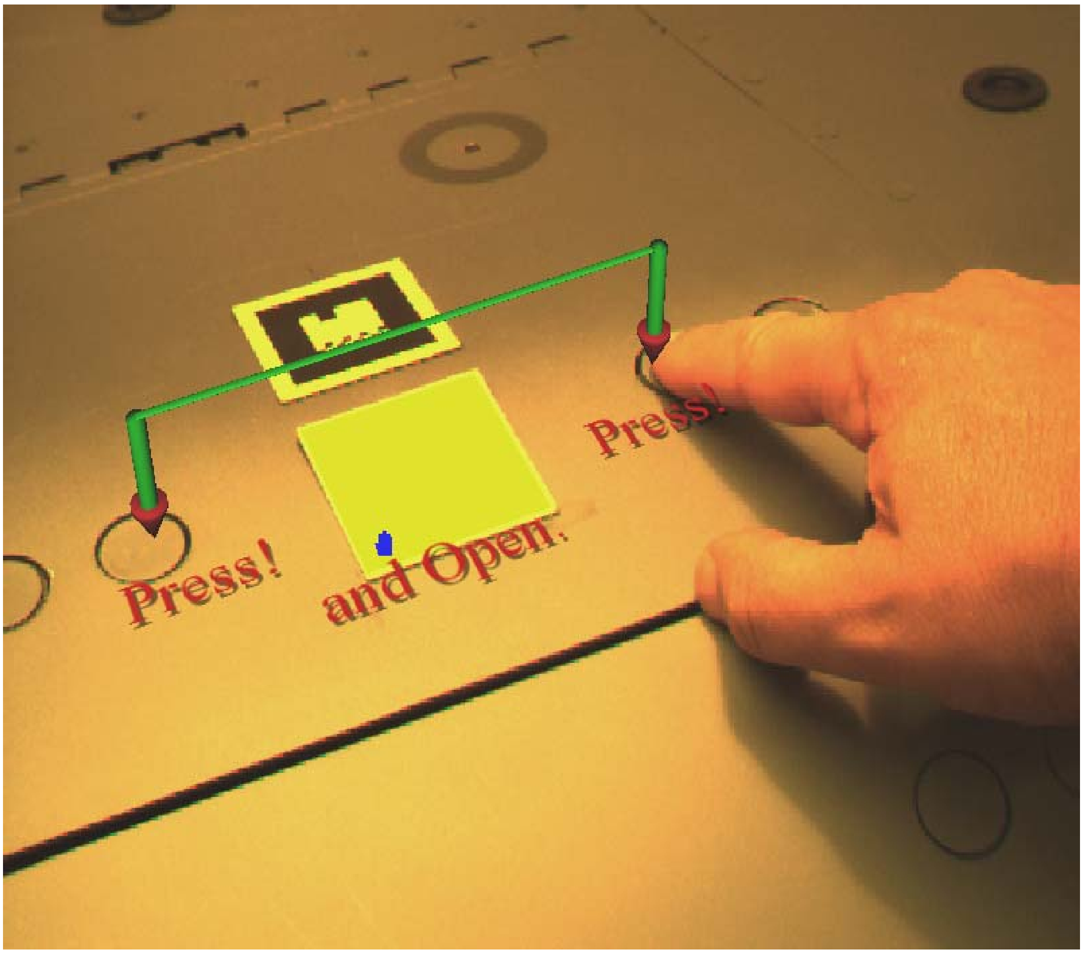
Mixed Reality Systems for Technical Maintenance and Gaze-Controlled Interaction

Torbjörn Gustafsson, FOI
Per Carleberg, FOI
Pär Svensson, LiU
Susanna Nilsson, LiU
Åke Sivertun, LiU

FOI
Linköpings Universitet



Mixed Reality: a mix of reality and virtual reality. Photo: FOI

FOI-R-- 1695 --SE
ISSN 1650-1942

Sensor Technology
User report

August 2005

FOI
Defence Research Agency
Sensor Technology
P.O. Box 1165
SE-581 11 Linköping

Tel:
Fax:

www.foi.se

Mixed Reality Systems for Technical Maintenance and Gaze-Controlled Interaction

FOI-R--1695--SE
ISSN 1650-1942

Sensor Technology
User report

August 2005

Issuing organization	Report number, ISRN	Report type
FOI – Swedish Defence Research Agency Sensor Technology P.O. Box 1165 SE-581 11 Linköping	FOI-R--1695--SE	User report
	Research area code 2. Operational Research, Modelling and Simulation 8. Human Systems	
	Month year August 2005	**Project no.** E3948
	Sub area code 23 Logistics	
	Sub area code 2 81 Human Factors and Physiology	
Author/s (editor/s) Torbjörn Gustafsson, FOI Per Carleberg, FOI Pär Svensson, LiU Susanna Nilsson, LiU Åke Sivertun, LiU	**Project manager** Torbjörn Gustafsson FOI, Åke Sivertun LiU	
	Approved by Lena Klasén	
	Sponsoring agency FMV	
	Scientifically and technically responsible Torbjörn Gustafsson FOI, Åke Sivertun LiU	

Report title

Mixed Reality Systems for Technical Maintenance and Gaze-Controlled Interaction

Abstract (not more than 200 words)

The purpose of this project is to build up knowledge about how future Mixed Reality (MR) systems should be designed concerning technical solutions, aspects of Human-Machine-Interactio (HMI) and logistics. The report describes the work performed in phase 2.

Regarding hardware a hand-held MR-unit, a wearable MR-system and a gaze-controlled MR-unit have been developed. The work regarding software has continued with the same software architecture and MR-tool as in the former phase 1. A number of improvements, extensions and minor changes have been conducted as well as a general update.

The work has also included experiments with two test case applications, “Turn-Round of Gripen (JAS)” and “Starting-Up Diathermy Apparatus”.

Comprehensive literature searches and surveys of knowledge of HMI aspects have been conducted, especially regarding gaze-controlled interaction. The report also includes a brief overview of other projects within the area of Mixed Reality.

Keywords

Mixed Reality, Augmented Reality, technical maintenance, gaze-controlled interaction, Human-Machine-Interaction, MR, AR, HMI

Further bibliographic information	Language English
ISSN 1650-1942	**Pages** 49 p.
	Price acc. to pricelist

Utgivare FOI - Totalförsvarets forskningsinstitut Sensorteknik Box 1165 581 11 Linköping	**Rapportnummer, ISRN** FOI-R--1695--SE	**Klassificering** Användarrapport
	Forskningsområde 2. Operationsanalys, modellering och simulering 8. Människa och teknik	
	Månad, år Augusti 2005	**Projektnummer** E3948
	Delområde 23 Logistik	
	Delområde 2 81 MSI med fysiologi	
Författare/redaktör Torbjörn Gustafsson, FOI Per Carleberg, FOI Pär Svensson, LiU Susanna Nilsson, LiU Åke Sivertun, LiU	**Projektledare** Torbjörn Gustafsson FOI, Åke Sivertun LiU	
	Godkänd av Lena Klasén	
	Uppdragsgivare/kundbeteckning FMV	
	Tekniskt och/eller vetenskapligt ansvarig Torbjörn Gustafsson FOI, Åke Sivertun LiU	

Rapportens titel (i översättning)

Mixed Reality för tekniskt stöd och blickstyrd interaktion

Sammanfattning (högst 200 ord)

Syftet med detta projekt är att bygga upp kunskap om hur framtidens Mixed Reality-system (MR) skall utformas beträffande tekniska lösningar, människa-system-interaktion (MSI) problematik och för användning i logistik-sammanhang. Rapporten beskriver det arbete som utförts i fas 2.

Beträffande hårdvara har en handhållen MR-enhet, ett bärbart MR-system samt en blickstyrd MR-enhet utvecklats. Arbetet med mjukvara har fortsatt med samma mjukvaruarkitektur och MR-verktyg som i den första fasen. Ett antal förbättringar, utvidgningar och mindre ändringar har genomförts liksom en generell uppdatering.

Arbetet har också inkluderat experiment med två relevanta tillämpningsexempel, ”Klargöring av Gripen (JAS)” och ”Start av diatermiapparatat”.

Omfattande litteratursökningar och kunskapsöversikt har genomförts där blickstyrd interaktion har specialstuderats. Rapporten innehåller även ett sammandrag över andra projekt inom forskningsområdet ”Mixed Reality”.

Nyckelord

Mixed Reality, förstärkt verklighet, tekniskt stöd, blickstyrd interaktion, människa-system-interaktion, MSI, MR, AR

Övriga bibliografiska uppgifter	**Språk** Engelska
ISSN 1650-1942	**Antal sidor:** 49 s.
Distribution enligt missiv	**Pris: Enligt prislista**

FOI 2005 Utgåva 12

Contents

1. Introduction

The present project started in 2003 and is a collaboration between the Swedish Defence Research Agency (FOI) and Linköping University (LiU). The main contractor is the Swedish Defence Material Administration (FoT6, “Physiology and Human System Interaction”, and FoT19 “Logistics”).

This report describes the work performed in phase 2. The work carried out in the previous phase, (phase 1), is described in the report “Mixed Reality for Technical Support”, [Gustafsson et al. 2004].

The purpose of the project is to generate knowledge about how future Mixed Reality (MR) systems should be designed concerning technical solutions, aspects of Human-Machine-Interaction (HMI) and logistics.

The research teams from FOI and LiU collaborate closely in the project. The research team from FOI focuses on the technical solutions and design while the research team from LiU focuses on Human-Machine-Interaction and logistics issuses.

To thoroughly study how Mixed Reality Systems could be used in the target contexts the research is mainly based on experiments with relevant test case application examples. The successive knowledge obtained when working with these application examples strongly influences the technical development of the MR-system in the project as well as the generation of knowledge regarding HMI aspects.

In the previous phase of the project the test case application example consisted of “Dismounting of a Machine Gun Mounting”. In the present phase the work is based on experiments in additional application examples, “Turn-Round of Gripen (JAS)”, (military aircraft), and “Starting-Up Diathermy Apparatus”, (medical equipment). These new application examples and related work are reported below.

When working with these test case applications, technical equipment (hardware and software) for the MR-system are being developed, comprehensive literature searches and surveys of knowledge of HMI aspects are being conducted as well as a number of practical tests with prototypical equipment of the MR-system.

The new test case application examples have forced the development of the MR-system to be more sophisticated regarding hardware as well as software. A hand-held MR-unit and a gaze-controlled MR-unit have been developed. Regarding software development the MR-software has been improved and new functions have been added. This development is also described herein.

2. Background

In the world of Network Centric Warfare all resources must be used in a flexible way. The technical systems will, at the same time, be more and more complicated and connected in complex structures. The technical personnel in combat service cannot be experts or updated on all technical systems at the same time. They will probably be more generally educated instead of one-system super experts. To achieve the needs for flexibility they need some kind of support and technical aid. They must have an easy-to-use and intuitive support system with access facilities to information databases, technical documentation, etc. They also need to communicate, online, with technical and system experts in industry and research institutes. Support systems must be useful in close combat situations, supporting and helping the technician to focus on his mission.

The approach described in this report is to base the future intuitive support system on Mixed Reality concepts and techniques. Mixed Reality, MR, is a relatively new, generic and highly interdisciplinary field of research, see figure 1.1, and MR concepts are applicable to a wide range of applications [Gustafsson and Carleberg, 2003]. MR could be described as a technical information system that allows human-machine interaction where real and virtually generated image information are mixed in varying degrees and presented in the field of view of a user. The virtually generated image information may arise from image sensors, databases and/or from models. The process of visualization is affected by a real-time interaction of several technologies where tracking of the user's position relative the surroundings is a key technique, see figure 1.1.

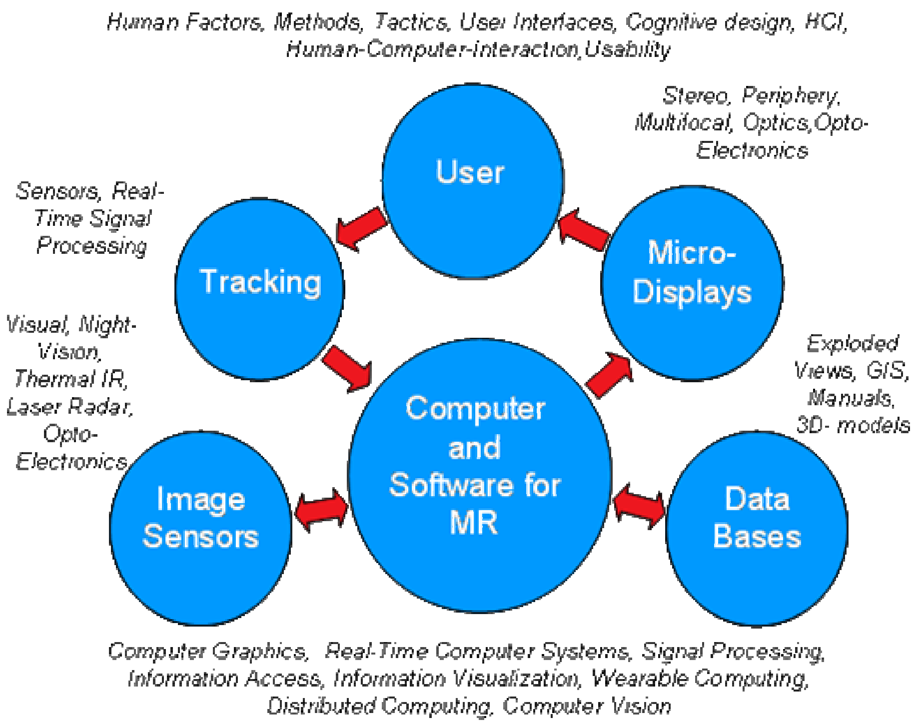


Figure 1.1 A general outline of an MR-system including involved fields of research.

Working with an intuitive support system the interaction process between user and system should be essential, especially the necessity to let the system know that an assignment/task has been completed or a goal has been achieved or to access information databases and technical documentation. During intensive work with a MR-system in the field both hands can be busy and not available for the interaction process, for example through a keyboard. Therefore, a gaze-controlled MR-system should be able to improve efficiency and system interaction in assisting military service personnel in troubleshooting, commissioning, maintenance and repair in the field.

Experience and knowledge obtained from the former phase 1 in combination with discussions with experts on logistics at the Swedish Defence Material Administration (FMV) have resulted in an granted patent regarding gaze-controlled Mixed Reality systems [Gustafsson et al. 2004].

A master thesis [Viberg 2005] has been performed in close co-operation with the project described in this report. This thesis describes the work of finding a frame of reference for interaction with a Mixed Reality-system for technical support and education. The work was commissioned by the Swedish Defence Materiel Administration, FMV, but has mainly been carried out in co-operation with the Swedish Defence Research Agency, FOI. The main goal of this thesis work was to find fundamental, basic and general concepts and tools for the Mixed Reality-system, which can be used as a platform for further development. The proposed methods and techniques should be suited, if possible, for gaze-control. The result of the thesis work was a frame of reference which proposes a new interaction paradigm for the Mixed Reality-system. There are also a number of proposals of techniques, methods and concepts for the design of the interaction with the system and for the tools of the system. The thesis work also resulted in three scenarios that describe how to apply and make use of Mixed Reality-technique in a military educational setting.

3. The development of the MR-system

3.1. Background

In the first phase, phase 1, of the project a first version of the MR-system was developed [Gustafsson et al. 2004]. Below in this chapter new versions of the MR-system are described.

An MR-system is a complex system. Therefore the technical development of the MR-system is divided into two parts, the hardware and the software part, see figure 3.1.

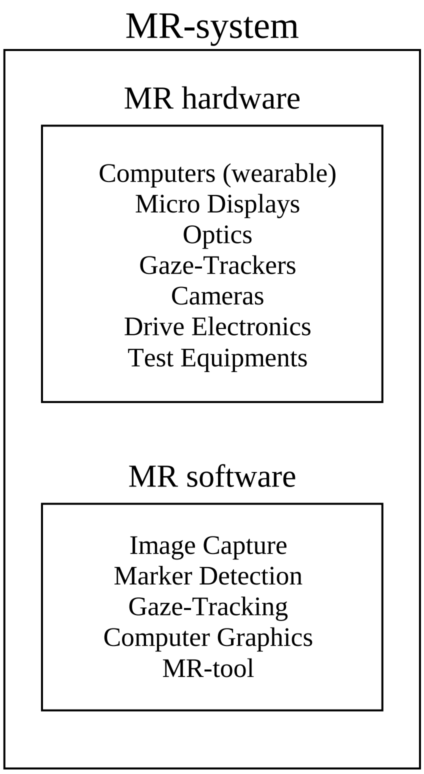


Figure 3.1 Block diagram of the MR-system.

The hardware part includes wearable computers, micro-displays, optics, gaze-trackers, cameras and drive electronics. These devices are called *MR-units*. As hardware also other relevant devices, for example, some test equipment, is included.

The software part includes an integrated set of software tools such as software for camera image capture, fiducial marker detection, gaze-tracking software, computer graphics software and also software developed specifically for MR-application scenarios, *the MR-tool.*

General demands upon the MR-system are real-time graphics, real-time tracking of gaze direction and real-time tracking of camera position relative the world.

3.2. MR hardware

3.2.1. Hand-Held Video-See-Through MR-Unit

As a complement to the head-wearable video-see-through MR-unit previously developed in phase 1, [Gustafsson et al. 2004], a hand-held MR-unit was developed in this present phase, see figure 3.2. The purpose with this hand-held unit was to achieve a unit easy to pick up, use and pass among several users. To be able to study multi-user situations a further handheld unit was also developed.

Figure 3.2 Video-See-Through MR-units. To the left a hand-held unit developed in the present phase and to the right the head-wearable unit developed in the previous phase.

The hand-held units consist of a handle with integrated drive electronics, micro display and micro camera systems. The unit has full colour 800(h) x 600(v) pixel resolution and a field of view of 27 x 20 degrees. The image presentation is bi-ocular, which means that the captured images from one camera are presented on two micro displays, one for each eye.

The micro camera system is based on FireWire colour compression, Bayern pattern, where the decompression in this case is performed by the computer. The image resolution is 640 x 480 pixels and the field of view is 31 x 23 degrees. Unlike the previously developed head-wearable unit this hand-held unit is equipped with a relatively high quality lens which makes a wider depth of field possible. The focal distance to objects presented on the display system is estimated to 1.2 m. The distance between the two micro displays is fixed to 63 mm. The camera and the micro-displays have different focal points dependent upon the bi-ocular solution and the difference in location regarding the camera and the displays. This causes some parallax between the camera and displays.

To be able to work under twilight conditions one hand-held unit is also supplied with two ultra compact light emitting diodes, see figure 3.3.

Figure 3.3 A hand-held MR-unit with lighting possibilities.

3.2.2. A wearable MR-system

In order to be able to perform practical experiments using MR-equipments a non-optimized first generation field test wearable MR-system has been developed. It consists of a helmet on which a micro display system, with integrated video-see-through camera, has been mounted. On the top of the helmet a drive electronic card for the camera has been integrated. The remaining components which are intended to be carried on the user's back are an electronic driver unit for the micro display system, battery unit and a wearable computer, see figure 3.4.

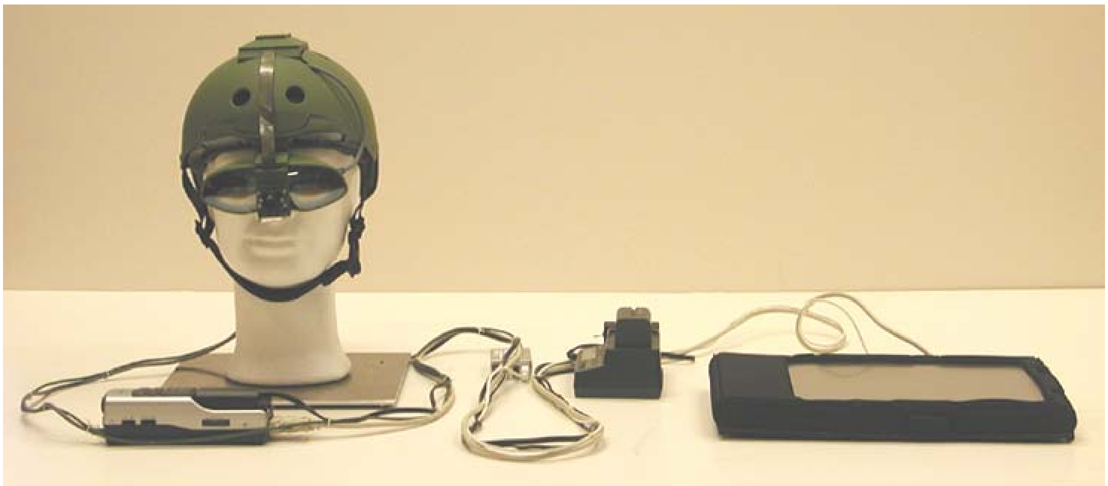


Figure 3.4 A wearable MR-system.

The micro display unit including the video-see-through camera is able to be turned up, see figure 3.5. The purpose for this is to be able to study different ways to work with and without the system. Further data of this wearable system is presented later on in this report.

Figure 3.5 Different modes, down or turned up, of using the video-see-through unit.

3.2.3. Gaze-Controlled Head-Wearable Video-See-Through MR-Unit

During the former phase in the project a technical development regarding gaze-controlled Mixed Reality was started. The purpose was to integrate a gaze-tracker into a head-wearable video-see-through MR-unit.

The gaze-tracking methods used were derived from experiences from former gaze-tracking projects funded by FMV and NUTEK, [Gustafsson and Carleberg, 2000] [Gustafsson and Carleberg, 2003].

The first prototype, developed in the previous phase, was based on high sensitive video cameras generating a standard video signal. This solution required a frame-grabber PC-card and in this manner also a desktop computer. The video camera solution caused a rather cumbersome design, see figure 3.6 to the right.

With the experience from the work with this first gaze-controlled prototype the development of a more sophisticated unit started. The new attempt was to base both the video-see-through MR-unit and the gaze-tracker on FireWire cameras and computer interfaces. Such a solution was able to be implemented in a laptop computer (wearable computer) in spite of this type of computer's hardware and software limitations.

Every new approach results in new types of problems due to integration of new technical components with, in this case, different characteristics than formerly used components.

After a few redesigns a second gaze-controlled prototype was finally developed. Due to the FireWire solution and the use of a new type of micro display system, the OLED, *Organic Light-Emitting Diode*, this prototype could be designed with a rather slim design compared to the previously developed prototype, see figure 3.6 to the left.

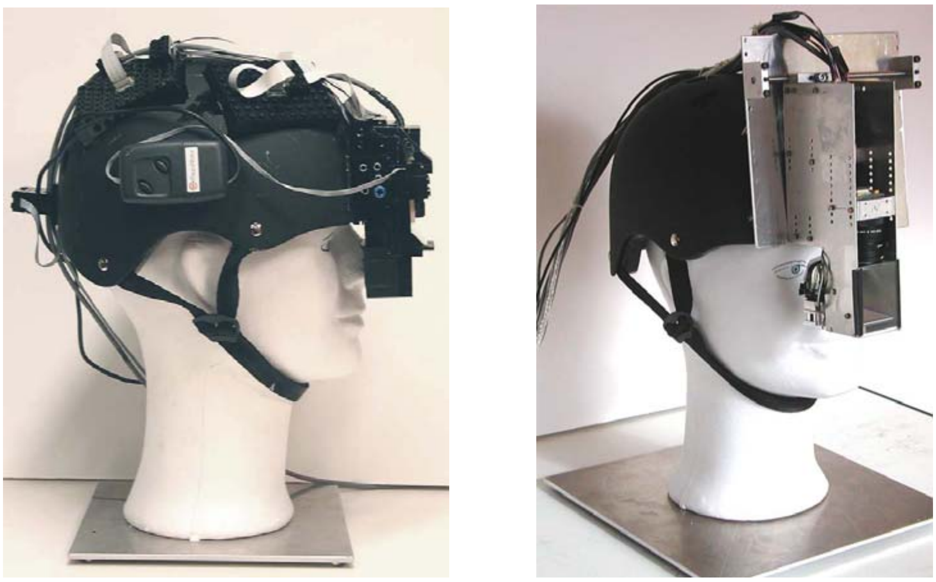


Figure 3.6 To the left a gaze-controlled head-wearable video-see-through MR-unit based on FireWire technology is shown. The system is still under development. To the right the previously developed prototype is shown.

One purpose of a gaze controlled MR-system is to allow the user to interact with the computer even when both hands are occupied and normal computer interaction with a mouse and a keyboard is not possible or not appropriate. In figure 3.7 images from the gaze-controlled MR system are shown. In the upper part of the image three interaction sub windows, buttons are located. Each button has a unique purpose. In this case the left button's purpose is to give the answer "NO", the right button is used to give the answer "YES" and the middle button is used for the command "ACKNOWLEDGE". In the figure, the image to the left shows when the user looks at a Mixed Reality object and the image to the right shows when the user looks at the right button with the aim to execute the "YES" command. Notice that the button frame changes from red to blue which indicates that the computer has recorded "the press" on the gaze-controlled button. "The press" of the button is determined by the user's eye gaze dwell time in the designated area.

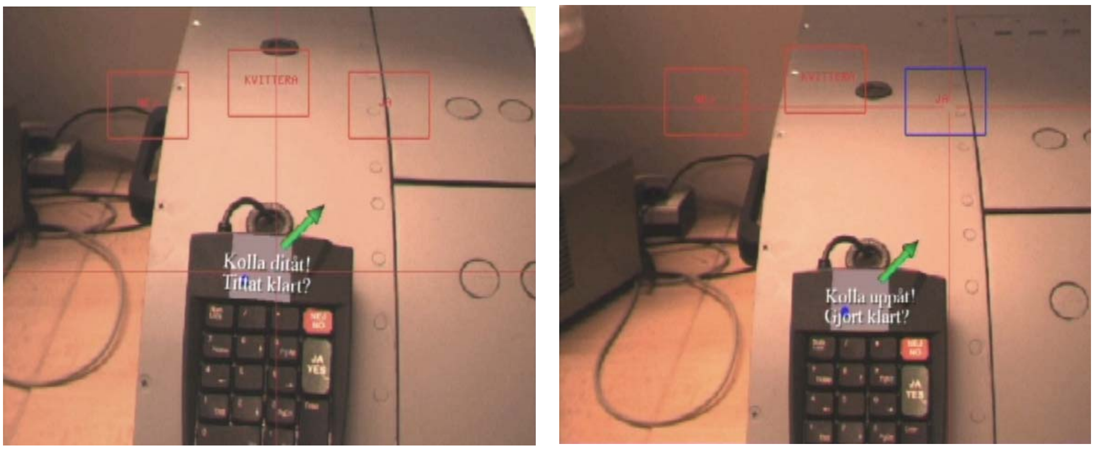


Figure 3.7 An example of gaze-controlled Mixed Reality System. The red cross shows the point of gaze.

3.2.4. The Turn-Round Gripen Mock-Up

In order to have a relevant test object when developing the MR-system a laboratory mock-up describing a part of the airfighter "Gripen (JAS)" was developed. The location of the part on the airfighter is shown in the figure 3.8 below.

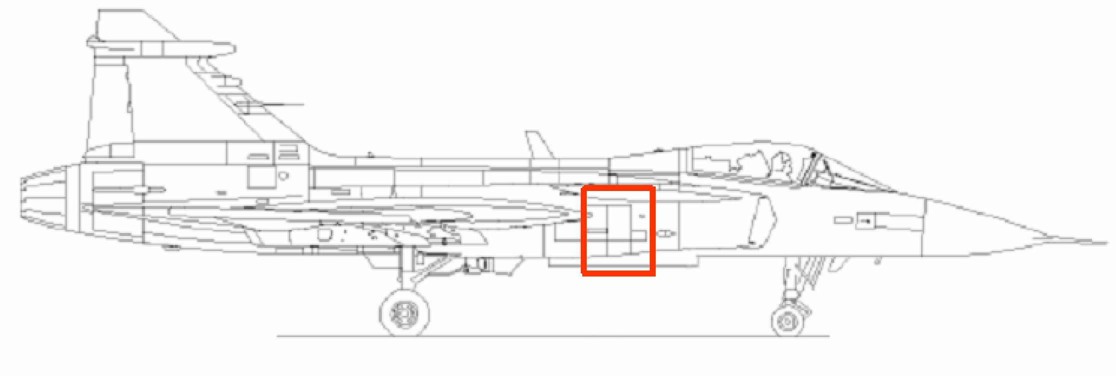


Figure 3.8 Location of the mock-up.

In figure 3.9 below the developed mock-up is shown to the left. The mock-up on the exhibition "Electronics/EP", February 1-3rd, 2005, is shown to the right.

At the exhibition an example of a Mixed Reality application was demonstrated. In figure 3.10 and 3.11 examples of Mixed Reality views from the exhibition example are shown.

Figure 3.9 The developed mock-up (left) and the mock-up at the exhibition "Electronics/EP" (right).

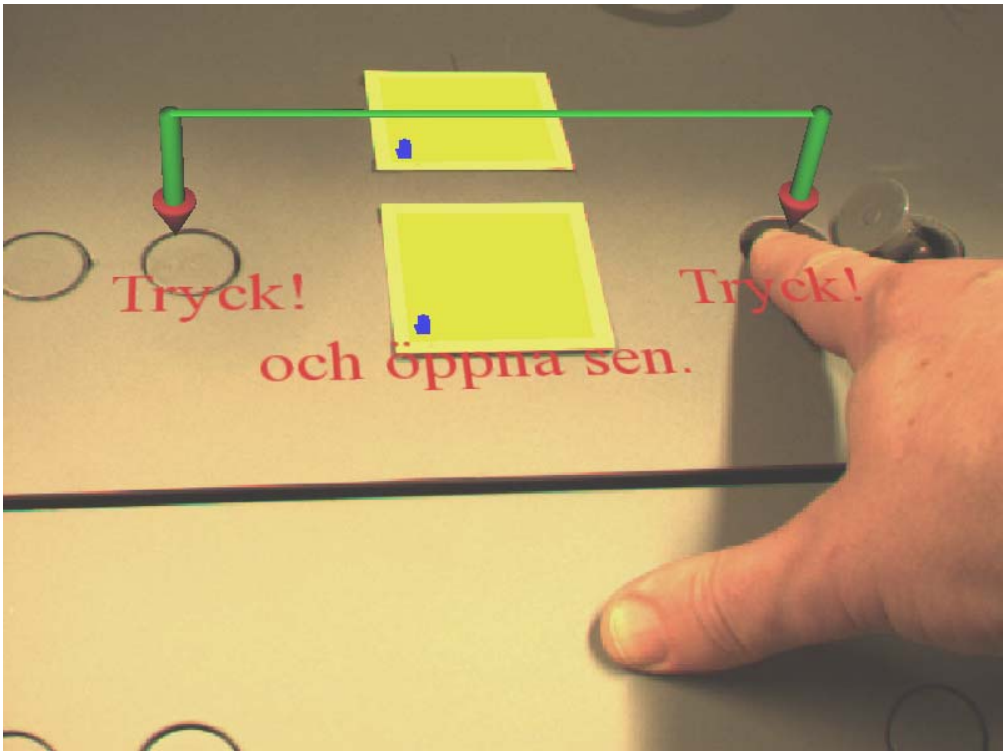


Figure 3.10 An example of a view from the exhibition application.

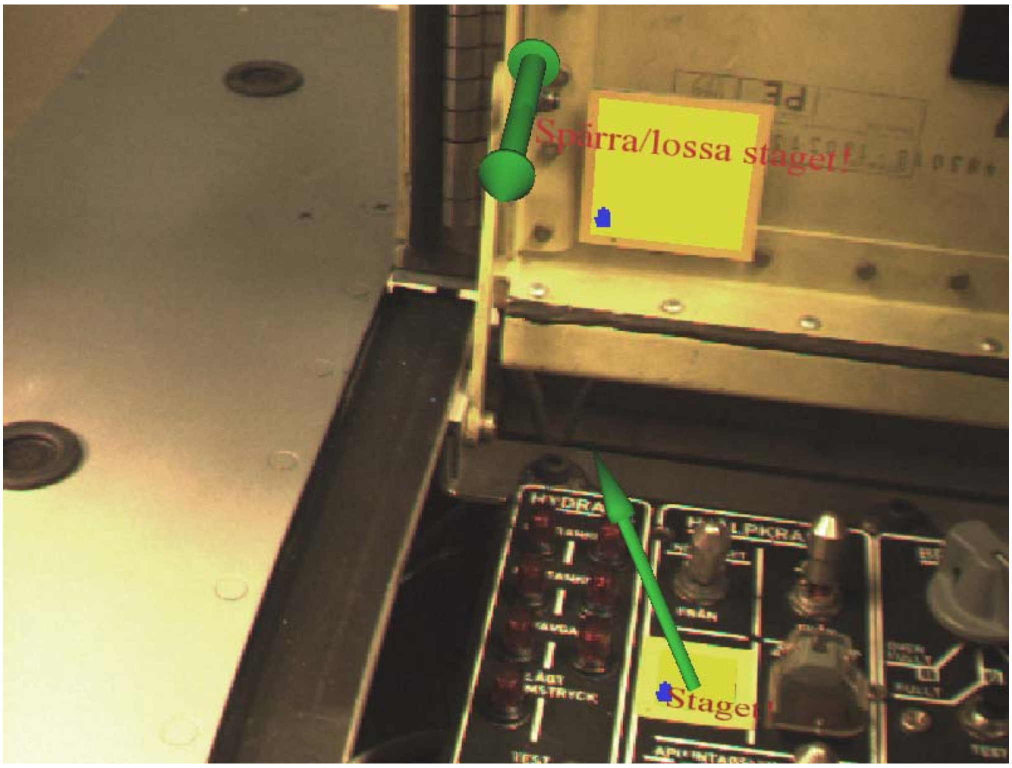


Figure 3.11 Another view from the exhibition application.

3.3. MR software

3.3.1 Recapitulation of phase 1

In phase 1 of the project an early decision was to use fiducial markers (indicates to the computer where to place the virtual information). A marker is simply a squared image printed on a piece of paper, see left image in figure 3.12 below. In the following the term marker will be used more as an object. A marker can for example have a certain pattern, be active, be acknowledged and for example show 3D text, as in the figure 3.12 to the right.

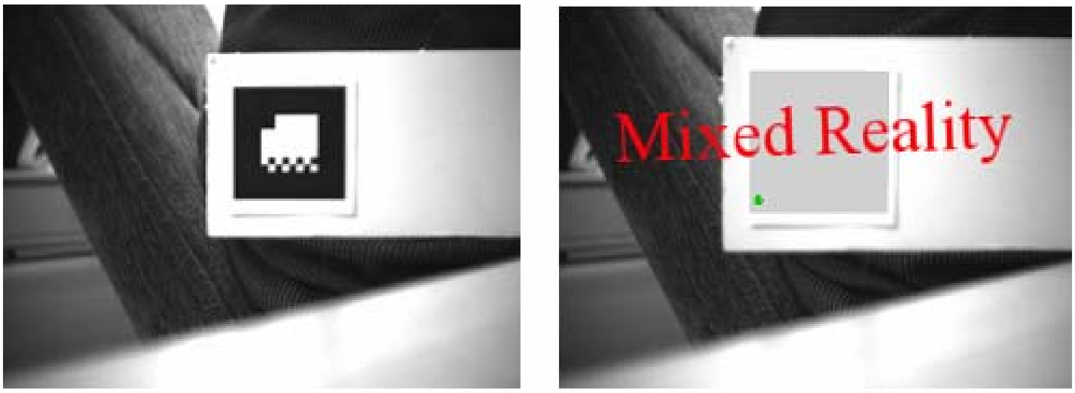


Figure 3.12 To the left a marker and to the right same marker with its overlaid information.

This approach with markers is much easier than general feature detection, and it has also proved to have several advantages compared to feature detection. When a marker is detected by the system, in the captured camera image, its size, position and rotation are determined relative to the camera position. With this information it is in the MR system possible to overlay the camera image with a 3D virtual model positioned at the marker location and rotated according to the marker rotation as in figure 3.12 above.

During phase 1 of the project lots of questions, such as exactly where to place the virtual model and how to operate/interact with the virtual model, came up. A list of such questions, requirements and demands was created in a “list of demands”. The list soon expanded and a priority had to be done. Each item has to be converted to a software XML-command, a scenario command, and its functionality must be added to the MR-tool, see chapter 3.1.

The approach requires that an application developer defines a scenario file, in XML-syntax. The scenario file is at runtime loaded by the MR-tool. One advantage of such an approach is that different scenarios in a rather simple way can be created and “programmed” outside the MR-tool. The main software development effort in phase 1 was therefore to both create a working system, the MR system, and a system that can load such a scenario-file, the MR-tool.

In figure 3.13 below we see a simple example of such a scenario file and an image of the result in the running MR-system.

```
<scenario>
  <name>example</name>
  <camerasettings>C:/MR/mr_camera_settings.d</camerasettings>
  <cameraparameters>C:/MR/camerap.d</cameraparameters>
  <threshold> 110</threshold>
   <patterndefinition>
    <filename>C:/MR/patterns/patt2.patt</filename>
    <width>20</width>
    <name>Tryck</name>
          <acknowledgedby>Ratten</ acknowledgedby >
          <prio>2</prio>
          <maxdist>700.0</maxdist>
          <action>
            <showimage>
                        <filename>
                C:/MR/images/tryck.bmp
                        </filename>
                <scale>100</scale>
            </showimage>
          </action>
   </patterndefinition>
</scenario>
```

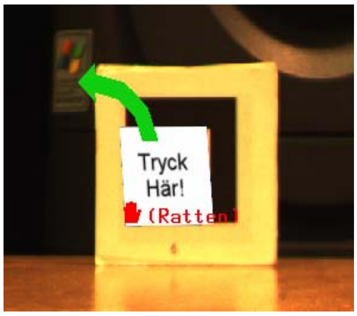


Figure 3.13 To the left an example of scenario file and to the right the result in the running MR system.

3.3.2 Phase 2

In this phase, phase 2, the software architecture and MR-tool as used in phase 1, have been further developed. A number of improvements, extensions and minor changes have been implemented as well as a general update. A number of new scenario commands have also been developed. As the gaze-tracker hardware has been rebuilt the MR software has been adapted to the hardware changes.

One fundamental issue in MR-tool is the concept of *acknowledgement* of a marker. If a marker is acknowledged it means that the user has executed the marker operation and can continue with new operations. If a certain marker is not acknowledged it will block other related markers from being activated.

To be able to work effectively, for example to acknowledge markers, an effective input device for interaction with the system is needed. As discussed in earlier reports input devices normally are handled by hands such as mouse, joystick or keypad However, it is also possible to interact with voice or gaze. A multi modal interaction could be beneficial. However, in this project the primary concern is a gaze-controlled MR-system. In a gaze-controlled system a quickly acknowledge can be done, for example, that a certain marker operation is completed. Furthermore, simple questions such as yes/no, go forward/backward, can quickly be answered.

This fundamental idea, to be able to interact in such a way and quickly acknowledge markers, is implemented in the MR-system. In systems not equipped with a gaze-tracker the keypad is used as a simple and first alternative to gaze-control.

3.3.3. Scenario commands

A large part of the MR software development concerns the development of new scenario commands and the coupling between them and between different markers.

The scenario commands are divided into the three types:

- *System setting commands* concern overall system setup parameters, such as camera parameters and camera calibration.
- *Marker definition/behaviour commands* concern marker naming and referencing, marker activation and deactivation, priorities, etc.
- *Marker display commands* concern which overlaid virtual information to display, for example 3D text, bitmaps, 3D models and how to display it.

3.3.4. System setting commands

The MR-tool has a settings file where system setup and camera parameters can be defined. Several additions have been made. One example is to match and tune the micro-displays field of view with the currently used camera lens. Another example is the possibility to change the overall marker detection sensibility (sensibility to find the right marker) and yet another is the possibility to stabilize the detected marker position.

Normally markers with a white background are used. The system is now adapted to also handle colored backgrounds, for example yellow. System setting commands are:

Command	Short description
camera	Defines the unique camera to use and its internal number.
exposure	Camera default exposure.
brightness	Camera default brightness.
blue	Camera white balance, blue part.
red	Camera white balance, red part.
autoexpose	Camera autoexpose, on/off.
take_one_camera	Use any available camera, on/off.
mrbg_leftright	Tunes the size of the background image for correct FOV.
mrbg_bottomtop	Tunes the size of the background image for correct FOV.
mrcamera	Internal camera number for background MR image.
brscamera	Internal camera number for gaze-tracker.
spycamera	Internal spy camera number.
livecamera	Internal live camera number.
fov	MR-camera field of view, for use by the graphics system.
adjustpos	Global adjustment of all markers position.
adjustrot	Global adjustment of all markers rotation.
nearclip	Near clip plane position.
farclip	Far clip plane position.
hidemarker	Hide marker mode on or off.
hidecolor	Defines colour of the “hidden” marker.
hideoffset	Defines how much of the marker to hide.
threshold	Default threshold in marker detection.
sense	Default marker sensibility.
edgesense	Camera edge sense filtering, on/off.
continous1	Use old marker position information, on/off.
continous50	Extended use of old marker position information.
ackmode	Start MR-system in acknowledgement mode, on/off.
slow	Run system in slow mode, on/off.
calibrationfile	Which camera calibration file to use.

3.3.5. Marker definition/behavior commands

The notion of active/inactive marker has been redefined since phase 1.

Currently, a marker is defined as active if it is identified by the system and if it is also displayed. An inactive marker is not active and not displayed.

Marker definition/behavior commands (or pattern definition commands) are:

Command	Short description
name	The marker internal name.
filename	The markers pattern definition file.
prio	The marker priority.
width	The physical marker width.
maxdist	The maximum distance from camera to marker (to make it active).
mindist	The minimum distance from camera to marker (to make it active).
maxangle	Maximum angle between marker and camera (to make it active).
minangle	Maximum angle between marker and camera (to make it active).
acknowledgeby	Defines which other markers that acknowledge this one.
acknowledgetext	Defines the default acknowledgement text for the marker.
acknowledged	Defines the default acknowledgement state of the marker.
active	Defines the default active state of the marker.
deactivate	Defines which markers to deactivate at marker acknowledgement.
activate	Defines which markers to activate at marker acknowledgement
connected	Which other (connected) marker to activate (at marker activation).
connectedpos	Position of connected marker relative to this marker.
acknowledgered	Defines color and size of acknowledgement icon (a hand).

3.3.6. Marker display commands

The intention with the list of demands is to test new ideas in new application scenarios. In new application scenarios a lack of functionality is often found, which in turn generates new demands to the list.

In phase 1 of the project one such lack of functionality was the possibility to have a number of yes/no questions on a single marker. With the new command "showquestions" (see below) this is now a possibility in the MR-tool.

Basic parameters for the display of a virtual model onto an active marker are:

scale	Size of the virtual model (default is 1.0)
position	Position of virtual model relative to marker centre.
rotation	Rotation of virtual model relative to marker orientation.

These parameters normally apply to all of the display commands below. Marker display commands (or action commands) and corresponding sub commands are:

Command	Subcommand	Short description
showimage		Displays a 2D bitmap image.
	filename	Bitmap filename.
	billboard	Billboard type (orientation to image).
showbook		Display a sequence of 2D bitmap images.
	filename	Bitmap filename (one page in book).
	startpage	Which page to display at start.
showvolume		Displays a 3D texture (a layered sequence of 2D bitmap images).
	filename	Filename of 3D texture.
	alpha	Alpha value at start.
showmodel		Displays a 3D model.
	filename	3D-model filename
	particlesystem	If the model is a particle system or not.
	nametoref	Reference name, for other markers to reference to the same model.
	refname	Reference to a 3D model loaded by another showmodel command.
	effect	Displays a graphics effect, for example bump mapping.
	animation	Defines an animation sequence of a 3D model.
	loop	At animation end, continue from start.
	swing	At animation end, go backwards in the animation sequence.
	speed	Animation overall speed.
	yaw	Animation direction
	offsetyaw	Direction offset.
	roll	Animation roll (rotation around main axes).
	point	Position of animated model.
	time	At this time use current position and rotation.
	chanim	Do character animation, time and cycle.
showtext		Displays 3D text.
	text3d	Text to display.
	color	Defines the text colour: red, green, blue and alpha.
	fontname	Text font.
	fontsize	Font size.
	shadow	Add text shadow.
	shadowpos	Position of shadow (relative text3d position)
	shadowcolor	Defines the shadow colour: red, green, blue and alpha.
showvideo		Displays a video stream onto the marker.
	filename	Name of video file (.avi).
	live	Connects the video image to a live camera.
showsound		Starts a sequence of 3D sounds.
	filename	Name of sound file (.wav).
	minvolume	Minimum volume level.
	maxvolume	Maximum volume level.
	startatactive	Start timer at marker activation.
	starttime	Start time of current sound.
	endtime	End time of current sound.
	looping	At sound file end, continue from start, yes/no.
showquestions		A mixed sequence of other display commands
	startquestion	Start item in the sequence of questions
	logic10	Defines the ”logic” behind each question, what to do if the question is acknowledged, answered with yes/no etc.
	algo	Help parameter in the question definition.
	activate	Activates another marker
	deactivate	Deactivates another marker.

In the figures below some different scenario functionalities from an MR example session are shown.

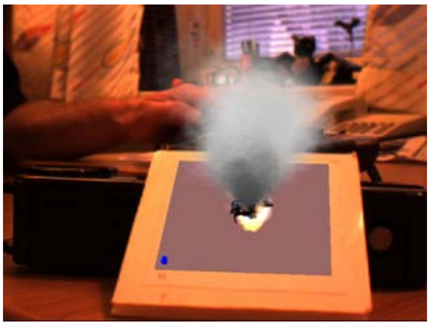


Figure 3.14 Example with a particle system.

Figure 3.15 Supporting questions, the user has in this example answered yes to the question to the left.

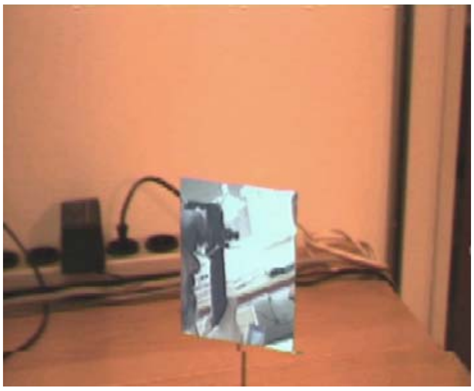


Figure 3.16 Live video onto a marker.

Figure 3.17 3D model and ordinary animation (the vehicle moves over the table).

3.3.7. Improvements of the MR-system

Marker display commands have as mentioned above a number of subcommands. A new subcommand is the subcommand <chanim> where full-featured character animations are included, see figure 3.18 below. In a character animation a more natural animation is achieved, in comparison with ordinary animation where only the models scale, position and rotation is changed. The idea behind this is to develop and investigate the possibilities with guided instructions both in technical maintenance applications as well as in “guided museum tours” applications. As a complement to this also a 3D sound has bee added.

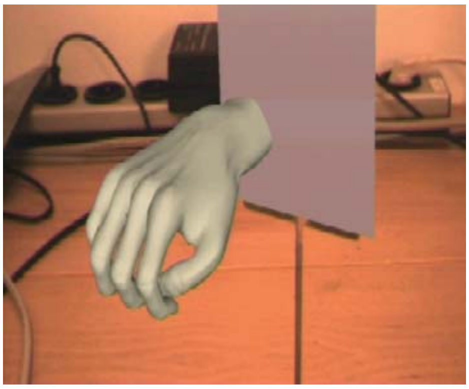


Figure 3.18 Character animation (the hand grasps naturally).

Another improvement of the MR-tool is the development of a new way of defining video sequences. With the new tool full control of video sequences, speed and start/stop is possible.

The system can now in real-time/on-line store a compressed AVI-file of an MR-session. This feature is useful in, for example, internal project discussions.

In the MR-tool it is now possible to do an internal restart of all acknowledged markers. In a next step it will be possible to move backward through a sequence of user operations.

Another new feature is the possibility to use a transparent overlay, se figure 3.19 below. The application designer creates an RGBA image to be transparently overlaid upon the ordinary camera image background. One idea behind this is to have available menu alternatives in this layer and also different kinds of information.

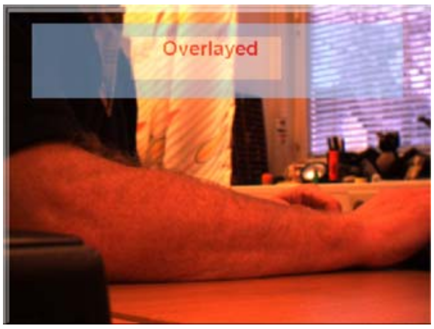


Figure 3.19 Example of a transparent overlay.

During the development work, tuning and marker stabilization issues have been studied and this work has resulted in a few new menu commands.

Volume visualization has been added to the MR-system, a sequence of layered 2D bitmaps is visualized via 3D texturing techniques. The data source is typically Magnetic Resonance data but also data from for example air flow or heat distribution computations. In another project, funded by VR (Vetenskapsrådet), these possibilities will be explored. The plan here is to integrate the MR-system with VTK (Visualization Tool Kit (VTK)).

Among other minor additions the possibility to use Swedish in showtext commands can be mentioned. This was earlier not possible due to the used XML parser.

Finally, the possibility to use two or more cameras in the system, have been added. One of the extra cameras can be used for capturing and displaying live images. Another camera can be used for marker detection outside the user's field of view. A fourth camera is used by the gaze-tracker.

3.3.8. MR-system architecture

The system architecture has the same layout as in phase 1 and will not be described here, see [Gustafsson et al 2004]. During application development ordinary Microsoft Windows menus, mouse and keypad input, are used. At runtime the user interacts with the system via the gaze-controlled functions or alternatively via the keypad.

One fundamental part of scenario development is the creation of 3D virtual models. In the project 3D Studio Max (3DStudio) have been chosen as the primary design tool. Models and animations can be imported from 3D Studio Max into the MR-system. Via the Cal3D software system (Cal3D), character animations from 3D studio can be imported. As secondary design tools AC3D (AC3D) is used for modeling and MilkShape (Milkshape) is used for modeling and character animation.

3.3.9. Gaze-tracker

Concerning the gaze-tracker, in phase 1 we transferred the core software parts from an earlier system to the MR-system. At that point only some gaze-controlled functions were transferred from the earlier system.

The gaze-tracker hardware has during phase 2 been rebuilt and is now equipped with a FireWire camera (in phase 1 an analogue video camera and a grabber was used). Most of the work has been concentrated on the physical rebuilding and miniaturizing of the gaze-tracker. To match the gaze-tracking software with the new device the software had to be adapted both to the new image capturing interface but mostly to the new physical layout, different placements of diodes, different camera resolution and different camera placement. Also the image processing software had to be adapted, changed and improved. Implemented gaze-controlled functions still only handles simple menu interactions.

3.3.10. Future techniques

In future MR-systems the user will be able to wear the whole MR-system, including the head-worn micro-display device with integrated gaze-tracker and video-see-through cameras. (This is in contrast to some definitions of “wearable computer” which refers to a small body-worn computer, e.g. user-programmable device, that is always on and always ready and accessible [Mann S, 1998].) Future systems will include wireless communication as well as a GPS (Global Positioning System) or other systems for positioning of the user.

To look into such wearable systems a handheld computer has been obtained, a personal digital assistant (PDA), the Dell Axim X50V [http://www.dell.com], see figure 3.20 below. The system weight is 170 g and includes, among other things, VGA output (to the micro display), wireless LAN and 3D graphics hardware support. This handheld unit is not powerful enough to run the whole MR-system but as a start minor parts of the MR-system have been transferred to it. The conducted work was solely for testing/learning purposes, what is possible and how to proceed in the next step towards a lightweight wearable MR-system. The MR system architecture looks a bit different on such machines. Full OpenGL [OpenGL] is for example not available. A limited version, OpenGL-ES (OpenGL for embedded systems) was used. The planned work with this handheld computer is not yet finished. In a next step investigation how to do wireless communication with this type of units will be performed.

Other products like the much more expensive OQO1, “the world’s smallest computer” [http://www.oqo.com] have also been studied. Unfortunately the OQO didn’t seem to be able to run the whole MR-system either. The new Mac mini [http://www.apple.com] weighs 1.5 kg which is too heavy for the purposes of this project. Further it does not run native Windows but seems otherwise to fulfill most of the other demands upon a wearable computer for Mixed Reality.

Figure 3.20 The handheld computer Axim X50V with VGA output, wireless LAN and 3D graphics hardware support.

3.3.11. Planned but not yet fulfilled activities

One earlier planned and prioritized improvement of MR-tool was to implement the possibility to interactively edit/change the scenario during runtime, for example tune the positioning/scaling of an object upon a marker. Included into this idea was of course the possibility to save the changes into a new scenario file.

As 3D Studio Max is used as the primary design tool, much better control upon matters such as position, size and color of a 3D mode, have been obvious. 3D Studio Max have all editing facilities needed, so therefore this planned item is not prioritized any more.

The improvements of the algorithms for marker detection, position and stabilization has been quit satisfactory. The partly planned idea to investigate other tracking possibilities such as to complement the marker detection with feature tracking of diodes has therefore not yet been tested. Marker detection is sensible to bad lighting conditions so nevertheless the planned idea to investigate other tracking possibilities is still appropriate.

3.4.12. Summary – MR software

The software development of the MR-system and the MR-tool for scenario definition has continued in phase 2 with the same overall software architecture as in phase 1.

Marker detection has improved and a number of new scenario commands have been developed as well as a general improvement of the system.

Concerning the gaze tracker and gaze-controlled functions most work has been on the adaptation to the new gaze-tracking hardware. Now we have a running system and we foresee more work concerning gaze-controlled functions in phase 3 of the project.

4. Gaze-controlled interaction in MR

Eye gaze interaction has been effectively used for over twenty years when it comes to eye typing, i.e. interacting with word processors through eye gaze instead of the traditional keyboard and mouse input [Hansen & Itoh et al 2004, Majaranta, P. et al 2004]. Over the last decade or so eye gaze interaction has also been adopted by other domains than that of system development for disabled people. However, gaze controlled interaction is very new to the Mixed Reality domain and the next step in this phase of the project is to implement gaze control in the system and before doing this, a limited survey of the research field has been done. The main focus of implementing gaze control into a helmet mounted mixed reality display is of course to make the interaction between the system and the user easier and more efficient. Eye gaze interfaces has the potential to become a very effective form of Human Computer Interaction [Ohno, 1998].

However, there is a second very important aspect to implementing gaze control into the system in this project – the ability to use the gaze control system itself for usability studies of the system. Eye gazing has been used to evaluate system usability as well as in psychological studies of human visual behaviour for a long time [see for example Hyrskykari, A. 1997, Oviatt & Cohen 2000, Qvarfordt 2004, see figure 4.1].

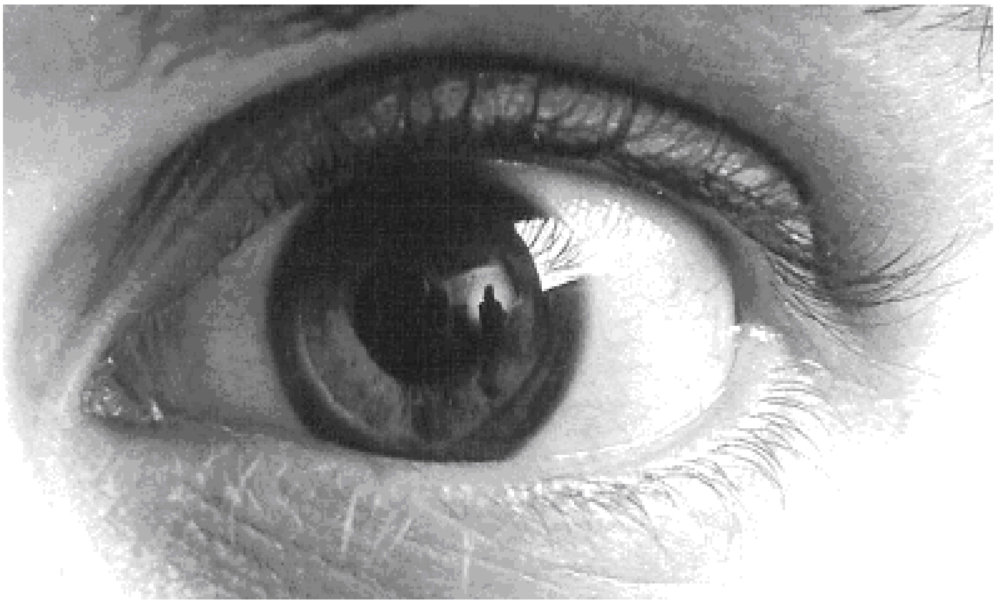


Figure 4.1 "Being a window to the mind, the eye and its movements are tightly coupled with cognitive processes" (Pernilla Qvarfordt, 2004)

Several studies have shown that eye gaze patterns differ between expert and novice users of a system or method [Law et al, 2004]. This implies that it is possible to use gaze patterns to see how a user interacts with a system and the changes in gaze patterns will also show how the users' skills improve during training. This would make it possible to use changes in gaze patterns to evaluate the effectiveness of a training method or the design of an interface. Usability testing methods presently involve a lot of trial and error when developing new interfaces. By using eye tracking of the users visual behaviour much time can be saved by directly giving answers to questions of layout, control functionality and visibility of different objects, [Karn et al, 1999].

The use of gaze control in the MR-system will also be useful during the development process of the interaction methods with the system and the design of the displays. Eye movements of the user can be used for interpretation of what the problem might be when the user does not handle the interface in the way it was intentioned [Karn et al 1999]. By analyzing the gaze patterns one might be able to see if the user has observed objects, if the gaze is distracted by other objects and so on.

However, for a gaze based interaction system to be useful, the gaze detection process should be implemented in a way that does not interfere with the user’s behaviour [Ohno & Mukawa, 2004]. A gaze-tracker should be able to work in different display configurations and in varying illumination situations and therefore the gaze-tracker often is a complex system. Many technologies are involved in a gaze-tracker, for example displays, optic devices, image sensors and real-time signal processing. Each technology, or its characteristics, naturally can affect the use of the tracker. For example, the gaze tracker’s accuracy and especially the systems transport delay are very important in real-time interaction applications. Another example could be the user’s sensitivity regarding near-infrared illumination that is used in most camera-based gaze-trackers. It is important to know the medical limits regarding this illumination of the eye, (Maximum Permissible Exposure, MPE). The near-infrared illumination can not be observed but high levels of illumination can give the user some sort of discomfort that may affect the user’s behaviour. Each user’s eye is unique and therefore a calibration process is essential for the tracker to work correctly. This calibration process should be simple and fast to perform not to put extra strain on the user.

Another area where gaze control is relevant is remote control where an operator controls a robot or other devices in areas that are not suitable for humans [Brooker et al 1999]. Similar functions can be found in distance expert help/support where an operator can get support from an expert somewhere else, who can see the same image as the operator sees.

Several studies discuss the potential problems of eye gaze based interaction [Vertegaal 2002, Ohno 1998, Zhai et al 1999, Zhai 2003]. One example of these problems is the ”Midas touch” problem that occurs when the user looks at an object, but does not mean to choose it but it activates anyway just by the first glance [Jakob 1991, Hyrskykari 1997, Ohno 1998]. Another problem occurs when using dwell time to determine whether a user has ”clicked” on an object or not, since it is hard to define what the exact time should be before the object is activated [Majaranta et al 2004]. If the time chosen is too short the ”Midas touch” problem occurs, if the time is too long the user might become annoyed. In both cases the user will probably consider the interface quite user unfriendly and will be reluctant to use it again [Ohno 1998, Jakob 1991]. If dwell time is used as a means of activating objects on the screen or in typing with eye gaze it is important that the feedback to the user is sharp and unambiguous [Majaranta et al 2004].

Another way to interact visually with the system is to use eye gestures such as winks, but this of course is very difficult since the system has to be able to differentiate between voluntary eye gestures and natural, involuntary eye gestures [Ohno, 1998]. However, in this application these problems are not the main issue. The use of eye gaze can be restricted both temporally and spatially where certain parts of the display will have the function, and only when there is a need for gaze interaction.

Another important limitation of using gaze movements to interact with a system is that gaze will probably illustrate visual interest, but not necessarily the cognitive interest of the user. It is one thing to determine if the user has had the information in the visual field of attention, a completely different issue to determine whether that information actually has been acknowledged and understood [Vertegaal 2002, Bates & Istance 2002]. In the MR-system in this project this issue has been solved by the use of “acknowledgements”, where the user has to give feedback to the system before the instructions move on to the next phase, thus forcing the user to at least reflect on the instruction before moving on. In the diathermy apparatus application (described in chapter 6.2) the instructions begin with three questions that have to be answered with a “yes” or “no” (at this stage this is done by pressing J or N on a keyboard) before the system gives further instructions and this will hopefully ensure that the user has acknowledged the information presented (see fig 4.2 and 4.3 below).

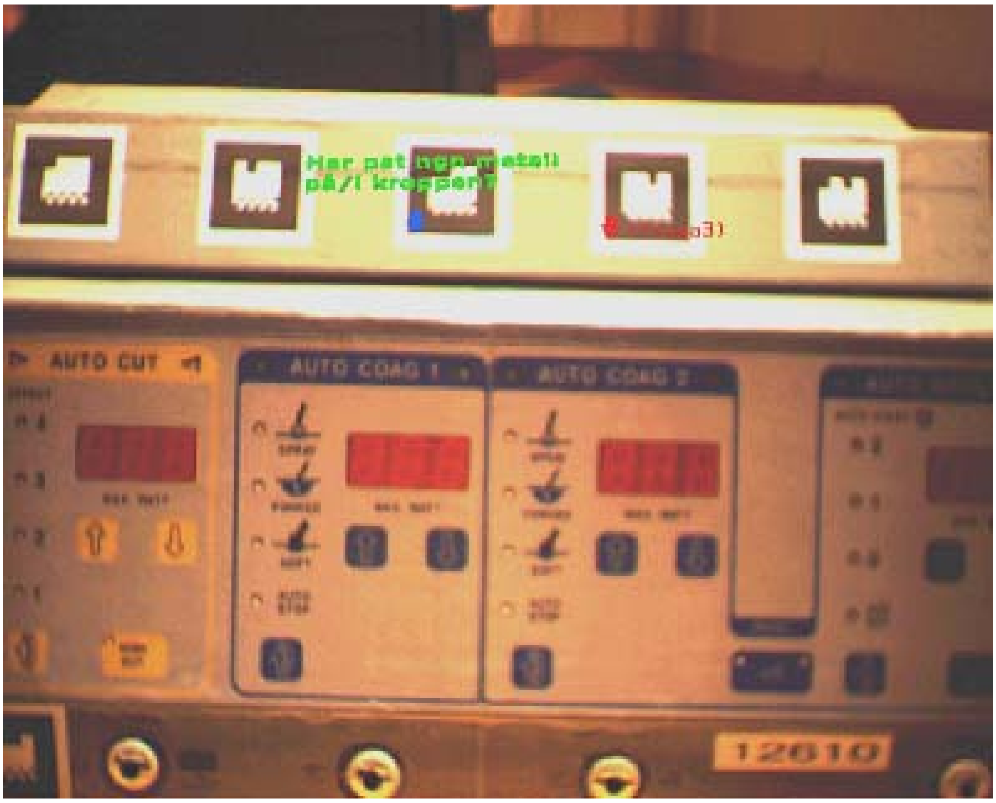


Figure 4.2 Image of the mock-up of the Diathermy Apparatus with markers. The blue hand below the question indicates that this question has to be answered before moving on. The marker next to it will only display a red message indicating that the previous step has to be completed before the next instruction will appear.

Figure 4.3 Close-up of the question to be answered by the user. A blue hand indicates that this question must be answered before moving on to the next instruction.

Users of gaze-controlled systems, in general, seem to appreciate this interaction technique – several studies have shown that even though it may not be perfect when it comes to accuracy etc users often prefer gaze and speech interaction to many other forms of interaction, such as hand based interaction (mainly keyboard and mouse) [Tan et al 2003, Oviatt 2000 & 2004, Qvarfordt 2004].

Eye gaze interaction is as a technique much faster than mouse interaction when subjects have to point at or, select objects on a display [Zhai et al 1999, Sibert & Jacob 2000]. When it comes to more complex tasks such as typing there have been no significant differences between gaze and hand input, but there appears to be a higher error rate for eye gaze [Hansen & Itoh et al 2004].

In the MR-system developed in this project there is so far no need for gaze typing tasks, but the gaze interaction would mainly be used for object or menu selection tasks where the positive effects already have been illustrated by previous studies [Ohno 1998, Zhai et al 1999, Sibert & Jacob 2000, Hansen & Itoh et al 2004]. These studies give incitement to continue the development of gaze based interaction in mixed reality settings, especially for use in situations where hand based interaction is not possible or not preferred, such as maintenance and repair of vehicles or surgery or other situations when the user's hands are occupied with other tasks. Eye gaze is also to prefer to speech interaction in situations where the environment is noisy or when speech is not appropriate for other reasons such as for example security.

5. Test case applications

A common feature of the applications chosen is that they involve equipment or material that demand a certain amount of knowledge and experience from the people working with it. The test case applications chosen for this phase of the project are first a set of basic routines involved in the “Turn- round of Gripen (JAS)” and secondly basic instructions for new and semi-new users of a surgical tool. The definition of semi-new users here is “users that knew how to use the equipment some time ago but because of not having used the equipment for some time are uncertain of their knowledge”.

The different applications have many common traits - the MR-system is in both cases used for user support and guidance through instructions. The instructions also include questions for the user to be sure she/he is solving the problem, or following the routines, correctly.

5.1. The Wearable Unit

To get closer to real performance a wearable unit was developed for the tests in phase 2. Another new thing is that the displays are assembled on hinges, so the user can have them in upper level when he/she does not need any virtual information (see figure 3.5). This makes the user more free to act and move without restrictions from the camera view.

The wearable unit consists of a Fujitsu Siemens Tablet Computer (Stylistic), a helmet with Sony Glasstron display II attached, Sony control unit, Sony AC adapter/charger with Sony standard battery pack, Belkin Fire Wire hub and Targus num pad (see table 6.1). The num pad, which is used temporarily during the gaze-tracker development phase, is for interacting. The computer and the accessories are carried in a small back pack. The num pad use Bluetooth for communication with the computer unit and is attached to the users arm. The approximate weight of the backpack with accessories is 3080 gram the num pad 125 gram and the helmet with mounted display 630 grams. This gives a total weight of 3835 gram for the mobile unit (see table 5.1). Those figures are for the prototype and will in the future probably be reduced as well as the size of the unit, due to technical advances. This holds in particular for the computer.

Component	Computer	Helmet	Control unit	Adapter	Hub	Batt Pack	Cables	Num Pad
Weight (gram)	1560	630*	335**	320**	45	315*	505	125

Table 5.1 The component and their approximate weights in gram.

* Not military class
** Inclusive standard battery package

The portable unit consists of four parts that need battery supply; those are the computer, the camera, the control unit and the num pad. The camera and the control unit use the same battery package. Most critical in respect to battery life is the control unit/camera with an operating time of 3 hours and 36 minutes. Least critical is the num pad with more than 4 hours of operating time* (see table 5.2).

Component	Computer	Control unit	Num Pad
Operating Time (h:m)	3:20	3:36	> 4:00*
Voltage (V)	16	8,4	2,4

Table 5.2 The different parts operating time with battery package.

* Test was interrupted after this time. The battery was still in good condition.

5.2. Markers

One new function in the MR-system is the possibility to present several sequential questions using only one marker. The demand of the function is a consequence of establishing that many instructions could be presented in series and/or done from one place.

With the two applications in phase 2 of the project, i.e. the diathermy and the JAS application, we start a new way of interacting by introducing the “multi level marker”. The multi level marker is one ordinary marker that is programmed to have several functions. Numerous multi level markers could be used in the same application. In our first application, the dismounting of a machine gun mounting, the system interacts with the user by giving instructions. When one instruction is fulfilled the user acknowledges the performance of the present instruction before getting the next instruction.

With the multilevel marker it is possible for the system to “ask” the user about different circumstances and give feedback depending on the answer. This possibility can be used in different manners, depending on how the instructions/questions are formulated. The user reacts to the instruction/question by answering with YES, NO or ACKNOWLEDGEMENT either by gaze control or by some kind of keyboard/num pad. In the current application there are three possible responses, however, there are no constraints on the number of responses other than those of usability and limited interaction space in the display.

The multi level-marker is used in different ways in the two applications. In the diathermy application two multi level markers is used. The first is for check of patient status, but also to control connections between patient and diathermy device. The second multi level marker is used for adjusting the device. In the JAS-application one multi level marker is used for the instructions and several common markers (i.e. one for each point of interest) for guidelines with text and animation. The main reason for using different implementations is the difference in size between the objects, the medical instrument and the aircraft, and to test different methods of multi level marker use.

5.3. "Turn-Round of Gripen (JAS)"

Before every flight operation an aircraft has to go through several check ups. The joint name for these check ups is "turn-round". The turn-round is performed by ground personnel; in the air force the ground personnel usually consists of conscripts (persons called for military service).

The aim of this application is to give the research team a better understanding of how the MR-system operates in a real situation and how the user experiences the MR-system. Does the system improve the learning ability and does it reduce the need of training for special purposes? What does the user think of wearing the MR-system; is it awkward and disturbing or does it fit into the normal equipment for a soldier in a convenient way? Those are some of the questions that this application tries to answer. Other questions are of more technical interest, such as problems and benefits of markers and their performance.

The application has so far, mainly been performed at the wing Malmen and at the nearby Air Force museum of Linköping, and with assistance from conscripts and fulltime employees at those establishments.

Several tests have been designed for the application and some are already in progress. The problem with most of these tests is the uncertainty of what is actually tested. In other words - is an unsuccessful result dependant on the system itself or does it mainly depend on the pedagogies used in the system? On the other hand, a successful result will rarely occur if one of these, the system or the pedagogy, has serious shortcomings. (A deeper discussion about the pedagogy follows further down.) The planned tests and those in progress are described below:

5.3.1. Learning ability

Is it possible to have totally inexperienced staff to perform technical work without giving them instructions except those given in the MR-system and if so, how well will these inexperienced people complete their tasks compared to people with relevant experience or knowledge of the tasks? To answer these questions the turn-round "pos 5" of Gripen (JAS) will be used.

Of course there could be circumstances that make it impossible to use inexperienced staff for the actual work e.g. safety rules or other rules that must be followed but the intention of the test is only to se if it is possible, not to change the working methods.

As test subjects our intention was to use conscripts, whose duties do not normally involve the "pos 5" of the Gripen (JAS) turn around. But different circumstances made this impossible. Instead we used students from the aero technical program of Anders Ljungstedts Gymnasium, Linköping. The students were placed into three groups with four persons in sub group A and B and three persons in group C. It should be emphasized that the test group consist of young (16 - 18 years) people with a technical and especially aero-technical interest. Only one person was a female. Each group performed the test twice and performance time as well as number of errors was recorded. After the tests each group was interviewed about the test and had the opportunity to give their view of the system.

5.3.2. Support ability

Could the MR-system be used as support for those with sufficient education and if so is it comparable to, or could it at least without disrupting the work, replace the human support of an instructor?

A main purpose for the MR-system in this project is to act as a “computerized instructor”, see figure 5.1. This function is especially important in the case where there is a physical distance between the person who needs assistance and the instructor. As in other educational or instructional related situations it is very important to achieve flexibility in the use of future technical systems as it can provide practical as well as economical advantages.

Figure 5.1 A member of the air force performs pos 5 with MR assistance.

5.3.3. Users’ experience

What do the users of the MR-system think about using it? Do they experience the system as awkward or do they think that the system is helping them in the completion of the task?

The user’s experience of the system is very important for the continuous development. Their opinions of the usability in the system are one of the most important quality aspects. If the users think that the system is difficult to use, they will avoid using it. At the same time the research group has to separate problems arising from the equipment from problems that have their ground in the interaction with the system.

It is the latter that is of highest importance for the research group to identify. Problems that have their ground in the equipment are of course of the same dignity but most of those are of a technical nature and in this phase of the project it is too early to concentrate on those questions. See figure 5.2.

Figure 5.2 One of the instructions from " pos 5" in the MR view. The text and the photo are virtual information.

A survey of the subjects' experience with the system will be performed. Questions that will be addressed in this survey include usability problems and how the user experiences the system and a comparison will be done between the novice and skilled users.

5.3.4. Test Result

The JAS application was tested on a group of 14 students from a civil aircraft technical education. 13 of the participators were male and 1 was female. The group was divided in 3 sub groups. Subgroup one performed the "pos 5" twice after that they had read a written instruction, sub group 2 first did the pos 5 after they had read the written instruction and then a second time using the MR system and without written instruction. The third sub group performed the "pos 5" twice, both times without written instruction. The results appear to be positive in the meaning that the test groups using the MR system did not find the system bothering or difficult to use. However they took longer time to fulfill the task when using the MR-system, but with fewer errors than the group that get standard instructions. When wearing the MR unit all test persons used the possibility to use the hinges attached between the helmet and the display.

Some problems occurred during the test. It was in the connection between the computer and the special Bluetooth num pad used during the test and had nothing to do with the MR-system itself.

The conclusion of the test is that the MR system, in a physiological aspect, does not bother the user when working with minor tasks but that it takes longer time to perform the task. This was also indicated in earlier tests in phase 1 of the project. We think that it would be a great advantage to make a test on a more complex task with several steps. This will make it harder for the person who performs the written task to remember each step and it probably also shows if there is any negative effect from using the MR system.

6. A proposed model of a Gaze-Controlled MR-System

How to best interact with a gaze-controlled MR system? To get a better understanding of in which direction the development of a new MR system should take, the following proposal is given, to be discussed under the development phase to get better insight into problems that might occur and possible solutions to those problems.

First of all, the gaze controlled MR-system has to be easy and intuitive to understand in respect to both handling of the equipment and the interaction with the system. Therefore, it is suggested that a head mounted part is mounted on some kind of hinges. When this part is in the upper position the system is shut off and when the part is in the lower position it is turned on. When the system is started, by placing it in the lower position, it should instruct the user on how to interact with it.

Nevertheless the system still has to give the trained user the opportunity to skip the learning process [Nielsen J, 1993]. This is probably easiest done by the use of a gaze tracked interaction point. If no interaction with the system has started in X time after the system has been started it begins to instruct the user how to use the system. It would be desirable if the user could both stop and start the instructions whenever she/he feels like it. The know-how that the system has to give the user when the system starts is, in our opinion, how to interact with it, alternately if the user knows the system, give the opportunity to go to one of the other modes (in this proposed model, the Passive System, the MR-System, the Gaze Controlled MR-System or the combined Gaze Controlled MR-System). See figure 6.1.

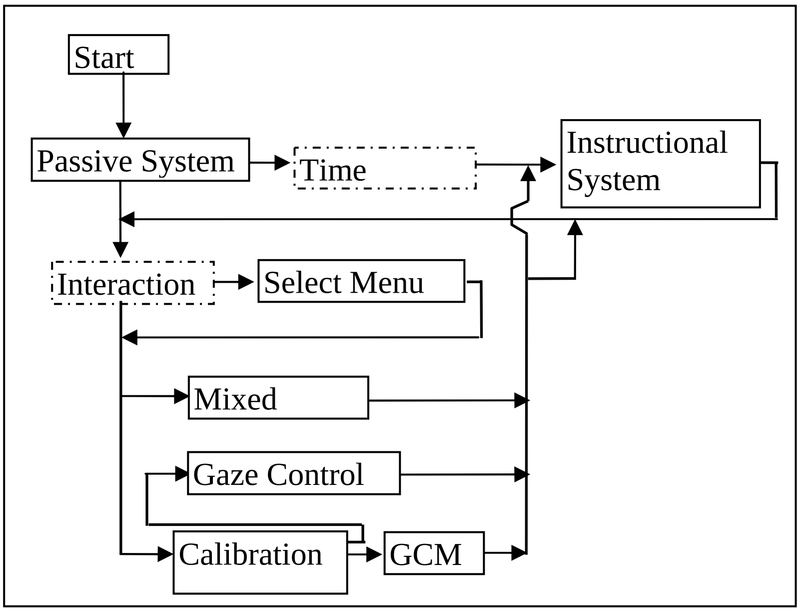


Figure 6.1 A schematic view of the different modes in a Gaze-Controlled MR-System. (GCM=Gaze-Controlled Mode)

6.1. The Passive System

The Passive system is the start up mode. In this mode no markers are detected and the only Gaze Control that is available is through the interaction point. This is also the mode that would be in use to get an overview over the surrounding environment without any augmented information displayed. Besides the overview possibility this mode is probably useful for adapting the eyes to “See-Through-Systems”.

6.2. The MR System

In this mode the system detects markers and shows the augmented information that is connected to those markers. The visualization connected to markers could be text (2D or 3D), object characters, pictures, animations, films, speech and other kinds of communication support. In this mode different visualization helpers called “object characters” are used. The object characters interact with or point to real objects. The characters can be divided into two specific areas; the standardized object character and the project specific character. A project specific character could for instance be a picture or a model of a certain building and an object character could for instance be an arrow pointing on an object in view. A proposal is that:

- Object Characters can exist on objects and in text, but not in menus.
- Object Characters can exist in different colours but not in the same colours as menu characters (se further Menu Characters).
- Object Characters are initiated by markers.
- Object Characters are of two types, the standardized object character and the project specific character.
- The same pattern could be used for standardized object characters throughout the whole project for the same type of interaction if not the project itself demands another solution

A library with suitable Object Characters ought to be made. Examples of suitable Object Characters beside those described below (see figure 6.2.) are different tools and urgent requests. The characters could be made in 2D or 3D, animated or non-animated.

The Object Character is closely connected to the marker but it could also include other types of sensor information.

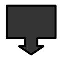
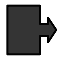
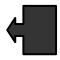
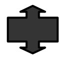
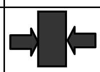
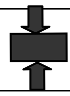
Denote	Object character
Rotation, clockwise	 alt. 
Rotation, counter clockwise	 alt. 
Lift, pull, pull down	 alt. 
Pull to left/right	 alt. 
Pull apart	 Alt 
Push together	 alt. 
Stop, Do not …	

Figure 6.2 Examples of Object Characters

6.3. The Gaze-Controlled System

The gaze-control can have several functions; mainly it should be used for interacting with the system (above described in The Instructional System). The user should also be able to use the gaze control to work within menus concerning the observed object or the activity in progress. The user should also be able to ‘point’ with his/hers eyes on different objects and a second person looking at a display in another place should be able to see which object the user is ‘pointing’ at. This means that the user could get external help from an expert.

The gaze-control demands a very precise measurement of the user’s eye throughout the whole operation. Therefore, the system has to be calibrated to the user’s eye before it could be taken in use. This calibrating is done once per user. In the Gaze Controlled mode there is no interaction with markers or other sensors.

6.4. The Gaze Controlled MR System

This mode is a combination of the MR-System and the Gaze Controlled System. All functions from both modes are available. The other modes are, as in all modes accessible by the interaction point.

6.5. The Need of System Interaction

When using a MR or a Gaze Controlled MR-system for instructional issues the user in some way has to inform the system that an instruction has been understood and fulfilled and that she/he is ready for the next one. This is done by some kind of feedback to the system. If the system does not get the feedback it will not allow the operator to go further (i.e. does not give any further instructions). There is one drawback with this feedback system. There is always a possibility that the user for one reason or another does not wish to start from the first instruction. In some cases the user just shuts off the system because he/she, for some reasons take a brake. The work might be resumed after a couple of minutes, hours or in a couple of days. Maybe it is another user with another set of equipment that should prepare the work. Anyway, the situation when the user may want to continue with or without the system occurs. It then will affect the user negatively if she/he has to go through all the steps in the MR instruction once again. One solution to this is a special menu that allows the operator to step to a selected instruction, *the select menu*.

6.6. Menu Characters

Menu characters can be used to navigate in the select menu, see figure 6.3. The menu is placed in a frame in the display area or as a transparent layer in the view and is used to move forward, back, up or down in menus, it could also be used in documents. The menu characters refer to the augmented information that does not directly relate to reality. The menu characters ought to be unique in form and colour, not to be mistaken for object characters.

Denote	Menu character
Back	←
Forward	→
Up	↑
Down	↓
Stop, Last	■ alt. ❙

Figure 6.3 Example of what Menu characters could look like.

A proposal is that menu Characters only exist in menus and documents.

- Menu Characters only exist in one unique colour
- Menu Characters are not allowed to have the same shape or colour as object characters.

There exists a possibility to use hidden interaction points instead of using menu characters. However, the use of hidden interaction points could, at least in the initial phase, make the user confused in several ways. Examples are -Where is the interaction point? Am I looking at the right point (if the user feels that it takes to long time to get feedback)? and interaction by mistake (i.e. the Midas Touch problem described in chapter 4).

As mentioned above there is a possibility to use transparent menu characters and place them inside the viewed image. Being transparent, they do not block information from the real world, but as soon as we place characters in the display area there is always a risk that they may be confused with the background. Therefore it is preferable to have a special area for the Menu Characters.

6.7. Symbols or text

When working with different applications such as the dismounting of a machine gun (phase 1), starting-up a diathermy apparatus (phase 2) and the turn-round of the SAAB air fighter “Gripen (JAS)” (phase 2) the written information or instructions have a tendency to be too long. Some kind of symbol language would be preferable. It is not only that it would require less space in the display area, it is also possible to have a worldwide recognition symbol system, hence follow the opportunity to use the same language everywhere and this would be preferable, especially as both civil and military applications more and more tend to be of international nature.

The use of a symbol language seems to be an important part of the development of an MR-system and its usability. A successful MR system is strongly related to pedagogy and one important part of that pedagogy is how the instructions are presented for the user.

7. Other research groups

This section gives a brief overview of projects within the Mixed and Augmented Reality area with some relevance to the project "Mixed Reality for Technical Support". The projects listed here are not a complete list of ongoing or completed AR/MR projects, but simply a selection of interest for our research group. In general there are a limited number of groups in this field, for example many of the projects listed below are conducted or have been conducted at Teknische Universität in Graz, Austria in the more general AR project called "Studierstube".

7.1. The ARVIKA project

A German project (funded by the German ministry of education and research and including sponsoring participants from Audi, Airbus, BMW, Zeiss, DamilerChrystler, Ford, Siemens and several more) during 1999 and 2003 [Friedrich 2004]: "The co-ordinating project **ARVIKA** aims to research into and to realise **Augmented-Reality-Technologies (AR)** which will support **development**, **production** and **servicing** with relation to complex technical products in a user-oriented and application driven manner. Through the visual superimposition of real objects on computer generated virtual objects, Augmented Reality technology makes it possible, to act appropriately in real, in the sense of this **extended reality**, working situations [...] ARVIKA achieved its goal after four years of project run time: The first mobile Augmented Reality Systems for industrial applications have been developed. More than 200 attendees and subsequently published press reports attested a positive response."
(http://www.arvika.de/www/index.htm last checked 2005-06-08)

Of relevance to "Mixed Reality for Technical Support": User-oriented, real-time technology, HMDs.
Main difference "MR for technical Support": Does not use gaze control, main focus on industrial applications.

7.2. ARTESAS

A project coordinated by Siemens, funded by the German Ministry for Education and Research and supervised by the DLR (German Aerospace Center). The main focus of the project is to develop markerless procedures for industrial environments with regard taken to ergonomical and usability questions. The goal is to implement the system in industrial application fields such as in service for vehicles and aircraft: "ARTESAS aims at the exploration and evaluation of Augmented Reality base technologies for applications in industrial service environments. The project is based on the results of the ARVIKA project".
(http://www.artesas.de/ last checked 2005-06-08)

Of relevance to "Mixed Reality for Technical Support": User-oriented, technical maintenance and support applications.

Main difference "MR for technical Support": Does not use gaze control and the main focus is on industry applications.

7.3. The AMIRE project, Authoring Mixed Reality

Amire was founded within EU:s 5th framework and the participants of the project were:

- Labein (Spain)
- Siemens Business Services C-lab (Germany)
- Fraunhofer AGC (Germany)
- Polytechnic University of Hagenberg (Austria)
- OMV AG (Austria)
- Guggenheim Bilbau Museum (Spain)
- Talent Code OY (Finland)
- Helsinki University of Technology (Finland)
- Institute for Applied Knowledge (Austria)

The project started in April 2002 and ended October 2004. The main goal of the project was to enable non-expert researchers to use MR for their applications and to create and modify these MR applications with the support of dedicated tools that foster an efficient authorizing process for MR. Within the project two demonstrators of different fields were developed, a museum application to enhance the visitor experience and a training application for an oil refinery. The most representative and innovative aspects of the project were (according to the project description);

- integration of different tracking systems (starting from ARToolKit)
- integration of open source 3D graphic libraries (OpenGL, OpenSceneGraph) and 2D graphic library (glGUI)
- XML based persistent implementation in order to export component information
- diversity of wearable hardware devices (Tablet PC, PDA …) for end users

One of the experiences from the AMIRE project was that interacting by speech is not enough, especially when the application is to be used in noisy places such as an oil refinery.

(http://www.amire.net/index.html. Last checked 2005-06-08)

Of relevance to "Mixed Reality for Technical Support": Human Factors, training and education.
Main difference from "MR for technical Support": Does not use gaze control.

7.4. The VIRTHUALIS project

A continuation of the AMIRE project is the VIRTHUALIS project. The overall objective of the VIRTHUALIS project is to reduce hazards in production plant and storage sites by addressing end-users' practical safety issues, such as training control room operators, designing proper alarm systems, training teams to cope with emergencies, assessing the impact of plant modifications on operators' reliability, helping managers to see the impact of their decisions on sharp-end operators' daily work, through the development of an innovative technology. The new technology will be achieved by merging Human Factors (HF) knowledge and Virtual Reality (VR) technologies. The VIRTHUALIS project starts in June, 2005.

Of relevance to "Mixed Reality for Technical Support": Human Factors, training and education.
Main difference from "MR for technical Support": To early to say as the project just started.

7.5. Educating Spatial Intelligence in Augmented Reality

Augmented Reality and Virtual Reality in mathematics and geometry education to improve learning by using human spatial abilities. The system used in the study is called Construct3D. (http://www.ims.tuwien.ac.at/research/spatial_abilities/ last checked on 2005-06-03)

Of relevance to "Mixed Reality for Technical Support": Educational issues of interest.
Main difference "MR for technical Support": Does not use HMDs or gaze control.

7.6. Mobile Collaborative Augmented Reality

"Our aim is to develop a mobile AR system that combines the following features:

- 3D stereographic display of computer generated images. This enhances the sensation of seeing virtual objects.
- Direct manipulation and interaction with the virtual objects.
- Instant collaboration with another user wearing another mobile kit.
- Application and data sharing between a mobile and a stationary user or two mobile users

We are using standard of-the-shelf equipment as well as industrial components to assemble our mobile computer kits. This allows simple and fast experiements with an abundance of user devices and to deploy them in new, unthought-of ways."
(http://www.studierstube.org/projects/mobile/ last checked 2005-06-03)

Of relevance to "Mixed Reality for Technical Support": Mobility.
Main difference from "MR for technical Support": Does not use HMDs or gaze control.

7.7. AR Furniture Assembly

Shown at Siemens Forum Vienna in 2003: “To facilitate this issue, AEKI generates a sequence of 3D graphics, showing a model of the furniture to be built in subsequent steps of construction. It then superimposes that model over the actual view of the scene as captured by a video camera. This sequence can be traversed forward and backwards (hence the name “AEKI”) as necessary.

Unlike printed assembly instructions, these 3D models can be viewed from an arbitrary angle, allowing for easier visualization and verification of the construction process. In addition, every single part has been outfitted with a rectangular marker pattern (*"fiducial"*) to make it recognizable by a computer vision algorithm. In this way the system can identify and highlight the correct piece required for the next assembly step.”
(http://www.ims.tuwien.ac.at/~thomas/siemensforum.html last checked 2005-06-03)

Of relevance to “Mixed Reality for Technical Support”: Instructions, use of ARtoolKit.
Main difference from “MR for technical Support”: Does not use HMDs or gaze control.

7.8. AR tracking technology

Research group focusing mainly on visual tracking and object recognition for use in AR and VR applications. (http://www.emt.tugraz.at/~tracking/ last checked 2005-06-03)

Of relevance to “Mixed Reality for Technical Support”: Object recognition interesting, but not relevant to our current project.
Main difference from “MR for technical Support”: Does not use gaze control, ARToolKit.

7.9. Handheld Augmented Reality

Another research project at the Graz Teknische Universität aiming to develop a truly mobile AR system: “Our solution provides the first fully autonomous AR implementation that works fully autonomous on a standard PDA, including video see-through 3D rendering and optical tracking with optional backend server support. The system features the following characteristics:

- an autonomous vision-based tracking system executing at interactive rates
- 3D scene rendering via a standardized graphics interface
- optional dynamic workload-sharing with a nearby backend server in selected areas, i.e., outsourcing the computationally expensive computer vision calculations to the server via a wireless network, if available, in order to speed up the overall process

- optional high-precision outside-in tracking of the PDA in a prepared environment
- integration of the handheld platform's software into our AR research framework Studierstube for mutual re-use of resulting software components between workstation/notebook and PDA-based AR".

Of relevance to "Mixed Reality for Technical Support": Mobility.
Main difference from "MR for technical Support": Does not use HMDs or gaze control.

7.10. Outdoor Collaborative Augmented Reality

The same research group as the handheld AR project described above: "A common application that can benefit from immediate access to information is navigation in general. In our work we try to support a pedestrian trying to navigate through a city. Instead of relying on maps and knowledge about the environment a user can enter the desired destination into the system and perceive a series of artificial waypoints blended into her view of the surroundings. Following these waypoints will automatically lead her to the destination.

Another interesting aspect of mobile computing systems is information display and visualization. Tourists can view interesting dates about buildings and places overlayed over their view of the objects of interest. Engineers and architects can perceive information about buildings, virtual models of buildings in place or other abstract information such as cell phone network strength and quality.
A central focus of our work is collaboration. We investigate user interfaces that can support several people equipped with mobile augmented reality setups in the tasks described above. Navigation tasks between several users can include following another user, establishing and finding a meeting point or guiding another user. Shared information displays allow users to highlight information for other participants or control their view of the presented data."
(http://www.ims.tuwien.ac.at/research/mobile/ocar/ last checked 2005-06-03)

Of relevance to "Mixed Reality for Technical Support": Mobility, collaboration, multi-user environment, similar hardware setup
Main difference from "MR for technical Support": Does not use gaze control

7.11. Bridging Multiple User Interface Dimensions with Augmented Reality

"*Studierstube* is an experimental user interface system, which bridges multiple user interface dimensions. At its core, it uses collaborative augmented reality to incorporate true 3D interaction into a work environment. This concept is extended to include multiple users interface dimensions into a single system:

- multiple users (for collaborative work, both co-located and remote)
- multiple networked hosts and platforms (IRIX, Windows, Linux)
- multiple display types (head-mounted display, Virtual Table, projection wall, ...)

- multiple concurrent applications (for multi-tasking work, drag and drop, ...)
- a multi-document interface for 3D data

All this happens almost totally transparent to the application programmer, so it is convenient to write applications that use all these dimensions. With this architecture, we can explore the user interface design space between pure augmented reality and the popular ubiquitous computing paradigm."
(http://www.cg.tuwien.ac.at/research/vr/studierstube/multidim/ last checked 2005-06-03)

Of relevance to "Mixed Reality for Technical Support": collaboration, multi-user environment,
Main difference from "MR for technical Support": Does not use gaze control.

7.12. Augmented prototyping for industrial design

Research conducted at the Fraunhofer Institut Graphische Datenverarbeitung on how to improve industrial design through overlaying virtual information on real prototypes: "Today, virtual reality and virtual prototyping techniques are essential to the modern product development process. Reviews and simulations using virtual products complete reviews of physical models." This project also uses haptics to give force feedback to the user interacting with the virtual prototype.
(http://www.igd.fhg.de/igd-a2/projects/MixedRealityAP/index.html/ last checked 2005-06-08)

Of relevance to "Mixed Reality for Technical Support": demonstrating and testing without an actual physical product.
Main difference from "MR for technical Support": Does not use gaze control

7.13. Augmented Reality Aided Surgery (ARAS)

A collaboration between Zentrum für Virtual Reality und Visualisierung Forschungs-GmbH and Universitätsklinik für Chirurgie and Universitätsklinik für Radiologie in Graz.
"To aid the surgeon during intra-operative planning by displaying computer-tomographic data and ultrasound data as three-dimensional objects. Furthermore we have to support a close collaboration between surgeon and radiologist by providing a video-link."
(http://www.vrvis.at/br1/aras/ last checked 2005-06-03)

Of relevance to "Mixed Reality for Technical Support": Very relevant - collaboration, multi-user environment, similar hardware setup, HMDs, surgery setting, visualisation in real-time 3D.
Main difference from "MR for technical Support": Does not use gaze control

7.14. TUG – Liver Surgery Planning System (LSPS)

A project adressing the obstacles in surgical planning of segment oriented liver resections based on CT data. Hardware in the system: stereoscopic see-through head-mounted displays (HMDs), optical tracking system, tracked input devices (tracked pencil and transparent plexiglass Personal InteractionPanel), rendering workstation and tracking workstation. The Augmented Reality System uses head-tracking for correct stereoscopic visualisation. The project is run by the Teknische Universität in Graz. (http://liverplanner.icg.tu-graz.ac.at/ last checked 2005-06-03)

Of relevance to “Mixed Reality for Technical Support”: Surgery planning, HMDs,
Main difference from “MR for technical Support”: Uses a see-through display system, tracked input devices other than mouse/keyboard. Does not use gaze control.

7.15. UNC Laparoscopic Visualization Research – AR Technology

Research project since 1992. “Emerging augmented reality (AR) technologies have the potential to bring the direct visualization advantage of open surgery back to its minimally invasive counterparts. They can augment the physician's view of his surroundings with information gathered from imaging and optical sources, and can allow the physician to move arbitrarily around the patient while looking into the patient. A physician might be able, for example, to see the exact location of a lesion on a patient's liver, in three-dimensions and within the patient, without making a single incision. A laparoscopic surgeon may be able to view the pneumoperitoneum from any angle merely by turning his head in that direction, without needing to physically adjust the endoscopic camera. Augmented reality may be able to free the surgeon from the technical limitations of his imaging and visualization equipment, thus recapturing the physical simplicity of open surgery.”
(http://www.cs.unc.edu/Research/us/laparo.html last checked 2005-06-03)

Of relevance to “Mixed Reality for Technical Support”: medical application, surgery planning, similar hardware setup.
Main difference from “MR for technical Support”: Does not use gaze control

8. References

BATES, R. & ISTANCE, H. *Zooming interfaces! Enhancing the performance of eye controlled pointing devices.* ACM (Association for Computing Machinery, Inc) 2002.

BROOKER, J. P., SHARKEY, P.M., WANN, J.P., PLOOY, A.M. *A helmet mounted display system with active gaze control for visual telepresence.* Mechatronics no 9 p. 703-716, 1999.

FRIEDRICH, W. *ARVIKA, Augmented Reality für Entwicklung, Produktion und Service*. Pubilcis Corporate Publishing, Erlangen, Germany 2004, in German.

GUSTAFSSON, T. AND CARLEBERG, P., *The Gaze-Direction Sensor - a Key Component in a Smart Display System*, FOA-R--0-01486-706--SE, 2000, in Swedish.

GUSTAFSSON, T. AND CARLEBERG, P., *Mixed Reality for Technical Support*, FOI--0857--SE, in Swedish 2003.

GUSTAFSSON, T. AND CARLEBERG, P., *The Gaze-Tracker. Development and Applications*, FOI-R--0826--SE, in Swedish 2003.

GUSTAFSSON, T., CARLEBERG, P., NILSSON, S., SVENSSON, P., SIVERTUN, Å., LeDUC, M., *Mixed Reality för tekniskt stöd*, FOI-R--1198--SE, in Swedish, 2004.

GUSTAFSSON, T., CARLEBERG, P., SOHLBERG, P., PERSSON, B., ALDRIN, C., *Interaktivt förfarande för att presentera information i en bild,* Patent Application No 0401582-2, Filing Date 04-06-18, in Swedish.

HANSEN, J.P., TÖRNING, K., SEVERIN JOHANSEN, A., ITOH, K., & AOKI, H. *Gaze typing Compared with input by Head and Hand.* ACM , 2004.

HYRSKYKARI, A. *Gaze Control as an Input device.* In RÄIHÄ, K-J. (ed), Proceedings of ACHI'97 (Advanced Course on Human Computer Interaction), University of Tampere, Report B-1997-4, p 22-27. 1997.

JAKOB, R.K. *The use of Eye Movements in Human-Computer Interaction Techniques: What You Look At is What You Get.* ACM Transactions on Informations Systems, vol 9. No 3, April p 152-169. 1991.

KARN, K.S, ELLIS, S. & JULIANO, C. *The Hunt for Usability: Tracking Eye Movements.* Proceedings of CHI'99 Workshops. ACM, 1999.

LAW, B., ATKINS, M.S., KIRKPATRICK, A.E., LOMAX, A.J. & MACKENZIE, C.L. *Eye Gaze Patterns Differentiate Novice and Experts in a Virtual Laparoscopic Surgery Training Environment.* ACM, 2004.

MAJARANTA, P., AULA, A. & RÄIHÄ, K-J. *Effects of Feedback on Eye Typing with a Short Dwell Time.* ACM, 2004.

MANN S, *Wearable computing as means for personal empowerment,* International Conference on Wearable Computing ICWC-98, Fairfax VA, May 1998

NIELSEN, J. *Usability Engineering*, Academic press, 1993.

SIBERT, L.E. & JACOB, R.J. *Evaluation of Eye Gaze Interaction.* CHI Letters vol 2 issue 1 2000.

SIGGRAPH, *Proceedings of ACM*, Siggraph 2004, August 2004

TAN, Y.K., SHERKAT, N. & ALLEN, T. *Error Recovery in a blended Style Eye Gaze and Speech interface.* Proceedings of the ICMI'03. ACM 2003.

OHNO, T. *Features of Eye Gaze Interface for Selection Tasks.* Proceedings of the Third Asian Pacific Computer and Human Interaction . IEEE Computer Society 1998.

OHNO, T. & MUKAWA, N. *A Free-head, Simple Calibration, Gaze Tracking System That Enables Gaze-Based Interaction.*, ACM, 2004.

OVIATT, S. Lecture at IDA, LIU nov. 18, 2004.

OVIATT, S. & COHEN, P. *Multimodal Interfaces that Process What Comes Naturally.* Communications of the ACM vol 43, no 3, March 2000.

QVARFORDT, P. *Eyes on Multimodal Interaction.* Linköping Studies in Science and Technology Dissertation no 893, Unitryck, Linköping 2004.

VERTEGAAL, R. (moderator) *What do the eyes behold for human-computer interaction?* Panel discussion, ETRA 2002.

VIBERG, L., *Mixed Reality for Technical Support and Education*, Master Thesis, Royal Institute of Technology, Stockholm, TRITA-NA-E05061, 2005, in Swedish.

ZHAI, S. *What's in the eyes for attentive input?* Communications of the ACM, Vol. 46, No.3 March 2003.

ZHAI, S., MORIMOTO, C. & IHDE, S. *Manual And Gaze Input Cascaded (MAGIC) Pointing.* Proceedings of CHI'99. ACM, 1999.

Resources:
http://www.erbe-med.de/englisch/index.html
http://www.public.kitware.com/VTK/
http://www.cal3d.sourceforge.net/
http://www.swissquake.ch/chumbalum-soft/index.html
www.discreet.com/
www.ac3d.org/

Other research groups:
http://www.arvika.de/www/index.htm
http://www.artesas.de/
http://www.amire.net/index.html
http://www.ims.tuwien.ac.at/research/spatial_abilities/
http://www.studierstube.org/projects/mobile/
http://www.ims.tuwien.ac.at/~thomas/siemensforum.html
http://www.emt.tugraz.at/~tracking/
http://www.ims.tuwien.ac.at/research/mobile/ocar/
http://www.cg.tuwien.ac.at/research/vr/studierstube/multidim/
http://www.igd.fhg.de/igd-a2/projects/MixedRealityAP/index.html
http://www.vrvis.at/br1/aras/
http://liverplanner.icg.tu-graz.ac.at/
http://www.cs.unc.edu/Research/us/laparo.html