The Optical Interface

for Augmented Reality in Computer-Assisted Navigation and 3D Visualization in Surgery

M. Eckholt, H. Bergmann

General Hospital Vienna
Department of Biomedical Engineering and Physics

January 1997

DESCRIPTION OF THE PROJECT

The Optical Interface

for Augmented Reality in Computer-Assisted Navigation and 3D Visualization in Surgery

Applicant:	*Prof. Dr. H. Bergmann*[1]
Co-applicants:	*Prof. Dr. K. Ehrenberger*[3], *Prof. DDr. R. Ewers, Dr. M. Truppe*[4]
Co-workers:	*Dipl.-Phys. M. Eckholt*[1], *Dr. F. Watzinger*[2], *Dr. M. Cartellieri*[3]

This project is a co-operation between the following departments of the General Hospital Vienna (AKH):

(1) Department of Biomedical Engineering and Physics
(2) Clinic for Cranio-Maxillofacial Surgery
(3) Department of Oto-Rhino-Laryngology

and the Austrian company

(4) ARTMA Biomedical Inc., Vienna [ARTMA].

Contents

1 DESCRIPTION OF THE PROBLEM

1.1 INTRODUCTION

The basic goal of Augmented Reality (AR) is to enhance the user's perception of the real world by providing additional information which the user's senses cannot perceive under normal conditions. When applying AR to surgery planning and intraoperative 3D navigation, the aim is to provide the surgeon with 'x-ray vision' by merging virtual 3D objects generated from CT images (or other sources, like MR, nuclear medicine or ultrasound images) with his or her view of the patient.

Rather than replacing reality completely, as Virtual Reality (VR) technologies do, AR supplements it. In VR environments, the user is cut off from any view of the real world that surrounds him or her. In contrast, AR allows the user to see the real world with superimposed virtual computer graphics or images.

This project focuses on the interface which combines views of the real and the virtual world and provides the surgeon with the superimposed augmented scene.

1.2 CURRENT STATE OF RESEARCH

1.2.1 Presentation of the augmented scene

AR is a relatively young field of research, most results having been published in the past four years [Azuma 96a]. Research into the use of AR in medicine focuses on surgery, including visualization in the operating room, pre-operative planning and also on training of novice surgeons. In the context of this project, one needs to consider which methods of presenting the augmented scene to the surgeon during an operation have been investigated so far.

A rather simple way of displaying AR information is to use a monitor [Drascic 93, Kikinis 94]. It shows the operating environment as seen by a fixed or mobile video camera overlaid (fused) with virtual objects (e.g. a tumor rendered from 3D CT data). Position and size of the latter must be calculated from the actual position of the camera and the patient, both tracked by attached sensors, and from landmark data within the image. Stereoscopic vision can be provided by using two cameras and a stereo monitor which interlaces the images from the 'right' and the 'left' camera. The surgeon then has to wear special glasses equipped with liquid crystal shutters which are electronically synchronized with the interlacing frequency. The obvious and most important disadvantage of this approach is that not the clinician's viewpoint but that of the camera is being displayed. The surgeon must direct his view to the monitor and is thus forced to look away from the patient.

A method that eliminates this problem in a simple fashion is the use of a large stand-mounted semi-permeable mirror fixed above the patient. This mirror serves as an optical merger of the view of the real world and virtual objects. The surgeon sees the patient through the partially transmissive mirror, but he or she also sees the reflection of a (stereo) monitor on which the virtual objects are displayed. Since sensors track the position of the clinician's head, the virtual objects can be made to follow his actual point of view and thus appear to remain in their correct position relative to the patient [MRCAS]. Of course, the surgeon cannot change his position as flexibly as usual, because he or she has to look through the mirror. Also, the mirror might restrict the surgeon's free access to the patient.

1.2.2 The head-mounted display (HMD)

The head-mounted display (HMD) is an interface that has received a lot of attention. It contains two miniature monitors close to the eyes of the users on which computer generated or video images can be displayed. As a head-worn device in the shape of a helmet or bulky 'sun-glasses' (figure 1) it offers a high degree of flexibility to the user. The first HMD was built in the mid-1960s [Sutherland 68]. Early devices were used in flight-simulators for combat or civil pilots, e.g. the VCASS system [Furness 86]. Some prototypes were much too heavy to be worn on the head, so they were installed externally above the user's head. These types are usually referred to as head-up displays (HUDs). An example is the stand-mounted BOOM display built by FakeSpace (Menlo Park, Calif.) based upon a NASA Ames (Mountain View, Calif.) prototype [Fisher 86, Fisher 89]. Military HMDs for flight simulation systems were made commercially available in 1986 by CAE Electronics [Barrette 92].

When used for VR purposes, HMDs are always of the closed-view type, because the philosophy of VR is to *replace* reality by a virtual world. AR, in contrast, leaves a connection between the user and his real environment, because it aims at *enhancing* reality. HMDs used in AR systems are of the *see-through* type: there is still visual contact between the user and his surroundings. Practically all see-through HMDs (STHMDs) are of one of the two following basic designs. Both have already been tested in clinical applications.

- The first type is a *video* STHMD. Although it is a closed-view HMD, it keeps the user in touch with his real environment by combining it with one or two miniature video cameras installed on top of the helmet. The image that is finally displayed to the user is created by blending the camera's view of the real world with virtual objects. Two cameras provide stereoscopic vision (depth perception) [Azuma 96a, Rolland 94].

- The second category forming the majority of STHMDs [Rolland 94] are *optical* STHMDs. They do not deprive the user of his natural view of the real world, but merge it with virtual objects by the aid of semi-permeable miniature mirrors ('sun-glasses') in front of the user's eye. The working principle of the semi-transparent mirror has been described above. Again, the virtual graphics are being displayed using miniature monitors integrated in the HMD. Stereoscopic vision is achieved by fitting the perspective of the virtual images to the viewpoint of the respective eye.

1.2.3 Research into video see-through HMDs

Significant research into both types of STHMDs has been performed at the Computer Science Department of the University of North Carolina (UNC CS) at Chapel Hill since the 1980s. A video STHMD, the VPL EyePhone (VPL Research, Mountain View, California), has been used to visualize a fetus inside the womb of a pregnant patient by using 3D ultrasound reconstructed images [Bajura 92, State 94, Azuma 96a]. The ability of video image overlay techniques to *occlude* real objects (the woman's belly) and replace them by virtual ones (the fetus) can enforce the illusion that the fetus is actually *inside*, and not only registered *upon* the womb. A detailed discussion of the VPL EyePhone can be found in [Robinett 91].

The most recent work at UNC CS that makes use of a video STHMD is to evaluate ultrasound-guided needle biopsies of breast tumors both on breast phantoms and on patients [Fuchs 96]. The ultrasound transducer is fixed to a mechanical arm for precise tracking. An off-the-shelf closed-view Virtual Research [VRS] VR4 HMD (intended for VR applications) equipped with two video cameras is used. When used intra-operatively, the performance suffers from a number of drawbacks [Fuchs 96], mostly due to the limitations of the tracking system. The video STHMD poses two more problems typical of this design: one is its weight (almost 3 kg), the other one is that the viewpoint of the video cameras installed on top of the HMD is about four inches higher than the surgeon's natural viewpoint. This is indeed a critical disadvantage of using video STHMDs in surgery and will be discussed later in more detail.

McDonell Douglas [Neumann 96] explore the usefulness of video STHMDs in aircraft manufacturing [Azuma 96a].

1.2.4 Research into optical see-through HMDs

Medical applications

At UNC CS Chapel Hill an optical STHMD system for craniofacial surgery planning has been developed. It was investigated by [Holloway 95] in order to identify and quantify the sources of registration errors which cause mismatches between real and virtual objects. Due to the high precision requirements of the intended application it has not been employed in operation planning yet [Rolland 94]. Another project aims at dynamically visualizing 3D anatomy of the arm registered upon the patient in real-time in order to teach radiology students bone dynamics during radiographic positioning [Kancherla 95, Rolland 94]. [Holmgren 92] built an optical STHMD employing off-the-shelf optical and mechanical components, including rather large LCD colour monitors (67 mm diagonal). Their size enforced a more complex design than small monitors would have (an extra fold mirror and additional optics). Furthermore, he included the option to set the virtual image focus anywhere between 35 cm and infinity. Therefore the monitors were mounted on a positioning stage. This design issue will be discussed in connection with the implementation intended for this project. For the above reasons the final weight amounted to almost 3 kg (47 ounces plus a three pound heavy counterweight). Modern designs of optical STHMDs are much lighter due to tiny flat panel monitors that are much easier to integrate into the HMD (e.g., the Virtual i-O [VIO] 'i-glasses!' model in figure 1 has two 0.7" full colour LCD monitors with a 30 degrees FOV). Ronald T. Azuma's work (now at Hughes Research Laboratories, California) focused on the investigation and reduction of registration errors [Azuma 93, Azuma 95, Azuma 96a] and the development of a long-range tracking system [Azuma 96b]. At UNC CS he combined an opto-electronic tracking system based on an optical STHMD with inertia sensors to improve registration by predicting future positions of the surgeon's head [Azuma 94].

In Austria, ARTMA Medizintechnik GmbH (ARTMA Biomedical Inc.) [ARTMA] uses a lightweight optical STHMD to intraoperatively register 3D CT, MRT or radiographic data upon a patient in real time [Truppe 95]. The surgeon is being guided towards a target (e.g. a tumor) that has been identified preoperatively. In co-operation with the Clinic for Cranio-Maxillofacial Surgery of the General Hospital (AKH) in Vienna the system has already been used for tumor surgery and analysis of the temporomandibular joint motion [Wagner 95, Wagner 96a, Wagner 96b]. Registration accuracy and resolution of the virtual images still need to be improved considerably. The same system has been employed by the ENT Department of the University of Innsbruck (Austria) for endoscopic surgery [Truppe 94, Gunkel 95, Truppe 96].

Technical applications

A project called "Studierstube" aimed at the development of a simple multi-user AR environment is being carried out by the Visualization and Animation Group of the Institute of Computer Graphics at the Vienna University of Technology [ICGV]. They use also the the Virtual i-O [VIO] 'i-glasses!' STHMD shown in figure 1.

An optical STHMD is used at the Columbia University Computer Science Department in New York [CUCS] for the maintenance of laser printers [Feiner 93] using tracker sensors attached to the printers' moving parts. Virtual inner components of the printer and related instructions (knowledge-based graphics) are displayed to the user (e.g. a mechanic).

Boeing (Seattle, Wash.) intends to use optical STHMDs in the assembly of airplanes. [Azuma 96a] believes that one of the first professional applications of HMD-based AR systems will be in this field. Computer-generated diagrams guide technicians in the construction of the airplane's electrical system, making the currently used costly layout boards obsolete [Caudell 92, Janin 93, Boeing 94, Sims 94]. Like with all current AR systems, sufficiently precise registration of real and virtual objects (especially over the wide range of distances required in airplane construction) is the biggest problem yet to be solved [Caudell 92, Rolland 94, Azuma 96a]. Also, virtual objects seem to lag behind their real counterparts during head movements, because tracking and image rendering cause a time delay. [Holloway 95] has found that the registration error caused by system delay is bigger than all others sources of error *combined.* Even for only moderately fast head movements (50 degrees/sec), the typical time delay of 100 ms encountered in most HMD-based systems [Azuma 96a] causes a registration error of 5 degrees (or 8.7 cm at a distance of 1 m). For a fixed head position, registration errors cause virtual objects to be misplaced or to 'swim' in their place [Azuma 96a, Holloway 95, Bajura 92].

The Computer Vision group of the Department of Computer-Assisted Geometry and Graphics (ICGG) at the University of Technology in Graz (Austria) is investigating the integration of photogrammetric 3D object reconstruction and 2D object interpretation and recognition. Together with the ICGG Object Reconstruction group, they work on the 'CyberCity' project which aims at establishing a photo-textured digital model of an entire city [ICGG]. Researchers of the Visualization and Animation Group of the Department of Computer Graphics [ICGV] at the Vienna University of Technology are applying their know-how of ray-tracing techniques and rendering of so-called directed cyclic graphs to VR applications to model natural objects such as plants and mountains. Several other AR and VR applications in various fields are listed in [Azuma 96a], [Rosen 96] and [Holloway 95].

1.2.5 Optical vs. video see-through HMDs

It is very instructive to take a look at the results which research into the two basic STHMD designs has brought forth in the past in order to decide which type will be more suitable for this project. Both types have been applied to surgery and each shows a number of advantages and drawbacks.

Quality of the real world view

The most important advantage of optical STHMDs is that they preserve the clinician's natural real world view in real time, essentially without any *distortion*, loss of *resolution* or restriction of the *field of view* (FOV) [Rolland 94]. This is due to the fact that there is only a thin semi-transparent parallel plate in front of the user's eyes. High spatial resolution is a safety issue in surgery, because the surgeon needs to perform exact instrument positioning and precise cuts. No current video camera comes close to the resolving power of the eye (about one minute of arc at the fovea [Azuma 96a]), and several authors state that most video HMDs suffer from too low

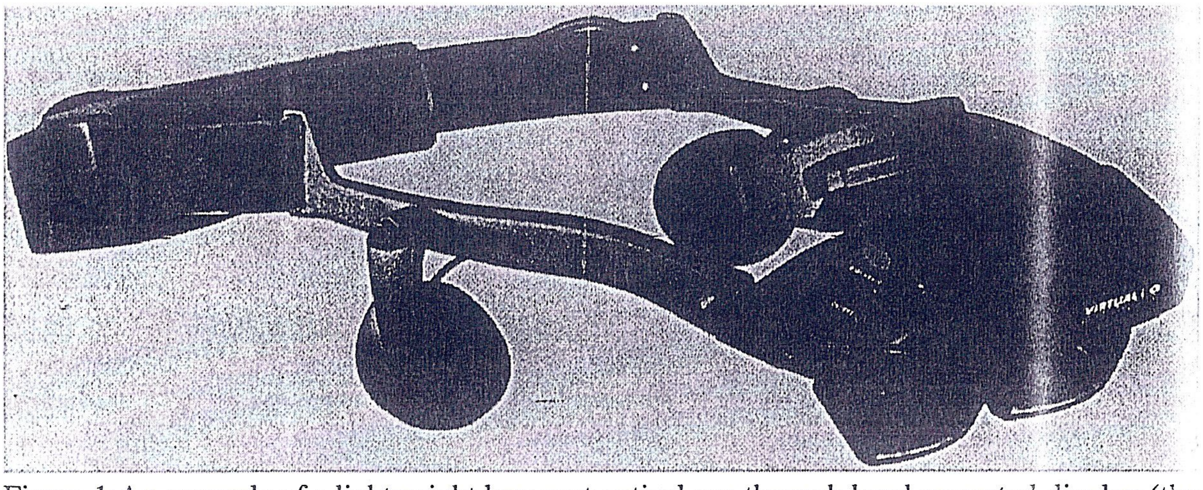

Figure 1:An example of a lightweight low-cost optical see-through head-mounted display (the 'i-glasses!' model manufactured by Virtual i-O [VIO] for video games)

resolution [Rosen 96, Rolland 94, Bajura 92]. Closed-view designs like video STHMDs need a rather large camera FOV ([Howlett 92, Yamaguchi 89] suggest at least 80 degrees), otherwise the user has the impression of looking through a window [Yoshida 95]. Therefore a small camera FOV is likely to reduce the surgeon's acceptance of the system as a replacement of his view of the real world. Also, a user confronted with a restricted FOV tends to make more head movements to get an overview [Rolland 94]. This contributes to larger tracking errors. Choosing a large camera FOV, on the other hand, enforces the need to compensate for image distortion. This can be done by image processing or by additional optics. The first method introduces a time delay that impedes real time image display and is likely to create a larger registration error than the distortion compensation removes [Holloway 95]. It also makes generation of the synthetic scene much more difficult [State 94]. The second method increases cost and weight of the HMD. The fact that optical STHMDs preserve the eyes' high resolution and wide FOV entails the advantage that the user is more likely to accept medium quality virtual images [Azuma 96a]. One could thus use tiny monitors of moderate resolution (especially for wire frame images) that cover a rather small FOV.

Safety

As far as *safety* is concerned, there are two more aspects to be considered. One is the closed view of the video STHMD. If the power supply fails, the surgeon is 'blind' - and safety considerations must assume that this could happen during a critical stage of the operation. This is not true for optical STHMDs which always provide full eye contact to the real world. The second safety issue is that the video cameras have a viewpoint that is several inches higher (when installed on top of the HMD) and/or farther to the front than the surgeon's natural viewpoint. Investigations have shown that human beings have a considerable ability to adapt to the new situation, but this needs some training [Rolland 94]. The performance quality of coordination tasks improves with increasing length of the training period. Negative aftereffects like overshoot in depth pointing tasks occur when the user takes the HMD off [Rolland 94], because he or she needs some time then to get used to the normal situation again. These facts represent a risk, since successful and safe surgery vitally depends on precise hand-eye-coordination at *any* time during an operation. Matching the viewpoints of the eyes and the cameras, e.g. by the periscope principle, requires additional optics and reduces the effective FOV [Edwards 93, Rolland 94].

Technical design

Optical STHMDs are simpler in *technical design* [Azuma 96a], because no video cameras have to be implemented, and there is only one video stream (providing virtual objects) to take care of. Image matching with higher precision can be achieved, because in contrast to video STHMDs the real world view remains unchanged and is therefore inherently precise. Furthermore, registration accuracy can be measured more precisely with optical STHMDs, because the digitized and processed real world video images cause an additional measurement error.

Real time display

As pointed out above, all AR applications suffer from the fact that, to some extent, the virtual objects lag behind their real counterparts during head movements. Optical STHMDs naturally provide *real time* information about the user's real environment. As a consequence, the time that elapses until the corresponding virtual graphics are presented to the user (typically several tens of milliseconds [Azuma 96a, Holloway 95]), must be minimized. According to [Azuma 94] this delay should be less than 60 ms. Video STHMDs, in contrast, offer only *near real time* display, because digitization (and perhaps processing) of the real world images takes some time. But on the other hand, they offer the possibility to deliberately delay the 'real' video stream against the 'virtual' one in order to enforce congruence. Still, this time delay must not be too large, since it causes a discrepancy between visual, vestibular and proprioceptive signals. This can result in 'motion sickness' [Held 87] which has often been encountered in flight simulations.

Advantages of the video see-through design

The fact that video STHMDs deliver the real world view in the form of digitized video images entails a number of advantages over the optical design. The most important one is that by using video mixing techniques it is possible to *occlude* real objects by virtual ones in a much more convincing way than can optical superposition [Azuma 96a], provided that a precise depth map of the real scene can be obtained [Rolland 94]. Virtual objects may look ghost-like and semi-transparent when superimposed optically onto real environments. This is important for depth perception, because occlusion is a very strong monocular depth cue. The fetus in a pregnant woman [State 94, Bajura 92] (see above) is a good example. Yet for *surgical* rather than diagnostic applications, it might be helpful or necessary that real objects be not completely obscured by virtual ones. Take the example of a tumor inside a patient. If the surgeon needs to cut the skin and navigate towards the tumor it is advantageous to see the skin and the surroundings of the tumor as well. Video-based systems also offer additional registration strategies that optical ones are not capable of, because the digitized real world video image provides a feedback on how closely the real and the virtual scene match. [Mellor 95a, Mellor 95b] used fiducial markers and demonstrated that structure recognition in real world video images can be used to fuse them automatically with virtual images after initial registration using a laser scanner. Finally, the large dynamic range of the eye (about six orders of magnitude due to its logarithmic sensitivity [Azuma 96a]) poses a brightness matching problem between the real and virtual images for optical STHMDs since video monitors have a limited dynamic range for brightness. So do video cameras, thus adjusting the brightness of the two video streams of a video STHMD is easier.

Conclusion

Taking into consideration the above advantages and disadvantages of the two alternative designs we have decided to concentrate on optical STHMDs in this project. The most important arguments for us to choose optical STHMDs are the advantages that come with the undegraded real world view and the safety considerations. From a human factors point of view we believe that an optical STHMD will be accepted more easily by surgeons, because they retain the natural high resolution of the real world view.

1.3 THE SCIENTIFIC PROBLEM ("THE STEP TOWARDS NEW SCIENTIFIC FRONTIERS")

1.3.1 Design and construction of the interface

The objective of this project is to concentrate on the essential optical components of the STHMD that provide the superposition of the virtual objects and the real world view. We will henceforth refer to this part as "the optical interface". Figure 2 illustrates its basic design. It comprises a miniature LCD or CRT video monitor on which the virtual objects are displayed, a background illumination for the monitor, and the optical system which merges the surgeon's real world view and the virtual scene. The optical parts consists of a lens system which produces a magnified image of the monitor. The semi-transparent beamsplitter plate serves as an optical combiner of this image and the real world view. Its working principle has been described above. We want to include the option to flip the beamsplitter plate out of the way, so that the user can see the unobstructed real surroundings. Especially when wearing the HMD, the surgeon then would not have to take it off, if he or she chooses not to make use of the virtual images for some time during the operation.

The above survey has shown that in most applications the optical interface as such has not received much attention, but has been viewed as part of a more complex interface like the HMD. We want to focus on the optical interface itself and investigate its usefulness for several surgical applications. To do so, we will connect the interface to several devices, namely optical STHMDs, binocular magnifying glasses, microscopes and endoscopes (see below). The aim is to investigate the effect of optical parameters on the image quality of the augmented scene presented to the surgeon. This includes providing stereoscopic 3D virtual objects. To reach this goal, we have to find out optimal performance parameters of the interface in these applications. The 'step towards new scientific frontiers' thus is the development and construction of an optimized optical interface that comes in the form of a flange and can be flexibly attached to the devices and instruments listed above.

The experience of other researchers teaches that special attention has to be paid to the weight and size of the interface (especially when integrated into a HMD). The dimensions of the interface should be of the order of one or two inches, and the complete HMD should not weigh

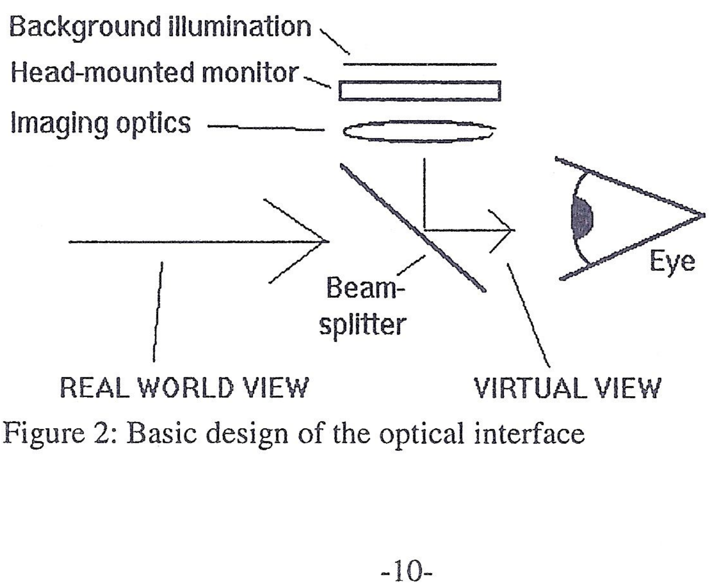

Figure 2: Basic design of the optical interface

much more than 500 g. One has to keep in mind that operations usually last several hours, and taking off the HMD makes a recalibration necessary. Lightweight optical STHMDs fulfilling these criteria are already commercially available and can serve as a 'proof of concept' (the Virtual i-O [VIO] 'i-glasses!' model in figure 1, for example, weighs 380 g).

From the point of view of physics, scientific research will focus on the transmission ratio and wavelength selectivity of the optical combiner and the adjustment of monitor brightness and of the projection distance of the virtual image. The role these parameters play in the visualization of the augmented scene will be discussed in more detail in connection with the implementation of the project.

1.3.2 Clinical applications

As a first step towards introducing AR by the STHMD to the operating room, this project will confine itself to craniofacial and ENT surgery applications, because the essentially rigid and partly even immobile segments of the skull are easier to track than soft tissue. They do not need to be scanned continuously (unlike a moving fetus, for example), but can be superimposed using preoperatively acquired static 3D computed tomography (CT) , magnetic resonance (MR) or nuclear medicine (NM) images.

An example of how craniofacial surgery can profit by AR is the intraoperative positioning of osteotomic skull and maxillary bone fragments. Virtual wire frame images that show the fragment's target position can be produced preoperatively from CT scans. During the operation, they will guide the clinician in placing the real bone fragment correctly.

An ENT surgery paradigm is navigation in trauma-sensitive regions of the head. The optimal minimally invasive path that should be taken, e.g. by the endoscope, can be found in preoperative planning. Virtual images that display this path superimposed on the clinician's view of the real scene will then guide him or her intraoperatively. This principle has already been demonstrated successfully by [Truppe 94, Gunkel 95, Truppe 96]. Also, the overlaid display of vitally important anatomical structures like nerves and blood vessels would be of great help [Ehrenberger 96].

Now we will consider important aspects of the various application options of the interface.

Binocular magnifying glasses

Binocular magnifying glasses are miniature telescopes attached to normal glasses (figure 3) or to head straps. They provide a moderately magnified view of the operation scene and are frequently used by surgeons, because they are easier to handle than a microscope. We believe that introducing AR to the clinician by simply attaching a flange to this see-through device he or she regularly uses can break grounds for more sophisticated applications that require accustoming, like the HMD. To our knowledge, no one has explored this application yet. As far as sensor calibration is concerned, binocular magnifying glasses have the advantage of being self-recalibrating, because unless they are positioned exactly in front of the user's eye pupil, the

FOV will be very restricted. In other words, their design automatically forces the user to align the optical axis of the magnifying lens system with that of his or her eye. Therefore the surgeon can put the glasses on and off again, without any recalibration being necessary. The simple set-up of this tool makes it a suitable testbed for the more complicated task of combining it with the HMD.

A second principal objective of the project is to make stereoscopic vision possible in order to enable the surgeon to *visually* perceive depth information from the virtual 3D images. This can be done by presenting slightly different views to each eye. We want to provide 3D vision for optical STHMDs. As outlined above, frequent head motions pose problems to sensoring and registration. Therefore, this might turn out to be quite a challenging task. Once accomplished, the HMD can be connected to a stereo microscope or an endoscope as an alternative display to the (stereo) monitor. This point will be discussed further in the following sections.

Endoscopy

A third interesting application of the interface is endoscopy. Most endoscopes are equipped with standard mounting rings or flanges, e.g. to connect a photographic camera to the ocular. By choosing standard mounting adapters, the flange containing the interface can be attached to many endoscopes already in use today. This would result in a cost-effective introduction of AR to endoscopy. As far as we know, no AR endoscopes are commercially available yet. The strongly distorted endoscopic view makes relatively great demands on the appropriate adaption of the virtual images. According to major endoscope vendors [Wolf 96, Olympus 96], the vast majority of surgeons prefers monitor display of endoscopic images to looking directly through the endoscope when performing routine tasks. But when in more complicated situations the danger of lesions increases (especially in endonasal surgery), most of them return to the see-

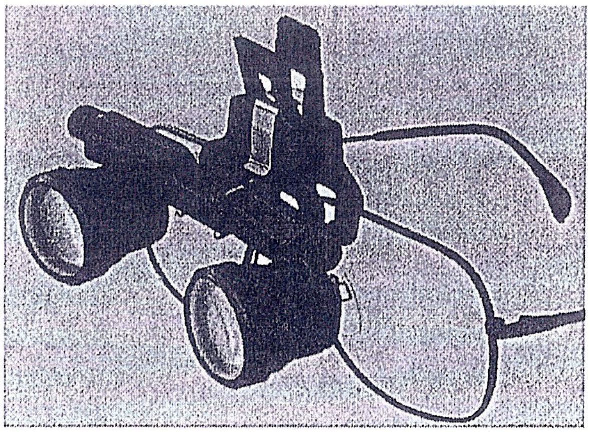

Figure 3: Binocular magnifying glasses

through option [Wolf 96, Olympus 96]. Therefore, we plan to apply AR to see-through mono-endoscopes. AR also offers the interesting opportunity to merge 3D virtual images with the surgeon's mono-endoscopic 2D video view. This becomes possible by replacing the video monitor by the HMD which can provide stereoscopic virtual images. We will confine ourselves for the time being to rigid endoscopes, because they are much easier to track than flexible ones. One cannot foresee the future development of stereo-endoscopy in craniofacial and ENT surgery at the moment. Currently they are hardly beyond the stage of being tested [Berger 96], but if they should prove superior to mono-endoscopes, they might gain wide acceptance. In this case they would be ideally suited to be combined with AR as well (but only the HMD option [Ehrenberger 96]).

Microsurgery

One final field of interest is microsurgery (of the temporomandibular joint, for example [Ewers 96]). Once a flexibly attachable flange has been constructed and the principles of stereoscopic virtual image creation have been investigated, it will be only a small step to progress to see-through stereo microscopy as far as optics are concerned. Zeiss already offer stereo microscope systems that are capable of overlaying virtual 2D objects with the magnified real world view. Since only one tube is used for the superposition, they do not provide a stereoscopic virtual scene. Yet they prove that the concept of using AR for microsurgery is realistic and that there is a demand for this application (about 25 copies of the MKM have been sold world-wide in the past two years [Zeiss 96]). Leica have developed a similar system (the 'ViewScope' model) that is capable of presenting stereoscopic images created pre-operatively from CT and MRT scans, but these are only of the wire frame type [Leica 96]. The manual and the computer-controlled Zeiss MKM version costs about three and seven million ATS respectively (280,000 and 650,000 US$). The price of the Leica ViewScope system is about four million ATS (375,000 US$) [Leica 96]. Upgrading microscopes to the MKM or SMN level will mostly be more expensive than buying a new system [Zeiss 96]. Only Leica microscopes can be upgraded to the ViewScope level, for about 2,600,000 ATS (240,000 US$) [Leica 96]. As in the case of endoscopy, the interface we plan to produce would be a low cost alternative, because it can be attached to practically every microscope already in use.

By fixing one video camera to each of the stereo microscope's tubes, the magnified 3D view can currently be displayed on a stereo monitor. This technique has already been described above. Unless the eyes' high resolution is required, many surgeons prefer this display option, because looking directly through the microscope forces them to bend over the patient, and sometimes remain in this positon for hours [Ewers 96]. But the need to turn the head away from the field of operation to look at the monitor can distract the surgeon's attention or cause unwanted counter movements of the hand. Therefore, the second option we plan to realize is to connect the microscope image to the HMD instead of the stereo monitor. A carefully designed HMD can provide both sources of information at a time, the augmented magnified scene as well as the field of operation, without making head movements necessary (see below: 'step 5').

1.4 OUR RELATED WORK

H. Bergmann has been working in the field of quantitative digital imaging for many years. More recently, he used 3-D visualization techniques, image fusion methods and statistical methods such as factor analysis for feature extraction. He is responsible for the image processing laboratory of the IBMTP. Since 1989 the group has a collaboration with the Biomedical Imaging Resource Unit (Director: Professor R.A. Robb) of the Mayo Clinic, Rochester, U.S.A., which deals with the application of sophisticated image processing software in the clinical environment. H. Bergmann has not worked on the use of head-mounted displays in AR yet.

DPhys. M. Eckholt worked at the optical laboratory of the Department of Medical Physics of the Medical Faculty of the University of Vienna during his graduate work on physics, performing optical experiments with low intensity infrared lasers. He is familiar with the optical components like beamsplitters and lenses. He has not worked in the field of AR yet.

Prof. DDr. R. Ewers, Director of the Clinic for Maxillo-Facial Surgery, and Dr. F. Watzinger, of the same department, have been using Computer Assisted Surgery (CAS) since 1994 for the treatment of selected patients, using the ARTMA system by Dr. Truppe. They will act as consultants during the design and construction phase, participate in the various in vitro tests of the optical interface and evaluate the new display during operations.

Prof. Dr. Ehrenberger, Director of the Department of Oto-Rhino-Laryngology, and Dr. M. Cartellieri of the same department, are at present evaluating commercial systems for CAS in oto-rhino-laryngology.They will act as consultants during the design and construction phase, participate in the various in vitro tests of the optical interface and evaluate the new display during operations.

Dr. M. Truppe, after having worked for several years as a surgeon at the Clinic of Maxillo-Facial Surgery in Vienna and in Freiburg, has developed a new system for visualizing and merging medical images with real-time video images. He demonstrated the usefulness of the system in co-operation with several clinical departments in Austria and other European countries. He will be providing his expertise and his software system for the project during the stage of preliminary clinical testing.

1.5 ANNOTATED BIBLIOGRAPHY

Azuma 93 — R. Azuma : *Tracking Requirements for Augmented Reality*, Communications of the ACM, 36, 7 (July 1993), 50-51.

Azuma 94 — R. Azuma, G. Bishop. *Improving Static and Dynamic Registration in an Optical See-Through HMD*, Proceedings of SIGGRAPH '94 (Orlando, FL, 24-29 July 1994), Computer Graphics, Annual Conference Series, 1994, 197-204 + CD-ROM appendix

Azuma 95 — R. Azuma: *Predictive Tracking for Augmented Reality*, Ph.D. Dissertation, (Technical Report TR95-007, Univ. of North Carolina at Chapel Hill, Dept. of Computer Science, Feb. 1992), on the WWW: http://www.cs.unc.edu/Research/tech-reports.html

Azuma 96a — R. Azuma : *A Survey of Augmented Reality* (updated version August 1996), to be published in 1997 in *Presence: Teleoperators and Virtual Environments*
On the WWW: http://www.cs.unc.edu/~azuma/azuma_publications.html

Azuma 96b — R. Azuma: *Wide Area Tracking, Navigation Technology of Head-Mounted Displays*, on the WWW: http://www.cs.unc.edu/~tracker

Bajura 92 M. Bajura, H. Fuchs, R. Ohbuchi: *Merging Virtual Objects with the Real World: Seeing Ultrasound Imagery Within the Patient*, in: [Taylor 96], 245ff. (previously publ. in Computer Graphics 26, 2 (July 1992): 203 - 210), also: Technical Report TR92-005, Univ. of North Carolina at Chapel Hill, Dept. of Computer Science, 1992, on the WWW : http://www.cs.unc.edu/Research/tech-reports.html

Berger 96 U. Berger, Department of Oto-Rhino-Laryngology of the General Hospital Vienna, personal communication (1996)

Campbell 57 F.W. Campbell: *The Depth of Field of the Human Eye*, Optica Acta 4 (4), Dec. 1957, 157-164

Caudell 92 T.P. Caudell, D.W. Mizell: *Augmented Reality: An Application of Heads-Up Technology to Manual Manufacturing Processes*, Proceedings of Hawaii International Conference on System Sciences, January 1992, 659-669

Drascic 93 D. Drascic, J.J. Grodski, P. Milgram, K. Ruffo, P. Wong, S. Zhai: *ARGOS: A Display System for Augmented Reality, Video Proceedings of INTERCHI '93: Human Factors in Computing Systems* (Amsterdam, 24-29 April 1993), also on the WWW (ETC-Lab, Univ. of Toronto): http://vered.rose.utoronto.ca/people/david_dir/CHI93/CHI93.html

Edwards 93 E.K. Edwards, J.P. Rolland, K.P. Keller: *Video See-Through Design for Merging of Real and Virtual Environments*, Proceedings of IEEE VRAIS '93 (Seattle, WA, 18-22 September 1993), 223-229

Ehrenberger 96 K. Ehrenberger, Prof. Dr., Head of the Department of Oto-Rhino-Laryngology of the General Hospital Vienna, personal communication (1996)

Ewers 96 R. Ewers, Prof. Dr., head of the Clinic for Cranio-Maxillofacial Surgery of the General Hospital Vienna, personal communication (1996)

Feiner 93 S. Feiner, B. MacIntyre, D. Seligmann: *Knowledge-Based Augmented Reality*, Communications of the ACM 36, 7 (July 1993), 52-62

Fisher 86 S. Fisher, M. McGreevy, J. Humphries, W. Robinett: *Virtual Environment Display System,* Proceedings of the 1986 ACM Workshop on Interactive Graphics (Chapel Hill, NC, 23-24 October 1986), 77-87

Fisher 89 S. Fisher: *Virtual Environments, Personal Simulation and Telepresence*, Implementing and Interacting with Realtime Microworlds (ACM SIGGRAPH '89 Course Notes no. 29)

Fuchs 96 H. Fuchs, A. State, E.T. Pisano, W.F. Garrett, G. Hirota, M. Livingston, M.C. Whitton, S.M. Pizer: *Towards Performing Ultrasound-Guided*

Needle Biopsies from within a Head-Mounted Display, Lecture Notes in Computer Science Vol. 1131, Visualization in Biomedical Computing (Hamburg, Germany, September 1996)

Gunkel 95 A.R. Gunkel, W. Freysinger, W. F. Thumfart, M. J. Truppe: *Application of the ARTMA Image-Guided Navigation System to Endonasal Sinus Surgery*, Proceedings Computer Assisted Radiology CAR '95 (Berlin, Germany, 21-24 June 1995), 1146-1151

Held 87 R. Held, N. Durlach: *Telepresence, Time Delay and Adaption*, NASA Conference Publication no. 10032, 1987

Holloway 95 R.L. Holloway : *Registration Errors in Augmented Reality Systems* - Ph.D. dissertation (Technical Report TR95-016,Univ. of North Carolina at Chapel Hill, Dept. of Computer Science, 1995), on the WWW : http://www.cs.unc.edu/Research/tech-reports.html

Holmgren 92 D.Holmgren : *Design and Construction of a 30-Degree See-Through Head-Mounted Display* (Technical Report TR92-030, Univ. of North Carolina at Chapel Hill, Dept. of Computer Science, 1992), on the WWW: http://www.cs.unc.edu/Research/tech-reports.html

Howlett 92 E.M. Howlett: *High-Resolution Inserts in Wide-Angle Head-Mounted Stereoscopic Displays*, SPIE Proceedings Vol. 1669, 193-203 (1992)

Janin 93 A.L. Janin, D.W. Mizell, T.P. Caudell: *Calibration of Head-Mounted Displays for Augmented Reality Applications*, Proceedings of IEEE VRAIS '93 (Seattle, WA, 18-22 September 1993), 246-255

Kancherla 95 A.R. Kancherla, J.P. Rolland, D.L. Wright, G. Burdea: *A Novel Virtual Reality Tool for Teaching Dynamic 3D Anatomy*, Proceedings of Computer Vision, Virtual Reality, and Robotics in Medicine CVRMed '95 (Nice, France, 3 - 6 April 1995), 163 - 169

Kikinis 94 R. Kikinis, P.L. Gleason, W. Lorenson, W. Wells, W.E.L. Grimson, T. Lozano-Perez, G. Ettinger, S. White, F. Jolesz: *Image Guiding Techniques for Neurosurgery*, SPIE Proceedings Vol. 2359: Visualization in Biomedical Computing (Rochester, Minnesota, 4-7 October 1994), 537-540

Leica 96 LEICA AG Heerbrugg, Switzerland, personal communication (1996)

Mellor 95a J.P. Mellor: *Enhanced Reality Visualization in a Surgical Environment*, MS Thesis, Dept. of Electrical Engineering and Computer Science, January 1995; Technical Report No. 1544 of the MIT Artificial Intelligence Laboratory, on the WWW: [Mellor WWW]

Mellor 95b — J.P. Mellor: *Realtime Camera Calibration for Enhanced Reality Visualization*, Proceedings of Computer Vision, Virtual Reality, and Robotics in Medicine CVRMed '95 (Nice, France, 3 - 6 April 1995), 471-475

Neumann 96 — U. Neumann, Y. Cho: *A Self-Tracking Augmented Reality System*, Proceedings of VRST '96 (Hong Kong, 1-4 July 1996), 109-115

Olympus 96 — OLYMPUS AUSTRIA GMBH Vienna, personal communication (1996)

Robinett 91 — W. Robinett, J.P. Rolland: *A Computational Model for the Stereoscopic Optics of a Head-Mounted Display* (Technical Report TR93-045, Univ. of North Carolina at Chapel Hill, Dept. of Computer Science, 1993), on the WWW : http://www.cs.unc.edu/Research/tech-reports.html

Rolland 93a — J.P. Rolland, T. Hopkins: *A Method of Computational Correction for Optical Distortion in Head-Mounted Displays* (Technical Report TR93-045,Univ. of North Carolina at Chapel Hill, Dept. of Computer Science, 1993), on the WWW: http://www.cs.unc.edu/Research/tech-reports.html

Rolland 93b — J.P. Rolland, C.A. Burbeck et al.: *Towards Quantifying Depth and Size Perception in 3D Virtual Environments* (Technical Report TR93-044, Univ. of North Carolina at Chapel Hill, Dept. of Computer Science, 1993), on the WWW: http://www.cs.unc.edu/Research/tech-reports.html

Rolland 94 — J.P. Rolland, R.L. Holloway, H. Fuchs: *A Comparison of Optical and Video See-Through Head-Mounted Displays,* SPIE Proceedings Vol. 2351: Telemanipulator and Telepresence Technologies (Boston, MA, 31 Oct - 4 Nov 1994), 293-307.
On the WWW : http://www.creol.ucf.edu/~rolland/hmd_optics.html

Rosen 96 — J. M. Rosen, A. Lasko-Harvill, R. Satava: *Virtual Reality and Surgery* In: [Taylor 96], 231ff.

Sims 94 — D. Sims: *New Realities in Aircraft Design and Manufacture*, IEEE Computer Graphics and Applications 14, 2 (March 1994), 91

State 94 — A. State, D.T. Chen, C. Tector, A. Brandt, H. Chen, R. Ohbuchi, M. Bajura, H. Fuchs : *Case Study: Observing a Volume Rendered Fetus Within a Pregnant Patient*, Proceedings of IEEE Visualization '94 (Wash. D.C., 17-21 October 1994)

Taylor 96 — R. H. Taylor, S. Lavallée, G. C. Burdea, R. Mösges (ed.): *COMPUTER-INTEGRATED SURGERY - Technology and Clinical Applications* (The MIT Press, Cambridge/Mass., 1996, ISBN 0-262-20097-X)

Truppe 94 M.J. Truppe: *The Artma Virtual Patient: 3D Monitoring of Endoscopic Surgery*, Medicine Meets Virtual Reality II (San Diego, CA, Jan 1994), 221

Truppe 95 M.J. Truppe, F. Pongracz, O. Ploder, A. Wagner, R. Ewers: *Interventional Video Tomography*, SPIE Proceedings of Lasers in Surgery (San Jose, CA, 4-6 February 1995), Paper #2395-34, 150-152
On the WWW: http://www.artma.com/html_files/IVT.html

Truppe 96 M. J. Truppe, W. Freysinger, A.R. Gunkel, W. F. Thumfart: *Teleassisted Stereotactic Endoscopic Surgery*, Proceedings Computer Assisted Radiology CAR '96 (Paris, France, 26 -29 June, 1996)
On the WWW: http://www.artma.com/html_files/CAR96.html

Wagner 95 A.Wagner, O. Ploder, G. Enislidis, M. Truppe, R. Ewers: *Virtual Image Guided Navigation in Tumor Surgery*, J. Craniomaxillofac. Surg. 1995; 23(5):271-273

Wagner 96a A.Wagner, O. Ploder, G. Enislidis, M. Truppe, R. Ewers: *Image Guided Surgery,* Int. J. Oral Maxillofac. Surg. 1996; 25:147-151

Wagner 96b A. Wagner, O. Ploder, J. Zuniga, G. Undt, R. Ewers: *Augmented Reality Environment for Temporomandibular Joint Motion Analysis*, Int. J. Adult Orthod. Orthognath. Surg. 1996; 11:127-136

Wolf 96 RICHARD WOLF AUSTRIA GMBH Vienna, personal communication (1996)

Yamaguchi89 H. Yamaguchi, A. Tomono, Y. Kobayashi: *Proposal for a Large Visual Field Display Employing Eye Movement Tracking*, SPIE Proceedings Vol. 1194, 13-20 (1989)

Yoshida 95 A. Yoshida, J.P. Rolland, J.H. Reif: *Optical Design and Analysis of a Head-Mounted Display with a High Resolution Insert,* SPIE Proceedings Vol. 2537 (1995)
On the WWW : http://www.creol.ucf.edu/~rolland/hmd_optics.html

Zeiss 96 CARL ZEISS AUSTRIA GMBH Vienna, personal communication (1996)

Postal and World-Wide-Web Addresses

ICGG Graz University of Technology, Institute of Computer-Assisted Geometry and Graphics, Münzgrabenstraße 11/II/IV/V, A-8010 Graz (Austria)
On the WWW: http://www.icg.tu-graz.ac.at

ICGV Vienna University of Technology, Institute of Computer Graphics,

Karlsplatz 13/186, A-1040 Vienna (Austria)
On the WWW: http://www.cg.tuwien.ac.at/research/

MRCAS — Center for Medical Robotics and Computer Assisted Surgery (MRCAS), Robotics Institute, Carnegie Mellon Univ., Pittsburgh (PA); on the WWW: http://www.cs.cmu.edu/afs/cs/project/mrcas/www/mrcas-home/overlay.html

CUCS — Columbia University Computer Science Department, 1214 Amsterdam Avenue, New York, NY 10027-7003, Phone (212) 939-7000, Fax (212) 666-0140, on the WWW: http://www.cs.columbia.edu

VIO — Virtual i-O, 1000 Lenora Street, Suite 600, Seattle, WA 98121, Phone: 206.382.7410, Fax: 206.382.8810, on the WWW: http://www.vio.com
European distribution: Virtual Products GmbH, Berliner Ring 89, D - 64625 Bensheim, Germany, phone +49-6251/80 22 00

VRS — Virtual Research Systems Inc., 2326 Walsh Ave., Santa Clara, CA 95051
Phone (408) 748-8712, Fax (408) 748-8714
On the WWW: http://www.virtualresearch.com//vr4.html

ARTMA — ARTMA Medizintechnik GmbH, Am Kanal 27, A-1110 Wien, Austria
Phone +43-1-7 40 40-300 or +43-664-205 89 30, Fax +43-1-7 40 40-740
On the WWW: http://www.artma.com

2 OBJECTIVES AND IMPLEMENTATION

2.1 OBJECTIVES

A) To *design and construct an optimized optical interface* that merges the surgeon's real view and virtual objects from miniaturized monitors and can be flexibly attached to optical STHMDs, binocular magnifying glasses, stereo microscopes and endoscopes. Figure 2 shows its basic mechanical design.

B) To provide *stereoscopic vision* of the virtual images for binocular magnifying glasses, optical STHMDs and stereo microscopes, because this will significantly facilitate surgical navigation.

C) As for microscopes and endoscopes, to adapt the HMD in such a manner that it can *replace the stereo monitor.*

D) To adapt *Software* to generate real time stereoscopic virtual images and 3D objects that correspond to the surgeon's actual viewing direction and his relative position to the (model) patient.

2.2 WORK PROGRAMME

We intend to realize the aforesaid aims by taking the following steps.

Step 1: Construction of a prototype interface

Following the basic design shown in figure 2 we will assemble a prototype interface that can be used to measure performance parameters (see step 2). In our opinion, the most cost-effective way to implement the prototype is to use a commercially available good quality lightweight optical STHMD and modify the optical interface in such a way that optical components can easily be exchanged and varied. We need e.g. to be able to exchange the beamsplitter plates in order to test several variants (see step 2). Also, the monitor brightness should be manually adjustable. The same is true for the distance at which the lens system forms the virtual images. For commercial HMDs this distance is fixed therefore we have modify the mechanical design (see step 2 c). This will be done by mounting the lens system (figure 2) into a screw thread. The HMD comes with a head strap or a light plastic helmet that allows for attaching tracking sensors and the interface in front of each eye (figure 1).

Step 2: Evaluating and improving the performance of the prototype

We want to examine the performance parameters listed below of the prototype interface. The objective here is to find out optimal values or value ranges. The clinical co-workers of the project will participate in the testing at this stage.

a) Transmission ratio of the beamsplitter plate

Beamsplitter plates consist of flat thin glass with dielectric reflection coating on one side which determines the transmission ratio of the plate. In our case, a high ratio means that the surgeon will see the real scene almost unattenuated with superimposed faint virtual images, very much like reflections in a normal window (transmission ratio of about 96 %). A very low ratio will give the impression of looking at the real world through dark sun-glasses and produce virtual objects of high intensity. Manufacturers of optical accessories offer standard beamsplitter plates with a wide range of transmission ratios. We want to test a range of them using suitable digital test images and anthropomorphic phantoms (e.g. the skull of a corpse) or subjects. Of course, the relative intensities of the real and virtual images the user subjectively perceives also depends on their relative brightness. In a well-lit room, for example, the virtual images will appear fainter than in a dark room and the user might then wish to increase the brightness of the monitor that displays the virtual images (see point c). It will also be necessary to test the effects of anti-reflection coatings on the backside of the beamsplitter plate.

b) Wavelength-selectivity of the beamsplitter

Another question we plan to investigate concerns the possibility to eliminate the disadvantageous 'sun-glasses effect' of normal beamsplitters. A transmission ratio of 70 percent means that the amount of light coming in from the real world will be reduced by 30 percent. Replacing the beamsplitter by a so-called notch filter as proposed by [Caudell 92, Azuma 96a] could be a remedy. In contrast to a usual beamsplitter, the dielectric reflection coating of a notch filter is wavelength-sensitive in the sense that only a narrow band around one specific wavelength is reflected. For this wavelength, a reflectivity of more than 90 % can quite easily be achieved. Virtually all the light from outside (except for the narrow wavelength band) will reach the user's eyes, too. A notch filter is like a glass plate that strongly reflects just one colour. By picking a colour for the virtual objects that is uncommon in the operation scene (e.g. blue) neither the virtual image nor the real world view would suffer significant intensity losses. Also, problems like chromatic aberration of the imaging optics do not occur with monochrome light [Caudell 92]. Drawbacks of this approach are the fact that the virtual images cannot be multi-coloured and the high cost of notch filters (they are usually custom made for a specific wavelength).

c) Monitor characteristics

We will investigate the monitor's brightness range, resolution and dynamic range that are necessary for our application. We have pointed out above that monitor brightness should be manually adjustable to adapt it to the brightness of the environment (operating rooms are usually well-lit). Although optical STHMDs monitors do not need as high a resolution as video types do (see the above discussion), the spatial resolution of the virtual images must clearly be higher than the required registration accuracy of the complete system (which should be one tenth of a degree). Also, virtual 3D objects with shaded surfaces will require a higher resolution than wire frame images. Thus one degree of the field of view should be covered by at least 20 pixels in each direction. The dynamic range determines the extent to which the monitor can display low and high intensity pixels at the same time, i.e. it quantifies the range of pixel intensities in a single image. Once again, complex images make higher demands than wire frames, and a convincing integration of virtual objects into real environments must take into account the very high dynamic range of the human eye [Azuma 96a]. For the examination we will use a monitor with a high dynamic range and a high resolution and vary the dynamic range and the resolution of the virtual images computationally.

d) Projection distance of the virtual images

Ideally, the virtual images the surgeon sees in front of him should be focussed to the very point (in the real world) he or she is currently looking at. Otherwise they will appear slightly blurred. This problem could theoretically be solved by an auto-focus mechanism (similar to photocameras). But we think that it is sufficient to implement a manual adjustment option (see step 1) for the following two reasons. Firstly, for the STHMD the distance between the patient and the surgeon can be assumed to be fairly constant (about an arm's length). Similarly, microscopes, endoscopes and binocular magnifying glasses produce real images at a constant (usually infinite) distance [Wolf 96, Olympus 96, Leica 96]. Secondly, the depth of field of the human eye and the brain's capability to 'process' blurred images should compensate for those deviations from the ideal case that actually occur. For a given accomodation distance of the eye, the depth of field marks the range in which objects still seem to be in focus, even though they produce a slightly blurred image on the retina. It depends strongly on the eye pupil's diameter. The depth of field for a two and four mm pupil is ±0.44 and ±0.24 diopters, respectively [Campbell 57]. For an object distance of 1 m and a four mm pupil this gives a clear focus range from 80.6 to 131.6 cm. Larger distances and smaller pupils increase this range, and a well-lit environment like the operating room causes eye pupils to be small. Especially the notch filter approach outlined in point b) promises to let in almost all the light that is available (see above). This step should answer the question whether manual adjustment of the image distance is sufficient.

e) Field of view (FOV) of the interface

The FOV denotes the horizontal and vertical angle under which the monitor images are seen by the user. It is given by the focal length f of the imaging lens system and the dimension d of the monitor according to

$$FOV = 2 \arctan (d/2f)$$

[Rolland 93b]. For d = 0.5" and f = 1", the FOV will be 28 degrees, a typical value for most low-cost HMDs. On one hand, the FOV should cover a rather large part of the operation scene and stereoscopic vision is possible only in the overlap area of the right and left interface FOV. On the other hand, optical STHMDs can do with a rather small FOV that, in addition, does not make any distortion compensation necessary (see the optical vs. video STHMD discussion). We plan to test several high-quality achromatic lenses of different focal lengths in order to find the best compromise.

We will work 'off-line' at this step and use static test images (like a grid) or objects (like a cube). As a consequence, sensoring and tracking problems will not occur. The set-up of the

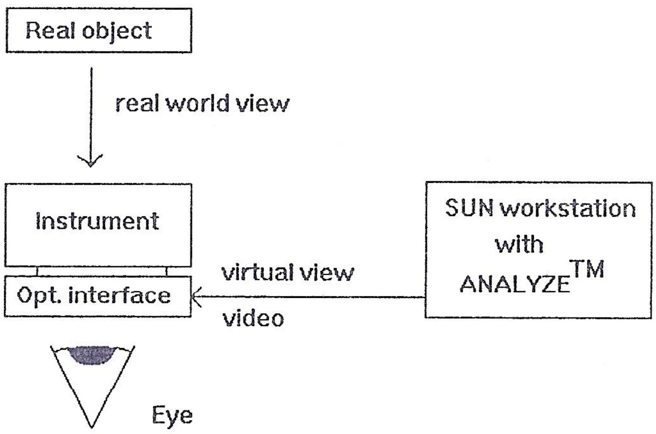

Figure 4: System set-up for step 2

system is shown schematically in figure 4. The 'instrument' is an HMD, a pair of binocular magnifying glasses or a microscope, to which the optical interface is attached. The user looks at a static real world object superimposed with matched virtual video images. The images are delivered by a SUN workstation using the ANALYZE™ programme. ANALYZE™, a product of the Biomedical Imaging Resource of the Mayo Clinic, Rochester, USA, is specifically suited to perform all the image processing options we need, like variation of brightness, colour, dynamic range and resolution.

Since the creation of endoscopic virtual images (see related project) is still under development, such images will probably not yet be available at this stage. Therefore, we will not employ endoscopes here.

At the end of this step we will construct two copies of an interface flange (one for each eye) that is optimized as far as size, weight and the physical parameters discussed above are concerned.

Step 3: Introduction of tracking sensors and stereo vision

Figure 5 shows the system set-up for this step. In contrast to step 2, tracking sensors are attached to both the optical interface and the object. During this step we will vary the position of the instrument and the object. Measurements will be carried out to check the quality of match between the real and the virtual images. Under static conditions, no real time computation of the correct virtual scene is needed in this phase.
The objects used here will be test images and an anthropomorphic phantom. Clinical images (CT, MR etc.) of that phantom will be processed using ANALYZE™ and displayed on the monitor of the optical interface. We will also investigate the display properties of rendered 2D and 3D objects (e.g. a bone structure indicated by a wire frame or shaded surfaces) from segmentation of the original images.

The tracking sensor data will be sent to the workstation and used by ANALYZE™ to calculate correct viewing angles and magnification of the 2D and 3D virtual images. We plan to use adequately transformed images from mono-endoscopes here, too (generated by another project [W. Backfrieder: *Virtual Endoscopy of Multi-modale 3D Datasets in Surgical Navigation*]).

The second goal of this step is to test stereoscopic virtual images for the user who is looking directly through the interface. Binocular magnifying glasses are a good testbed for this application for the reasons given above. After having tested the set-up and established basic performance parameters, we will proceed to the HMD as a stand-alone system. The third device to be considered here is the stereo OR microscope. We will use the ANALYZE™ programme to produce the stereoscopic test images.

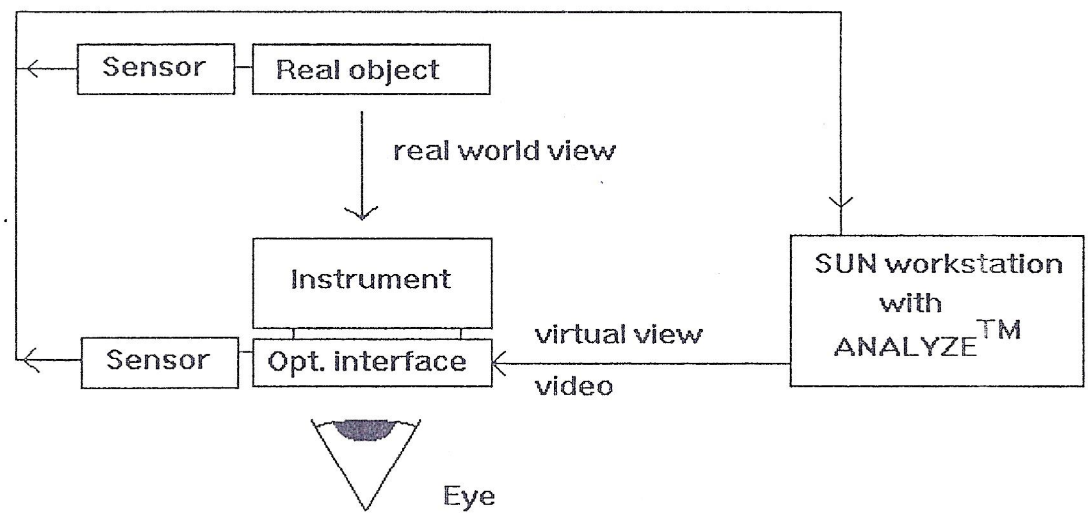

Figure 5: System set-up for step 3

Throughout this step, clinicians will be consulted regularly to give their advice concerning the quality of the overlaid and stereoscopic images.

Step 4: Simulation of an operation

The system set-up for this step is shown in Figure 6. In contrast to the previous step, we drop the restriction of fixed final positions of the 'object' and 'instrument'. The position of the virtual 3D object images must now be calculated *in real time* from tracking sensor data of the user's head (in the case of the HMD and binocular magnifying glasses) and the object's position. During this step we will use the ARTMA Virtual Patient ® System [ARTMA] to produce the virtual images in real time.

Additional software must be produced to create virtual 3D objects that correspond to the actual viewing direction of the surgeon as far as viewing angle and size are concerned.

First we will use simple test objects and images, and then proceed to an anthropomorphic phantom with preoperatively acquired images (CT, MR, etc.). From segmentation of these images, virtual 3D objects (e.g. a tumor) of the wire frame type or with shaded surfaces can be rendered in preoperative planning. This facilitates real time generation of the virtual scene during the operation due to the substantial reduction of the amount of data to be processed.

It is the second goal of this step is to also generate *stereoscopic* virtual images in real time. New software has to be developed to this end. This will be done in collaboration with ARTMA Biomedical Inc. [ARTMA].

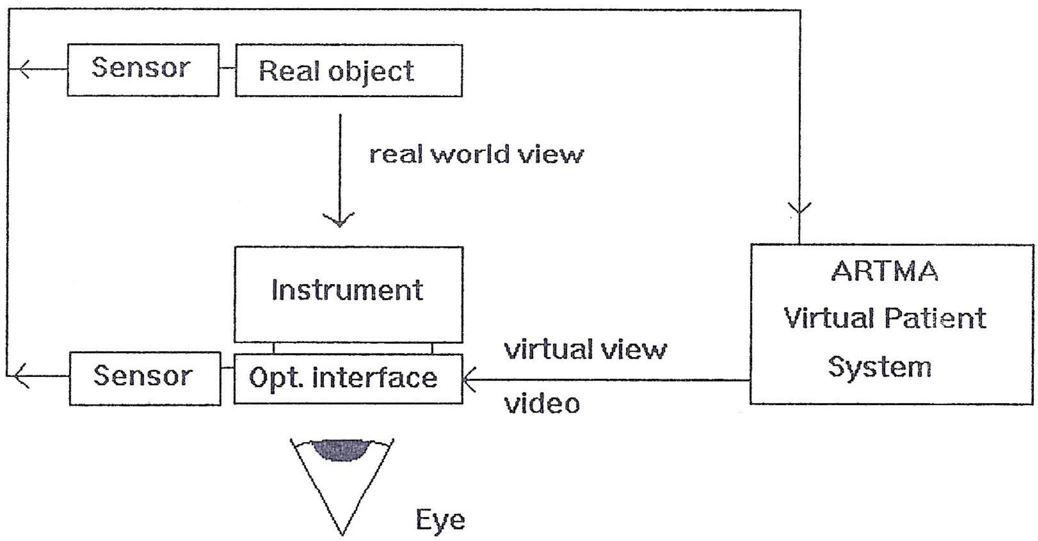

Figure 6: System set-up for step 4

Step 5: Replacement of the stereo monitor by the HMD

As discussed above, we intend the HMD to be used as an alternative to the stereo monitor normally used by the surgeon. This part concerns the stereo microscope and mono-endoscopes (and possibly stereo-endoscopes as well). After having succeeded in providing stereo vision for the see-through mode, it will be straightforward to take this step. The computer generated virtual images produced by the ARTMA Virtual Patient ® System will be merged with the video images provided by the video cameras attached to the respective device. The merged images are then displayed on the HMD's monitors.

In the case of the mono-endoscope, the single video camera will deliver only a 2D real world view. But even for this device, the virtual objects will be presented stereoscopically to give a depth clue and facilitate navigation. This application requires a slightly modified arrangement of the optical parts of the HMD. The surgeon needs to see the operation scene unobstructedly and must be able to look at the merged images as well. So the interfaces containing the monitors will not be placed directly in front of the clinician's eye pupils, but at a suitable distance above the normal viewing direction. Mounted in this way, the surgeon only needs to move his eyes (and not the head !) to change from the operation scene to the monitor display. Since the optical system of the interface produces an image of the monitor at the clinician's working distance, no eye accomodation is necessary. Furthermore, this application does not need the see-through option of the interface, so the beamsplitter can be replaced by a simple mirror.

Specific tasks of the personnel

The personnel involved in the project will participate in the following way. The applicant Prof. Dr. H. Bergmann and Dipl.-Phys. M. Eckholt are responsible for the management of all scientific, technical and organizational problems throughout the project. Technical support and advice as well as the real time generation of mono- and stereoscopic virtual 3D objects and images will be provided by Dr. M. Truppe. The clinical co-workers will give medical and surgical advice during steps 2 to 5 and participate in the evaluation and testing at regular times. They will then test the final product during operations.

2.3 TIME SCHEDULE

A period of *two years* will be needed to complete the project. An estimated breakdown of the time allotted for the steps described above is given in the following table. Step 2 requires the completion of step 1, and so does step 3 with respect to step 2. Steps 3, 4 and 5 can be carried out concurrently. This seperation is indicated below by horizontal lines.

Step 1	:	4	months
Step 2	:	5	months
Step 3	:	6	months
Step 4	:	6	months
Step 5	:	3	months
Total		*24*	*months*

2.4 ESTIMATION OF COSTS

Duration of the project: 2 years

Expenses		*1st year*	*2nd year*	*Total*
Personnel	ATS	328.000	328.000	*656.000*
Devices	ATS	922.200	150.036	*1.072.236*
Material	ATS	80.000	80.000	*160.000*
Travel	ATS	50.000	50.000	*100.000*
Total	ATS	*1.380.200*	*608.036*	***1.988.236***
				Total amount in ATS

3 RESEARCH LOCATION, PERSONNEL AND EQUIPMENT

3.1 RESEARCH LOCATION

Development of the interface and in vitro testing (steps 1 - 4) will be done at the Department of Biomedical Engineering and Physics at the General Hospital Vienna. It provides mechanical and electronical workshops and an optical laboratory furnished with basic optical measurement equipment. It is fully equipped with hard- and software for image processing (appendix 1). An overview of the available electronic and mechanical environment is shown in appendix 2.

Tests of pre- and intraoperative situations will be made in surgery planning rooms and operation theatres at the Department of Oto-Rhino-Laryngology and the Clinic for Cranio-Maxillofacial Surgery, both located in the General Hospital Vienna.

3.2 PERSONNEL

Existing personnel

Applicant:	Prof. Dr. H. Bergmann[1]
Co-applicants:	Prof. Dr. K. Ehrenberger[3],Prof. Dr. R. Ewers[2],Dr. M. Truppe[4]
Co-workers:	Dipl.-Phys. M. Eckholt[1], Dr. F. Watzinger[2],Dr. M. Cartellieri[3],

This project is a co-operation between the following departments of the General Hospital Vienna (AKH)

(1) Department of Biomedical Engineering and Physics
(2) Clinic for Cranio-Maxillofacial Surgery
(3) Department of Oto-Rhino-Laryngology

and the Austrian company

(4) ARTMA Biomedical Inc., Vienna [ARTMA].

Required personnel

1 Ph.D student (physics) ('Dienstvertrag', Dipl.-Phys. M. Eckholt)	ATS	328.000,- p.a.

3.3 EQUIPMENT

Existing equipment

The optical laboratory of the Department of Biomedical Engineering and Physics is furnished with basic optical equipment to carry out optical measurements. A binocular Zeiss microscope is available, but none of the specific optical components needed for this project.

A suitable workstation (SUN Ultrasparc1 170 MHz) and the ANALYZE™ programme is available and will be provided by the Department of Biomedical Engineering and Physics.

The Department of Oto-Rhino-Laryngology and the Clinic for Cranio-Maxillofacial Surgery will provide surgical equipment, especially endoscopes, microscopes and video cameras for short test periods. But none of these devices can be made available exclusively for our use for the duration of the project.

ARTMA Biomedical Inc. [ARTMA] will provide the Virtual Patient ® System software. The additional software needed for the real time generation of stereoscopic virtual 3D objects and images (see step 4) has still to be developed.

Required equipment

1.	A state-of-the-art lightweight optical see-through HMD with high-resolution miniature monitors (see below)	ATS	740.000,-
2.	A rigid see-through mono-endoscope, diameter 4 mm	ATS	30.036,-
3.	A pair of binocular magnifying glasses	ATS	16.600,-
4.	Two high-quality notch filters with an anti-reflection coating (on the backside)	ATS	45.600,-
5.	Development of specific software for the ARTMA Virtual Patient ® System (see step 4)	ATS	240.000,-
		ATS	1.072.236,-

Ad 1: This is the HMD we want to use throughout the project. Due to the rapid development of the VR and AR market the state of the art of HMDs has and will continue to change rapidly. Several manufacturers are expected to come up with significantly improved models in 1997. Therefore we cannot specify a specific HMD model at the moment, but will have to select the most suitable one at the beginning of the project. The price given above is for an optical STHMD equipped with two CRT monitors with a 1" diagonal and a resolution of 1280 x 1014 pixels. The field of view diagonal for each eye amounts to 52 degrees. The interpupillary distance can be individually adjusted. A typical quotation is enclosed.

The generic specifications are as follows. We need a professional high quality product adequate for the precision and resolution required in surgery. We have already pointed out in step 2 c) that at least 20 monitor pixels should cover one degree of the field of view (FOV). A FOV of at least 40 degrees is needed. This leads to a minimum monitor resolution of 800 pixels in each direction. Available low-cost HMDs (like the Virtual i-O [VIO] 'i-glasses!' model in figure 1) suffer from poor resolution and inadequate performance, because important parameters cannot be individually adjusted (e.g. the user-dependent interpupillary distance). In order to develop an interface that has optimal state-of-the-art performance parameters when the project is finished, we need to start out with a HMD of sufficient quality to be able to assess the optical performance parameters relevant for practical work and the software associated performance. Both are expected to be distinctly different from low quality HMDs. Above all, we need to evaluate the performance of achievable delay times for the updating of images under high resolution and real time as required for practical work.

Ad 2 and 3: We will need these, since none of the clinical project partners can provide such equipment for the duration of the adaptation and testing, and limited test periods will not be enough to adapt the instruments for our requirements and carry out the tests.

Ad 4: These are needed in step 2 b). Optical manufacturers produce notch filters only made to order (for a specific wavelength).

4 REQUIRED MATERIAL

We need several optical components (a set of beamsplitter plates with a variety of transmission ratios, achromatic lenses of different focal lengths and mirrors) as well as mechanical parts (e.g. for the option to flip the beamsplitter out of the way, and flange mounting rings), auxiliary components for the electronics of the interface and possibly an extra head-strap or plastic helmet for the HMD.

The optical components have to be shaped into the required form. Grinding, cutting or coating of optical parts and the construction of some special mechanical components (like a flange) cannot be done in our workshops, custom-made parts will have to be purchased form specialized providers.

Together with consumables this will sum up to an estimated amount of ATS 80.000,- per year.

5 REQUIRED TRAVEL EXPENSES

For project-related international travelling we will need ATS 50.000,- per year.

6 MISCELLANEOUS

Nothing needed here.

7 ADDITIONAL REMARKS

No other body has granted or been asked for support.