MASTER

Performance evaluation of a multiple-user eye-tracking system for 3D displays

van der Ven, R.C.G.M.

Award date:
2009

ɪe universiteit eindhoven

/ department of technology management

TU/e technische universiteit eindhoven

Institut
Nachrichtentechnik
Heinrich-Hertz-Institut

Performance evaluation of a multiple-user eye-tracking system for 3D displays.

R.C.G.M. van der Ven
0553796

Master's Thesis
December 2008

Graduation committee:
First supervisor TU/e: Wijnand Ijsselsteijn
Second supervisors TU/e: Marc Lambooij
Maurice van Beurden
Supervisor HHI: Klaus Hopf

Eindhoven University of Technology –
Department of Industrial Engineering and Innovation Sciences
Educational Program Technical Innovation Science
Master's Program Human Technology Interaction

Preface

The following report describes the results of an internship of six months at the Heinrich-Hertz-Institute in Berlin. This internship was part of my Master thesis at the Eindhoven University of Technology at the Department of Technology Management with master Human Technology Interaction. During my internship I received support from several people to whom I want show my appreciation here.
First I would like to thank the Heinrich Hertz Institute as a company for providing me with the opportunity to spend six months in Berlin. At Heinrich Hertz I would like to thank Klaus Hopf and Frank Neumann for their guidance.
Also I thank dr. Wijnand IJsselsteijn, ir. ing. Maurice van Beurden and ir. ing. Marc Lambooij, for their feedback during and after the internship.
Ans last but not least, special thanks goes out to my parents, for the support they gave me while I was abroad and while writing this report in the Netherlands. Their encouragements have helped me to bring my education in Technology Innovation Sciences to a good ending. Also the help and advice of a good friend of mine, Roderick Hoek, may not be left unmentioned.

Ralf van der Ven, October/November 2008

Abstract

The Heinrich Herz Institute has developed a multiple user eye tracking system for a 3D display within the EU-funded MUTED project. To become commercially feasible, extensive testing of the system is needed to assure that the performance is at least comparable to an existing 2D display. The MUTED display employs video based tracking of the viewers' eyes. For an optimal functioning of the tracking, it is of importance to know what reasonable environmental circumstances may be expected under which the system will be used.

This research consists of two parts. The first part focusses on retrieving a set of environmental factors under which the performance of the MUTED multiple user tracking system will be tested. The factors that are under consideration are lighting conditions, viewing distance, number of users and user movements. By performing an observational study, the different ranges for those environmental factors were retrieved. The results of the observational study are used for the second part of this research, the performance study. This study was carried out to test the accuracy of the MUTED system under the different ranges for the environmental factors. The accuracy was judged by the error of the system. This error was calculated as the difference between the eye positions found by the MUTED system, compared to the eye position found by a reference system. This reference system was regarded to detect the exact eye positions. The larger the error between both systems, the less accurate the system works. During the analyses of the data files that resulted from the performance study it turned out that the user movements and number of users could not be analysed due to problems with the data. Also it turned out that there was a significant difference between the error of the left eye and the error of the right eye. Therefore both eyes where analysed separately. Regarding the lighting conditions the accuracy of the MUTED system is not dependent on an increase or decrease of lighting. Viewing distance only has an effect on the z-direction of the right eye. Also for every condition the error that is outside the critical limits of +/- 5 mm for the x-values and +/- 10 mm for the z-values was calculated. It turned out that for this moment especially the 5-50 lux condition does not function accurate at a viewing distance of 30-100 cm. Since this viewing distance and lighting conditions correspond to the values found for the medical environment this could be problematic. Further research is necessary to determine if the user is tracked incorrectly for a small part of the trial (for example at the beginning) or throughout the whole trial. Also analysis of the spread of the eye positions as found by the MUTED system can help to determine if, even if an eye is found on the incorrect position, the MUTED systems holds this position within the defined critical limits. This study can be used as an important guideline for further research.

Index

1. Introduction

1.1 3D view in a display

In the last century a lot of progress has been made in the area of displays. Displays have become technologically more advanced and commercially more attractive to the consumer. The next development in the generation of displays will be 3-Dimensional (3D) displays.
Within the Advanced Three-dimensional Television System Technologies (ATTEST) project several participants contributed in the development of a single user 3D display. In the EU-funded Multiple-User 3D Television Display (MUTED) components of ATTEST are used and extended to create a multiple user 3D display that can be introduced into the market.
A develop duch a display it needs to comply with various requirements (Surman, Sexton, Hopf, Lee, Bates, IJsselsteijn & Buckley, 2006). The viewer must have the ability to move freely within the room, while still being able to see 3D and multiple viewers must be supported. Also, the display has to be auto-stereoscopic, which means the viewer does not have to wear any viewing aids to see a 3D view.

1.1 Definition of the problem

In general, the introduction of the MUTED 3D display can only be a success if the viewing comfort of the display is at least comparable to an existing 2D display (Ijsselsteijn, Seuntiens & Meesters, 2002). This inherently means that the multiple user tracking system used in the MUTED 3D display must provide accurate tracking of all viewers in three dimensions under all environmental circumstances. So far some research has been done on environmental aspects that could be of influence on the accuracy of the multiple user tracking system, such as lighting or movement, (Liu, Conomis, Zhu & Pastoor, 2001; Zhu & Ji, 2005; Jones, Lee, Holliman & Ezra, 2001; Quante & Hopf, 2003). There also is some information available about the different values of those aspects within different environments (Beauchemin & Hays, 1996; Berguer, 2007; Krotz & Eastman, 1999). But so far no research has been done to investigate under which environments the MUTED 3D display will be used and determine the different values of the aspects that are of interest.

1.2 Objective

The main objective is to investigate the influence of different environmental factors on the accuracy of the multiple user tracking system. This accuracy is judged by the error of the multiple user tracking system. The error is calculated as the difference between the real eye position and the tracked eye position. The larger the error, the less accurate the multiple user tracking system functions.
To achieve a representative set of environmental aspects under which the accuracy can be tested, this research focuses on the two environments where it is most likely that the MUTED 3D display will be introduced first; a medical environment and an entertainment environment. In the first part of this research an observational study was conducted within both environments to determine different ranges for the environmental aspects. This set of representative ranges is used in the second part of this research, the performance study. By examining the influence of the different ranges for the environmental factors on the error of the multiple user tracking system, it is possible to determine under which circumstances the MUTED system functions accurate. Neumann, Hopf & Przewozny (2007) have defined first target values for the error to be +/- 5 mm for the x and y-axis and +/-10 for the z-axis. Also the average time it takes a person to notice a difference in its view is 50 ms (Adelstein, Lee & Ellis, 2003). So, for this research it is defined that the systems does not work accurately when the error is larger than the defined limits for longer than 50 ms. This inaccuracy would be perceived as discomforting due to the visual strain (Ijsselsteijn et al., 2005).

1.3 Structure of this report

This report describes the research that has been carried out to determine the influence of different environmental factors on the accuracy of the MUTED multiple user tracking system. The background of the MUTED display is discussed in Chapter 2. Several tracking principles are reviewed in Chapter 3, after which the implementation of those principles used in this research are discussed. Chapter 4 describes the aim of this project. The observational study is considered in chapter 5. This study uses a qualitative approach to retrieve a set of values for the environmental factors of interest. The procedure used for this study is explained and the results are presented. This chapter finishes with a discussion. Chapter 6 considers the performance study, in which the results of the observational study are used to test the accuracy of the MUTED multiple user tracking system under the found values. The used method and results are presented after which this chapter ends with a discussion. Finally, chapter 7 discusses some general limitations and recommendations.

2 The MUTED display

There are three methods for creating an auto-stereoscopic view in a display: holographic, volumetric and direction multiplexed. It is discussed that volumetric and holographic methods potentially give the best 3D experience (Meesters, Ijsselsteijn & Seuntiens, 2004; Halle, 1997). In volumetric displays the image is formed within a certain space, while in holographic displays the image is formed by a reconstructed wavefront that is identical to the original scene (Surman et al., 2006). The advantage of volumetric and holographic techniques is that they inherently support the requirements mentioned earlier: auto-stereoscopy, freedom of movement and multiple viewers. But both techniques also have drawbacks such as the amount of data that is needed for holographic images and limited resolution for volumetric images (Meesters et al., 2004; Benzie, Watson, Surman, 2007). Direction multiplexed displays may have their limitations that generally can be related to inconsistent depth cues which cause visual strain (Ijsselsteijn, Seuntiens & Meesters, 2005). Also direction multiplexing displays do not inherently support freedom of movement and multiple viewers. Despite those drawbacks, for this moment direction multiplexing is most feasible and commercially attractive for creating an autostereoscopic view in a display (Meesters et al., 2004). This technique is therefore used in the MUTED display.

2.1 Direction multiplexing

Direction multiplexing refers to a process where two separate signals are sent to the appropriate eye and thereby creating a 3D view. This can be achieved by positioning the viewer on a fixed position. Due to those fixed positions, the eye positions are known and the display can send two different images to exactly those eye positions. This technique restricts the freedom of movement of the viewer and is therefore unsuitable for our purposes (Dodgson, 2005). Another possibility is to locate the viewers' head position and use this position to direct two different views regions, called exit pupils, to both eyes. This locating of the user position is done with a so-called tracking-system. There are different ways to do this and every company or institute involved in research at 3D display techniques have their own approach. Interested readers are referred to Holliman (2005) for an outline of some of those approaches. Within the MUTED display the steering of the different views is done by spatial multiplexing. Another method to steer the exit pupils is using parallax barriers, which directs the light by using thin, black illumination lines. This technique suffers from the limits that the light throughput from the display is reduced to approximately 50% and it therefore puts large constraints on the viewing experience (Surman et al., 2006a). For this reason the parallax barriers technique is unsuitable for our purposes and spatial multiplexing is used.

2.2 Spatial multiplexing in general

In spatial multiplexing the light source is diffracted to create different exit pupils. This exit pupil is formed by placing a lens in front of a vertical light source and thereby creating a real image at the centre of the exit pupil (Figure 1). The light source is then seen within the shaded regions by the observer across the complete screen area (Surman et al, 2006a). To achieve a 3D view an additional exit pupil can be created next to the existing one, whereby each exit pupil corresponds to one eye. But doing this causes an important drawback: the exit pupils cannot be formed across the complete viewing area, which is required for multiple user systems (Surman, Sexton, Hopf, Bates & Lee, 2006).

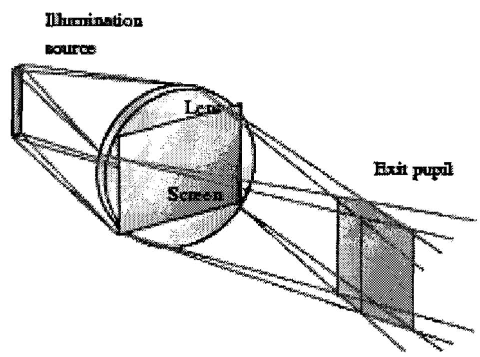


Figure 1 – Formation of exit pupils (Surman et al., 2006b)

This problem can be overcome by using a lens array that is placed in front of the illumination source (Figure 2). The alternate pixel columns each have a slightly different perspective and each lens focuses on the image of the pixel column behind it. By directing this information towards the different exit pupils a 3D view is created (Halle, 1997).

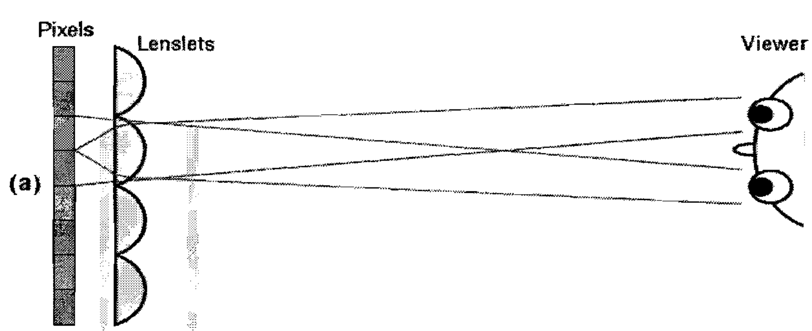


Figure 2 – Principle of spatial multiplexing (Dodgson, 2005).

To accommodate to the viewing distance of the viewer the lenses are vertically aligned with a certain pitch of a magnitude that is approximately twice the horizontal LCD pitch (Benzie et al., 2007). Besides this benefit of viewing freedom, spatial multiplexing is commercially attractive due to their manufacturability and low costs (Holliman, 2005). A disadvantage of the technique used in the

MUTED display is that motion parallax, the change of view of an image caused by movement, is not supported. This could be solved by adjusting the views of each scene to every movement a viewer makes. But according to Benzie et al. (2007) "the presentation of two images only can be a positive advantage in a television system as it enables the simplest capture and transmission, and places the smallest demands on the display in terms of the amount of information that has to be displayed."

2.3 Spatial multiplexing in the MUTED display

The MUTED display as illustrates in Figure 3 uses a holographic laser projector which is controlled by a multiple user tracking system. Two optical lens arrays, each corresponding to one eye, convert the input from the laser into exit pupils. Those exit pupils are used by the LCD to form a viewing image (Surman, Sexton, Hopf, Lee, Buckley, Jones & Bates, 2008). Therefore the tracking system is a critical element in the MUTED display. Incorrect tracking would result in exit pupils that are created on the wrong position and the viewer would not be able to see a proper 3D image.

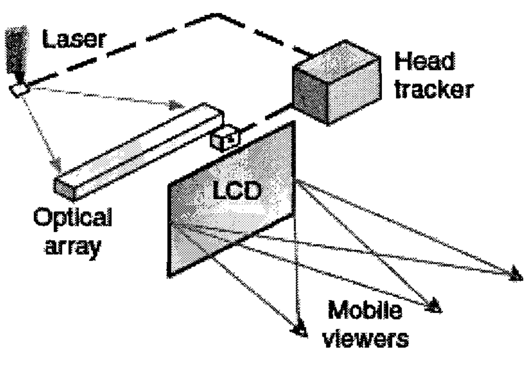


Figure 3 – Schematic diagram of the MUTED display (Surman et al, 2006b)

It must be noted that with 'holographic' the method of projection is meant, and not the projected image (Surman et al., 2006). The holographic laser projector produces a series of spots which correspond to the required illumination pattern. This pattern is spread over the complete width of the elements of the optical array by a diffuser.

Figure 4 illustrates a single element of an optical array. The Fresnel lens directs this light beam towards a soft aperture. The exit pupils are formed by the collimated exit beam, which results from the light beam that is channeled from the soft aperature towards the curved front surface. This surface functions as a vertical diffuser (Buckley, Corbett, Surman & Sexton, 2008).

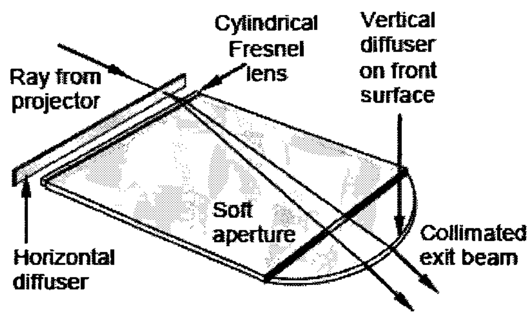


Figure 4 – Single element of the optical array (Buckley et al., 2008)

One series of illumination spots for the LCD display as shown in Figure 5 is produced by a single element of the optical array. Both horizontal and vertical diffusers of the optical element determine the spatial intensity (Buckley et al., 2008).

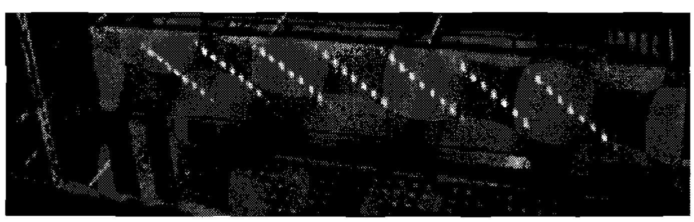


Figure 5 – Illumination spots produced by the projector (Surman et al., 2008)

A possible drawback from using a holographic projector could be inconsistent illumination due to the aberrations in the optical array. These aberrations are caused by manufacturing tolerances in the construction of the optical array. This problem is solved in the MUTED display by controlling the illumination spots and correction of field dependent phase errors (Buckley et al., 2008). This principle of the holographic projector has the advantage that only the necessary illumination spots which can be seen by the involved viewers are illuminated (Surman et al., 2008).
As mentioned the tracking system is critical for the functioning of the display. Therefore the multiple user tracking system used in the MUTED display is discussed in more detail in the next chapter.

3 The MUTED multiple user tracking system

There are different principles available that could be used for the tracking of the viewer. This chapter discusses those, after which the principle that is implemented in the multiple user tracking system of the MUTED display is explained in more detail.
To determine if the multiple user tracking system used in the MUTED display tracks the eye positions of the viewer accurately, those eye positions need to be compared to another system that measures the exact eye position. This is the 'reference-system', which is the participant of the final part of this chapter.

3.1 Tracking principles

There are different principles to track the positions of the users: electromagnetic, electromechanical, acoustic, inertial, optical and video based tracking. They all have their own unique approach, advantages and disadvantages.

3.1.1 Electromagnetic tracking

Electromagnetic tracking is one of the least expensive and simplest technologies (Baratoff & Blanksteen, 1993). It uses transmitter coils that are attached to the target. A magnetic field is then generated by sequentially sending current through these transmitter coils. The receiver then measures this electromagnetic field. The position of the target is determined by comparing the strength of the received signal to the strength of the sent signal and the orientation is measured by comparing the received signals (Kindratenko, 1994). Environments that contain metal or other devices that produce magnetic fields can disturb this system very easily. Therefore it is advised to avoid this tracking principle for accuracy sensitive applications (Milne, Chess, Johnson & King, 1996). Another drawback is that the system is restricted to a few cubical meters (Kidratenko, 1994).

3.1.2 Electromechanical tracking

This principle uses a mechanical connection between a reference point and the target. By placing encoders at the joints of this mechanical connection and measure those encoders relatively to a reference point, the position and orientation of the target can be determined (Baratoff & Blanksteen, 1993). An advantage of the system is they are cheap and highly accurate (Gibbs, Arapis, Breiteneder, Lalioti, Mostafawy & Speier, 1998). Their main disadvantage is the limited range of this system and the mechanical arm may limit motion (Stone, 1996).

3.1.3 Acoustic tracking

Acoustic tracking uses ultrasound sources for position measurements and orientation of the target object. This can be done by comparing the phase of a sent signal to that of a reference signal, the so-called phase-coherence tracking (Stone, 1996; Vallidis, 2002). This approach works by periodic updates of position instead of continuously measuring absolute position, which could cause error accumulation over time (Baratoff & Blanksteen, 1993).

Another way is by measuring the time of flight of a sound. Sound has a relatively fixed speed and the position and orientation of the target can therefore be calculated by multiplying the time of flight by the speed of sound (Vallidis, 2002). This principle is also referred to as time-of-flight tracking. The main disadvantage of this approach is the susceptibility to noise in the environment. Aspects such as temperature, barometric pressure, and humidity could have a big influence in the tracking results (Baratoff & Blanksteen, 1993; Stone, 1996; Vallidis, 2002).

3.1.4 Inertial tracking

The principle of inertial tracking uses special sensors, accelerometers, which measure the acceleration of the device compared to a reference point to determine the position of the target. The orientation is measured with rate-gyros, which measure how the angular velocity of the target changes (Foxlin, Harrington & Altshuler, 1998). The inertial tracking approach has several advantages: low latency, unrestricted range and no occlusion problems (Foxlin et al., 1998). A disadvantage could be ascribes to the fact that noise on one of the sensors can not be distinguished from the signal. This could result in an error in the measurement of the calculated position and orientation, even if there is no motion (Vallidis, 2002).

3.1.5 Optical tracking

Optical tracking can be done in several ways. It can use a head mounted device, which has one or more cameras mounted on top. By placing a set of makers on fixed positions within the environment, the position and orientation of the target can be calculated (Baratoff & Blanksteen, 1993). The other way is to place the cameras on fixed positions in the room and place the markers on fixed positions on the head mounted device (Gibbs et al., 1998). Advantages of optical trackers are the short lags and the high update rate (Baratoff & Blanksteen, 1993). Also, compared to other tracking approaches they are less prone to noise and less restrictive (Gibbs et al., 1998). But when an obstacle stands between the sensor and source this will seriously influence the performance of the tracking (Baratoff & Blanksteen, 1993). Also, the real world coordinates of the markers must be known to measure the position and orientation of the target (Stone, 1996).

3.1.6 Video based tracking

There are various approaches for video based tracking, such as template-based, feature-based or appearance-based approaches (Lipton, Fujiyoshi & Patil, 1998; Neumann et al., 2007; Zhu & Ji, 2005). The template-based method a generic eye model is designed first. Then template matching is used to scan the image for the eyes (Zhu & Ji, 2005). In the feature-based approach the target is tracked by comparing the different video frames compared to find regions that have been changed (Lipton, Fujiyoshi & Patil, 1998). The appearance-based approach uses the photometric appearance of the eyes for detection (Zhu & Ji, 2005).

An advantage is that for all three approaches no viewing aids are required. Disadvantages are the computational complexity, problems due to occlusion and the limitations of camera optics, such as noise in images or limited depth of focus (Li, Taskiran & Danielsen, 2007; Yilmaz, Javed & Shah, 2006; Neumann et al., 2007).

3.2 The MUTED display

One of the conditions of importance for a 3D display is that no viewing aids are needed to view a 3D image (Surman et al., 2006a). Therefore the MUTED display is based on a technical setup a video based multiple user tracking technology. The technical implementation and the way in which disadvantages of the video based tracking technique are addressed are discussed next.

3.2.1 Hardware and software set-up

As mentioned in paragraph 3.1.6, one of the disadvantages of using the video based tracking technology is the limitations based on camera optics. Those limitations could restrict the detection areas of the multiple user tracking system (Neumann et al., 2007). Since freedom of movement must be supported, this could be problematic for our purposes.

For this reason the camera system used in the MUTED multiple user tracking system has a configuration with six cameras. The cameras are grouped into three different pairs: one left pair, one middle pair and one right pair. Each camera pair is focussing on a different maximal user distance. The principle is illustrated in Figure 4.

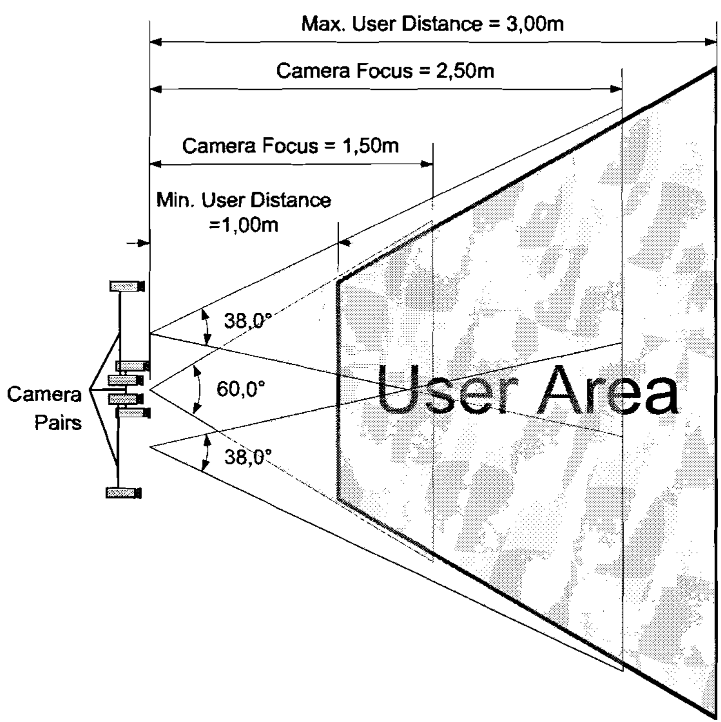


Figure 4 – Multi-camera approach (Neumann et al, 2007)

The focus of the middle camera pair is adjusted to the closer user distance. The focus of this camera pair comprises the whole user area. The left and right camera pairs are near the maximal user distance. These camera pairs are only covering a part of the user area.

Optical limitations could cause stereo distortions in the stereoscopic image. Differences in the stereoscopic image compared to the actual viewing of the scene cause stereoscopic distortion (Woods, Docherty & Koch, 1993). To address this, the camera system must be configured correctly. Therefore the video streams of the camera pairs are synchronized. This reduces stereoscopic distortion to a minimum by controlling for consistency between the video streams of the camera pairs (Neumann et al., 2007).

The rough modular concept of the tracking software is illustrated in Figure 5. The main modules are image capturing, image processing, stereo analysis and data combination. The system works in a data driven fashion. The system addresses the other limitation of the video based tracking technology, computational complexity, by using a multithreading approach that is easy manageable by the software engineer (Neuman et al., 2007).

The data capturing module handles the data source management, by assigning a reference to the data input when a data processing unit needs to access a specific data source.

The data processing unit handles all data processing, such as face detection, feature detection and the data-handling of external devices like the reference system. Next, the stereo analysis module generates the 3D data by combining the 2D data from the several data processing views. Finally the 3D tracking data of the view groups is collected by the data combination module. This module generates the tracking output.

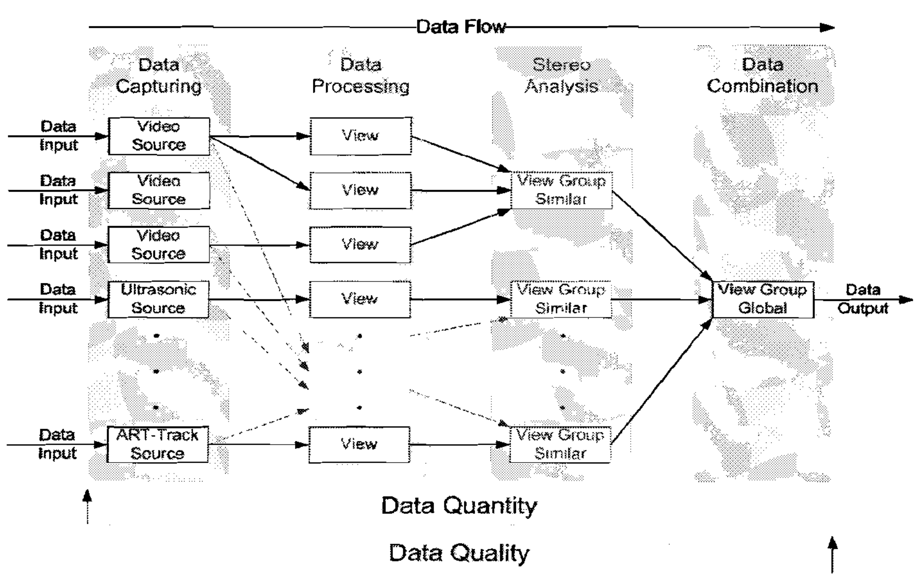


Figure 5 – Concept of the tracking software (Neumann et al., 2007)

3.2.2 Tracking in the MUTED system

The tracking technology is based on the detection of different facial features and the computation of the positions of the eye from those features. With the hardware and software setup explained, the technical implementation of this tracking technology needs some more attention.

The MUTED multiple user tracking system uses the appearance-based approach for initial head detection. This means that the system detects the head based on its photometric appearance (Zhu & Ji, 2005). After the initial head detection, the tracking system applies facial feature detection with an adaptive block matching algorithm on the initially found head for specific eye location measurements.

Figure 6 – Example of a block matching algorithm of the ATTEST single user system (Surman et al., 2006a)

The facial feature detection uses the reference eye patterns that resulted from an initialization phase, in which general characteristics of the human eye from various people are learned (Talmi & Liu, 1999). After the initialization the tracking system has training data with different eye patterns of various sizes. Then the adaptive block matching approach compares this training data with the current image (Surman et al., 2006a).

This results in squares that are positioned on the eyes as in Figure 6. As can be seen in those three photos, the algorithm used in the ATTEST single user tracking system works under different illumination and with different backgrounds. Also wearing of glasses is supported. Although the MUTED multiple user tracking system will be able to automatically adjust the brightness and gain of the view to the lighting conditions in the room, the system has not yet been tested on the aspects of lighting conditions and glasses yet.

Although Figure 6 shows only single users within one view, the same principle for facial feature detection with a block matching algorthm is applied in the multiple user tracking system to detect more then one user.

3.2.3 Calibration of the MUTED system

Within the coordinate system of the MUTED multiple user tracking system the calibration point (0,0,0) lies in the left top of the rack where the cameras are placed on, seen from the position of the participants. The z position is the distance from the eyes to the calibration point of the tracking system. The x position represents how far the eyes are situated left or right from the calibration point (right gives a plus value, left a minus value). The height of the users eyes compared to the calibration point can be evaluated by the y-value. A minus value means the user is sitting under the (0,0,0) point of the system. All is illustrated in Figure 8.

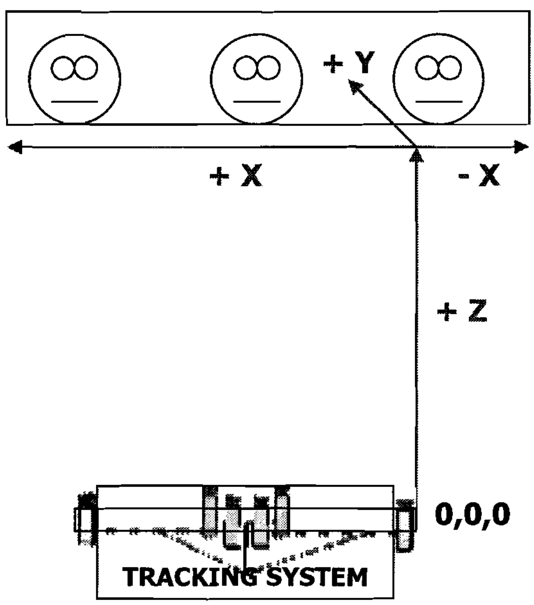


Figure 8 – Coordinate system of the MUTED multiple user tracking system

3.3 The reference system

To judge the performance of the MUTED multiple user tracking system, it has to be compared to a reference system which is regarded to resemble the exact eye positions of the viewers. This reference system uses the principle of optical tracking, because this technique has short lags, high update rates and is less prone to noise compared to other techniques. The next part discusses the technical setup and shows how the disadvantages of optical tracking, occlusion problems and the fact that the real world coordinates must be known to measure the position and orientation, are addressed.

3.3.1 Technical set-up

The technique used for the reference system is developed by the German Advanced Realtime Tracking (A.R.T.) Company. Figure 7 shows a schematic overview of the system.

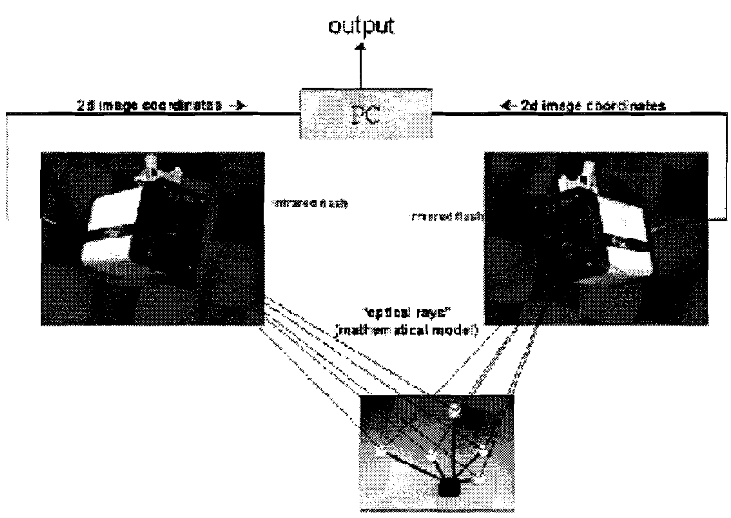


Figure 7 – Overview of the reference system

The user that shall be tracked is equipped with 5 light reflecting markers, which are placed in a different configuration for every user. This is done to discriminate the different users in the tracking phase. The tracking cameras scan the room with an infrared flash, detect the light that comes from the markers and calculate the 2D image coordinates with an accuracy of 0,01 mm (according to the D-Track manual).

These 2D coordinates are handed over to the stereo analysis module of the MUTED multiple user tracking system, which calculates the coordinates that describe the position of the body carrying the markers.

3.3.2 Calibration of the reference system

During calibration the software fixes a local coordinate system for the room and the different body marker configurations within this room. To measure the exact position of every viewer involved in this coordinate system, a special "marker-hat" was developed. As can be seen in Figure 8, every "marker-hat" has its own configuration of reflective markers, which makes it possible to uniquely identify the different users. During pre-testing it turned out that the measurements of the MUTED multiple user tracking system are not influenced by the wearing of those "marker-hats". Also it seems that short occlusion, for example one viewer crossing the view of another viewer, did not have an influence on the tracking data. Long occlusion was resolved by placing the "marker-hats" on the head of each viewer in such a way that the reflective markers are always visible to the reference system.

Figure 8 – The "marker-hats", each with a different configuration of the retro-reflective markers to distinguish the different users.

Figure 9 – Special "marker-stick" to define the position of the eyes.

With the "marker hats" the position of the viewer can be measured, but the exact position of the user's eyes in relation to this "marker hat" still is unknown. This problem is resolved with another calibration procedure. With a "marker-stick" the position of the eyes was measured. The end points of this "marker-stick" are placed on the middle of each eye (Figure 9). In the stereo analysis module, a transformation algorithm is calculated to determine the position of the markers on the "hat" in relation to the markers on the "stick". With this transformation algorithm, the relative position of the eyes compared to the "hat" can be calculated without the "stick". Those eye positions are transformed to the same coordinate system as used by the MUTED multiple user tracking system. The calibration with the "stick" must be done before and after the measurement, to ensure that the "hat" has not been moved during the measurement.

4 Environmental factors

Defining the requirements for the MUTED multiple user tracking system to decide which aspects need further development is an important step towards building a high quality 3D display system. This should result in a detailed description of those aspects, "sufficiently articulated to provide technical guidance for system specification" (Carroll, Rosson, Chin & Koenemann, 1998). To reach this description, first the factors of the medical and entertainment environments that need to be investigated are chosen. Then the available literature of those factors in both environments is discussed.

4.1 Factors of interest

The MUTED tracking system is developed to support the tracking of multiple users (Surman et al., 2008). Therefore the first factor of interest for this research is the average number of users in both mentioned environments.

Also freedom of movement must be supported (Surman et al., 2006a). To acknowledge this, it is of importance to have a good description of the different movements users make while using the MUTED system (Liu et al., 2001).

Zhu & Ji (2005) and Liu et al. (2001) concluded that under low illumination, it is possible that the appearance based tracking approaches as used in the MUTED multiple user tracking system may not work very well. It is possible that the camera images are so dark that the tracking software is unable to distinguish the necessary facial features. The same way of thinking could be applied with high illumination and images that are too bright. Exact lux values are not provided. The single user system works correctly under different illumination circumstances (Surman et al., 2006) and the MUTED multiple user tracking system will be able to correct for lighting circumstances by automatically adjusting the brightness and contrast of the cameras. Although it is assumed that based on these automatic adjustments, the MUTED system will function better; this has not been tested yet.

As concluded in paragraph 2.3, the spatial resolution of the display is controlled by the magnitude of the lens pitch. This lens pitch is steered by the MUTED multiple user tracking system. According to Pastoor (1993) the spatial resolution is directly influenced by the viewing distance, which is therefore also of interest for this research.

4.2 Available literature

The environmental factors that are considered in this research are: number of users, user movements, lighting circumstances and viewing distance. This paragraph gives an overview of the available literature regarding those environmental factors. For the medical environment the focus lies on the operating room and the medical imaging room, since these are the main areas for the application of the 3D technology (Kitasaka, Mori & Suenaga, 2005).

Number of users

The maximum number of users defined for the MUTED multiple user tracking system is four (Neumann et al., 2007). Information is available about the different actors that are present within an operating room, on average one or two surgeons, one or two nurses and one anaesthesiologist (Berguer, 1999; Hajdukiewicz, Vicente, Doyle, Milgram & Burns, 2001).

But since the nurses do not necessarily have to watch the 3D display, and therefore have to be tracked by the tracking system, three seems a reasonable number of users. For the medical imaging room a maximum of three users should be supported (Naggy, Siegel, Hanson, Kreiner, Johnson & Reiner, 2003). The average number of users within an entertainment environment is unknown, but a typical family size of four seems reasonable.

Lighting circumstances

Beauchemin & Hays (1996) conclude that on a cloudy day a bright hospital room was measured at 500 lux and a dull room at 200 lux. On a bright day with light clouds those values where measured at 500 and 1700 lux respectively. Artificial light usually ranges between 2500 and 10000 lux.

For a medical imaging room a lighting level of 150-400 lux is required when using a 3D display, with lower light levels improving the contrast of the screen compared to the environment (Foord, 1999). In a room where a minimally invasive surgery takes place there must be as little lightning as possible, to give a high contrast with the display and avoid depth cues that are provided by shadows on the display from off-axis illumination (Boppart, Deutsch & Rattner, 1999; Hanna & Cuscheiri, 2001). Specific lux values for a medical imaging room are not available.

Viewing distance

The MUTED system has a viewing range of 1-3m (Neumann et al., 2007). For the medical environment the viewing distance is partly dependent on the size of the monitor to avoid neck and eye strain (Foord, 1999). But in general a viewing distance of 50-100 cm seems reasonable for both the medical imaging room and the operating room (Zheng, Janmohamed & MacKenzie, 2003; Prabhu, Gandhi & Goddard, 2005). Pastoor (1993) concluded that a distance of between three and four times the picture height seems to be advisable for an entertainment environment.

User movements

For the operating room, Berguer (1999) concludes that the work postures during surgery are for more than 70% static. The movements during the other 30% are left unmentioned. For a medical imaging room also most movements seem static, except when a colleague is consulted or while teaching (Naggy et al., 2003). Exact descriptions of movements in these situations are not given. Also for an entertainment environment information about possible user movements is scarce. As Crabtree and Rodden (2004) conclude that "there is at least an order of magnitude more papers about offices and workplaces than about homes and consumers (and the latter totals only a few dozen publications over the last decade).

From the above it is evident that, although some information is available, extra work needs to be done to complete the set of requirements that are necessary to test the system.

5 Aim of this project

Before the MUTED display can be successfully introduced in commercial applications, extensive testing is needed. The display must be functioning at least comparable to an existing 2D display, so it has to function properly within the different environments where 2D displays are used.
For the MUTED display one of the most important parts that need to be tested is the multiple user tracking system. This system detects the eye pairs of all involved users, after which the system can send two different signals to all involved eye pairs. This inherently means that this tracking of the eyes must be as accurately as possible. Therefore first it needs to be defined what is regarded as accurate. The accuracy is evaluated according to the error of the system. The error is defined as the difference between the measurement and the true value. In this research the error therefore is calculated as the difference between the real eye position and the tracked eye position. When this difference is larger then +/- 6mm (x and y-axis) or +/- 10 mm (z-axis) for longer then 50ms, the MUTED multiple user tracking system is regarded to function inaccurately (Neumann et al., 2007; Adelstein et al., 2003). The longer the system works inaccurately, the larger the chance that visual strain will occur, such as crosstalk (when the view of one eye leaks through the view of the other eye), shear distortion (when the viewer changes his viewing position and the image seems to follow this movement) , the picket fence effect (when the black mask between the columns of pixels in an LCD create a vertical banding in the image) and image blur (when the left and right image differ in color, luminance or sharpness) (Ijsselsteijn et al., 2005).
To evaluate the accuracy of the MUTED multiple user tracking system it is of importance to retrieve, for both environments, a set of values for all environmental factors that are of interest in this research: viewing distance, user movement, lighting circumstances and number of users. With these values it can be evaluated under which conditions the system works accurately. Then it can be checked if the MUTED system works properly within the environments where it will likely be introduced first, the medical and the entertainment environment.
So far some research has been done about the influence of illumination on the accuracy of eye tracking (Zhu & Ji, 2005; Liu et al., 2001; Surman et al., 2006). Under low illumination, the appearance based approach as used for the MUTED system may not function properly. What is not known so far are concrete values for the environmental factors within a medical and an entertainment environment. Also the influence of those values on the accuracy of the tracking system is an issue that has not been researched yet.

6 Observational study

6.1 Introduction

The observational study described in this chapter was performed to retrieve a set of values for the environmental factors of interest. The factors that followed from the theory to be relevant for this research are: viewing distance, user movements, number of users and lighting conditions. The focus of the observations lies on the environments of interest: the medical and the entertainment environment.

6.2 Procedure

Explorative research was conducted for all environmental factors with a qualitative analysis of the data. The information of the entertainment environment was obtained by observing five living rooms on an afternoon and evening. The participants are sampled on behaviour, which means that only attention is paid to someone who shows behaviour that is relevant for our research. This is also known as ad libitum sampling (Sande, 1999). For this study this means the user movements are retrieved by focussing on participants that interact with the display or with each other while viewing the display. The viewing distance was measured with a measurement lint from the spot between the participants' eyes to the middle top of the display. The lux values are measured with a lux meter from the eye positions of the participants, with the lux meter aimed towards the display. This is further referred to as the 'lux-value' of a participant. Besides the viewing distance, lux values, number of users and user movements, the observer also noted the placement of furniture. Although this did not come forward from the literature as a relevant environmental factor, it was decided that it would give a more complete impression of the room and the situations under which the other environmental factors are evaluated. The observations are done with use of a checklist which can be found in Appendix 1. Before the observations started, the observer asked the participant to watch the display as normal. The participants were not informed which aspects would be studied. As an addition to the observations, a short interview was conducted afterwards. This method still has the disadvantage that the participants could show reactivity; that is, the behaviour and answers would be influenced by the knowledge that it is being used for research. We tried to get as accurate descriptions as possible by observing and interviewing more than one person in every place visited. This has the advantage that participants can correct each other. For the medical environment two interviews were conducted. The interview questions are found in Appendix 2.

6.3 Results

6.3.1 Entertainment environment

For two of the places visited, the household consisted of two people without any children. In the other three household, the family consisted of two people with one or two children. Although there was some variability between the observations of the entertainment environments, for example the number of users or the placement of furniture, it was possible to distinguish some commonalities for the environmental factors within the observed living rooms.

Number of users

In the evening mostly 3 participants are watching. In the daytime the television isn't watched that often. Maximally two participants are watching and they pay the same amount of attention to the television as in the evening.

Lighting circumstances

In the evening luminance lies between a value of 10 and 30 lux. There is some variation within the same space, for example when the lux-value of a participant is around 60 lux and the other participant, who is sitting more close to a window or another light source, has a lux-value of around 200.

Viewing distance

Viewing distances are between 300 cm and 250 cm for the participants sitting the furthest away in the daytime scenario. The participants that are sitting 250-200 cm from the display are only viewing in the evening.

User movements

While watching television, participants don't tend to move much. In two of the observations one of the observed participants was switching his attention between the television and a laptop standing next to him. This movement corresponds with participants looking at each other while speaking and then turning back to the television. In more then one occasion, someone stood up to walk around the room. This was done in two different ways: the focus was kept on the television (for instance during an exciting soccer-match) or the focus was lost when the person stood up (for instance when receiving a phone call). Also at some point some additional people came walking in the room. They did not participate in the observation, but just walked by and watching the display while passing.

Placement of furniture

Although the differences regarding the placement of furniture where sometimes large between some of the observed living rooms, in general the furniture was placed as illustrated in Figure 10.

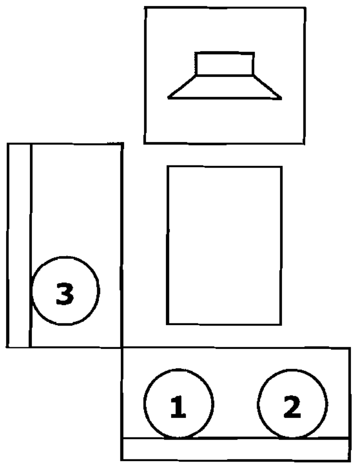


Figure 10 – Placement of furniture in the entertainment environment

In one occasion the couch where participant 3 is sitting on is replaced by a comfortable chair and was placed on the right side of participant 1 and 2. In the daytime participant 1 and 2 are watching as in Figure 10. In the evening all participants are positioned as illustrated.

6.3.2 Medical environment

Results for the medical environment are obtained through interviews at the Amsterdam Medical Centre (AMC) in Amsterdam. Two persons were interviewed: a radiologist, Dr. Stoker, and a surgeon, Dr. Bemelman. During the interviews both mentioned two rooms of interest for this research: an operating room during minimally invasive surgery and a medical imaging room. A 3D display is likely to be introduced here, since depth perception is seen as a potential benefit here.

Number of users

During minimally invasive surgery 2 users are watching one display at the same time. During medical imaging the number of users varies between 1 (when making a diagnosis) and 2 (when consulting a colleague or teaching a student).

Lighting circumstances

The lighting in the operating room can be divided into two different scenario's, based on the sort of operating that is done. In case of an open operation there needs to be a high lighting level, so an operating lamp is used. A minimally invasive surgery will acquire different lighting then during an open operation. Since there is a high focus on the screen it is of critical importance that the content on the screen can be viewed as good as possible. So for that the contrast with the rest of the environment must be very high. To create this effect it is important that the ambient lighting is around 10-50 lux.

Viewing distance

Viewing distances are for a medical imaging room 100-50 cm for the person sitting in front of the display and 100-150 cm for the person that walks towards the display to give advice or an opinion. Viewing distances during minimally invasive surgery are not mentioned.

User movements

In the operating room and the medical imaging room the focus lies on the screen and movements are static. The only movement a surgeon makes is watching left or right when a new surgery tool is handed over. The only movements in a medical imaging room are made when a person comes up towards the screen, focuses on the screen, gives his opinion and walks away.

Placement of furniture

The placement of furniture in the medical imaging room is illustrated in Figure 11 and the placement of furniture during minimally invasive surgery is shown in Figure 12.

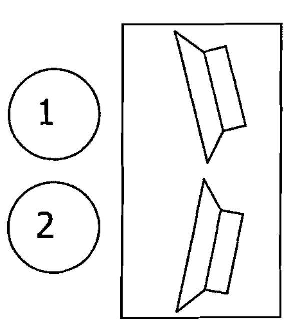


Figure 11 – Placement of furniture in a medical imaging room

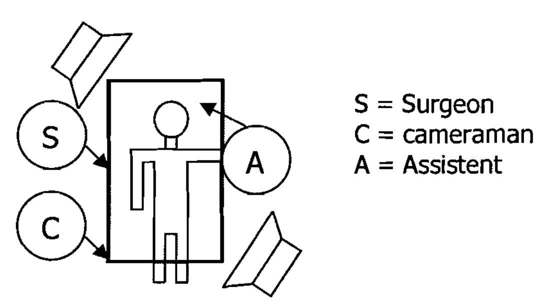


Figure 12 – Placement of furniture during minimally invasive surgery

6.4 Discussion

The interviews conducted for the medical environment confirmed the literature which stated that the medical imaging room and operating room are the places where a 3D display will be introduced first. The 3D technique will give the potential benefit of depth perception, which is currently judged by different depth cues in the 2D image. Therefore 'reading' depth in an image has to be learned. By using the 3D technology this learning process can be made easier and more precise depth estimations can be made.

From the interviews combined with the theory, the values of the different environmental factors came forward for the medical imaging, minimally invasive surgery and entertainment environment. In the following section it is discussed how those values are used to come to the conditions under which the MUTED multiple user tracking system will be tested and to which of the environments those conditions correspond.

6.4.1 Number of users and user movements

Regarding the number of users, at most 3 participants were watching the display. When combining this with the user movement of a participant walking away, it was possible to create a situation when only two people are watching (corresponding to the medical imaging room, the minimally invasive surgery situation and the daytime entertainment situation). Also a situation can be created that more than 3 participants are watching the display, by letting two people walk by and focus on the display shortly (corresponding to the evening entertainment situation when people are walking by). When no movement took place, the participants are static and focussed on the display. To summarize, the user movements resulted from this observational study are shown in Table 1.

Table 1 – User movements

Movement	Environment
Participant looks to his right	Minimally invasive surgery, open surgery, entertainment
Participant shortly walks away, without watching the display	Entertainment
Participant shortly walks away while watching the display	Entertainment
Participant walks towards the display, while watching the display	Medical imaging
Two participants walk by, while shortly watching the display	Entertainment

6.4.2 Lighting conditions

The different lighting condition that emerged from the observations and interviews are shown in Table 2. It must be noted that the observations are conducted in winter and in all daytime situations it was clouded. In correspondence with the research of Beauchemin & Hays (1996) it is reasonable to assume that on a sunny day the lux-values in a room can go from approximately 200 lux up to 400 lux.

Table 2 – Lighting conditions

Lighting conditions	Environment
5-50 lux	Minimally invasive surgery, medical imaging, entertainment (evening)
50-100 lux	Entertainment (daytime in a dull room on a cloudy day)
100-200 lux	Entertainment (daytime in a dull room on a sunny day)
200-400 lux	Entertainment (daytime in a bright room)
>400 lux	Open surgery

The illumination values for the medical imaging room did not correspond to the values of 150-400 lux as found in the literature. This can be explained by the fact that it is also discussed that illumination lower then 150-400 lux gives an improved contrast for the display compared to the environment, so more accurate depth estimations can be made (Foord, 1999). Since a lighting of 5-50 lux is relevant for all environments, this is regarded the most important lighting condition.

6.4.3 Viewing distance and placement of furniture

The placement of furniture is shown in Figure 13. By combining this figure with the different viewing distances as in Table 3, situations similar to the different environment are created. For example, when the viewing distance is approximately 50 cm, the display is so close to person 1 and 2, that the situation is similar to that of a medical imaging room. Although the viewing distances of 200-150 cm did not directly correspond to one of the environments, it is still mentioned to complete the set of ranges.

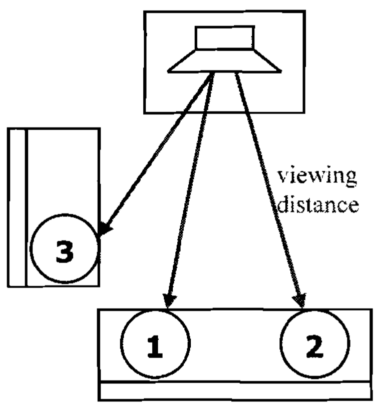


Figure 13 – Placement of furniture

Table 3 – Viewing distances

Viewing distance	Environment
250-300 cm	Entertainment
200-250 cm	Entertainment
150-200 cm	-
100-150 cm	Medical imaging
30-100 cm	Medical imaging, minimally invasive surgery, open surgery

7 Performance study

7.1 Introduction

The performance study was conducted to test MUTED multiple user tracking system under the values for the environmental factors as found in the observational study. With the results from the performance study it can be evaluated under which values the MUTED system works accurately and if those values correspond to the different environments in which the MUTED system is most likely to be introduced first. The results can also help to find possible other application areas for the system. For this study the research question is formulated as: do numbers of viewers, lighting circumstances, movements of the viewers and viewing distance have an influence on the accuracy of the tracking system? This research question is tested in the experiment as described below.

7.2 Method

7.2.1 Design

An study was designed to measure the accuracy of the MUTED multiple user tracking system as the dependent variable. The independent variables viewing distance, lighting conditions and user movements were varied independent of each other in a 5 x 5 x 2 (viewing distance, lighting conditions, user movements) design, resulting in 50 conditions. The values for the independent variables are summarized in Table 4. The user movements are combined into one 'dynamic' condition. In this condition, the user movements are placed sequentially behind each other. Before analysis, the data files will be split into parts to distinguish the different movements. The situation of two people that walk by is also used to analyse if this would distort the tracking of the three primary participants. In the same way the number of users is analyzed; when a user walks away, one or two are left within the view. This can then be compared to the static conditions with three users.

Table 4 – Conditions used in the experiment

Environmental factor	Values
Viewing distance	250-300cm, 200-250cm, 150-200cm, 100-150cm, 30-100cm
Lighting conditions	5-50 lux, 50-100 lux, 100-200 lux, 200-400 lux, >400 lux
User movements	Dynamic (Looking the the right, walking away without watching the display, walking away and watch the display, walking towards the display, two people walk by and watch the display shortly) Static

The experiment had a blocked design. First the lighting is held constant and the other two variables are varied. After a short break, in which the lighting settings are changed, the following block is conducted.
The independent variables are not randomised, since the performance of a mechanical system is tested. It is assumed that a system doesn't have sequence or order effects. Any sequence or order effects of the participants are cancelled out by pre-defining all the user-movements as precise as possible.

7.2.2 Stimuli

In the experiment, 5 participants voluntarily participated by accepting a personal invitation. All of them were employees working at the department of Interactive Media - Human Factors of the Fraunhofer Heinrich-Hertz-Institut. Four were men; one women. Two of them wore glasses; two wore contact lenses.

7.2.3 Equipment

The MUTED multiple user tracking system uses 6 Sony XC-999P cameras. The cameras are arranged in three pairs, whereby one pair was adjusted for observing the close distances to a focal length of 6mm and a camera basis of 120mm. These cameras can track the user up to 1,5m. The two pairs for the further distances are adjusted to a focal length of 12mm and a camera basis of 250mm. They are able to track the user up to a distance of 2,5m. The frame rate of the MUTED system is not constant.
For the optical reference system the ARTtrack2 system as developed by Advanced Realtime Tracking GmbH was used. This system has a frame rate of 60Hz and is able to track the user up to 4m.
The viewing distance was measured with a measurement lint and for the lux-values a lux-meter was used. For the lighting conditions 4 light sources are used. The average lux-values produced by these light sources are 120-140 lux (twilight lamp), 80-100 lux (twilight lamp), 690-730 lux (dimmable lamp) and 700-2000 lux (fluorescent lamp).

7.2.4 Procedure

The experiment was conducted in a relaxing room at the Heinrich Herz Institute. The placement of furniture in the room was set up according to the results from the observational study (Figure 13). The participants are placed in front of the display as shown in Figure 14. The optical reference system was placed on top of the camera rack for the MUTED system, as shown in Figure 15.

Figure 14 – Setup of the room

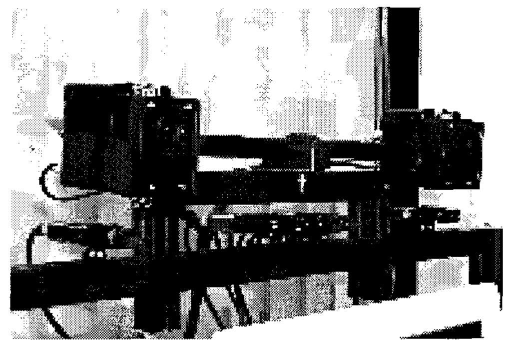


Figure 15 – Setup with the reference system

To make sure the movements are as identical as possible between the experimental conditions, every participant received a short training before the experiment started. In this training the different movements are practiced under full light. Every participant was given one or two movements of Table 4. The participants were also asked to keep the movements as similar as possible. So when someone is asked to walk away and afterwards come back, this is done in the same manner and in the same speed every time. The participants needed to keep focused on the screen, also when the instruction is read for that person. Before the instruction was read out, the name of the person was said and next the instruction followed. After a sign on the exact time as seen in Table 5 the instruction was carried out.

Table 5 – Instructions

Time	Person	Instruction
0.05	1	Look to your right
0.10	1	Look back to the screen
0.20	2	Walk to the right without watching the display
0.25	2	Walk back and look at the display when sitting down
0.35	3	Walk to the right while watching the display
0.40	3	Walk back to his seat while looking at the display
0.50	1	Stand up and walk towards the display, turn and walk back. When sitting down focus on the display
1:00	4 and 5	Walk to the right and stand next to person 1. When in place focus on the display

After these instructions, the participants had to undergo a calibration procedure as described in paragraph 3.3.2, to determine the eye positions compared to the marker-hat of the reference system. The lighting circumstances are set so that lighting in the room falls between 5-50 lux. The viewing distance, as measured from the middle of the rack of Figure 15 to the middle of the eyes of the two participants sitting in front of the display, was set to 300cm. Then the position of the third participant was measured, to keep the relative distance with the other two participants the same every time the viewing distance of 300-250cm is measured.
All participants had to watch the display for 15 seconds without any movement. This was the 'static' user movement. Then in the 'dynamic' user movement condition, all movements as mentioned in table 4 are placed sequentially behind each other with a 5 second break between each movement. This procedure was repeated for all viewing distances under the 5-50 lux lighting condition. After this there was a short break of approximately 10 minutes.
During this break the configuration of the light sources in the room was changed to create the next lighting condition. Also, the brightness and gain of the cameras are adjusted. This will be done automatically in the future, but since this function was not implemented at the moment this experiment took place it was decided to make those adjustments manually. This procedure for the user movement and viewing distances was repeated for all lighting conditions. The experiment ended with the same calibration procedure as the initial one. This was done to control for a change in the position of the "marker hat" during the experiment.

7.3 Results

The MUTED tracking software produces a text-file for each experimental condition, containing the data of both the MUTED multiple user tracking system and the reference tracker. The first columns contain the timestamp and frame index of the reference system and the measurement sensor for the calibration. The next 18 columns represent the values of the reference system for the x, y and z positions of both eyes for the three primary users. The last part of the data file contained the timestamp, frame index and x, y and z positions of both eyes for the three primary users of the MUTED multiple user tracking system. Before analysis could be performed on the data, first the data files had to be prepared to calculate the error between the reference system and the MUTED system for all x and z values of each condition. The y-position was left out of the data-analysis, since it was discussed in paragraph 2.2 that for creating the different exit pupils correctly; the x- and z values are of more importance than the y-values.

7.3.1 Preparing the data

In the data files, the reference system and the MUTED multiple user tracking system are in a different coordinate system, so the eye positions of the reference system have to be transposed. To reach this, Matlab R2007b was used to write a script to transform the data so that both systems are in the same coordinate system. The result from this script is further referred to as the 'transformed data file'. Another script was written to calculate the error of each frame for all conditions by substracting the x- and z-values of the reference system from the x- and z-values of the MUTED multiple user tracking system. With use of those error values per frame an average error for the x and z values of every user was calculated for every condition. The values for the measurement sensors showed that the 'marker hats' of the reference system did not move during the experiment.

While analysing the data some problems emerged. First the data of the MUTED system contained significant numbers of missing data where a user is not found, represented by 0-values. The average error for the x and z values is calculated for every condition over the frames where a person is found. Second, the timestamps of the MUTED system are different from the timestamps of the reference system. The timestamps of the MUTED system are not monotonically increasing. This was done, since different video streams are processed in parallel and the operating system is responsible for managing the different threads. Output is generated by the data combination module of the MUTED system when new tracking data from a video stream pair is received. This parallel computed data is written sequentially to the output file. Therefore the data can only be normalized on the timestamps of the MUTED system. These timestamps are not filtered and do not have a constant frame rate. Third, the data of a person is not fixed to one series of columns in case of the MUTED system. The reference system can uniquely identify all three persons, due to the 'marker-hats'. But the MUTED system does not fix every person to one column. Every camera pair of the MUTED system focuses on a different user distance. Movement between these different areas could result in the fact that the user is lost and found again by another camera pair, placing this data into another column in the data file. For this reason the data of the MUTED system was manually altered for the static movement conditions to make sure that all positions of one person are in the same column. It was discussed that this would also be theoretically possible for the dynamic user movements. However, several problems would have to be dealt with in combination. Due to the known position of every person, it would be possible to distinguish which of the columns belong to which person. But to make some analyses of the different movements, the data files need to be divided into different parts. Every part would then represent one of the movements made. Because of the variable frame rate, the fact that not every data file starts at the same frame as the original reference tracker file and the amount of missing data it is nearly impossible to divide those data files into different pieces correctly. Therefore analysis of the dynamic conditions would be very susceptible to error and further analysis was done on the static conditions. Person 1 is found in 8 out of the 25 conditions and within those conditions only for a short time. Person 2 is found in all of the 100-30 cm, 200-150 cm and the 250-200 cm distance conditions. Within the conditions where person 2 is found 3 data files also contained a reasonable amount of missing data. When comparing the data of person 3 with the data of the conditions where person 2 is found no trend emerged between both persons. Although this reduces the generalizability of the findings, person 1 and 2 are left out of the analyses. The analysis is therefore done on the data of person 3, who is found in every condition.

7.3.2 Difference between the reference and the MUTED system

It was examined if there is a significant difference between the reference system and the MUTED system. Therefore the transformed data files are analyzed, with the 0-values regarded as outliers. Both systems did not have a normal distribution. Although, the data of the MUTED system has heterogeneous variances, the reference system did not. Since during pretesting it turned out that the measurement of the reference system does not influence the measurements of the MUTED system, the Wilcoxon rank-sum test for two independent samples was performed. This is a non-parametric test for assessing whether two samples of observations come from the same distribution
For the x-values, the position measured by the reference system (M= 1083.51) significantly differ from the positions measured by the MUTED system (M= 1086.83), Ws= 3836772, p<0.001, r= -0.32. For the z-values, the reference system (M= 1906.9) measured significantly different positions than the MUTED system (M= 1892.02), Ws= 2915066.5, p<0.001, r= -0.39. So there is a significant difference between both system and further analysis is needed to determine if this error is dependent on lighting conditions or viewing distances. The technical relevance of this finding is researched further in paragraph 7.3.6.

7.3.3 Difference between both eyes

Next, the possible difference in average errors between the eyes in every condition is examined. This was done to determine if further analysis can be done over the complete data set, or that both eyes have to be analyzed separately.

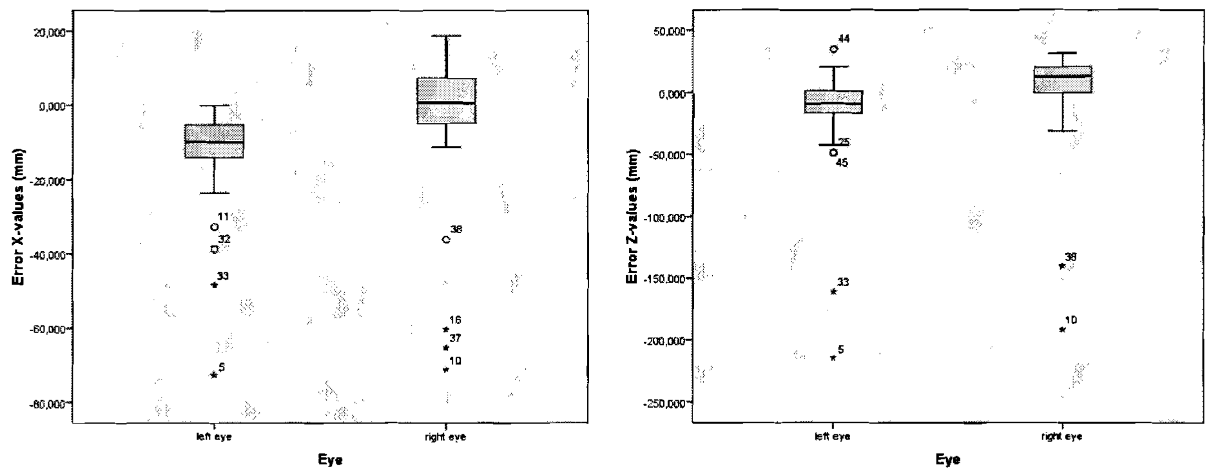


Figure 16 – Boxplots of the error of both eyes separately

When examining the box plots as shown in Figure 16, it seems that the median of the right eye is higher than the median of the left eye. The spread of the right eye seems higher for the x-values although, due to the relatively large number of outliers, this is discussable.

Since we are interested if the medians of the populations from which samples are drawn are identical, the nonparametric median test is performed. For the x-positions the average error for the left eye (*M*= -14.98) is significantly different from the right eye (*M*= -6,53), *Ws*= 460, *p*<0.001, *r*= -0.48. The z-positions also showed a significant difference between the left eye (*M*= -22.2) and the right eye (*M*= -4.6), *Ws*= 491, *p*<0.001, *r*= -0.4. Therefore, the influence of lighting conditions and viewing distance is analyzed for both eyes separately.

7.3.4 Effect of lighting conditions

The box plots of Figure 17 show that the spread is higher within the 200-400 lux conditions for both eyes. Although the effect is less in the case of the z-values this condition can possibly differ from the other lighting conditions. Besides the 200-400 lux condition, it seems that the error of the x-values of the left eye is higher than is the case for the x-values of the right eye.

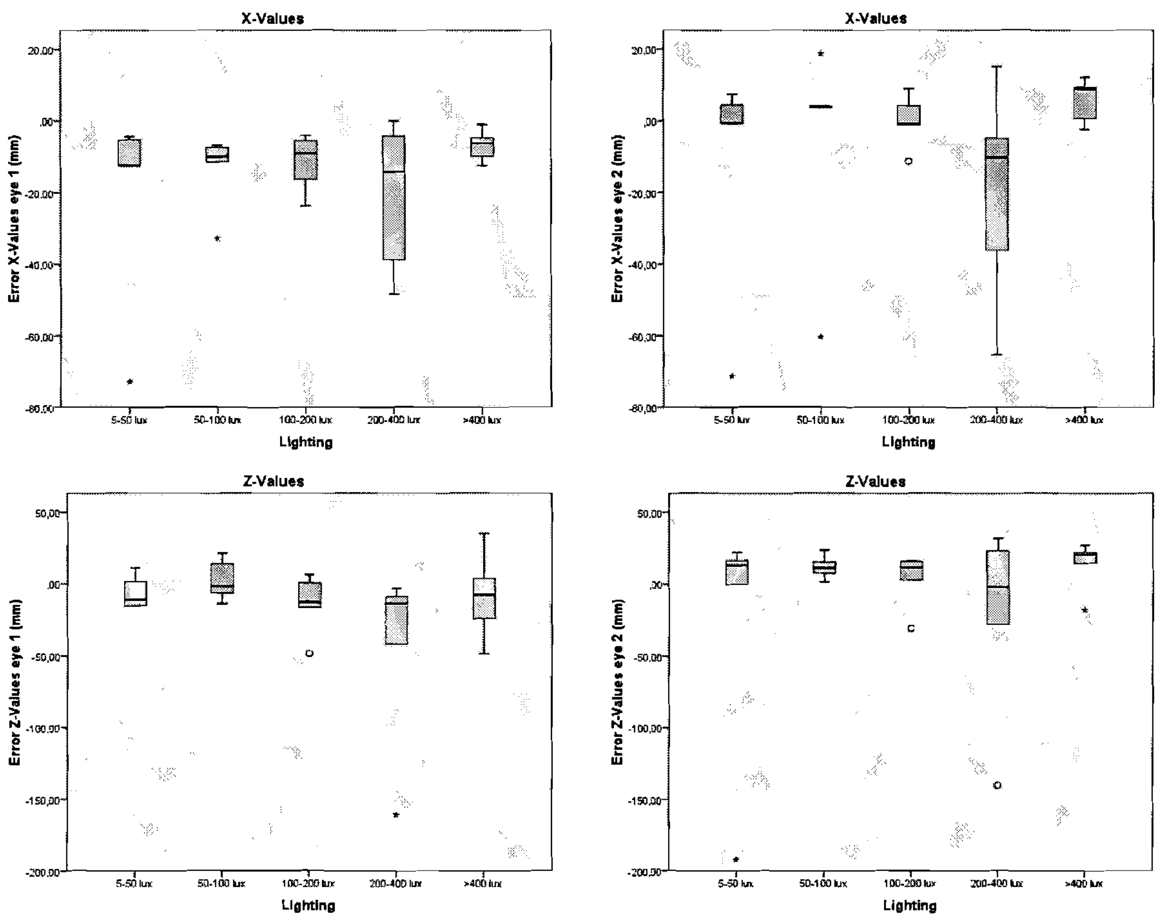


Figure 17 – Boxplots for the lighting conditions

Especially the medians for the z-values stay fairly close to 0. It must be noted that the box plots are made over only 5 averages per condition. Due to the number of outliers and the wide spread of the boxplots in some cases, interpretation is hard and susceptible to error. To analyze the effect of lighting conditions the Kruskall-Wallis test for several independent groups is used. This is a non-parametric test for for testing equality of population medians among groups. The test shows that there is no significant effect for lighting condition on the x-values of the left eye ($H(4)=1.15, p>0.1$), or the x-values of the right eye ($H(4)=3.759, p>0.1$). Also for the z-values of the left eye ($H(4)=3.729, p>0.1$) and the z-values of the right eye ($H(4)=1.543, p>0.1$) the lighting conditions do not show a significant effect.

7.3.5 Effect of viewing distance

The box plots of Figure 18 show that the spread is higher in the 30-100cm conditions than for the other viewing distances. Also for the x-values of the left eye the spread of the 250-300cm distances is larger then the other conditions. The 30-100cm condition seems to differ from the other condition in case of the z-values for both eyes. Since the median gets larger with an increasing viewing distance for the z-values of the right eye, a positive trend of viewing distances on the error is assumed. Beside the 30-100cm viewing distance, the medians for the other distances are fairly close to 0 in case of the x-values of the right eye and the z-values of the left eye. Again, it must be noted that the box plots are made over only 5 averages per condition. Due to the number of outliers, interpretation is hard and susceptible to error.

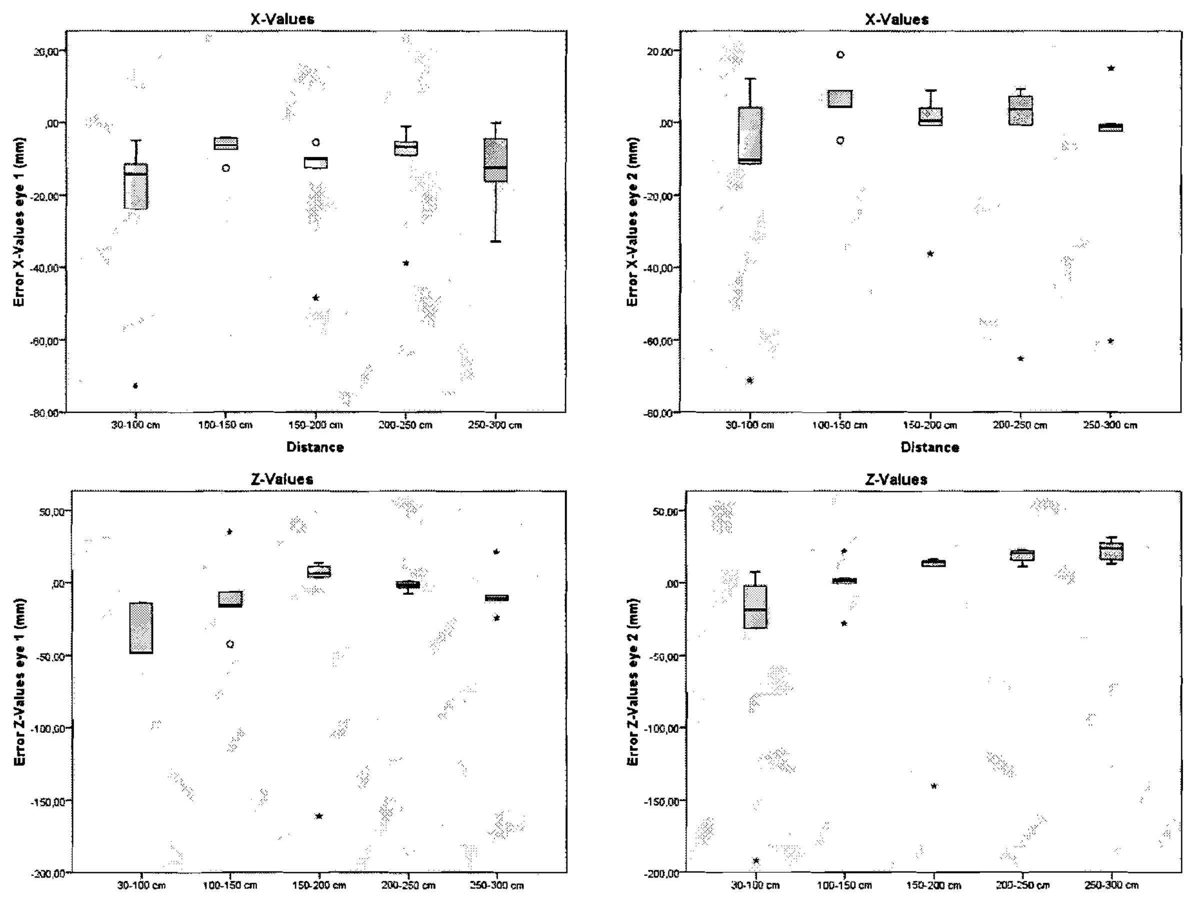


Figure 17 – Boxplots for the viewing distances

To analyze the effect of viewing distance the Kruskall-Wallis test for several independent groups is used. The test shows that there is no significant effect for viewing distance on the x-values of the left eye (H(4)=4.076, $p>0.05$), the x-values of the right eye (H(4)=2.585, $p>0.05$) and the z-values of the left eye (H(4)=8.714, $p>0.05$).

The viewing distance has a significant effect on the z-values of the right eye (H(4)=14.215, p<0.01). Wilcoxon rank-sum tests were used as follow up to test whether two samples of observations come from the same distribution. A Bonferonni correction was applied to ensure that Type I errors don't build up to more than 0.05. Therefore the critical value of 0.05 is divided by the number of tests that are conducted, so all effects are reported at a 0.005 level of significance (Field, 2005). It appeared that there was a significant difference between the 200-250cm viewing distance (Ws= 15, r= -0.83) and the 250-300cm viewing distance (Ws= 15, r= 0.83) compared to the 30-100cm viewing distance. Jonckheere's test for non parametric testing of trends in the data revealed a significant trend of the z-values of the right eye: an increasing viewing distance shows a positive trend on the error, J= 125, z= 4.064, r= 0.83.

7.3.6 Examination of critical limits

As described by Neumann et al. (2007) the critical limits for the error are +/- 5mm for the x-values and +/- 10 mm for the z-values. The error that is outside this critical limit is calculated for every condition by extracting the corresponding critical limit (5 mm for x-values and 10 mm for z-values) from the average of absolute errors. If the error outside the critical limits is between 0 mm and -5 mm for the x-values and 0 mm and -10 mm for the z-values, the system works accurately. Figure 18 shows the results for both eyes. Appendix 3 contains a complete overview.

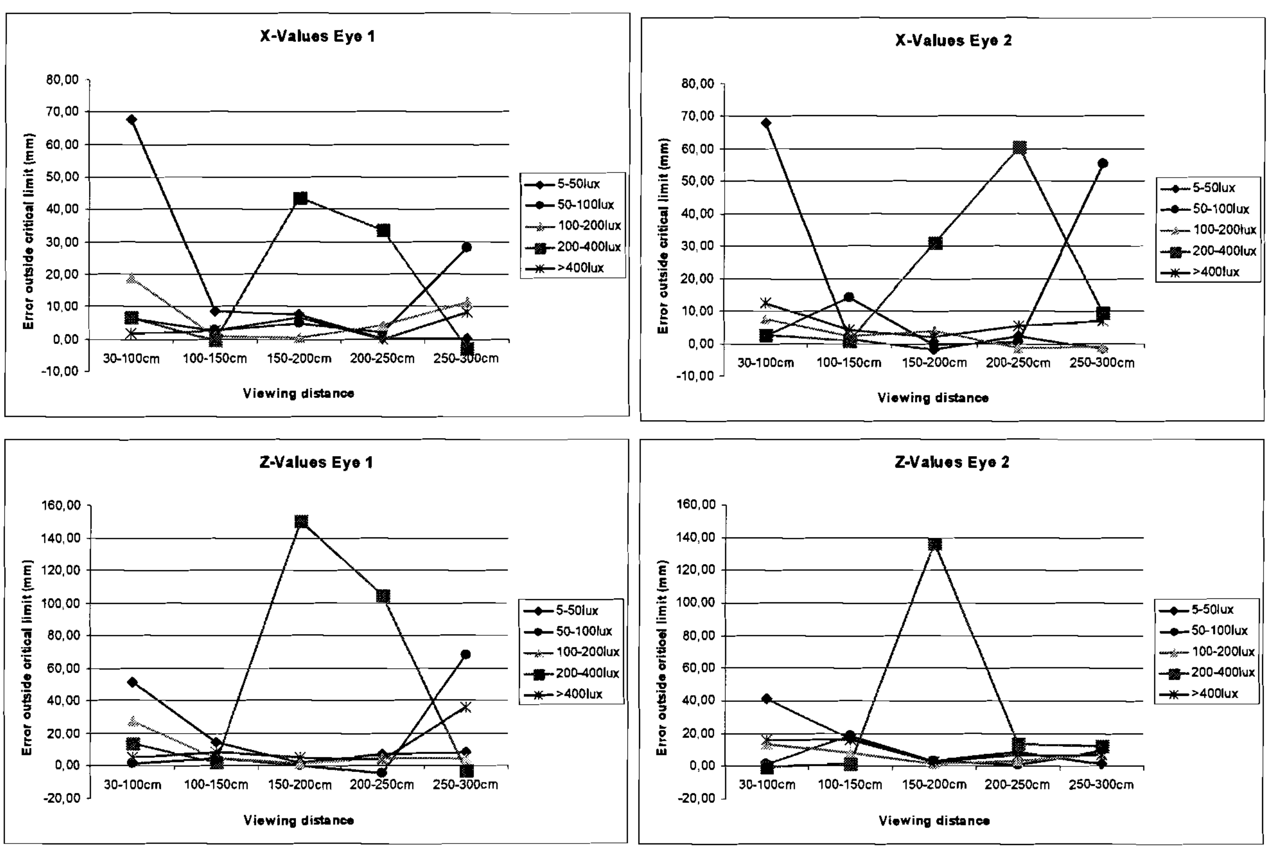


Figure 18 – Percentages of data between the critical limits

The percentiles of the descriptive statistics show that in order to let 95% of the error for the x-values fall between the critical limits, these limits should be between -65.46 and -0.3575 for the left eye and between -69.43 and 17.67 for the right eye. For the z-values these errors should be between -198.16 and 30.9 for the left eye and between 176.33 and 30.31 for the right eye. In order to let 75% of the error for the x-values fall between the critical limits, these limits should be between -5.21 and 5 for the left eye and between -7.54 and 8.08 for the right eye. For the z-values these errors should be between -20.38 and 2.65 for the left eye and between -1.25 and 21.28 for the right eye.

7.4 Discussion

One of the main conclusions that can be drawn from this experiment is that, besides for the z-values of the viewing distance of the right eye, it seems that lighting conditions and viewing distance do not have an influence on the accuracy of the MUTED multiple user tracking system. It must be noted that the factors that are used for statistical analysis, viewing distance and lighting conditions, are analysed for only one person due to problems with the data files. This significantly reduces the generalizability of the results. Therefore the conclusions have to be interpreted as points of interest for future research rather than findings that can be generalized for a large population.
The statistical results show that there is a significant difference between the error of the left and the error of the right eye. For the x-values the left eye is detected more to the left as seen from the user, while for the z-values the right eye is detected further away from the tracking system then the left eye. It is possible that this is due to the fact that only one transformation matrix was used for both eyes for every participant to transform the data of the reference system to the same coordinate system as the MUTED system. It is also possible that, since only one person is analyzed, the difference between both eyes is due to some facial features of that person that make detection difficult (Liu, Conomis, Zhu & Pastoor, 2001).
In contrast to the research of Zhu & Ji (2005), which concluded that an eye tracking system performs worse under low illumination, the statistics show that lighting has no influence on the accuracy of the tracking system. This could be explained by the fact that the MUTED multiple user tracking system corrects for low illumination by adjusting the brightness and contrast settings of the cameras.
Viewing distance has an effect on the z-values of the right eye, but this effect was just slightly significant. This would suggest that the 30-100cm condition does not cause a larger error, although the technical implementation of the system has defined an optimal viewing range of 1m to 3m.
To evaluate the technical relevance of these findings, the error that is outside the critical limits of +/- 5 mm for the x-values and +/- 10 mm for the z-values is examined. It seems that the 200-400 lux lighting condition works inaccurately at a viewing distance of 150-200cm and 200-250cm. Interesting at this point is that for the other viewing distances the 200-400 lux lighting condition stays within or is

only slightly outside the critical limits. Also the 5-50 lux lighting condition gives higher error results at a viewing distance of 30-100cm and the 50-100 lux lighting condition at a viewing distance of 250-300 cm. In general, the other conditions seem to stay within reasonable values, although those values are larger than the currently defined critical limits.

Assuming that the reference tracker detects the eye positions correctly, the critical limits also show an interesting finding. In order to let 75% of the error for the z-values fall between the critical limits, these limits must be between -20.38 and 2.65 for the left eye and -1.25 and 21.38 for the right eye. This suggests that the error stay within the defined critical limits of +/- 10 mm, but there is an offset on the data. For the x-values those critical limits already correspond to the defined limits for the left eye, while the critical limits for the right eye are slightly larger but reasonable.

It is possible that a person is tracked incorrectly at the first part of the data file, after which for the remaining part of the data file the error stays within the critical limits. Another possibility is that the incorrect tracked positions are spread over the complete data file. If those parts of incorrect tracked positions throughout the data file are large enough to be noticeable, it is assumed that this causes more visual strain (Kooi & Toet, 2004).

8 Conclusion and discussion

In the introduction of this report it was argued that for a successful introduction of 3D displays into the market, the viewing comfort of the display has to be at least comparable to an existing 2D display. This inherently means that the multiple user tracking system used in the MUTED 3D display must provide accurate tracking of all viewers in three dimensions under all environmental circumstances. The main objective of this report is to investigate the influence of different environmental factors on the accuracy of the multiple user tracking system. This accuracy is judged by the error of the multiple user tracking system. The error is calculated as the difference between the real eye position and the tracked eye position. The larger the error, the less accurate the multiple user tracking system functions. To achieve a representative set of environmental factors under which the accuracy can be tested, an observational study was conducted within the two environments where it is most likely that the MUTED 3D display will be introduced first; a medical environment and an entertainment environment.

The environmental factors that resulted from the literature are viewing distances, lighting conditions, number of viewers and user movements. A representative set of ranges for those environmental factors was retrieved from the observational study and used in the second part of this research, the performance study. By examining the influence of the different ranges for the environmental factors on the error of the multiple user tracking system, it is possible to determine under which circumstances the MUTED system functions accurate. Neumann, Hopf & Przewozny (2007) have defined first target values for the error to be +/- 5 mm for the x and y-axis and +/-10 for the z-axis. Although the viewing distance and lighting conditions do not seem to have a significant effect on the accuracy of the multiple user tracking system, the technical relevance of this finding is discussable. Since only one person is taken into analysis, the results are not generalizable and have to be interpreted as points of interest for future research. From the observational study it was concluded that for the environments of interest the 5-50 lux lighting condition is of importance for the medical imaging and the minimally invasive surgery environment. For the entertainment environment all lighting conditions, except the >400 lux condition are of interest. The 30-100 cm viewing distance is of importance for the medical imaging and the minimally invasive surgery environment, while the 250-300 cm and 200-250 cm viewing distances for the entertainment setting. Linking those results to the findings of the performance study can be concluded that the system needs extra development regarding the 200-400 lux lighting condition and 200-250 cm viewing distance for the entertainment environment and regarding the 5-50 lux lighting condition and 30-100 cm viewing distance for the medical environment.

9 Further research

There are a few limiting factors, which are most caused by practical limitations. An experiment cannot be expanded infinitely in time, so choices have to be made. An effort is made to justify most choices. In spite of solid justification, limitations still have an influence on the degree to which results can be generalized. From the results of this research it is evident that much more work needs to be done. For future research, the limitations of this research can be used to learn from. This part will discuss the problems that have to be solved before further analysis on the data is possible. It also gives some suggestions for future work.

In general it is advised to start with testing with one participant in a static condition at a fixed viewing distance and a fixed lighting condition. If the system works accurately, the lighting conditions or viewing distance can be varied. As a next step another person can be added to the design. This can help to clearly distinguish possible shortcomings of the system and point out specific problems that need to be solved before further testing can be done properly.

One of the most limiting aspects during the analysis of the performance study was the amount of missing data in the data files. To do some reasonable analysis on the data the missing values where deleted. For future research it would be wise to keep the data fixed at the last found position when a person is lost. In this way it is possible to distinguish real missing data from data where a person is found, but lost for a short moment from data where a person is not found at all.

Since the experiment was recorded on video stream and analyzed afterwards, it was not clear how the data files would be build up exactly at the moment the experiment took place. It was decided to make all movements of one condition within one video stream. Afterwards those files would be split to distinguish the different movements. In the data files some of the data switches between the columns. To do some comparison between the MUTED system and the reference system it is of importance that different users can be distinguished by different columns. For this research the data files of the static conditions were altered manually, but to make some analysis for user movements it is essential to record different movements into different video streams. It is possible to distinguish movements from one file afterwards, but this would be very bothersome and a possible cause for error.

The participants all are employees of the Heinrich Hertz Institute and worked in the same department. They all have different facial features which could influence the ability of the system to track that person. Within this research it is therefore discussable if person 3 has facial features that the system can recognize better compared to person 1, since person 1 wore glasses. This could be the reason that person 1 is found only incidentally and person 3 is found in every condition. Therefore the MUTED system has to be tested with more people then was done during this experiment.

One transformation matrix was used for both eyes to transform the data of the reference tracker to the same coordinate system as the MUTED system. It is advised to use different transformation matrices for each eye. This could possibly eliminate the difference between the errors of both eyes.
For the system to be useful in a medical environment, where the data will be critically appraised, it must be as accurately as possible (Bulpitt, 1987). In further research it is therefore advised to test the performance of the MUTED system with a 99% confidence interval. This will overcome some ethical issues that otherwise could disrupt the introduction of such a system in a medical environment.
Other recommendations for further research are the examination of the spread of the MUTED system and the examination of the frames. The spread can help to determine if, although the eye is not found on the correct position, the MUTED system holds on to this incorrect position within the critical limits. The examination of the frames can help to determine if the MUTED system tracks the user incorrectly for a small part of the trial (for example at the beginning) or throughout the whole trial, since when those parts are larger than the just noticeable difference of 50 ms as defined by Adelstein et al. (2003) the latter seems to be more crucial regarding visual strain (Kooi & Toet, 2004).

References

Adelstein, B.D., Lee, T.G. & Ellis, S.R. (2003). Head tracking latency in virtual environments: psychophysics and a model. *Proceedings of the human factors and ergonomics society 47th annual meeting*, pp. 2083-2087.

Baratoff, G., & Blanksteen, S. (1993). Tracking Devices. *Encyclopedia of Virtual Environments*. Received September 10th 2008 on http://www.hitl.washington.edu/scivw/EVE/I.D.1.b.TrackingDevices.html

Beachemin, K.M. & Hays, P. (1996). Sunny hospital rooms expedite recovery from severe and refractory depressions, *Journal of affective disorders* **40**: 49-51.

Benzie, P., Watson, J. & Surman, P. (2007). A survey of 3DTV displays: techniques and technologies, *IEEE Transactions On Circuits And Systems For Video Technology* **17**: 1647-1658.

Berguer, R. (1999). Surgery and ergonomics, *Arch Surg* **134**: 1011-1016.

Boppart, S.A., Deutsch, T.F. & Rattner, D.W. (1993). Optical imaging technology in minimally invasive surgery - Current status and future directions, *Surg Endoscopy* **13**: 718–722.

Buckley, E., Corbett, A., Surman, P. & Sexton, I. (2008) Multi-viewer autostereoscopic display with dynamically-addressable holographic backlight. *SID Conference May 2008*. Received May 5th 2008 on http://lightblueoptics.com/images/downloads/251.pdf.

Carroll, J.M., Rosson, M.B., Chin Jr., G. & Koenemann, J. (1998) Requirements development in scenario-based design, *IEEE Transactions on software engineering* **24**: 1156-1170.

Crabtree, A. & Rodden, T. (2004) Domestic routines and design for the home, *Computer Supported Coorperative Work: The Journal Of Collaborative Computing* **13**: 191-220.

Dodgson, N.A. (2005) Autostereoscopic 3D Displays, *Computer* **38** (8): 31-36.

Ezra, D., Woodgate, G.J. & Omar, B. (1992). Autostereoscopic directional display apparatus. *European Patent Application EP-A-0 602 934.*

Field, A. (2005). *Discovering Statistics Using SPSS*. London, England: SAGE Publications Ltd., pp. 556.

Foord, K.D. (1999). PACS workstation respecification: display, data flow, system integration, and environmental issues, derived from analysis of the conquest hospital pre-DICOM PACS experience, *European Radiology* **9**: 1161-1169.

Foxlin, E., Harrington, M. & Altshuler, Y. (1998). Miniature 6-DOF inertial system for tracking HMDs. *In Proceedings of AeroSense 98 Conference on Helmet- and Head-Mounted Displays* **3362** (3): 1-15.

Gibbs, S., Arapis, C., Breiteneder, C., Lalioti, V., Mostafawy, S. & Speier, J. (1998). Virtual studios: An overview, *IEEE Multimedia* **5** (1): 24–35.

Hajdukiewicz, J.R., Vicente, K.J., Doyle D.J., Milgram, P. & Burns, C.M. (2001) Modeling a medical environment: an ontology for integrated medical informatics design. *International Journal of Medical Informatics* **62**: 79–99

Halle, M. (1997). Autostereoscopic displays and computer graphics, *Computer Graphics. ACM Siggraph* **31** (104): 58-62.

Hanna, G., Cuschieri, A. (2001). Image display technology and image processing, *World Journal of Surgery* **25**: 1419-1427.

Holliman, N. (2007) 3D Display Systems, *Handbook of Optoelectronics, Institute of Physics Press.*

Ijsselsteijn, W.A., Seuntiens, P.J.H. & Meesters, L.M.J. (2002). State-Of-The-Art in human factors and quality issues of stereoscopic broadcast television. *Technical report ATTEST-IST-2001-34396.*

IJsselsteijn, W.A., Seuntiëns, P.J.H., & Meesters, L.M.J. (2005). Human factors of 3D displays. *In: Schreer, O., Kauff, P., & Sikora, T. (eds.), 3D Videocommunication - Algorithms, concepts and real-time systems in human-centred communication. John Wiley & Sons, Ltd.*, pp. 219-234.

Jones, G., Lee, D., Holliman, N. & Ezra, D. (2001). Controlling perceived depth in stereoscopic images. *Proceedings of SPIE* **4297**: 42-53.

Kidratenko, V. (1999). Calibration of electromagnetic tracking devices, *Virtual reality* **4**: 139-150.

Kitasaka, T., More, K. & Suenaga, Y. (2005). Development of Advanced Image Processing Technology and Its Application to Computer Assisted Diagnosis and Surgery. *Advances in Multimedia Information Processing - PCM 2004* **331**: 514-521.

Kooi, F.L. & Toet, A. (2004). Visual comfort of binocular and 3D displays, *Displays* **25** (2): 99-108.

Krotz, F. & Eastman, S.T. (1999). Orientations toward television outside the home, *Journal of communication* **49**: 5-27.

Li, R., Taskiran, C. & Danielsen, M. (2007). Head pose tracking and gesture detection using block motion vectors on mobile devices, *Proceedings of the 4th international conference on mobile technology, applications, and systems and the 1st international symposium on Computer human interaction in mobile technology, September 10-12, 2007, Singapore,* pp. 572-575.

Lipton, A., Fujiyoshi, H., & Patil, R. (1998). Moving target detection and classification from real-time video, *Proceedings of IEEE Workshop Application of Computer Vision (WACV '98)*, pp. 8-14.

Liu, J., Conomis, C., Zhu, Z. & Pastoor, S. (2001). Real-time head tracking in a multimodal interface with changes in geometry and illumination, *IASTED International Conference "Visualization, Imaging, and Image Processing"*, pp. 608-612.

Meesters, M.J., Ijsselsteijn, W.A. & Seuntiens, P.J.H. (2004). A survey of perceptual evaluations and requirements of three-dimensional TV, *IEEE Transactions On Circuits And Systems For Video Technology* **14**: 381-391.

Milne, A.D., Chess, D.G., Johnson, J.A. & King, G.J.W. (1996). Accuracy of an electromagnetic tracking device: A study of the optimal operating range and metal interference, *Biomechanics* **29**: 791-793.

Naggy, P., Siegel, E., Hanson, T., Kreiner, L., Johnson, K. & Reiner, B. (2003). PACS reading room design, *Seminars in roentgenology* **38**: 244-255.

Neumann, F., Hopf, K., Przewozny, D. (2007). MUTED Deliverable 5.2. First Generation Head Tracker. *IST-5-034099.*

Prabhu, S.P., Gandhi, S., Goddard, P.R. (2005) Ergonomics of digital imaging, *British Journal of Radiology* **78**: 582-586.

Pastoor, S. (1993). Human factors of 3D displays in advanced image communications, *Displays* **14** (3): 99-108.

Quante, B. & Hopf, K. (2003). Evaluation of a single user autostereoscopic display - an example of a user centred design cycle. *Proceedings of the 19th International Symposium on Human Factors in Telecommunication.*

Sande, J.P. (1999). *Gedragsobservaties. Een inleiding tot systematisch observeren,* Wolters Noordhoff.

Stone, R.J. (1996). Position and orientation sensing in virtual environments. *Sensor Review* **16**: 40–46.

Surman, P., Sexton, I., Hopf, K., Bates, R. & Lee, W.K. (2006). Head tracked 3D displays, *Lecture notes in computer science* **4105**: 769-776.

Surman, P., Sexton, I., Hopf, K., Lee, W.K., Bates, R., IJsselsteijn, W.A., & Buckley, E. (2006). Head Tracked Single and Multi-user Auto stereoscopic Displays. *IET Conference on Visual Media Production 2006, Savoy Place, London.*

Surman, P., Sexton, I. Hopf, K, Lee, W.K., Buckley, E., Jones, G. & Bates, R. (2008). European research into head tracked autostereoscopic displays, *3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video*, pp. 161 – 164.

Talmi, K. & Liu, J. (1999). Eye and gaze tracking for visually controlled interactive stereoscopic displays, *Signal Processing: Image Communications* **14**: 799-810.

Vallidis, N. (2002). WHISPER: A Spread Spectrum Approach to Occlusion in Acoustic
Tracking, *PhD thesis, University of North Carolina at Chapel Hill, Department of Computer Science.*

Woods, A., Docherty, T. & Koch, R. (1993). Image distortions in stereoscopic video systems, *Proceedings of the SPIE* **1915**: 36-48.

Yilmaz, A., Javed, O. & Shah, M. (2006). Object tracking: a survey, *ACM Comput. Surv*ey **38** (4), Received 11th October on http://doi.acm.org/10.1145/1177352.1177355

Zhu, Z. & Ji, Q. (2005). Robust real-time eye detection and tracking under variable lightning conditions and various face orientations, *Computer Vision and Image Understanding* **98** (1): 124-154.

Zheng, B., Janmohamed, Z., MacKenzie, C.L. (2003). Reaction times and the decision-making process in endoscopic surgery, *Surgical Endoscopy* **17**: 1475–1480.

Appendices

Appendix 1 – Checklist observations

- Introduction.

- Ask if the main objective of the research is clear.
  Main objective: Make an observational study at the activities of surgeons (themselves and between each other) in a pre-surgery planning room. What are the behaviours of those surgeons and interactions between them? Where does the furniture stands in the room and what is the distance between them? With the results of this research we can perform a performance evaluation of the eye-tracker that is used for the new 3D display under different scenario's to decide which aspects need further development.

- (medical setting) Tell the observation is focused on a pre-surgery planning, because this is at this moment the most likely scenario where our 3D technology can be applied.

- (medical setting) Give the interview.

- Make a drawing of the room. ***Don't forget to measure the size of the display.***

- Measure the distance of the furniture.

- Instruct the person questioned to think specifically at situations where more then one person was sitting in the room while answering the following questions.

- (entertainment setting) Ask how often within a week the television is watched at that time period (afternoon/evening).

- (entertainment setting) Ask how many people are on average in the room at the moment the television is watched in that time period (afternoon/evening).

- (entertainment setting) Ask where those people have there 'own' place in the room.

- Measure the distance of those places towards each other and towards the display.
  (People are measured from the (approximate) position of the eyes; the display is measured from the middle top. This is also the place where the lux-measurement needs to be done.)

- Measure the light intensity. ***Don't forget to note the kind of weather for the daytime measurement.***

- Ask for at least two typical movements these people make themselves.

- Ask for at least two typical interactions these people have with each other.

Appendix 2 – Interview questions medical environment

Introduction

Introduce ourselves and ask if it is possible to do the interview in a room where normally a pre-surgery planning takes place.

Background information about the project
Make a study of the use of a display in a pre-surgery planning room. With the results of this research we can make some basic scenarios to perform a performance evaluation of the eye-tracker that is used for the new 3D displays. The end goal is to decide which aspects need further development.

The interview

(I) Background information of the surgeon

1. Would you please tell us your age?

2. How many years of experience do you have within radiology?

3. Could you tell us a bit more of your background in radiology?

4. What is your experience of displays within the radiology sector?

(II) Sort of activities and involved people in a pre-surgery planning room

1. What sorts of activities are done in a room where pre-surgery planning takes place?
   (Possible answers: emergency, planning a surgery, informing the patient – if one of those is not asked, ask about it!)

2. What is the aim of these activities?
   (If necessary call them by name using the answers of question III - 1)

3. Which people are normally involved during these activities?
   (If necessary call them by name using the answers of question III - 1)

4. What are the functions of those people within a specific activity?
   (If necessary call them by name using answers of question III - 1 and III - 3)

5. Do those people have an 'own' place in the room during these activities? And if so, could you please describe or show that place?
   (If necessary call them by name using the answers of question III - 3)

(III) Interactions of people between each other and with the display

1. Can you please give two typical movements those people make during an activity?

2. Is there a typical place where someone stands regarding the other people involved?

3. Could you please give some typical movements and interactions that people make them or between each other, for example two surgeons turning to each other to discuss something and then together turn to the radiologist to get him involved?
   (Use the answers from questions II - 1 and II - 3)

4. Which of those people are involved in the use of a (3D) display?
   (Use the answers from questions II - 1 and II - 3)

5. Are they looking at the (3D) display for longer periods or not?
   (Use the people that resulted from question III - 4)

6. Approximately how far do they stand from the display?
   (Use the people that resulted from question III - 4)

(IV) Lighting conditions in a pre-surgery planning room

1. Which kind of light sources are used within the pre-surgery planning room?

2. Where are those light sources placed within the room?
   (Use the answers from question IV-1)

3. In a range of low-medium-bright, how would you describe the brightness of every light source?
   (Use the answers from question IV-1)

4. On the spot of the display, how would you describe the brightness of the light source there?

(V) Difference daytime and evening

1. Is there a difference between daytime or evening regarding the factors (and if so what are these differences):
   a. Number of people involved?
   b. Movements of people involved?
   c. Attention paid to the display?
   d. Distance to the display?
   e. Light sources?
   f. Brightness of the different light sources?

(VI) Placement of furniture in a pre-surgery planning room

1. What kind of furniture is standing in a pre-surgery planning room?

2. Does this furniture have its own typical place in the room? If yes: where?

(VII) Difference between hospitals

1. Is there much difference between hospitals regarding the kind of furniture in a pre-surgery room?

2. Is there much difference between hospitals regarding the sources of lighting in a pre-surgery room?

3. Is there much difference between hospitals regarding the number of people involved?

4. Is there much difference between hospitals regarding the way people act in a pre-surgery room?

5. Is the attention to the display very different between hospitals?

6. Is the distance from the display very different between hospitals?

(VIII) Observations

1. Would it at this moment be possible to see such a pre-surgery room and make some observations ourselves?
   (Only if it wasn't possible to do the interview there)

2. Is it possible to actually see such a pre-surgery planning (or a staged one)?
   (Only if we need some extra visual data to answers to the questions above)

Ending of the interview

Ask if he/she has other contacts where similar interviews/observations can be done.

Thank you for your time!

Appendix 3 – Errors outside critical limits

	x-values eye 1 **5 - 50 lux (average)**	**50 - 100 lux (average)**	**100 - 200 lux (average)**	**200 - 400 lux (average)**	**> 400 lux (average)**
30-100cm	67,78	6,41	18,67	6,84	1,89
100-150cm	8,70	2,69	0,91	-0,46	2,73
150-200cm	7,53	5,00	0,49	43,37	6,85
200-250cm	0,30	2,12	4,12	33,76	-0,24
250-300cm	0,10	28,13	11,23	-2,82	8,35

	x-values eye 2 **5 - 50 lux (average)**	**50 - 100 lux (average)**	**100 - 200 lux (average)**	**200 - 400 lux (average)**	**> 400 lux (average)**
30-100cm	67,87	2,35	7,46	2,66	12,45
100-150cm	1,39	14,12	2,42	0,94	4,10
150-200cm	-1,87	-0,39	4,00	31,13	2,08
200-250cm	2,35	0,62	-1,29	60,62	5,28
250-300cm	-1,64	55,33	-1,10	9,33	6,94

	z-values eye 1 **5 - 50 lux (average)**	**50 - 100 lux (average)**	**100 - 200 lux (average)**	**200 - 400 lux (average)**	**> 400 lux (average)**
30-100cm	51,00	1,23	27,56	13,96	5,07
100-150cm	14,52	4,83	4,42	2,37	8,55
150-200cm	1,98	0,40	1,79	150,34	5,09
200-250cm	7,32	-4,78	4,43	104,66	3,18
250-300cm	8,56	67,72	4,49	-2,66	35,70

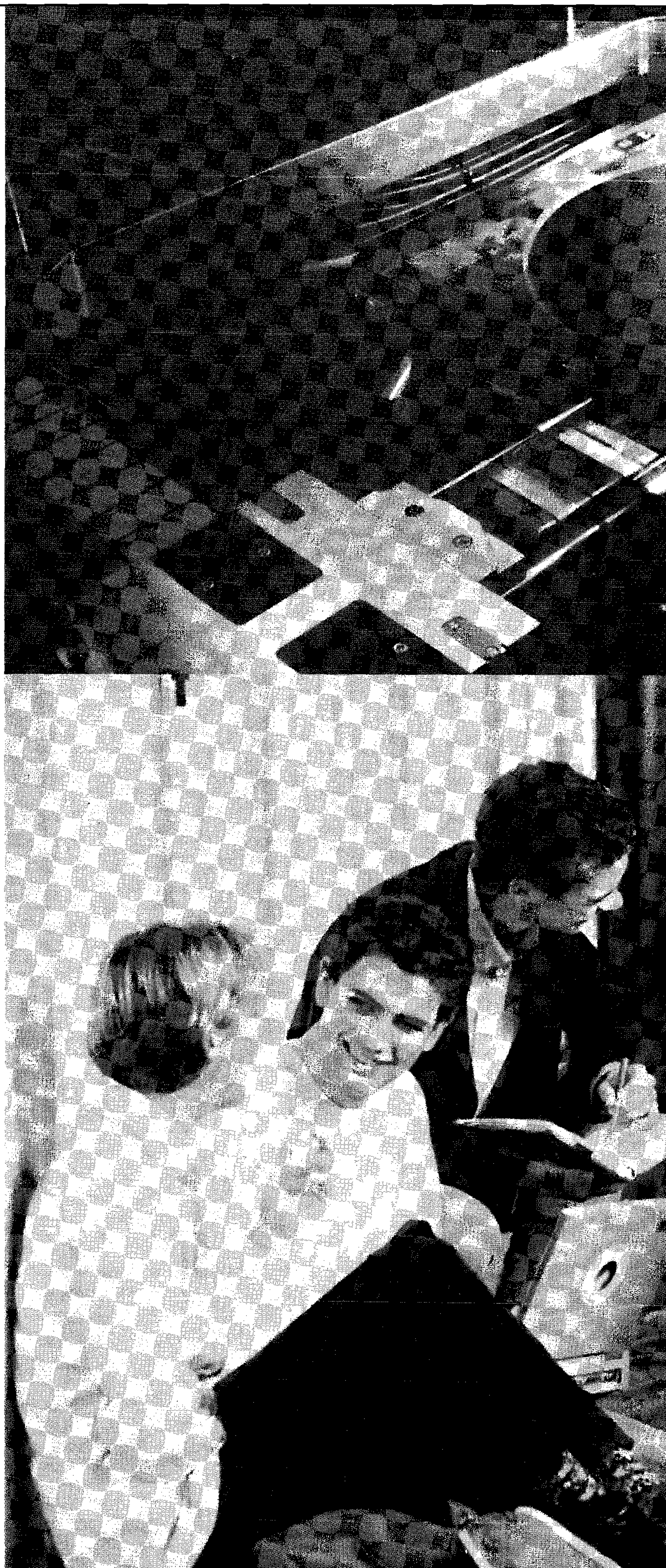
	z-values eye 2 **5 - 50 lux (average)**	**50 - 100 lux (average)**	**100 - 200 lux (average)**	**200 - 400 lux (average)**	**> 400 lux (average)**
30-100cm	41,01	1,18	13,39	-0,49	15,95
100-150cm	16,72	18,29	8,07	1,57	16,29
150-200cm	3,17	3,11	1,69	137,03	2,48
200-250cm	8,95	1,07	3,40	13,79	6,65
250-300cm	1,47	9,30	7,57	12,63	7,20

TU/e technische universiteit eindh

Technische Universiteit Eindhoven
Department of Technology Management
Tema 1.32
P.O. Box 513
5600 MB Eindhoven
The Netherlands

Telephone: +31 (0)40 247 26 35
Fax: +31 (0)40 246 85 26
E-mail: fac@tm.tue.nl
www.tue.nl/tm

Photograph:
Norbert van Onna, Eindhoven



/ department of technology management