, Bernd Fröhlich, Heinrich Schüth, Wolfgang Strauss, and Gerold Wesche. The responsive workbench: A virtual work environment. *IEEE Computer*, pages 42–48, July 1995.

[15] Wolfgang Krüger and Bernd Fröhlich.The Responsive Workbench, *IEEE Computer Graphics and Applications*, vol. 14, no. 3, pp. 12-15, May/June, 1994.

[16] Thomas B. Moeslund, Moritz Störring and Erik Granum. Vision-Based User Interface for Interacting with a Virtual Environment. *9th Danish conference on pattern recognition and image analysis,* 2000.

[17] www.fas.org/spp/military/docops/usaf/2020/app-v.htm (May 20th 2006)

[18] Computer Science Department at the Naval Postgraduate School in Monterey, California www.cs.nps.navy.mil/people/faculty/capps/4473/projects/mag-track/full.html (May 10th 2006)