

Optical See-Through HMD Calibration: A Stereo Method Validated with a Video See-Through System

Yakup Genc[†] Frank Sauer[†] Fabian Wenzel[†] Mihran Tuceryan[‡] Nassir Navab[†]

[†]Siemens Corporate Research
Imaging and Visualization Department
Princeton, NJ 08540, USA
{ygenic,fsauer,fwenzel,navab}@scr.siemens.com

[‡]Dept. of Computer and Information Science
Indiana Univ. Purdue Univ. Indianapolis (IUPUI)
Indianapolis, IN 46202, USA
tuceryan@acm.org

Abstract

This paper presents initial results from an ongoing work to calibrate optical see-through head-mounted displays (HMDs). We have developed a method to calibrate stereoscopic optical see-through HMDs based on the 3D alignment of a target in the physical world with a virtual object in user's view. This is an extension of the Single Point Active Alignment Method (SPAAM) [28] developed for monocular HMDs.

Going from the monocular to the stereoscopic optical HMDs for calibration purposes is not straightforward. This is in part due to the perceptual complexity of stereo fusion process bringing up completely new challenges including the choice of the shape of the virtual object, the physical target and how to display the virtual object without any knowledge of the characteristics of the HMD and eye combination, i.e. the projection model.

In this paper we have addressed these issues and proposed a solution for the calibration problem which we have validated through experiments on the see-through HMD system described in [25]. By experimenting, we have found the appropriate type of virtual objects and physical features to be used in the 3D alignment. Furthermore, we have measured how good the alignment and calibration actually are.

1 Introduction

Augmented reality is a technology in which a user's view of the *real* world is enhanced with additional information generated from a computer model. The enhancements may include labels, 3D rendered models, or shading and illumination changes. Augmented reality allows a user to work with and examine the physical world, while receiving additional information about the objects in it. Some target application areas of augmented reality includes computer-aided surgery, repair and maintenance of complex engines, facilities

modification, and interior design.

In a typical augmented reality system, the view of a real scene is augmented by superimposing computer generated graphics on this view such that the generated graphics are properly aligned with real world objects as needed by the application. The graphics are generated from geometric models of both virtual objects and real objects in the environment. In order for the graphics and the video to align properly, the pose and optical properties of the real and virtual cameras must be the same. The position and orientation of the real and virtual objects in some world coordinate system must also be known. The locations of the geometric models and virtual cameras within the augmented environment may be modified by moving its real counterpart. This is accomplished by tracking the location of the real objects and using this information to update the corresponding transformations within the virtual world. Once these capabilities have been brought together, real objects and computer-generated graphics may be blended together, thus augmenting a dynamic real scene with information stored and processed on a computer.

In order for augmented reality to be effective the real and virtual objects must be accurately positioned relative to each other and properties of certain devices must be accurately specified. This implies that certain measurements or *calibrations* need to be made. These calibrations involve measuring the pose of various components such as the trackers, cameras, etc. What needs to be calibrated in an augmented reality system and how easy or difficult it is to accomplish this depends on the architecture of the particular system and what types of components are used.

For calibration purposes, there are two major modes of display which determine what types of technical issues that need to be addressed: (i) video see-through HMD systems and (ii) optical see-through

HMD systems. The calibration issues in a video see-through HMD system are described in detail elsewhere [27]. We define an optical see-through HMD system as the combination of an optical see-through HMD and a person’s eyes.

In this paper, we look at the calibration issues in an augmented reality system equipped with an optical see-through HMD. In particular, we concentrate on the camera calibration of a stereoscopic optical see-through HMD and describe a method of calibration of such a system. Furthermore, we use a calibrated video see-through HMD system as a development and test environment. This enables us to readily perform quantitative tests to assess the feasibility of a proposed method.

Following the idea proposed by Tuceryan and Navab [28] for the calibration of a monocular optical see-through HMD, we propose a novel approach to calibrate a stereoscopic optical see-through HMD. Instead of simply calibrating the two eyes separately, we align a virtual object (such as a solid disk) with a known physical target. This alignment simultaneously yields two sets of constraints which in turn can be used to calibrate the optical see-through display.

The stereo alignment process is not as straightforward as it is in the monocular case due to the complexities associated with the stereo perception of human eye system. In fact, in the monocular case the alignment is done in the image whereas in the stereo case the alignment is done in 3D. The proposed alignment process works as follows: *The two eyes are presented with the image of a marker with some disparity. This causes the user to see a virtual object that is at some distance away. Then the user aligns this virtual object with a physical target in the scene.* This alignment process brings up completely new challenges including the choice of the shape of the virtual object and the physical target and how to display the virtual object without any knowledge of the characteristics of the optical see-through HMD and the eye combination, i.e. the projection model.

To address these problems, we have used a stereoscopic video see-through HMD system described in [25] as our testbed. This system provided us the means to find out a feasible way to do the stereo alignment without the perception difficulties. In fact, through experiments we have found out that instead of aligning a virtual cross-hair cursor with a physical target on a solid object, it is much easier to align a virtual disk with the center of a circular hole carved on a planar surface. This makes it easier to align the virtual disk without any confusion in depth perception. The video see-through system has also given us

the opportunity to numerically test the accuracy of our alignment process as well as the accuracy of the calibration by providing us with the ground truth.

Once we have tested and fine tuned the proposed stereoscopic optical see-through HMD calibration method on the video see-through system, we calibrated our optical see-through HMD system. Note that while we can test the results of the calibration of the optical HMD only visually, our video see-through system provides us the means to do error analysis on our algorithm.

This paper is organized as follows: After reviewing the previous work done in the area, we provide an overview of the two types of systems we have used in our experiments. Then, we describe the calibration algorithm along with our experiments and conclude with a summary of our findings and planned future work.

2 Background

Research in augmented reality is a recent but expanding activity. We briefly summarize the research conducted to date in this area. Baudel and Beaudouin-Lafon [3] have looked at the problem of controlling certain objects through the use of free hand gestures. Feiner et al. [8] have used augmented reality in a laser printer maintenance task in which the augmented reality system aids the user in the steps required to open the printer and replace various parts. Bajura et al. [2] have used augmented reality in medical applications in which the ultrasound imagery of a patient is superposed on the patient’s video image. Once more, the various registration issues, realism, etc. are open research questions which need to be studied and improved. Lorensen et al. [21] use augmented reality system in surgical planning applications.

Calibration has been an important aspect of research in augmented reality, as well as in other fields, including robotics and computer vision. Camera calibration, in particular, has been studied extensively in the computer vision community (e.g., [22, 29, 20]). Its use in computer graphics, however, has been limited. Deering [6] has explored the methods required to produce accurate high resolution head-tracked stereo display in order to achieve sub-centimeter virtual to physical registration. Azuma and Bishop [1], and Janin et al. [17] describe techniques for calibrating a see-through head-mounted display. Janin’s method comes closest to our approach in terms of its context and intent. Recently, Kato and Billinghurst describe an interactive camera calibration method [19] that uses multiple points on a grid. Gottschalk and Hughes

[11] present a method for auto-calibrating tracking equipment used in augmented and virtual reality. Gleicher and Witkin [10] state that their through-the-lens controls may be used to register 3D models with objects in images. Grimson et al. [12] have explored vision techniques to automate the process of registering medical data to a patient's head.

Some researchers have studied the calibration issues relevant to head mounted displays [2, 5, 1, 15, 18]. Others have focused on monitor based approaches [27, 4, 13, 14, 23, 24]. Both approaches can be suitable depending on the demands of the particular application.

3 Overview of Our Two HMD Systems

In this section we review the two systems we have used in our development process and experiments. First we review the optical see-through system for which we are developing the calibration algorithm. Then we review the video see-through system which is used as a development and verification platform (see Figure 1 for pictures of these systems).

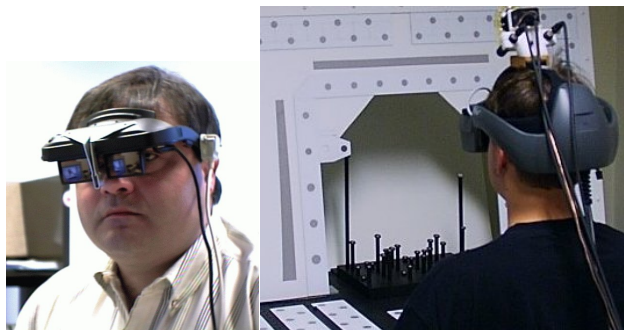


Figure 1: The pictures of the two HMD systems.

3.1 The Optical See-Through System

Our optical see-through system is illustrated in Figure 2. In this configuration, the head-mounted display (HMD) used is a pair of *i-glasses*TM. Although it can be used both as immersive displays as well as see-through displays by removing a piece of opaque plastic from the front of the display screens, we have been using these HMD's as see-through displays permanently. The graphical image is generated by the workstation hardware and displayed on the workstation's monitor which is fed at the same time to the *i-glasses*TM over a VGA port. A 6-degrees-of-freedom (6-DOF) magnetic tracker, which is capable of sensing the three translational and the three rotational degrees of freedom, provides the workstation with continually updated values for the position and orientation of the

tracked objects which include the *i-glasses*TM and a 3D mouse pointing device.

The software is based on the Grasp system that was developed at ECRC (the European Computer-Industry Research Centre). We have added the calibration capabilities to the Grasp software and tested our methods in this environment. The Grasp software was implemented using the C++ programming language.

3.2 The Video See-Through System

The video see-through system we have used is developed by Sauer *et al.* and described in detail in [25]. Here we shortly review this system for the sake of completeness. Figure 3 describes the system that is based on a Kaiser ProView XL35 HMD. Two Panasonic GP-KS1000 color cameras provide the stereo images, a b/w Sony XC-77RR is used for tracking. Two SGI visual PCs, one 320 and one 540, process the three video streams. The system runs in realtime at a framerate of 30 frames per second and exhibits a low latency of only about 2 frames. The software was developed under Windows NT and now runs under Windows 2000.

4 Calibration for Stereoscopic HMDs

An augmented reality system equipped with a see-through HMD needs to be calibrated. Calibration is the process of estimating the parameters of camera and objects models in order to match the virtual objects with their physical counterparts. These parameters may be the optical characteristics of a physical camera as well as position and orientation information of various entities such as the camera, the trackers, and the various objects.

The calibration requirements of a video see-through augmented reality system have been described elsewhere [27]. The video see-through system described in Section 3.2 is calibrated using Tsai's calibration algorithm [26].

A calibration procedure for a monocular see-through HMD system is described in [28]. A straightforward extension of the monocular algorithm to the stereo case is to calibrate two eyes independently. However this will double the effort on the user's side to collect the data for the calibration. Instead, we propose a novel procedure for calibrating the stereoscopic HMDs requiring the same amount of effort as the monocular case.

This approach simplifies both the mathematical model of the camera and the user interaction. First, the user interaction required to collect the data for the calibration is a streamlined process and does not impose a great burden on the user. The user's collection of the necessary data to calibrate the display is a very

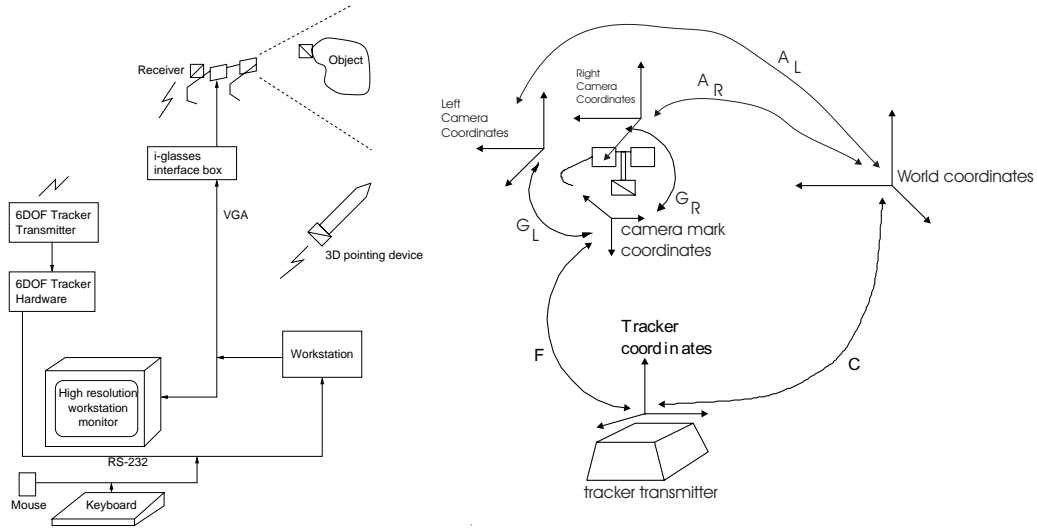


Figure 2: The hardware diagram for the optical see-through system used in our experiments. The see-through displays we use are from *i-glasses*TM, and have a limited resolution (640×480). A simplified version of the coordinate systems that are relevant for the camera calibration of stereo optical see-through systems is shown on the right.

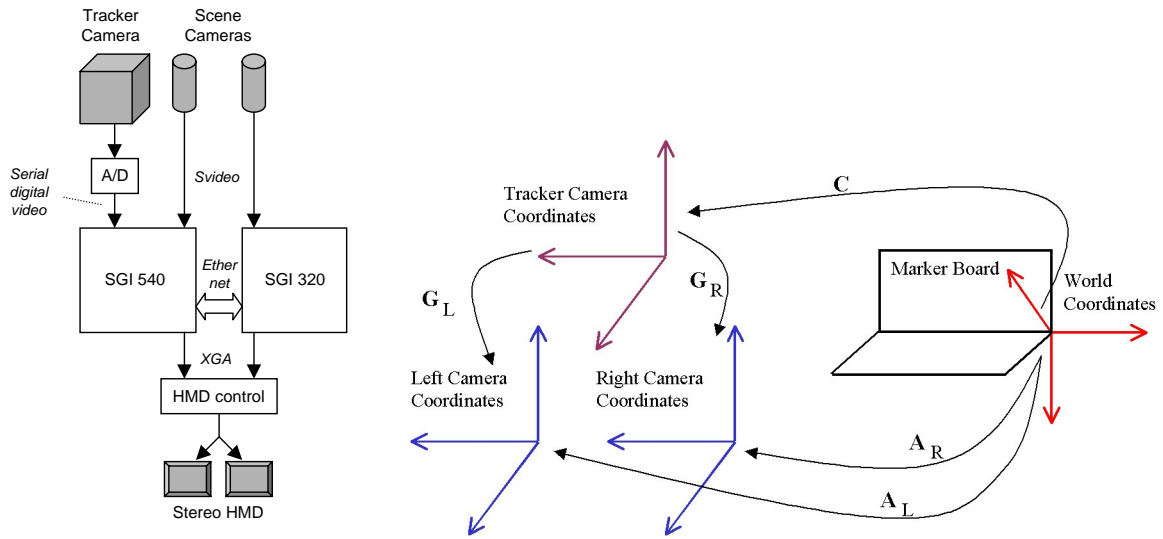


Figure 3: The hardware diagram for the video see-through system used as a testbed. Along with a custom made vision-based tracking system, a Kaiser ProView XL35 HMD with XGA resolution (768×1024) is used. The relevant coordinate systems are shown on the right.

quick and easy process. During this process, the user is not required to have his head fixed and is allowed to move. Second, the camera model is simplified by not insisting on recovering the intrinsic and extrinsic camera parameters separately. Third, the collection of calibration data for both eyes is done in a single step in contrast to the usual practice of calibrating for each eye independently.

In the following, we briefly describe the camera model we are using which defines the parameters to be estimated for the monocular case. We then explain the calibration procedure for the stereoscopic HMDs.

4.1 Camera Model

A simple pinhole model is used for the camera, which defines the basic projective imaging geometry with which the 3D objects are projected onto the 2D image plane. This is an ideal model commonly used in computer graphics and computer vision to capture the imaging geometry [16, 7]. It does not account for certain optical effects (such as non-linear distortions) that are often properties of real cameras but can be ignored in most cases.

The camera can be modeled by a set of intrinsic and extrinsic parameters. The intrinsic parameters are those that define the optical properties of the camera such as the focal length, the aspect ratio of the pixels, and the location of the image center where the optical axis intersects the image plane. One last intrinsic parameter is the skew of the image plane axes. The extrinsic parameters define the position and orientation (pose) of the camera with respect to some external world coordinate system. The 3D points in the world coordinate system get projected onto the image plane of the camera to form the image points.

The transformation that maps the 3D world points into the 2D image coordinates can be characterized by writing the transformation matrices for (i) the rigid transformation matrix defining the camera pose and (ii) the projection matrix defining the image formation process. Let R and T represent the camera pose in the world coordinate system, in which R is a 3×3 rotation matrix and T is a 3D vector representing the translation.

The rigid transformation can also be written as a 4×4 homogeneous matrix

$$T_{\text{pose}} = \begin{pmatrix} R & T \\ \mathbf{0} & 1 \end{pmatrix} \quad (1)$$

The perspective projection that projects the 3D points onto the image plane can be modeled by a 3×4 pro-

jection matrix given by

$$T_{\text{proj}} = \begin{pmatrix} \alpha_u & s & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (2)$$

where α_u and α_v are horizontal and vertical scales in the image plane, u_0 and v_0 are the center of the image and s is the parameter that models the skew in the image plane coordinate axes. Note that T_{pose} describes the extrinsic parameters of the camera and T_{proj} describes the intrinsic camera parameters. The overall camera transformation is then given by the product

$$T_{\text{cam}} = T_{\text{proj}}T_{\text{pose}} \quad (3)$$

Given a set of 3D points and their images, the transformation T_{cam} can be recovered using linear least-square methods.

4.2 Calibration Formulation

The overall projective transformation defined by the camera can be written by a 3×4 matrix T_{cam} and the entries of this projection matrix can be estimated directly instead of the actual extrinsic and intrinsic camera parameters.

The estimation of the 3×4 projection matrix is a standard technique often used in computer vision [7]. The calibration proceeds by collecting a number of 2D image coordinates of known 3D calibration points, and the correspondence between the 3D and 2D coordinates defines a linear system to be solved in terms of the unknown projection matrix entries which can be solved for using linear least-squares.

There are some modifications to the traditional camera calibration method in our approach which allows the user to move his head during the calibration procedure. In our method, we have a tracking system that measures the rigid transformation from the world coordinate system to the tracker coordinate system. Because the tracker (the receiver for magnetic tracking case and the tracker camera for the vision-based tracking case) is attached rigidly to the HMD, the camera can be defined and calibrated with respect to the tracker coordinate system. Therefore, taking this approach, we have the camera transformation fixed and unaffected by the head motion. *This is the reason that the head is allowed to move freely during the calibration procedure.*

The entire setup is summarized in Figures 2 and 3 which shows the coordinate systems that are relevant for calibration of a stereo see-through system. In these figures, we see six transformations (A_L , A_R , FC , G_L , and G_R) that need to be estimated. The transformations A_L and A_R are the 3×4 projective camera

transformation with respect to the world coordinate system that needs to be estimated. The matrix product FC is the 4×4 homogeneous transform matrix that defines the world to tracker rigid transform (F is the identity matrix for the vision tracker case). Finally, G_L and G_R are the 3×4 projection matrices that define the camera transformations with respect to the tracker coordinates. The figure can be summarized by the equations

$$\begin{aligned} A_L &= G_L FC \\ A_R &= G_R FC \end{aligned} \quad (4)$$

To keep the notation simple, we will refer to the transformations A_L and A_R as A , collectively. Similarly, we will use G for G_L and G_R . The equations apply equally to both eyes for the calibration estimation. The contribution of the paper comes in collecting the calibration data for both eyes in a single step.

The transformation A varies as the HMD moves about, and this movement is captured by the transformation FC which can be obtained from the tracking system. The other transformation, G is fixed and does not change and can be estimated by other calibration procedures.

To summarize, in order to calibrate the camera (i.e., to estimate the transformation A) we need to get the image coordinates of known 3D points in the world coordinate system. But A is not fixed and varies as the user moves his head. Therefore, unless we want to force the user to keep his head static we need to estimate the camera transformation another way. This we do by calibrating the camera in the tracker coordinate system (i.e., estimate the transformation G which does not change. In order to accomplish this we take the known 3D calibration point and transform it into the tracker coordinate system, then perform the standard camera calibration procedure on the new point. Let $P_W = [x_W, y_W, z_W, 1]^T$ be the homogeneous coordinates of the known 3D point in the world coordinate system. Let $P_I = [u, v, s]^T$ be the homogeneous coordinates of its image point. Then first we transform the world coordinates to tracker coordinates by

$$P_M = FCP_W \quad (5)$$

Then we use the P_M and its image P_I to estimate the transformation G . The standard projective camera calibration is set up as follows. Let there be n calibration points whose image coordinates we measure. There are 12 parameters of the 3×4 projection matrix we need to estimate. But the projection matrix is defined up to a scale factor, therefore, there are only 11 independent parameters that need to be estimated.

So, there should be at least $n \geq 6$ calibration points to be measured.

4.3 Calibration Procedure

In order to get a practical calibration procedure for the stereoscopic HMDs, the above formulation needs to be converted to a user-friendly procedure. We have implemented the calibration procedure as follows:

1. The world coordinate system is fixed with respect to the tracker coordinate system (by defining the world coordinate system on the tracker transmitter for magnetic tracking or any point in the scene for vision-based tracking). We assume that the internal calibration for the tracker has already been performed.
2. A *single* point in the world coordinate system is used to collect the calibration data. This single point in the world coordinate system is mapped to many distinct points in the tracker coordinate system as the user's head moves.
3. The user is presented with a marker (i.e., a solid disk) on the display for each eye placed randomly in the 2D image plane. The disks for the two eyes are slightly offset creating a disparity. The user's brain automatically fuses these and the user perceives the disk in 3D at a particular location in depth. The user collects the calibration data by moving his head until he aligns the perceived disk in 3D with the 3D physical calibration point (see Figure 4). The user then clicks a button on the mouse and the data is collected for calibration that consists of the image coordinates of the the disk and the 3D coordinates of the calibration point in tracker coordinates.
4. These collected points are then fed into the linear system which is used to estimate the camera parameters. Notice here that the more of the tracker volume the user's head covers, the more of possible systematic errors in the tracker measurements will be taken into account in the optimization process. The user is encouraged to move his head around the tracker transmitter as much as possible while collecting the calibration data.

The estimated projection matrices can be integrated into OpenGL in order to render the graphics with the correct perspective as described in [28].

4.4 Stereo Alignment

Figures 4 and 5 explain the 3D stereo alignment process. As a user is observing the physical world through an HMD, a marker is displayed in left and right eyes with a disparity. When viewed in stereo, these markers induce a virtual object (or marker)

which is perceived at some distance away from the user. The user’s job the is to align the perceived virtual marker with a physical target in the scene. When the alignment is done, the user’s head will be at some distance away from the physical marker proportional to the disparity.

The distance at which the virtual marker is perceived is proportional to the amount of disparity between the images of the marker in the two eyes. For any choice of disparity, the alignment can be done from many different positions keeping the user’s head at some fixed distance away from the physical target as long as the disparity remains the same.

The choice of the virtual marker is critical. We first used the cross-hair marker and tried to align its perceived virtual counterpart with a surface mark in the scene. However, the cross-hair did not provide enough visual clues in order the user to perceive the depth. We instead used a solid disk which overcame this problem.

Another critical choice is that of the physical target. Alignment of the virtual disk and a surface mark turned out be problematic. The problem stemmed from the fact the human brain is confused while the virtual object is still visible at the same time it is perceived behind the surface of a solid object. We have avoided this by picking the physical target as the center of a circular hole cut on a planar surface. This way the user was able to move back and forth to align the depth correctly.

5 Experimental Verification

A serious problem with the verification of an optical see-through display calibration is that it is not possible to show how well the model corresponds with the object for a human viewer. We have implemented the calibration procedure described in this paper in both the video see-through and the optical see-through systems described earlier in Section 3. We have tested the proposed algorithm on the optical see-through system and tried to verify the results visually. A better way is to use the video see-through system allowing us to collect quantitative data and do an error analysis on the algorithm.

One of the most important issues we were interested in addressing was to find out the degree of accuracy of the stereo alignment process described in Section 4.4. Since our video see-through system provided us with the complete tracking and calibration parameters, we conducted the following experiment to assess the accuracy of the alignment process. We first set the disparity between the left and right images of the virtual marker such that it is at some distance away from the

Table 1: Reprojection errors for reconstructed projection matrices using the proposed algorithm on the video see-through system. Since optical see-through system does not provide any means of measuring the reprojection errors, the video see-through system becomes a good testbed for validating the performance of the proposed algorithm. The ground truth corresponds to the projection matrix obtained using the calibration and tracking results from the system. See the text for more detailed interpretation of these results.

Projection Matrix		Error in Reprojection [Pixels]	
		Mean	Std. Dev.
Ground 1	M_L	1.31859	0.37578
	M_R	1.33312	0.35230
Stereo 1	M_L	1.53801	0.47569
	M_R	1.58263	0.50138
Ground 2	M_L	1.33705	0.32332
	M_R	1.35228	0.30009
Stereo 2	M_L	1.41955	0.39548
	M_R	1.38305	0.32819
Ground 3	M_L	1.28706	0.39539
	M_R	1.29697	0.36024
Stereo 3	M_L	1.45690	0.52982
	M_R	1.23107	0.44084
Ground 4	M_L	1.32917	0.41486
	M_R	1.34742	0.40013
Stereo 4	M_L	1.49881	0.42945
	M_R	1.31938	0.33465

user. Using the calibration parameters obtained for the two cameras providing the stereo image stream for the HMD, we computed the position of the virtual object in the tracker coordinate system. We then let the users do the alignment from different positions for the same disparity many times. As we have pointed out earlier, with the fixed disparity the users can move their head on a sphere centered at the physical target. Therefore, when the virtual marker and the physical target are aligned, the different positions of the user’s head should be at the same distance to the physical target. We recorded some 15-35 such alignments for 2 people for two different disparity values and found out that with a distance of about 750mm (resp. 720mm) to the scene, our alignment resulted in an average distance very close to the measured ones with less than 1mm (resp. 2mm) standard deviation. Figure 6 shows two graphs of the collected data plotted as the distance from the physical target to the user’s head. As can be observed from the graph, the alignment is very consistent.

We also conducted some experiments to assess the reprojection errors for the estimated projection matrices. The results are presented in Table 1 which

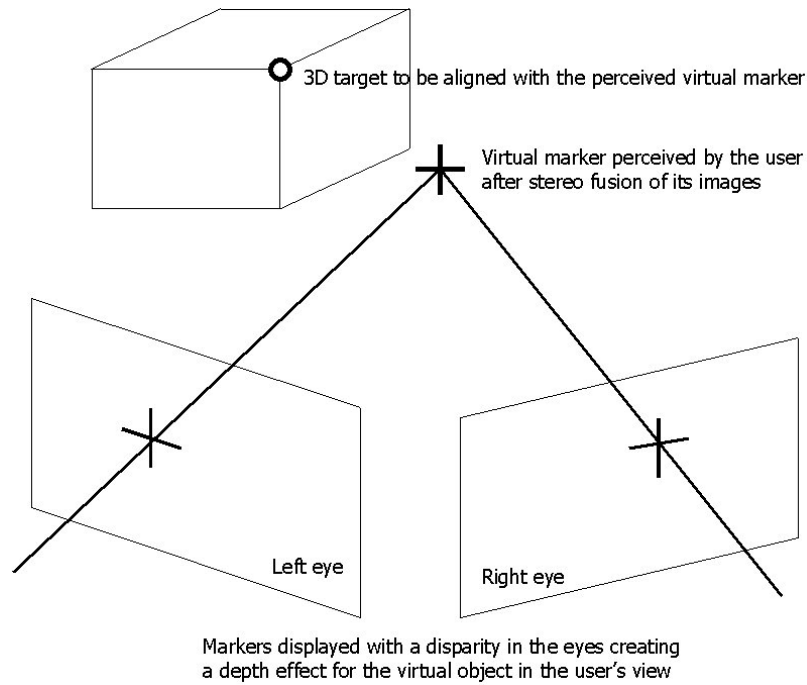


Figure 4: Aligning a virtual marker with a physical target in 3D: When markers with a disparity in left and right eyes are displayed in stereo, the user perceives a virtual marker at some distance away. The user then aligns the virtual marker with a physical target in the scene.

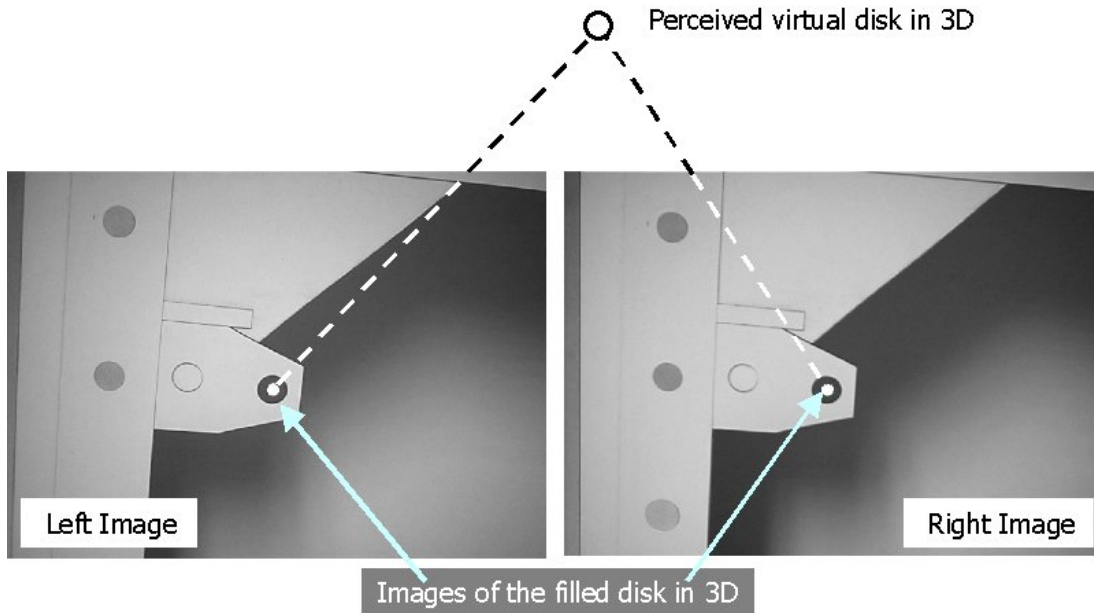


Figure 5: An example of alignment in stereo (with the video see-through system): the left and right images show the physical target (the center of the hole) augmented with the virtual disk. The hole provides strong visual clues for the eyes allowing the right depth perception. A cross-hair marker provides weaker visual clues for the depth perception.

shows the results for four different experiments (two people repeating the experiments twice). To be able to compare, we have computed the projection matrices for each camera from the calibration parameters of the video see-through system which is obtained by Tsai's camera calibration algorithm [26] and regard these as *ground* truth values. The reconstructed projection matrices using the proposed algorithm is labeled as *stereo*. As it can be seen from this table, the reprojection errors for both the ground truth and the reconstructed ones are quite low, less than 1.5 pixels on the average,

5.1 Remaining Issues

Although we have experimentally validated our calibration algorithm on a video see-through system, to date, we have not validated it for the optical see-through system. Here we discuss some of the issues we are currently investigating that is relevant to optical see-through HMDs.

First of all, in our stereo alignment process we assumed that only the lateral disparity will be needed to perceive a 3D virtual object in the user's view. We believe that human eye will still fuse the virtual object if the vertical disparity is not large. However, since we do not know the geometry of the eye and HMD, we cannot claim it for sure. We plan to investigate this problem in detail to find out what kind of disparities are needed for the right perception in depth.

A second problem is associated with the rendering of the virtual markers on displays for the eyes. For best perception, the marker images should reflect the perspective and distortion effects that comes from the eyes and the HMDs. However, we have noticed that human eye is very good at fusing the virtual markers even without these effects.

Finally, our current alignment process requires that the physical target be viewed from frontal views. In order to cover a larger area we need to modify our target in such a way that it can be seen at an oblique angle without any difficulty in alignment. We are currently investigating alternative targets that overcome this issue.

6 Conclusion

In this paper, we presented a camera calibration procedure for *stereo* optical see-through head-mounted displays for augmented reality systems. Since we do not have direct access to the image produced on the retina, the procedure needs to use indirect methods to do the calibration. The method presented in this paper uses an interactive method to collect calibration data for both eyes in one step and

it does not require that the user keep his head still during the data gathering process.

The inability to measure the accuracy of any calibration procedure for an optical see-through system prompted us to use a well calibrated video see-through system developed by Sauer *et al.* [25] as a testbed. The use of the video see-through system allowed us to verify the accuracy of our stereo alignment process which is in fact quite accurate. Furthermore, the testbed provided a fast development environment in which we have fine-tuned our choice of the virtual features that is used in the alignment process. Finally, the data collected from the video see-through system allowed us to do error analysis on our calibration algorithm.

Our current work is directed towards resolving the remaining issues discussed in Section 5.1. Further research plans include more experiments for a complete error analysis of our algorithm.

References

- [1] R. Azuma and G. Bishop. Improving static and dynamic registration in an optical see-through display. *Computer Graphics*, July 1994.
- [2] M. Bajura, H. Fuchs, and Ohbuchi. Merging virtual objects with the real world: Seeing ultrasound imagery within the patient. *Computer Graphics*, July 1992.
- [3] M. Baudel and M. Beaudouin-Lafon. Charade: Remote control of objects using freehand gestures. *Comm. of the ACM*, 37(7):28–35.
- [4] F. Betting, J. Feldmar, N. Ayache, and F. Devernay. A framework for fusing stereo images with volumetric medical images. In *Proc. of CVRMed'95*, 1995.
- [5] T. Caudell and D. Mizell. Augmented reality: An application of heads-up display technology to manual manufacturing processes. In *Proc. of the Hawaii Int. Conf. on System Sciences*, 1992.
- [6] M. Deering. High resolution virtual reality. *Computer Graphics*, 26(2):195–202, 1992.
- [7] O.D. Faugeras. *Three-Dimensional Computer Vision*. MIT Press, 1993.
- [8] S. Feiner, B. MacIntyre, and D. Seligmann. Knowledge-based augmented reality. *Comm. of the ACM*, 36(2):53–62, 1993.
- [9] A. Fournier. Illumination problems in computer augmented reality. In *Journée INRIA, Analyse/Synthèse D'Images*, pages 1–21, January 1994.
- [10] M. Gleicher and A. Witkin. Through-the-lens camera control. *Computer Graphics*, pages 331–340, July 1992.
- [11] S. Gottschalk and J. Hughes. Autocalibration for virtual environments tracking hardware. *Computer Graphics*, pages 65–72, August 1993.

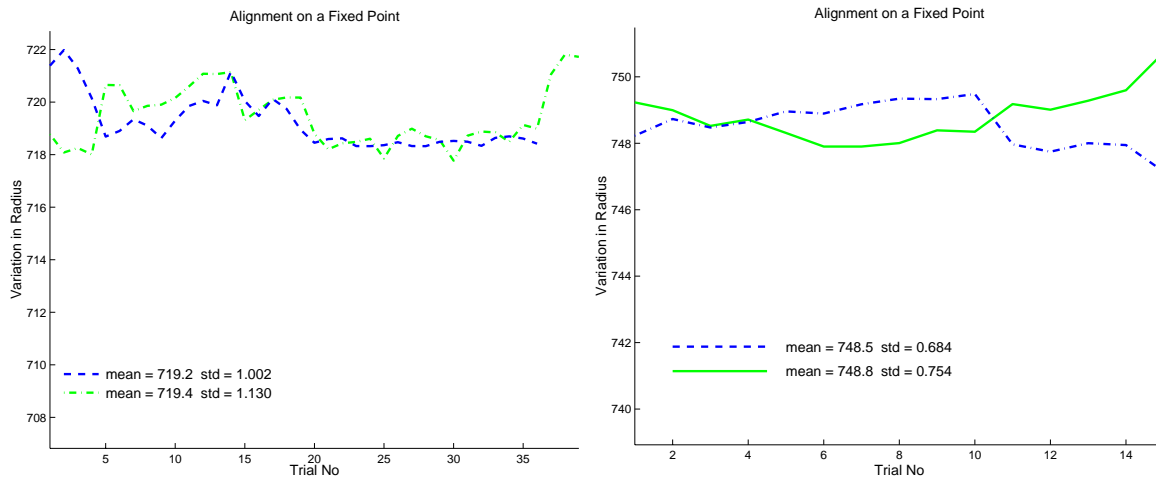


Figure 6: Experimental results of measuring the accuracy of the depth judgment for the perceived 3D disk. For each experiment, two stereo pairs of images of the disk are aligned with respect to the center of a hole in the scene as explained in the text. Each pair restricts the position of the camera centers to lie on a sphere centered at the 3D scene point.

- [12] E. Grimson, T. Lozano-Perez, W. M. Wells, G. J. Ettinger, S. J. White, and R. Kikinis. An automatic registration method for frameless stereotaxy, image guided surgery, and enhanced reality visualization. In *Proc. of CVPR'94*, 1994.
- [13] W.E.L. Grimson, G. J. Ettinger, S.J. White, P.L. Gleason, T. Lozano-Perez, III W. M. Wells, and R. Kikinis. Evaluating and validating an automated registration system for enhanced reality visualization in surgery. In *Proc. of CVRMed'95*, pages 3–12, 1995.
- [14] C. J. Henri, A.C.F. Colchester, J. Zhao, D.J. Hawkes, D.L.G. Hill, and R. L. Evans. Registration of 3D surface data for intra-operative guidance and visualization in frameless stereotactic neurosurgery. In *CVRMed'95*, 1995.
- [15] R. Holloway. *An Analysis of Registration Errors in a See-Through Head-Mounted Display System for Craniofacial Surgery Planning*. PhD thesis, U. of North Carolina, 1994.
- [16] B.K.P. Horn. *Computer Vision*. MIT Press, Cambridge, Mass., 1986.
- [17] A. Janin, D. Mizell, and T. Caudell. Calibration of head-mounted displays for augmented reality applications. In *Proc. of VRAIS'93*, pages 246–255, 1993.
- [18] A. R. Kancherla, J. P. Rolland, D. L. Wright, and G. Burdea. A novel virtual reality tool for teaching dynamic 3D anatomy. In *CVRMed'95*, 1995.
- [19] H. Kato and M. Billinghurst. Marker tracking and hmd calibration for a video-based augmented reality conferencing system. In *Proc. of the 2nd Int. Workshop on Augmented Reality '99*, 1999.
- [20] R. K. Lenz and R.Y. Tsai. Techniques for calibration of the scale factor and image center for high accuracy 3-d machine vision metrology. *IEEE PAMI*, 10:713–720, 1988.
- [21] W. Lorensen, H. Cline, C. Nafis, R. Kikinis, D. Altabelli, and L. Gleason. Enhancing reality in the operating room. In *Proc. of IEEE Conf. on Visualization*, pages 410–415, 1993.
- [22] S. J. Maybank and O. D. Faugeras. A theory of self calibration of a moving camera. *Int. J. of Comp. Vision*, 8(2):123–151, 1992.
- [23] J. P. Mellor. Real-time camera calibration for enhanced reality visualizations. In *CVRMed'95*, 1995.
- [24] O. Peria, L. Chevalier, A. François-Joubert, J. P. Caravel, S. Dalsoglio, S. Lavallee, and P. Cinquin. Using a 3D position sensor for registration of spect and us images of the kidney. In *Proc. of CVRMed'95*, 1995.
- [25] F. Sauer, F. Wenzel, S. Vogt, Y. Tao, Y. Genc, and A. Bani-Hashemi. Augmented workspace: Designing an AR testbed. In *Int. Symp. for Augmented Reality*, 2000.
- [26] R.Y. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras. *IEEE J. of Robotics and Automation*, 3(4):323–344, 1987.
- [27] M. Tuceryan, D. Greer, R. Whitaker, D. Breen, C. Crampton, E. Rose, and K. Ahlers. Calibration requirements and procedures for a monitor-based augmented reality system. *IEEE Trans. on Vis. and Comp. Graphics*, 1(3):255–273, 1995.
- [28] M. Tuceryan and N. Navab. Single point active alignment method (spaam) for optical see-through hmd calibration for ar. In *Int. Symp. for Augmented Reality*, 2000.
- [29] J. Weng, P. Cohen, and M. Herniou. Camera calibration with distortion models and accuracy evaluation. *IEEE PAMI*, 14(10):965–980, 1992.