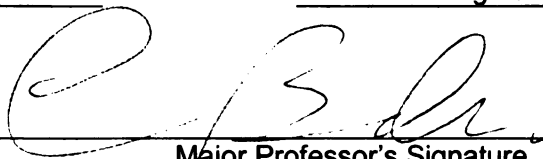This is to certify that the
dissertation entitled

CALIBRATION OF OPTICAL SEE THROUGH HEAD
MOUNTED DISPLAYS FOR AUGMENTED REALITY

presented by

JI ZHOU

has been accepted towards fulfillment
of the requirements for the

Ph. D.　　　degree in　　　Computer Science and
　　　　　　　　　　　　　　　　Engineering

Major Professor's Signature

12 - 14 - 2007

Date

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE

6/07 p:/CIRC/DateDue.indd-p.1

CALIBRATION OF OPTICAL SEE THROUGH HEAD MOUNTED
DISPLAYS FOR AUGMENTED REALITY

By

Ji Zhou

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSIPHY

Department of Computer Science and Engineering

2007

ABSTRACT

CALIBRATION OF OPTICAL SEE THROUGH HEAD MOUNTED DISPLAYS FOR AUGMENTED REALITY

By

Ji Zhou

Optical See-Through Head-Mounted Displays (OSTHMDs) superimpose computer-generated imagery on the user's view to enhance the user's visual perception and task performance in the Augmented Reality (AR) applications. Calibration is a crucial step for OSTHMD-based AR systems to achieve the required accurate registration between the virtual and physical objects in the composite environment presented to the user. The OSTHMD calibration is a challenging problem because of the complexity of handling the indispensable human factors in the computational model and operational procedure. The existing techniques are either unstable or inaccurate because of excessive dependence on the user procedure or lack of user adjustment design.

This dissertation presents the Display Relative Calibration (DRC) method, a flexible two-phase technique which provides an effective solution to satisfy the accuracy and usability requirements for the OSTHMD calibration. The DRC method is based on a virtual camera model which establishes the computational relationships of the OSTHMD imaging system involving the user's eyes. The analysis of the virtual camera parameters

allows the design of a two-phase approach which gives a close estimation of all parameters in the first phase and refines the user-dependent parameters in the second phase. To produce stable results without relying on the user performance, the DRC phase I involves no user interaction by using a calibrated camera in place of the user's eye. The DRC phase II provides multiple options of user procedures to accommodate different AR applications and user groups. This dissertation also describes the implementation of the DRC virtual camera model and presents analyses of the sensitivity and errors. Furthermore, an experimental user study is conducted to evaluate the performance of the DRC method in contrast to the existing OSTHMD calibration techniques. The design and results of this experiment are described in detail.

ACKNOWLEDGEMENTS

First and the foremost, I extend my sincere appreciation to my academic advisor, Dr. Charles B. Owen, for his continuous help, patience and valuable guidance throughout my Ph.D. program.

Next, I want to thank my guidance committee members, Dr. George Stockman, Dr. Juyang Weng, and Dr. Frank Biocca, for their precious advice and comments on my work.

My gratitude goes to the faculty and staff member of the Department of Computer Science and Engineering for their kind help and support.

I am indebted to Dr. Cynthia M. Okolo in the Department of Counseling, Educational Psychology, and Special Education and Dr. Shu-Guang Li in the Department of Civil and Environmental Engineering, for their support during my graduate program.

My thanks go to researchers and students in the METLAB and MIND Lab for their suggestion and discussion, and my friends in MSU for their lots of help.

Furthermore, I would like to express my grateful acknowledgements to all those who helped me complete this dissertation. Without their support and encouragement, I could never have made it happen.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANOVA   Analysis of Variance

API   Application Programming Interface

AR   Augmented Reality

HMD   Head-Mounted Display

DOF   Degree Of Freedom

DCT   Discrete Cosine Transform

LCD   Liquid Crystal Display

MSE   Mean Squared Error

OSTHMD   Optical See-Through Head-Mounted Display

RMS   Root Mean Square

SMC   Stylus-Marker Calibration

STHMD   See-Through Head-Mounted Display

SPAAM   Single Point Active Alignment Method

VR   Virtual Reality

VSTHMD   Video See-Through Head-Mounted Display

Chapter 1 Introduction

Optical See-Through Head-Mounted Display (OSTHMD) is an image compositing device widely used in Augmented Reality (AR) applications to enhance the user's visual perception with computer-generated imagery. Most AR applications demands accurate registration, which implements spatial and visual consistency between the augmented virtual elements and the physical scene. The OSTHMD calibration is a crucial step for OSTHMD-based AR systems to achieve registration.

This chapter presents an introductory overview of the OSTHMD calibration for AR, including the concepts and techniques of AR, the OSTHMD-based AR systems, the AR registration requirements, and the motivations of the OSTHMD calibration. The thesis statement and a brief summary of the research contributions are also given in this chapter.

1.1 Background

The OSTHMD calibration is used for the implementation of AR systems and applications. To better understand this problem, it is necessary to introduce the concepts, components and techniques of the AR and the registration requirements for AR.

1.1.1 Augmented Reality

Augmented Reality (AR) refers to computer systems which supplement a user's

sensory perception with computer-generated elements [1-3]. As a variation of Virtual Reality (VR), AR presents the user a composite environment with coexisting virtual and physical elements instead of a complete computerized synthetic world produce by a VR system[4]. Although AR is a generic human-computer interaction paradigm which can be applied to any human sense such as vision [5], hearing [6, 7], touch [8] and even smell [9], most AR research and applications have been focused on vision augmentation which overlays computer-generated graphics or text on the human user's view of the surroundings. A typical vision-based AR system superimposes graphical objects into the user's visual field such that the user observes the virtual and physical objects simultaneously. Figure 1.1 shows a composite view presented by an augmented vision system where the physical objects are superimposed with graphical elements. Without special note, the AR system mentioned in this thesis always refers to vision-based AR.
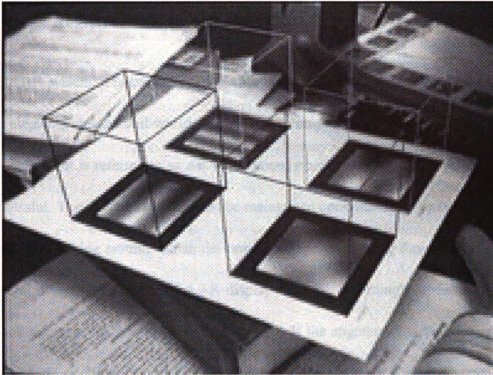
Figure 1.1 An AR composite view with coexisting graphical elements and real objects

By supplementing the user's vision with computer-generated information, AR systems can enhance the user's situational awareness [10], improve the task performance [11] and support better interaction with the surroundings [12]. AR provides numerous possibilities for new types of computer-human interaction and richer user experiences [3, 13-15]. AR has been applied to many applications in different fields such as medical visualization [16-18], industrial planning and manufacturing [19-24], battlefield operation support [25-27], aircraft pilot display [28-32], media [33, 34] and entertainment [35, 36].

1.1.2 AR Registration

In an AR system, it is essential to maintain visual and spatial consistency between the virtual elements and the real-world objects in the composite view observed by the user. This requirement is referred to as AR registration, which is crucial for an AR system to be successful. The AR registration must be maintained continuously, even when the user is moving or looking around. When the user's position or viewing direction is changed, the virtual objects rendered by the AR display must be accordingly updated such that they look the same as their physical counterparts. If the augmented virtual elements are spatially static or "floating around" upon the real objects, such an AR system cannot provide more information than a normal computer monitor display can do. The AR registration enables the virtual elements to be seamlessly integrated with the physical environment. Most AR applications demand accurate registration, especially for the medical or military applications where the registration is critical for the system user.

1.1.3 Components and Techniques

As shown in Figure 1.2, a typical AR system is composed of three major enabling components: tracking, rendering and compositing. The achievement of AR registration depends on proper implementation and configuration of these components and the relevant techniques.
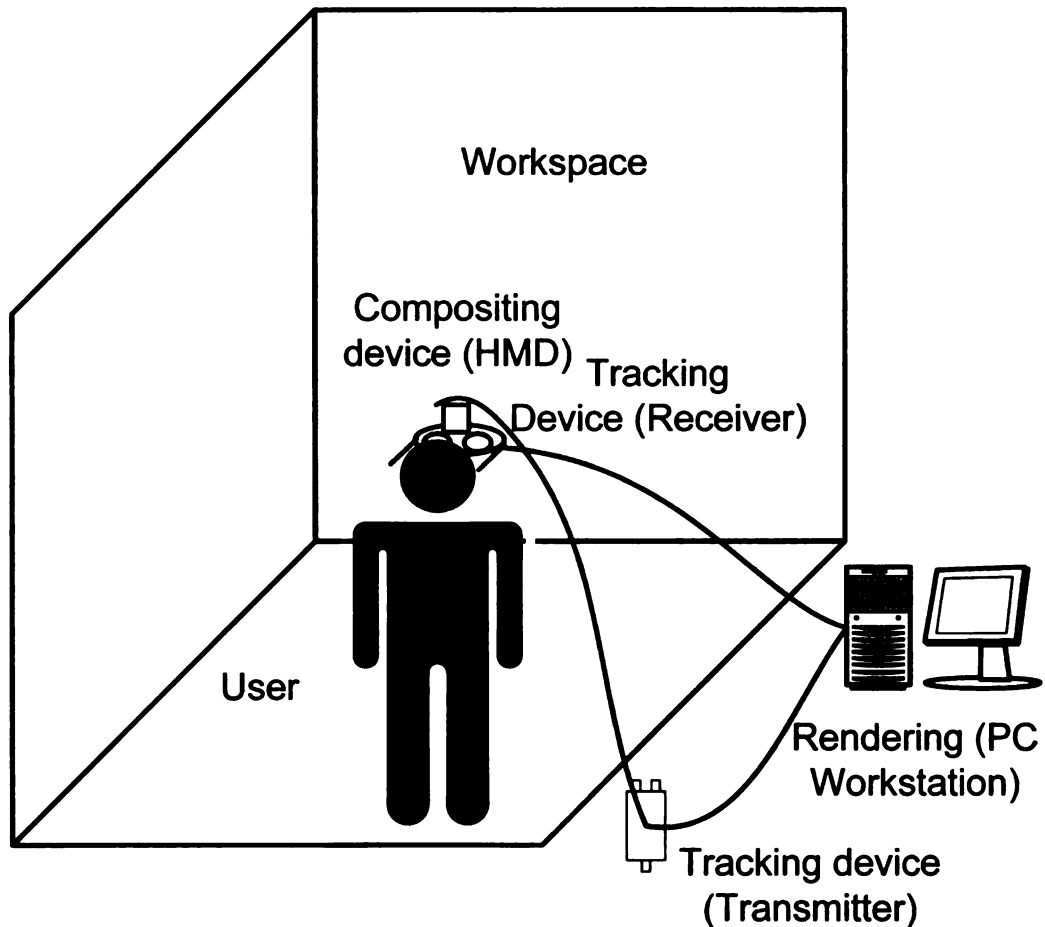
Figure 1.2 Components of a typical HMD-based AR system

In AR systems, the registration is implemented dynamically in accordance to the user's location and viewing direction such that the rendered imagery can be updated when the user moves or looks around. To obtain the real-time position and orientation of the user's eye, a tracker is usually mounted on the user's head or the Head-Mounted Display (HMD). The tracker cannot directly track the user's eye point, but such information can be obtained by combining the tracker data and the tracker calibration data which gives the rigid transformation between the tracker and the user's eye. Azuma discusses the tracking requirements for AR, including high-accuracy, low latency, and

long working range [37]. For outdoor AR applications, GPS or vision-based tracking can be applied [36, 38, 39]. The trackers used in AR systems can be built with various technologies including inertial sensing, magnetism, acoustics, optoelectronics, and mechanical linkages [40-42]. Different trackers provide different levels of precision in their applicable environments. Vision-based camera tracking has been widely used in AR applications because of its simple setup and decent accuracy [43-46]. Based on whether fiducials are used, vision-based tracking can be categorized as fiducial tracking [47-49] and markerless tracking [50-54].

The rendering module generates the imagery supplemented to the user's view. The produced imagery can be any graphics or text, but most AR applications produce 3-D image such that the user perceives a composite environment with coexisting graphical virtual objects and physical objects. When properly configured with correct computational model and rendering parameters, the 3-D imagery can be seamlessly integrated into the user's view of the surroundings such that the perceived 3-D virtual objects look as if they are part of the actual scene. To achieve this goal, the 3-D rendering module needs to simulate the human eye's viewing process and visual perception. A variety of 3-D graphics Application Programming Interfaces (APIs) can be used for AR rendering and OpenGL [55, 56] has been a popular and versatile 3-D API supported by many graphics hardware and systems. OpenGL converts 3D elements to display graphical primitives by utilizing a series of modeling, viewing and projection

transformations to simulate the view of a camera or human eyes [57, 58]. For the AR

system under investigation, these transformations and the relevant parameters can be

defined by establishing a computational model for the imaging process. When all the

required parameters are determined by calibration or by tracker data, the computational

model of the AR system can be implemented using OpenGL to render the 3-D imagery

which satisfies the registration requirements.

Compositing refers to the imaging devices and processes which combine the

computerized and physical elements to present the user an integrated view in AR systems

[59]. Categorized by the compositing techniques, the AR systems include See-Through

Head-Mounted Display (STHMD)-based [4], monitor-based [22, 60] and

projection-based [61-64]. Patick et al. conduct an empirical study to investigate

differences in spatial knowledge learned for a virtual environment presented in STHMD,

projection screen, and desktop-monitor [65]. STHMDs are more popular and utilized in

more AR systems and applications because the STHMDs are more intuitive and flexible.

The concept of the STHMD has been initiated by E. Sutherland, a pioneer in computer

graphics, HMD and AR research, as "presenting the user with a perspective image which

changes when he moves" [66].

There are two types of STHMDs, including Video See-Through Head-Mounted

Displays (VSTHMDs) [67-69] and Optical See-Through Head-Mounted Displays

(OSTHMDs) [70-73]. As their names implies, OSTHMDs rely on optical merging

mechanism such as a "half-silvered" mirror to combine the virtual and physical scene, while Video See-through HMD uses digital video technique to combine two video streams, one generated by computer and another coming from a video camera mounted on the user's head. Azuma and Rolland et al. compare the advantages and disadvantages of OSTHMDs and VSTHMDs [3, 70]. OSTHMDs have many advantages over VSTHMDs and other AR displays in that the OSTHMD user can view the real scene without added latency and there is no limitation for the resolution of the view of the physical environment [24, 70].

1.1.4 OSTHMD Calibration

Regardless of the many advantages, OSTHMDs have not shown extensive utilization and development for AR systems and applications as would be expected. One of the major technical challenges for OSTHMD systems is the calibration, which determines the computational model of the OSTHMD-based AR system and initializes the required parameters to achieve spatial registration of augmentations. Whereas the video combining approaches used by monitor-based and VSTHMD-based AR systems allow for direct capture and analysis of the registered scene, OSTHMDs present the combined image to the retina of the user's eye. As shown in Figure 1.3, an OSTHMD typically utilizes a beam combiner, i.e. a half-silvered mirror, for combining virtual and physical scenes. When placed in front of the eye, the beam combiner reflects the computer-generated image into the user's line of sight, and also allows lights from the

surrounding environment to pass through. The OSTHMD's lens system transforms the rays of the miniature LCD screen to form a virtual display converged with proper size at a comfortable distance. The combination and registration of the virtual and physical scenes are observed by the specific OSHMD user. Due to the discrepancy of the eye position, different user of the same OSTHMD system might experience variance of the combined scene. A stereo HMD view can be achieved if separate imaging systems and corresponding computer graphics are provided for each eye.
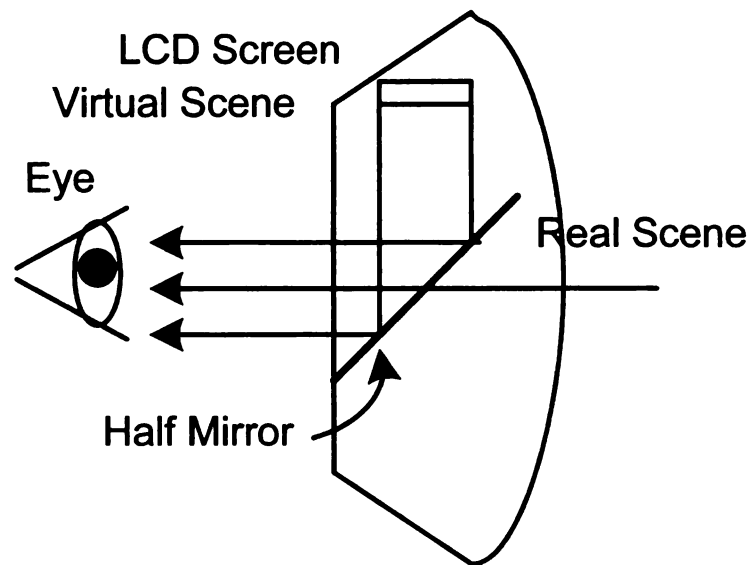
Figure 1.3 Conceptual imaging system of OSTHMD

The OSTHMD calibration establishes and solves the computational model for the above imaging process such that the rendered virtual scene is properly aligned with the real scene. The special optical design of OSTHMDs makes the calibration difficult because the actual combined scene cannot be accessed and user interaction is required to indicate the registration observed by the user. Compared with camera calibration, the

9

OSTHMD calibration is more challenging because of the complexity of handling the user factors together with system computation in the design and implementation of the calibration method. Many OSTHMD-based systems need to be calibrated such that the registration between virtual and real objects can be achieved with reasonable accuracy suitable for the applications and the users.

1.2　Motivation

Extensive research has been conducted on the OSTHMD calibration problem ([45, 74-79] to cite a few). However, as far as registration accuracy and usability are concerned, the existing approaches still have considerable drawbacks and ignored issues which require further investigation. This thesis seeks to provide correct explanations and effective solutions to the following questions which have not been properly resolved in the previous work:

1)　Since the user is an indispensable part of the viewing process, how can we establish a computational model of the OSTHMD imaging system to take into account the user factors and to support a reasonable solution?

Many existing calibration methods treat the OSTHMD imaging system and the user's eye together as a whole camera model for calibration. This kind of approach is straightforward but the role of the user's eyes in the computational model is usually unclear or ignored. Without knowledge of how the user's eyes work in OSTHMD

viewing process, the existing approaches usually fail to properly handle the user factors in the calibration procedure. It is necessary to conduct further investigation on the OSTHMD computational model based on analyses of the use's eyes and the relevant user factors in the calibration process.

2) What would be the best way to solve the user-dependent OSTHMD computational model without resorting to complicated user procedure or excessive user efforts?

Most existing OSTHMD calibration approaches require complicated user procedure because there are large numbers of unknown parameters to be solved. Meanwhile, the unstable user performance makes the calibration results unreliable. An effective solution is to reduce the number of unknown parameters, i.e. degrees of freedom (DOF), which demands an in-depth study on the computation and the user procedure design.

3) Is it possible to design a reliable user procedure which works for most OSTHMD users, even those having difficulties to complete calibration? Meanwhile, can we make the calibration method versatile such that different users can reach their expected registration accuracy through similar operational steps?

It is important for the OSTHMD calibration to always reach acceptable time rather than unusable data, while some existing methods might fail due to data error caused by poor user performance. To avoid calibration fail, a possible solution is to design a

two-step approach which seeks to improve an acceptable calibration results instead of pursuing calibration from scratch. Meanwhile, most existing calibration methods are not suitable for all the users due to the static design of user procedure. To make the calibration flexible, an option is to design multiple user procedures such that the user can choose the suitable one.

4)   How can we estimate the registration accuracy and evaluate the performance of an OSTHMD calibration method?

In most previous research work, there is no theoretical analysis of the calibration accuracy to be expected. The estimation of the accuracy depends on the system and user errors as wells as the computational limitations. These analyses are necessary for an OSTHMD calibration method to be successful. Meanwhile, very few user studies have been conducted to comparatively evaluate the performance of an OSTHMD calibration method in contrast to other techniques. Considering that both registration accuracy and usability are essential for the OSTHMD calibration, the evaluation should be an experimental study including comparable measurements of user operability.

1.3   Problem Definition

The OSTHMD calibration problem is challenging due to the following human factors introduced by the OSTHMD system:

1)   Both the registration and the imaging are user-dependent processes related to the user's eye position and viewing direction. The same point on the OSTHMD image plane might overlay on the different world positions for different users.

2)   There is no way to directly access the virtual image plane observed by the user. The registration data required for calculating the calibration parameters can only be indicated by the user himself/herself for the best accuracy.

Taking into account the above human factors, the design of OSTHMD calibration should satisfy the requirements for both registration accuracy and usability, including the following criteria:

1)   Accurate: The image rendered using the computational model and the calibration results should be precisely registered with the physical counterparts in the user's view.

2)   Simple: The calibration algorithm can be easily implemented. The system setup should be simple without any special hardware involved.

3)   User-adjustable: The user interaction should be included in the calibration procedure and the solved parameters should be adjustable for a specific user.

4)   Reliable: The calibration should not excessively depend on the user procedure to avoid unreliable results due to user performance.

5)   Flexible: The user procedure should be flexible to accommodate different user groups and applications. Variable options should be available such that an inexperienced

user can obtain acceptable accuracy using a simple approach while a professional user can achieve very accurate registration using a more complicated schema.

The goal of the OSTHMD calibration problem is to present an effective calibration method which satisfies the above criteria. There are two tasks to be fulfilled: First, to establish and solve a computational model for the OSTHMD imaging system such that the AR registration can be achieved, and next, to design and implement a user-friendly operational scheme and the relevant computation processes. To help the readers implement the proposed method in their applications, the presentation of the proposed method should be algorithm-oriented including the system setup, detailed steps, and the implementation. Furthermore, a user experiment is necessary to evaluate the performance of the proposed method and compare it with the existing techniques.

1.4   Thesis Statement and Contributions

The following statement will be explained and proved in this thesis:

To solve the OSTHMD calibration problem, it is crucial to consider the user factors when establishing the computational model and designing the operational procedure. By properly integrating the user into a virtual camera model, the Display Relative Calibration (DRC) method proposed in this thesis provides effective and accurate results using a two-phase approach. The DRC method produces a close estimation using a calibrated camera in place of the user's eye and allows the user-dependent parameters

to be refined through user procedure. With supporting of optional multiple user procedures, the DRC method is suitable for various OSTHMD-based AR systems and different users.

This thesis makes the following contributions to the research and development of OSTHMD-based AR:

1) This thesis presents Display Relative Calibration (DRC), an effective and flexible calibration method suitable for various OSTHMD-based AR applications and different users. Compared with the existing approaches, DRC is unique because it is the only OSTHMD calibration method which demonstrates good performance in both accuracy and usability. By supporting simplified operational procedures and accurate calibration results, the DRC method is expected to promote the utilization of the OSTHMD systems for AR applications.

2) The two-phase strategy used in DRC which combines offline image-based computation and online user procedures is an innovative and effective solution to OSTHMD calibration. This design is different from the widely-used two-stage algorithms and the same idea might be applied to other similar problems which are involved with the user factors.

3) The design of multiple user procedure options in DRC is also a first of its kind creation which provides the user the flexibility to choose the most suitable procedure based on application requirements or user groups. On the contrary, the existing

approaches support only a single fixed user procedure, which limits their applicable applications and users.

4)   The modeling and computation of the OSTHMD curved image plane benefit both optical designer and HMD system developer. This is the first research to investigate and determine the characteristics of OSTHMD virtual image plane because most previous work assumes a flat image plane and ignores the errors introduced by optical aberration.

5)   The conducted user study takes more user subjects than any other previous experiments. Thus, the experimental results are more effective in revealing the relationships among the registration accuracy, the completion time and the task load of the investigated OSTHMD calibration procedure. Such information would be helpful for the design of the user interaction in the OSTHMD calibration.

1.5   Outline

The remainder of this thesis is organized as the following:

Chapter 2 reviews the existing OSTHMD calibration approaches, which can be categorized as offline and online methods based on whether active user interaction is involved. The methods in each category are further classified into and their advantages and drawbacks are described.

Chapter 3 presents the computational model for the OSTHMD imaging system and

viewing process based on the registration requirements of AR.

Chapter 4 introduces the first phase of the new two-phase DRC method. The DRC phase I solves the parameters of a virtual camera model which uses a calibrated in place of the user's eye. The OSTHMD virtual image plane is also calibrated in this phase using triangulation.

Chapter 5 describes the phase II of the DRC method which solves the user-dependent parameters of the virtual camera model using multiple options of DRC phase II. These options provide different levels of accuracy and user efforts to satisfy requirements of different AR applications and user groups.

Chapter 6 provides details to implement the established OSTHMD computational model and also gives a sensitivity analysis which estimates the registration error ranges due to the user's eye offsets.

Chapter 7 analyzes the major sources of errors related to the OSTHMD calibration to give an estimation of the registration accuracy that can be expected.

Chapter 8 discusses options to optimize the OSTHMD calibration, including computational methods and user interface design.

Chapter 9 describes the user experiment which evaluates the accuracy and usability of the proposed DRC method in contrast to other selected calibration methods. The

experimental design, system setup and the procedure are given in details. The experiment

results are analyzed using statistical methods.

Chapter 10 concludes the research results and proposes the topics for future research.

Chapter 2   Related Work

This chapter reviews the existing approaches of the OSTHMD calibration and summarizes the advantages and drawbacks of these methods. Based on whether the user interaction is required, the existing OSTHMD calibration methods can be categorized into offline and online approaches. Online calibration method requires certain user input or interaction during the calibration procedure while offline can be done without any user participation.

2.1   Offline Methods

User interaction is usually not required for the offline approaches because the parameters are obtained either by directly measurement or by calibrating a camera located in the eye position of the OSTHMD. Without user interaction, the offline methods are simple and fast but usually unable to reach very high accuracy since they cannot reproduce the exact viewing process experienced by the user. The offline methods include the following approaches:

1) Offline calibration without user interaction

Gao et al. [80] use a calibrated camera to simulate human eye and compute the parameters of the OSTHMD imaging system based on the point correspondences extracted using image-based feature detection algorithms. Involving no user procedure, this kind of approaches is fast but actually incomplete because the results are inaccurate

and unadjusted for a specific user.

2) Measuring eye position using special hardware

Caudell's method [19] uses a special registration frame kinematically mounted with the display to measure the user eye position. Summers et al. [81] estimate the center point on the OSTHMD glasses as the eye position. With the help of specially-designed calibration hardware, the user's eye position can be measured through a simple user procedure. These techniques cannot achieve high accuracy because other parameters such as the image center and the focal length are not calculated and adjusted for each user.

2.2   Online Methods

In online calibration methods, the user is usually required to perform alignment of the virtual and real points or other kind of operations such that the collected data can be used to solve the virtual camera model. Therefore, if the user can provide fairly precise point correspondences, the online methods can adapt different users to the display and provide much more accurate calibration results than those obtained by offline methods. The disadvantage of this category is that the user interaction methods used to collect point correspondence are usually complicated, especially for those inexperienced users. Based on the alignment schemes for collecting virtual-real correspondences, the online methods can be further divided into the following types:

1) Static alignment

These methods require the user to keep the head still during the user procedures which

usually request the user to align the perceived virtual points, which is not a simple task without help from external tool such as a head rest or a bite bar. The calibration results may be unstable when the user's head is moved. Klinker et al. [82] compute the offset from the use's eye to a tracking camera. The user is requested to indicate misregistration by drawing lines from the observed markers to the real ones. A head rest is used to keep the head still during the procedure. Oishi et al. [79, 83] calibrate the transformations caused by the mechanical misalignments and displacement of the user's eye. The user is required to align and record the position of a virtual cursor on a movable plate. In the calibration method by Hua et al. [84, 85], the user needs to align a large number of observed image point with their corresponding points on a calibration frame.

2) Dynamic alignment of virtual-world objects

These techniques require the user to dynamically align a custom virtual object with their physical counterpart. Although there is no need to keep the head still, the user procedure is usually complicated and hard to accomplish. McGarrity et al. [74] propose an interactive scheme to dynamically estimate the parameters by using mouse to align the image points with their corresponding physical   points. The method by Grassat et al. [86] requests the user to match a fixed image to the real object by moving the head. Azuma's method [87] requires the user to perform boresight operations to calibrate the transformation from the tracker to the eye and the field of view (FOV).

3) Dynamic alignment of virtual-world point correspondences

The methods in this category request the user to align a number of virtual image

21

points with the physical points such that the calibration parameters can be solved based on the collected point correspondences. The accuracy calibration results might be unstable due to the excessive dependence on the user's alignment performance. The requirement of the large number of alignments makes the user procedure tedious and fallible. Janin [20] combines direct measuring the transformation and projection parameters using instruments and an optimization procedure in which twenty calibration points are manually matched to a fixed virtual crosshair. Tuceryan et al. [88, 89] propose the Single Point Active Alignment Method (SPAAM), which requests the user to move the head to align different 2D virtual markers with a single fixed world point. SPAAM is a popular OSTHMD calibration method because the procedure is straightforward. At least 6 point correspondences are required to solve the virtual camera model of the OSTHMD imaging system. Genc et al. propose several extensions to the original SPAAM, including a stereo calibration method [90], an approach using a vision-based tracker [45] and a two-stage method [91]. Baillot et al. apply the SPAAM calibration to mobile AR platforms when the transformation from world to base is unknown [92]. Fuhrmann et al. [75-77] proposed stylus-based calibration, in which the user aligns a virtual marker with a real marker on a tracked stylus. The advantage of the stylus-based approach is that the manual dexterity of the hands is considerably greater than that of the neck. A similar approach proposed by Kato et al. uses vision-based tracking to capture point correspondences [49].

2.3   Summary

An effective OSTHMD calibration algorithm should achieve accurate registration for

each specific user without excessively relying on the performance user procedure. Both the offline and online calibration methods have their advantages and disadvantages, but an obvious fact is that the offline approaches are apt to increase user operability and the online approaches prefer accuracy. Thus, none of the existing OSTHMD calibration approaches can reach the requirements of both accuracy and usability.

Ideally, the offline and online methods should be combined into a two-stage adaptable calibration approach, like the two-stage method proposed by Genc et al. [91]. But Genc's method is actually an extension of SPAAM calibration [89], which only decreases the number of required point correspondences but still needs considerable user interaction.
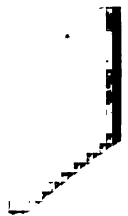
Chapter 3   Computational Model

The goal of OSTHMD calibration is to establish and solve a computational model which can be implemented by rendering software to generate imagery for achieving AR registration. Besides the AR registration requirements, the established OSTHMD computational model should be implemented by the OSTHMD rendering module to generate the expected imagery. The computational model proposed by Robinett et al. [93] [94] is complicated but it is hard to be implemented using generic graphics software.

Accurately modeling the OSTHMD imaging process is a prerequisite to the success of calibration. This chapter describes the establishment of the OSTHMD computational model. The OSTHMD imaging system and viewing process are analyzed to embody the registration requirements which are the basis of the computational model. The OSTHMD optical model is explained to account for the special feature of the OSTHMD virtual image plane. Based on these analyses, a virtual camera model is created to model the OSTHMD imaging system composed of the user's eyes and the OSTHMD optics.

3.1   Imaging System and Viewing Process

Figure 3.1 illustrates a conceptual diagram of the imaging system of a typical OSTHMD. A 3-D virtual object perceived by the user is rendered as a 2-D image on a miniature LCD screen using 3-D graphics library such as OpenGL [55, 57]. The OSTHMD optical system transforms the rays of the LCD screen to form an enlarged virtual image plane which converges at a certain distance in front of the user's eyes. This

virtual image plane is designed as a flat rectangle plane which looks half-transparent to the user. Compared with the original LCD screen, this virtual image plane keeps the same number of pixels in both horizontal and vertical dimensions but the distance between adjacent pixels becomes larger. The special design of the OSTHMD imaging system enhances the user's visual experience by allowing the user to observe the real scene overlaid with the 3-D virtual (graphical) objects.
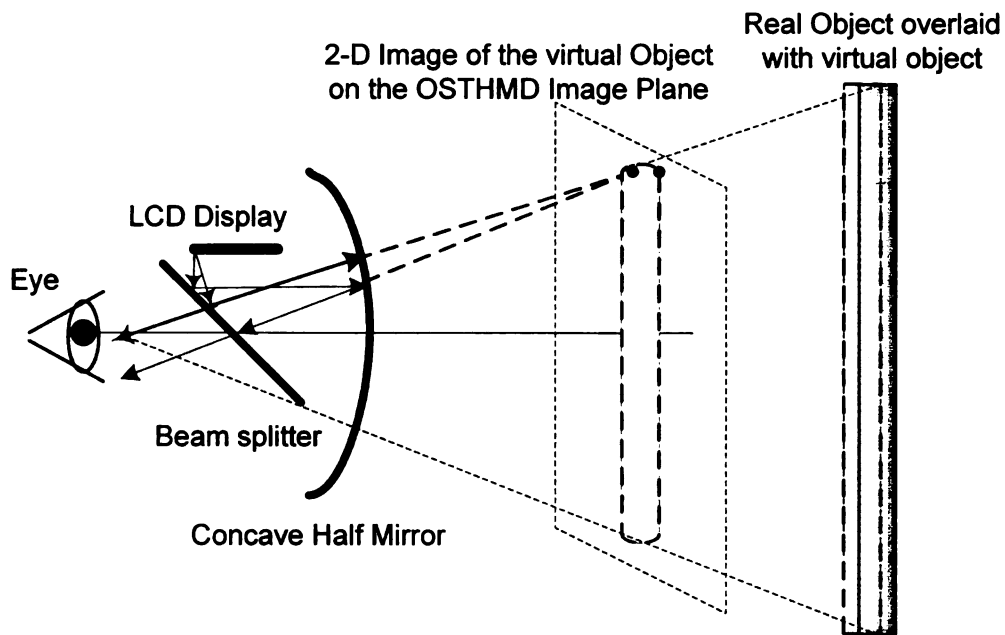
Figure 3.1 Imaging system and viewing process of OSTHMD (monocular configuration)

As shown in Figure 3.1, the 3-D virtual objects perceived by the user are displayed as 2-D imagery on the OSTHMD virtual image plane. To achieve registration between a virtual and physical object, the 2-D imagery of the virtual object must exactly overlay on its physical counterpart such that it looks the same as the 2-D imagery of the physical object observed from the user's eye. In most 3-D graphics libraries including OpenGL, the viewing system of the human's eye can be equivalently modeled as a virtual camera

with perspective projection. To satisfy the registration requirements, the OSTHMD rendering system must precisely reproduce the same modeling, viewing and projection transformations as experienced by the user's eye, including the following conditions:

1)   The point coordinates of a virtual object must be the same as those of their physical counterparts relative to a same predefined world coordinate system.

2)   The center of projection and the viewing direction of the virtual camera must be the same as those of the user's eye.

3)   The intrinsic parameters of the virtual camera define the same viewing volume as the one established by the user's eye.

3.2   Optical Model

The above description is based on paraxial analysis which assumes a constant focal length and a pure flat virtual image plane. Actually, as shown in Figure 3.2, the field curvature of the concave mirrors typically used in the OSTHMD results in a shortening of the effective focal length as the distance from the center to the virtual image plane increases. Note that the theoretical focal length of a concave mirror $f_0 = \dfrac{R}{2}$ (R is the radius of the curvature) only holds for paraxial analysis which assumes an infinitesimally small aperture and only applies to a bundle of rays close to the optical center. According to [95], the focal length affected by the field curvature can be given as:

$$f = \frac{R}{2} - \frac{(1+K)r^2}{4R} - \frac{(1+K)(3+K)r^4}{16R^3} - \ldots \tag{3.1}$$

where $r = \sqrt{x^2 + y^2}$ is the radial distance from a point on the mirror to the optical axis,

R is the radius of curvature, and K is the conic constant of the conic section, $K = -e^2$ (e is the eccentricity). Considering the variation of the focal length f, the depth value z of the virtual image plane can be expressed as:

$$z = z_0 + \frac{2s}{(R-2s)^2} \cdot r^2 - \frac{4(1+K)}{(R-2s)^2 R^2} \cdot r^4 - \ldots \tag{3.2}$$

where s is the distance from the object to apex and z_0 is the depth computed from the

conventional equation $\dfrac{1}{f} = \dfrac{1}{s} - \dfrac{1}{z}$ under the condition of paraxial reflection.
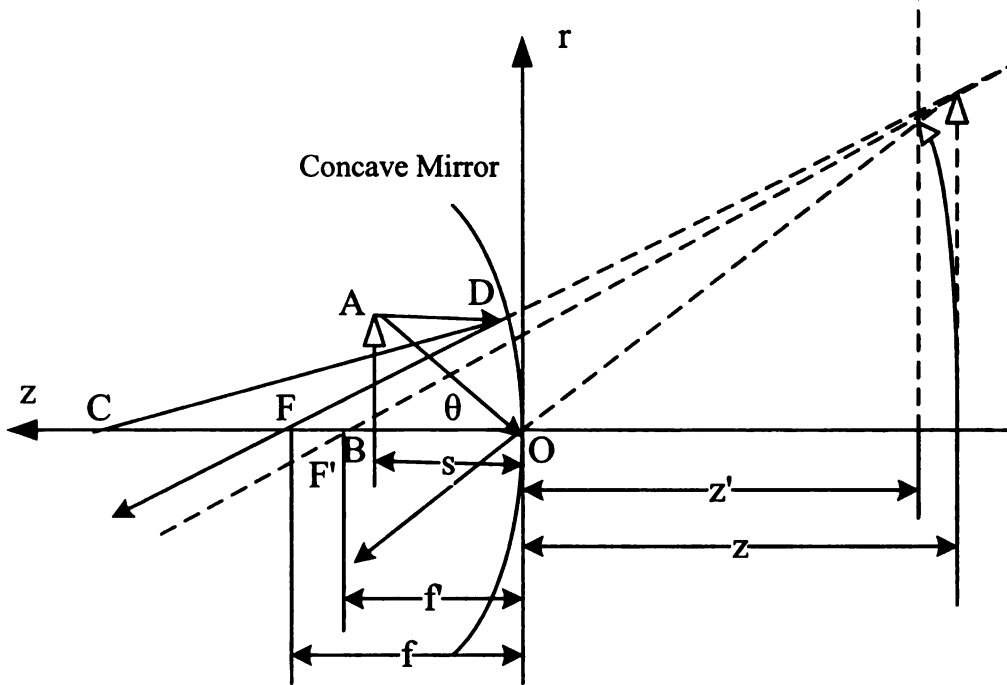
Figure 3.2 OSTHMD optical geometry with field curvature

Thus, the depth of points on the virtual image plane can be modeled as

$$z = z_0 + K_2 \left(\left(x - a_x \right)^2 + \left(y - a_y \right)^2 \right) \qquad (3.3)$$

where z_0 is the maximum depth of the plane (correspondent to the depth of apex), K_2 is an aberration parameter, $\left(a_x, a_y \right)$ is the apex of the concave mirror in the DRC coordinates, and $\left(x, y \right)$ refers to the coordinates of a point on a normalized virtual display plane with depth $z = -1$.

3.3   Virtual Camera Model

Although the above analysis indicates that the OSTHMD virtual image is curved instead of flat, it does not significantly change the observed view when the user's eyes are on or close to the optical axis. Without loss of generality, here the computational model of the OSTHMD imaging system is given based on the paraxial analysis with a flat image plane and the effects of curved plane on the modeling and viewing will be discussed later.

Based on the registration requirements, the OSTHMD imaging system can be modeled as a virtual camera whose center of projection coincides with the user's eye. To solve the OSTHMD virtual camera, its intrinsic and extrinsic parameters need to be defined and associated with the coordinate systems and the transformations of an OSTHMD system shown in Figure 3.3. Among these coordinate systems, the world coordinate system is predefined and fixed in the working space and the tracker coordinate

system is usually associated with the tracker mounted with the OSTHMD. The virtual camera coordinate system originates at the user's eye center, and its z axis coincides with the optical axis, which is perpendicular to the OSTHMD image plane. The x, y axes are parallel to the horizontal and vertical edges of the OSTHMD image plane, respectively. Points in these different coordinate systems are related with each other by rigid transformations, including T_{WtoT} from the world coordinate system to the tracker coordinate system and T_{TtoV} from the tracker coordinate system to the virtual camera coordinate system. An image coordinate system can be defined on the OSTHMD image plane. It is centered at the principle point $O'(u_0, v_0)$ where the optical axis of the OSTHMD virtual camera intersects with the image plane and its x, y axes are chosen to be parallel to the horizontal and vertical edges of the virtual image plane, respectively.
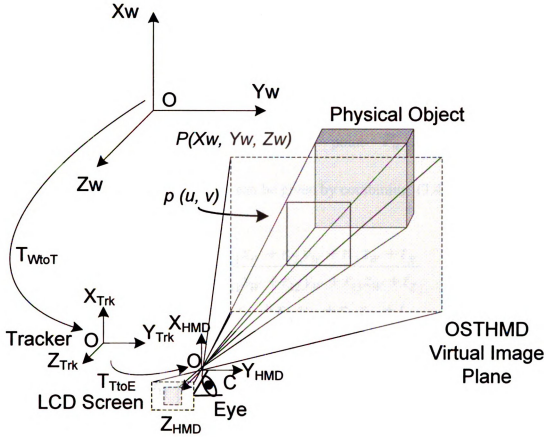
Figure 3.3 Coordinate systems and viewing transformations in an typical OSTHMD system

A point $P_W(x_W, y_W, z_W)$ defined in the world coordinate system can be transformed to the point $P_V(x_V, y_V, z_V)$ in the virtual camera coordinate system by:

$$P_V = T_{TtoV} T_{WtoT} P_W \qquad (3.4)$$

Denote the effective focal length of the virtual camera by f and the pixel aspect ratio by s_x, the image point coordinates $p(u, v)$ of a point $P_V(x_V, y_V, z_V)$ in the virtual camera coordinate system are given by:

$$u - u_0 = s_X f \frac{x_V}{z_V}$$

$$v - v_0 = f \frac{y_V}{z_V}$$

(3.5)

Thus, the relationship between a world point $P_W\left(x_W, y_W, z_W\right)$ and its

corresponding image point $p\left(u, v\right)$ can be given by combining (3.4) and (3.5) as:

$$\frac{u - u_0}{s_x} = f \frac{r_{11}x_W + r_{12}y_W + r_{13}z_W + t_X}{r_{31}x_W + r_{32}y_W + r_{33}z_W + t_Z}$$

$$v - v_0 = f \frac{r_{21}x_W + r_{22}y_W + r_{23}z_W + t_Y}{r_{31}x_W + r_{32}y_W + r_{33}z_W + t_Z}$$

(3.6)

where $R = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}$ and $T = \begin{pmatrix} t_X \\ t_Y \\ t_Z \end{pmatrix}$ are the rotation matrix and translation

vector components of the rigid transformation T_{WtoV} from world coordinate system to

virtual camera coordinate system. T_{WtoV} can be expressed by combining the

transformations T_{WtoT} from the world to tracker coordinate system and T_{TtoV} from

the tracker to the virtual camera coordinate system as $T_{WtoV} = T_{TtoV}T_{WtoT}$.

The equations relating the 3-D world coordinates to the 2-D image coordinates given

in (3.6) are the same for any pinhole camera model without consideration distortion. For

OSTHMD systems, previous research [96] shows that it is necessary to compensate the

radial distortion introduced by the optics. Thus, the radial distortion can be modeled by

31

applying a first-order factor k_1 to the above distortion-free calibration model as:

$$u_d\left(1+k_1\left(u_d^{\,2}+v_d^{\,2}\right)\right)=u$$
$$v_d\left(1+k_1\left(u_d^{\,2}+v_d^{\,2}\right)\right)=v$$

(3.7)

where $\left(u_d,v_d\right)$ and $\left(u,v\right)$ are the distorted (observed) image coordinates and undistorted (ideal) image coordinates, respectively.

The complete OSTHMD virtual camera model which relates a point $P_W\left(x_W,y_W,z_W\right)$ in the world coordinate system to its corresponding distorted image point $p_d\left(u_d,v_d\right)$ on the virtual image plane can be given by combining (3.6) and (3.7) as:

$$u_d\left(1+k_1\left(u_d^{\,2}+v_d^{\,2}\right)\right)=f\frac{r_{11}x_W+r_{12}y_W+r_{13}z_W+t_X}{r_{31}x_W+r_{32}y_W+r_{33}z_W+t_Z}$$
$$v_d\left(1+k_1\left(u_d^{\,2}+v_d^{\,2}\right)\right)=f\frac{r_{21}x_W+r_{22}y_W+r_{23}z_W+t_Y}{r_{31}x_W+r_{32}y_W+r_{33}z_W+t_Z}$$

(3.8)

Note that the curved virtual image plane denoted by (3.3) does not affect the modeling and solving of the OSTHMD virtual camera unless the center of projection (i.e. the user's eye) is significantly off the axis of the parabolic surface. When the user's eye is on the axis of the parabolic surface, the curved virtual image plane does not change the positions of the points on the virtual image plane perceived by the user. In such a case, the user cannot notice the curved image plane. This is because that the OSTHMD user is actually observing a virtual image plane with unknown depth and pixel distance. The points on the virtual image plane are determined by the direction of rays while the depth of the points does not matter to the user's view. The modeling of a flat virtual image plane is a

valid assumption for the convenience of handling and solving the special optical design of the OSTHMD imaging system. In this paper, the user's eye is considered very close to the axis of the parabolic image surface such that any further distortion introduced by off-axis viewing can be ignored. The sensitivity analysis of the effects of the user's eye position or the off-axis viewing on the registration accuracy is given in Section 6.2.

Chapter 4   Display Relative Calibration Phase I

A two-phase method called Display Relative Calibration (DRC) is proposed to address the challenges of OSTHMD calibration. Instead of solving the OSTHMD virtual camera model directly, the DRC method obtains a close estimation of the virtual camera without user interaction in the first phase and then seeks to refine the results through user procedure in the second phase. Since most parameters will be solved in the first phase, the user procedure can be greatly simplified and the calibration results can be more reliable than those obtained by the existing calibration approaches.

This chapter describes the steps and computations of the DRC phase I, the offline procedure which uses a calibrated camera as a substitute for the user's eye to estimate the parameters of the OSTHMD virtual camera. Although the virtual camera composed of the OSTHMD optics and the camera is different from the one including the user's eye, the phase I calibration results are very close to the actual user calibration data. For OSTHMD-based AR applications with low-accuracy requirements, the DRC phase I results can be directly applied to obtain acceptable registration. The DRC phase I also determines the characteristics of the OSTHMD virtual image plane.

4.1   System Setup

Figure 4.1 shows the system setup used in the DRC phase I. This configuration improves the initial design of the DRC method by eliminating the specially designed calibration jig [97]. The OSTHMD is fixed on a tripod during the whole procedure. A

non-coplanar calibration jig printed with sufficient fiducials is placed in front of the OSTHMD at a distance approximately equal to the OSTHMD focal length (or the distance of the virtual image plane) listed in the specification. The display and the calibration jig must remain fixed relative to each other during this process. A color gel filter can be used to cover the eye pieces of the OSTHMD for the single-image calibration described later in section 4.2.4. A spotlight can be applied optionally to give more lighting on the calibration jig when necessary.
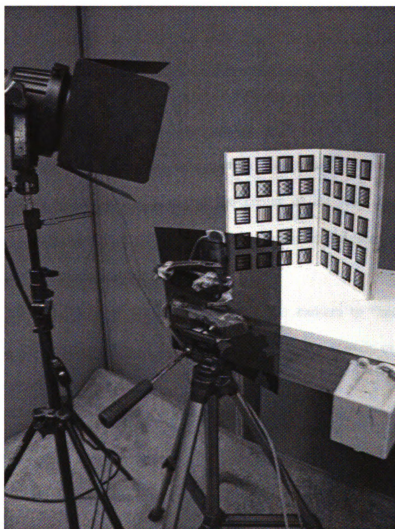
Figure 4.1 DRC phase I system setup

4.2   OSTHMD Virtual Camera Calibration

To calibrate the virtual camera model composed of a camera and the OSTHMD optics, the camera needs to be calibrated at first. Since the projection center of the camera coincides with that of the OSTHMD virtual camera, the calibration becomes a restricted problem which can be solved from a picture of the virtual camera image taken by the calibrated camera.

4.2.1   Operational Steps

The virtual camera calibration includes the following steps:

1)   Place the camera at a position close to one of the OSTHMD exit pupils where the user's eye is supposed to be. Adjust the camera's position and orientation to make the observed OSTHMD virtual image plane at the center of its view and the viewing direction of the camera approximately perpendicular to the virtual image plane. Fix the camera in a tripod when adjustments are done.

2)   Turn off the power of the OSTHMD. Use the camera to "see-through" the OSTHMD lens and take a picture of the calibration jig. This picture is used to calibrate the camera as described in Section 4.2.2.

3)   Keep the camera's position and orientation unchanged. Turn on the OSTHMD and display a full-screen image with same width and height dimensions of the OSTHMD screen. Figure 4.2 shows a sample image with known corner coordinates for each fiducial pattern are known. Use the camera to take a picture of the OSTHMD virtual image plane. The calibration jig can be covered using a black cloth or paper to avoid interference of

two kinds of fiducials. The picture of the OSTHMD virtual image plane is used for the virtual camera calibration described in Section 4.2.3.
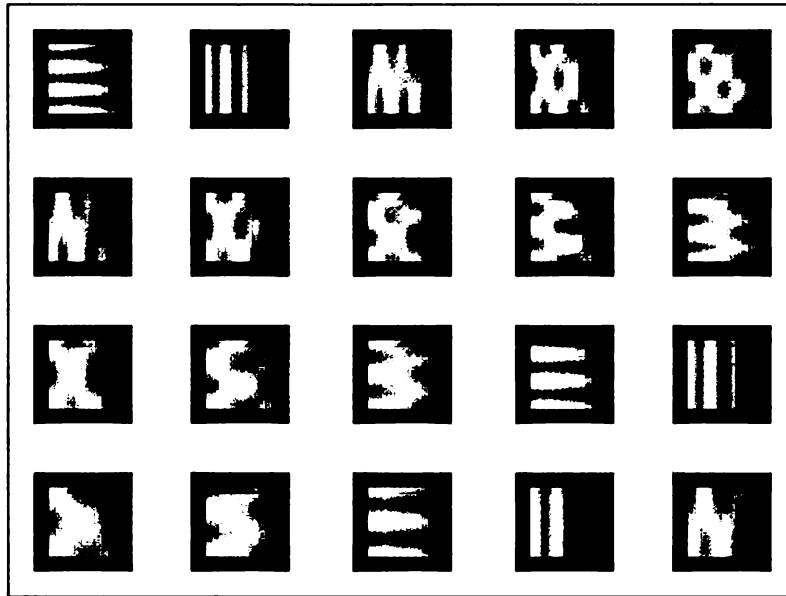
Figure 4.2 Fiducial image displayed on the OSTHMD for virtual camera calibration

    4)   Repeat the above steps for another side of the OSTHMD eyepieces. The left and right eyepiece of the OSTHMD need to be calibrated separately.

Note that instead of taking two pictures, the calibrations for the camera and the virtual camera can be done from a single image of the calibration jig overlaid with the virtual image plane. To separate the two images, the colors of the OSTHMD virtual image and the calibration jig should be different such that a color-based image segmentation algorithm can be applied. Section 4.2.4 describes the use of color gel filter to achieve this goal. The advantage of this approach is to avoid any errors introduced by the change of the camera's position or orientation between taking the two pictures.

4.2.2 Camera Calibration

The camera in place of the user's eye can be calibrated from a picture of the calibration grid using Tsai's calibration algorithm [98, 99]. Given sufficient points on the calibration grid with known coordinates in the world coordinate system and the corresponding image points, the intrinsic and extrinsic parameters of the camera can be solved, including focal length f_C, image center (u_{C0}, v_{C0}), aspect ratio s_{Cx}, radial distortion factor k_{C1}, and the rigid transformations from the world coordinate system to the digital camera coordinate system T_{WtoC} which is composed of a rotation matrix R_{WtoC} and a translation vector t_{WtoC}.

This step determines the center of projection of the OSTHMD virtual camera model because it coincides with the center of projection of the digital camera. Meanwhile, the calibrated camera acts as a direction sensor which is used to measure the direction of rays from the center of projection to the image points in the OSTHMD virtual camera calibration.

The relationship between an image point pc $p_C = (u_{Cd}, v_{Cd})$ and the direction of the ray $\left(\dfrac{X_C}{Z_C}, \dfrac{Y_C}{Z_C}\right)$ from the center of projection to the corresponding point $P_C = (X_C, Y_C, Z_C)$ in the camera coordinate system can be given by the calibrated camera as:

$$\frac{u_{Cd}\left(1+k_1\left(u_{Cd}{}^2+v_{Cd}{}^2\right)\right)-u_{C0}}{s_{Cx}f_c}=\frac{X_c}{Z_c}$$

$$\frac{v_{Cd}\left(1+k_1\left(u_{Cd}{}^2+v_{Cd}{}^2\right)\right)-v_{C0}}{f_C}=\frac{Y_c}{Z_c} \tag{4.1}$$

where P_C is transformed from the corresponding point P_W in the world coordinate

system by $P_C{}^T = R_{WtoC}P_W{}^T + t_{WtoC}$.

4.2.3   Virtual Camera Calibration

Once the camera is calibrated, the OSTHMD virtual camera composed of the camera

and the OSTHMD imaging system can be solved. When the center of projection is known,

the OSTHMD virtual camera calibration becomes a restricted problem which can be

solved based on a single picture of the virtual image plane taken by the calibrated camera.

The calibrated camera has twofold functions as a direction sensor for calibrating the

virtual camera as well as a component of the virtual camera. The intrinsic and extrinsic

parameters of the digital camera must keep unchanged to keep the center of projection the

same as the one solved in the camera calibration step.
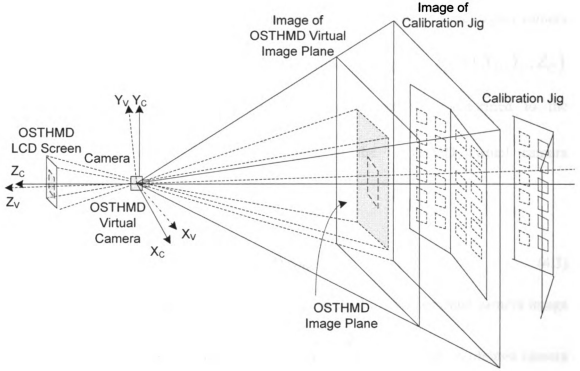
Figure 4.3 DRC phase I calibration

To proceed, display a 2-D planar fiducial pattern with sufficient known calibration points on the OSTHMD virtual image plane and take a picture of it using the calibrated camera as shown in Figure 4.3.

An image point $p_V = (u_V, v_V)$ on the OSTHMD virtual image plane is related to a 3-D point $P_v = (X_V, Y_V, Z_V)$ in the virtual camera coordinate system by:

$$u_V - u_{V0} = s_{xV} f_V \frac{X_V}{Z_V}$$

$$v_V - v_{V0} = f_V \frac{Y_V}{Z_V}$$

(4.2)

As the center of projection of the OSTHMD virtual camera coincides with that of the

calibrated camera, the virtual camera coordinate system is related to the digital camera

coordinate system by a rotation matrix R_{CtoV}. Thus a 3-D point $P_C = (X_C, Y_C, Z_C)$

defined in the calibrated camera coordinate system can be transformed to the

corresponding 3-D point $P_v = (X_V, Y_V, Z_V)$ in the OSTHMD virtual camera

coordinate system by:

$$P_V = R_{CtoV} P_C \qquad (4.3)$$

Based on the OSTHMD virtual camera computational model, a virtual camera image

point p_V can be related to the direction of the ray $\left(\dfrac{X_C}{Z_C}, \dfrac{Y_C}{Z_C} \right)$ in the calibrated camera

coordinate system by:

$$u_V - u_{V0} = s_{xV} f_V \frac{r_{11} X_C/Z_C + r_{12} Y_C/Z_C + r_{13}}{r_{31} X_C/Z_C + r_{32} Y_C/Z_C + r_{33}}$$

$$v_V - v_{V0} = f_V \frac{r_{21} X_C/Z_C + r_{22} Y_C/Z_C + r_{33}}{r_{31} X_C/Z_C + r_{32} Y_C/Z_C + r_{33}} \qquad (4.4)$$

where the $\left(\dfrac{X_C}{Z_C}, \dfrac{Y_C}{Z_C} \right)$ associated with each image point p_c on the camera picture

taken can be determined by (4.1).

Each pair of the point correspondences produce the following two linear equations

derived from (4.4):

$$\left(u_V - u_{V0}\right)r_{31}\frac{X_C}{Z_C} + \left(u_V - u_{V0}\right)r_{32}\frac{Y_C}{Z_C} + \left(u_V - u_{V0}\right)r_{33}$$

$$- s_{xV}f_V r_{11}\frac{X_C}{Z_C} - s_{xV}f_V r_{12}\frac{Y_C}{Z_C} - s_{xV}f_V r_{13} = 0$$

$$\left(v_V - v_{V0}\right)r_{31}\frac{X_C}{Z_C} + \left(v_V - v_{V0}\right)r_{32}\frac{Y_C}{Z_C} + \left(v_V - v_{V0}\right)r_{33} \qquad (4.5)$$

$$- s_{xV}f_V r_{21}\frac{X_C}{Z_C} - s_{xV}f_V r_{22}\frac{Y_C}{Z_C} - s_{xV}f_V r_{23} = 0$$

Given four or more image points and their corresponding points on the picture of OSTHMD virtual image plane, the intrinsic parameters $u_{V0}, v_{V0}, s_{Vx}, f_V$ and rotation matrix R_{CtoV} of the OSTHMD virtual camera model can be estimated by solving the following over-determined equations by linearizing (4.5) as described in [100-102]:

$$AW = 0 \qquad (4.6)$$

where

$$A = \begin{pmatrix} \dfrac{X_{Ci}}{Z_{Ci}} & \dfrac{Y_{Ci}}{Z_{Ci}} & 1 & 0 & 0 & 0 & -u_{Vi}\dfrac{X_{Ci}}{Z_{Ci}} & -u_{Vi}\dfrac{Y_{Ci}}{Z_{Ci}} & -u_{Vi} \\ 0 & 0 & 0 & \dfrac{X_{Ci}}{Z_{Ci}} & \dfrac{Y_{Ci}}{Z_{Ci}} & 1 & -v_{Vi}\dfrac{X_{Ci}}{Z_{Ci}} & -v_{Vi}\dfrac{Y_{Ci}}{Z_{Ci}} & -v_{Vi} \end{pmatrix} \qquad (4.7)$$

and

$$W = \begin{pmatrix} W_1 \\ W_2 \\ W_3 \end{pmatrix} = \begin{pmatrix} s_{Vx}f_V R_1 + u_{Vi}R_3 \\ f_V R_2 + v_{Vi}R_3 \\ R_3 \end{pmatrix}. \qquad (4.8)$$

The column vectors R_1, R_2 and R_3 in (4.8) correspond to the three rows of matrix

R_{CtoV} .

To solve W in the above linear homogeneous equation (4.6), a temporary constraint $\|W\| = 1$ can be applied. The solution W' is well known to be the vector corresponding to the smallest singular value of the SVD of A or equivalently the eigenvector of $A^T A$ with the least eigenvalue. Note that the image and model points data are required to be normalized before solving (4.6) to increase numerical stability [103].

Once W' is computed, W can be obtained by applying $W = \dfrac{W'}{\|W'_3\|}$ because W_3 is the last row of the rotation matrix R_{CtoV} with $\|W_3\| = 1$. Now the parameters of the OSTHMD virtual camera model can be initially estimated by the following equations:

$$
\begin{aligned}
\overline{u_{V0}} &= W_1^T W_3 \\
\overline{v_{V0}} &= W_2^T W_3 \\
\overline{f_V} &= \|W_2 - v_{V0} W_3\| \\
\overline{s_{xV}} &= \frac{\|W_1 - u_{V0} W_3\|}{f_V} \\
\overline{R_1} &= \frac{W_1 - u_{V0} W_3}{s_{xV} f_V} \\
\overline{R_2} &= \frac{W_2 - v_{V0} W_3}{f_V} \\
\overline{R_3} &= W_3
\end{aligned}
\tag{4.9}
$$

The estimated $\overline{R}$ can be further improved by applying orthonormality requirement $R^T R = 1$. A closed-form solution is given in [100, 104] for computing an orthonormal

R which minimizes $\left\| \overline{R} - R \right\|$. The parameters $u_{V0}, v_{V0}, s_{Vx}, f_V$ can be accordingly improved by replacing W_3 with R_3 in (4.9).

Note that the solution to (4.6) is optimum in a least-squared sense, but the optimality is not defined by the re-projection error, but rather by the least squared fitting of minimizing AW. The solution does not include the radial distortion introduced by the OSTHMD optics. Though sub-optimal, this calibration is near to optimum and can be further refined through nonlinear optimization.

To take the radial distortion into account, the virtual camera parameters f, C_X, C_Y, R_{VtoC} including the radial distortion factor k_1 can be obtained by a nonlinear optimization which minimizes the following sum of the distances between the detected and re-projected points as:

$$\min \sum_i d \left(p_{Vi}, p_{Vi}{}'\left(k_1, f, C_X, C_Y, R_{VtoC}, T_{WtoC}, P_{Ci} \right) \right) \qquad (4.10)$$

To impose the orthonormality constraint, the rotation matrix R_{VtoC} is parameterized as three Euler angles. Minimizing (4.10) can be solved by a modification of the Levenberg-Marquardt algorithm [105].

Note that the implementation for virtual camera with distortion is different from the non-distorted model. Although OpenGL does not support distorted model directly, the radial distortion can be compensated for OSTHMD rendering. More details of implementing the OSTHMD virtual camera will be given in the Section 6.1.

In practice, the parameters of the OSTHMD virtual camera model can also be directly estimated to obtain an initial solution. The rotation matrix R_{CtoV} from the camera coordinate system to the OSTHMD virtual camera coordinate system can be estimated as an identity matrix. The principal point of the virtual camera model can be initialized at the center of the virtual image plane, and the initial value of the focal length f_V can be estimated from two known points $p_{V1}(u_{V1}, v_{V1}), p_{V2}(u_{V2}, v_{V2})$ on the LCD image plane and their corresponding image points $p_{C1}(u_{C1}, v_{C1}), p_{C2}(u_{C2}, v_{C2})$ on the camera picture based on the following derivation.

The two vectors $\overrightarrow{O_C P_{C1}}, \overrightarrow{O_C P_{C2}}$ which originate from the in the digital camera coordinate system corresponding to p_{C1}, p_{C2} can be calculated from (4.1) as $\left(\dfrac{X_{C1}}{Z_{C1}}, \dfrac{Y_{C1}}{Z_{C1}}, 1\right)$ and $\left(\dfrac{X_{C2}}{Z_{C2}}, \dfrac{Y_{C2}}{Z_{C2}}, 1\right)$ by setting $Z_C = 1$. The angle between $\overrightarrow{O_C P_{C1}}$ and $\overrightarrow{O_C P_{C2}}$ can be given by:

$$\theta = \arccos\left(\frac{\overrightarrow{OP_{C1}} \cdot \overrightarrow{OP_{C2}}}{\left\|\overrightarrow{OP_{C1}}\right\|\left\|\overrightarrow{OP_{C2}}\right\|}\right) \tag{4.11}$$

where $\overrightarrow{OP_{C1}} \cdot \overrightarrow{OP_{C2}}$ is the dot product of $\overrightarrow{O_C P_{C1}}$ and $\overrightarrow{O_C P_{C2}}$, and $\left|\overrightarrow{OP_{C2}}\right|$, $\left|\overrightarrow{OP_{C2}}\right|$ are the norm of $\overrightarrow{O_C P_{C1}}$ and $\overrightarrow{O_C P_{C2}}$, respectively.

The focal length f_V is defined as the length from the origin of the virtual camera

coordinate system to the virtual image plane. Since f_V is defined in pixels, it can be computed as:

$$f_V = \frac{|p_{V1} - p_{V2}|}{2\tan(\theta/2)} \qquad (4.12)$$

where $|p_{V1} - p_{V2}|$ is the distance between points p_{V1}, p_{V2} and θ is the angle between their corresponding vectors given by (4.12).

Under the DRC phase I system setup, the above initial solution is very close to the one computed by using the linear procedure (4.5) - (4.9), and the final results after nonlinear optimization procedure given in (4.10) are comparable to the first approach as well. Since this estimation method is easier to implement than the linear solution, it can be served as a simple and fast approach to obtain effective initial solution for OSTHMD virtual camera.

4.2.4   Calibration using a Single Camera Picture

The above calibrations of the camera and the OSTHMD virtual camera can be solved from a single camera picture where the OSTHMD virtual image plane is overlaid on the calibration jig. This can be achieved by separating the two sub-pictures using color-based image segmentation.

To allow color-based image segmentation, a red color gel filter is covered on the OSTHMD eye piece and the OSTHMD virtual image plane is rendered with blue background. When the colors and the lighting are properly configured, the image

segmentation can be done based on simple RGB values. For example, a typical configuration is a picture with the blue virtual image plane overlaid on the red calibration jig. Figure 4.4 and Figure 4.5 present the separated pictures of the calibration jig and the OSTHMD virtual image plane, respectively.
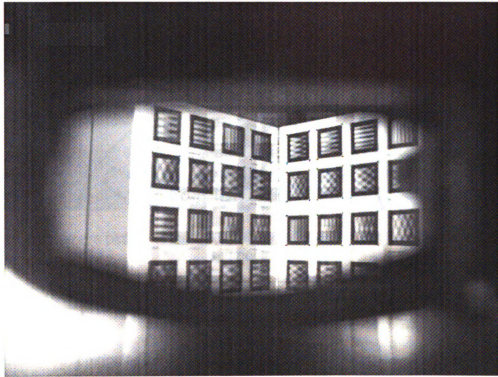
Figure 4.4 Camera picture of the calibration jig obtained using image segmentation
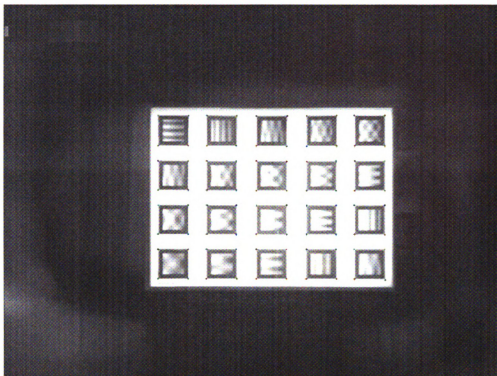
Figure 4.5 Camera picture of the OSTHMD virtual image plane obtained using image segmentation

4.3   Calibration of OSTHMD Virtual Image Plane

In the above solution to the OSTHMD virtual camera model, the effective focal length f_v is calculated in pixels as a relative value. The actual distance from the center of projection of the OSTHMD virtual camera model to the virtual image plane is unknown. Although this distance is inessential to rendering or registration, knowledge of this distance is crucial to the DRC phase II which updates the virtual camera parameters based on the displacement of the center of projection from the calibrated camera optical center to the user's eye. Meanwhile, the distance from the virtual plane to the projection center is essential to the estimation of registration error. As shown in Figure 4.6, a closer distance from the eye to the OSTHMD image plane leads to a larger registration error.
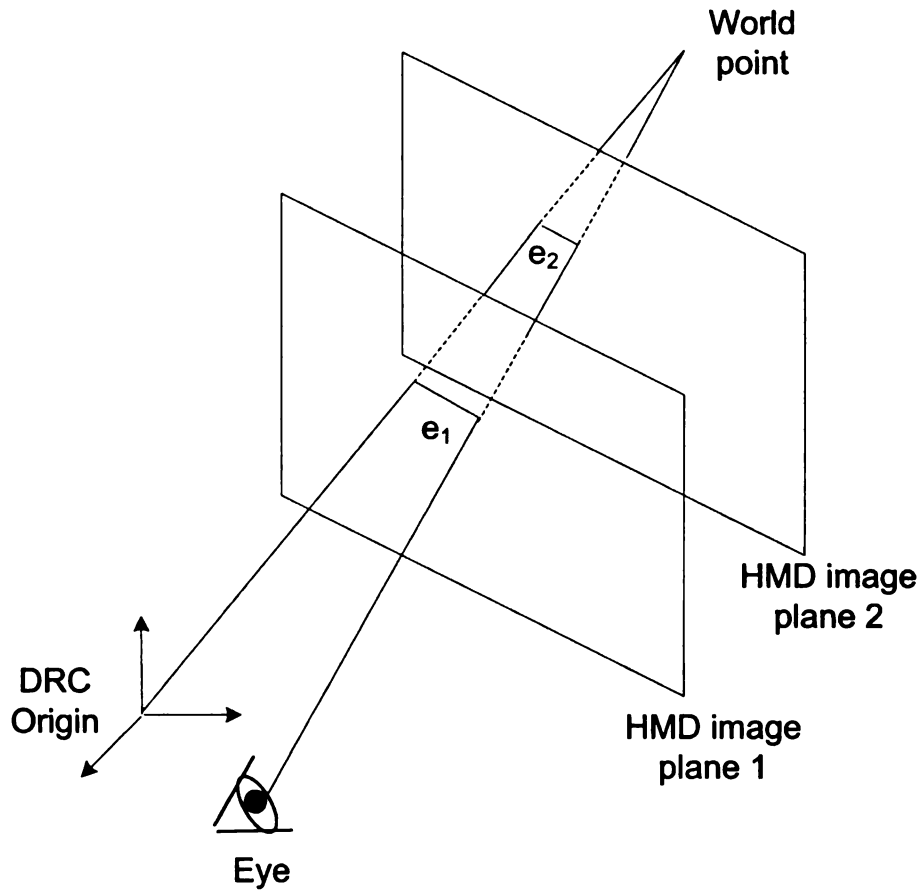
Figure 4.6 Effect of the OSTHMD image plane depth on the registration error

4.3.1 Operational Steps

The calibration of OSTHMD virtual image plane follows right after the virtual camera calibration, including the following steps:

1) Slightly change the location and orientation of the camera in the previous step of virtual camera calibration to take another set of pictures (or a single picture when image segmentation is used) of the calibration pattern and the OSTHMD virtual image plane. Repeat this procedure to take several more sets of pictures.

2) Use multiple viewing position calibration of Tsai's algorithm [99] to calibrate the camera based on the pictures of the calibration pattern taken in different viewing position.

The solved extrinsic parameters of the camera are different from view to view, but the intrinsic parameters remain the same.

3) Compute the fundamental matrix F for the pairwise set of the same calibrated camera matrices with different viewing positions [103, 106].

4) Apply Hartley's algorithm [107] to correct the image points subject to the epipolar constraint.

5) Use the mid-point method [106] to compute the 3-D position of the points on the OSTHMD virtual image plane.

6) Fit the curved OSTHMD image plane equation (3.3) using least-square method based on the computed virtual image plane points.

4.3.2 Fitting Virtual Image Plane

To demonstrate the process of fitting OSTHMD virtual image plane,

Figure 4.7 shows an example of all the computed 3-D points on the virtual image plane obtained from pairwise triangulation of five camera pictures. These points are noisy with varying depths, and they do not form a flat image plane as indicated by the OSTHMD documentation and as intuitively expected.
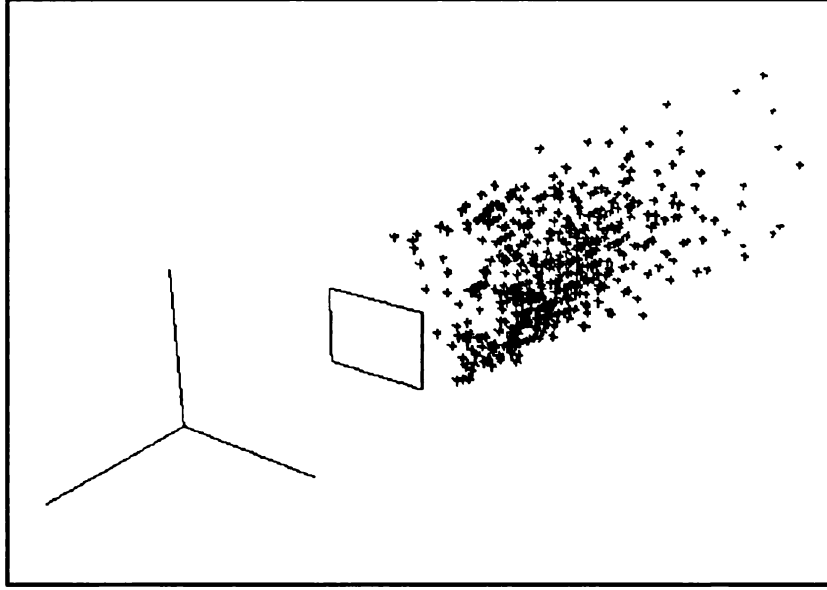
Figure 4.7 Calculated 3-D points of the OSTHMD virtual image plane

Instead, the computed 3-D points of the virtual image plane appear to be distributed on a parabolic surface with aberrations due to noise. Based on the analysis of the OSTHMD optical model presented in Section 3.2, this curved image plane is due to the field curvature [108] of the concave mirror used in the OSTHMD imaging system. The highly curved virtual image plane is an astonishing finding that none of previous research work has mentioned, and validates the deficiency of previous OSTHMD calibration methods based on pure projective 2D-3D point correspondences.

A radially symmetrical parabolic plane is fitted to the virtual image plane by choosing the optimal parameters for the OSTHMD virtual image plane optical model as specified in (3.3). The parameters z_0, K_2, a_x and a_y are computed based on the intersection points data using non-linear least-squares minimization such as the Levenberg-Marquardt method [105]. Figure 4.8 shows an example of a fitted display plane with

$z_0 = -1681mm$, $a_x = 1.56$, $a_y = 0.44$, and $K_2 = 19.5$. Although the plane

looks quite curved, the range of the eye locations within the DRC is comparatively very small. This optical aberration appears to the user as a slight pincushion distortion when the eye moves away from the apex. This pincushion distortion has been well known for these displays, but was previously assumed to be due to errors in the design of the optics. The choice of a spherical mirror in the design of these displays does introduce this large spherical curvature, but also allows for variation in the eye location. A parabolic mirror could be designed that would cancel the curvature, but only for a single viewpoint. Hence, the display would be less general in its applicability.
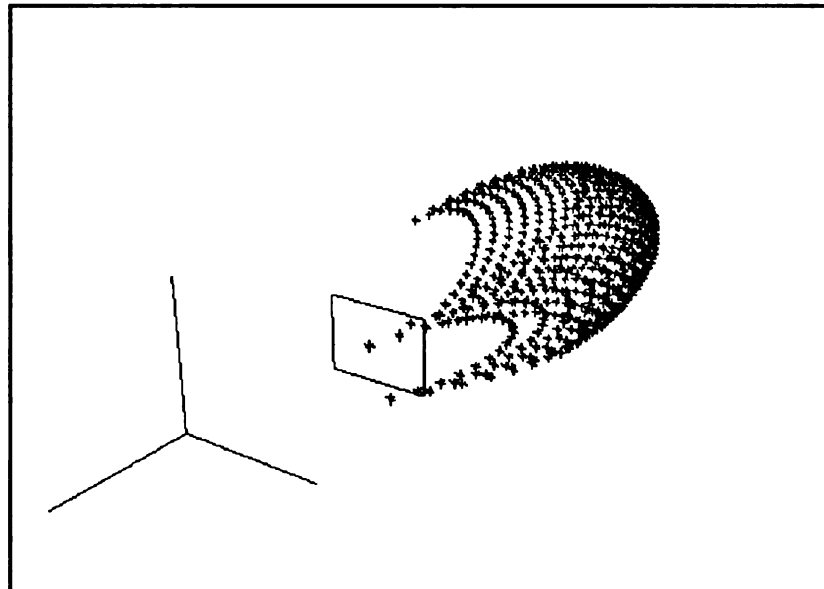
Figure 4.8 Fitted OSTHMD virtual image plane

4.4 Summary

The DRC phase I solves the OSTHMD virtual camera model using a calibrated camera in place of the user's eye using offline algorithm without user interaction. Although the virtual camera parameters solved in DRC Phase I is an approximation of the virtual camera model composed of the user's eye, most of the computed parameters do

not significantly depend on the user's eye. To refine the calibration parameters for a specific user, only those user-dependent parameters need to be improved. The DRC phase I effectively simplifies the user procedure and improves the reliability of the calibration results. Furthermore, the DRC phase I reveals that the calibration of the OSTHMD virtual image plane converges as a curved plane due to the optical design. This is a new finding that has been ignored in previous research and it is important to take it into account for accurate registration.

Chapter 5   Display Relative Calibration Phase II

This chapter describes the computations and user procedures of the DRC phase II, which seeks to improve the phase I calibration results by adjusting the user-dependent parameters for a specific user. The phase I calibration solves an OSTHMD virtual camera model relative to the center of position and viewing direction of a calibrated camera. The use of the calibrated camera simplifies the calibration procedure and produces reliable results without any user interaction. However, the OSTHMD virtual camera model composed of the user's eye is different from the one calibrated by the digital camera in the phase I. The solved phase I parameters need to be adjusted in the phase II to achieve registration for a specific user. Therefore, the DRC phase II needs to solve the new virtual camera composed of the user's eye and the OSTHMD imaging system to achieve registration for the user wearing the OSTHMD. Instead of recalculating all the parameters from scratch, the calibration of the changed virtual camera can be estimated from the phase I results. To make the DRC method applicable to different applications and users, the phase II includes multiple user options with different accuracy and user efforts.

5.1   User-dependent Parameters

The main difference between phase I and phase II of the DRC calibration method is that the user's eyes replace the camera such that a new imaging system is created together with the OSTHMD imaging system. Figure 5.1 illustrates the coordinate systems and viewing transformations of the DRC phase I and phase II virtual camera models. The

major change of phase II virtual camera model is the displacement of the center of projection from the digital camera projective center to the user's eye position. Although both the intrinsic and extrinsic parameters of the OSTHMD virtual camera are affected by the shift of center of projection, only the parameters significantly changed from the phase I need to be recalculated.
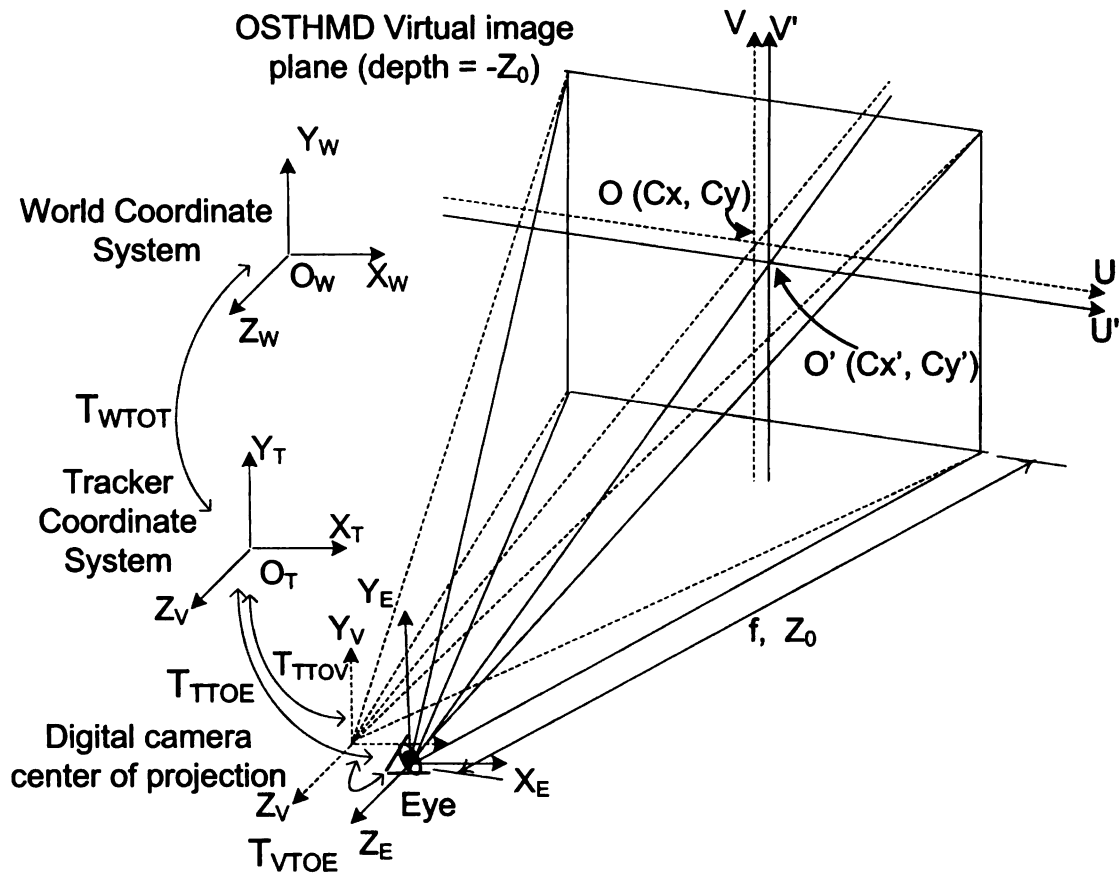
Figure 5.1 Coordinate systems and transformations of the DRC phase I and II

In the DRC phase II, the origin of the virtual camera coordinate system coincides with the user's eye position. Same as the definition in the phase I, the x, y axes of the phase II virtual camera coordinate system are parallel to the horizontal and vertical side of the OSTHMD virtual image plane, respectively, and the z axis is perpendicular to the virtual

55

image plane. The relationship between a point (x_V, y_V, z_V) defined in the phase I virtual camera coordinate system and the same point (x_E, y_E, z_E) can be expressed as:

$$
\begin{pmatrix} x_E \\ y_E \\ z_E \end{pmatrix} = R_{VtoE} \begin{pmatrix} x_V \\ y_V \\ z_V \end{pmatrix} + T_{VtoE} \tag{5.1}
$$

where R_{VtoE} is the rotation matrix defining the orientation change of the phase II virtual camera coordinate system relative to that of the phase I, and $T_{VtoE} = (t_X, t_Y, t_Z)^T$ is the translation vector defining the displacement of the virtual camera center of projection from the phase I to the phase II.

Based on the definition of the OSTHMD virtual camera model, the virtual camera coordinate systems in the phase I and the phase II have the same orientation relative to the OSTHMD virtual image plane. When the OSTHMD is properly worn on the user's head, the user's viewing direction should be perpendicular to the virtual image plane as well. To pursue a fast and effective calibration for the phase II, we can first assume that the orientation of the phase II virtual camera coordinate system is the same as that of the phase I such that R_{VtoE} can be set as a 3×3 identity matrix. Therefore, we have:

$$
\begin{pmatrix} x_E \\ y_E \\ z_E \end{pmatrix} = \begin{pmatrix} x_V \\ y_V \\ z_V \end{pmatrix} + \begin{pmatrix} t_X \\ t_Y \\ t_Z \end{pmatrix} \tag{5.2}
$$
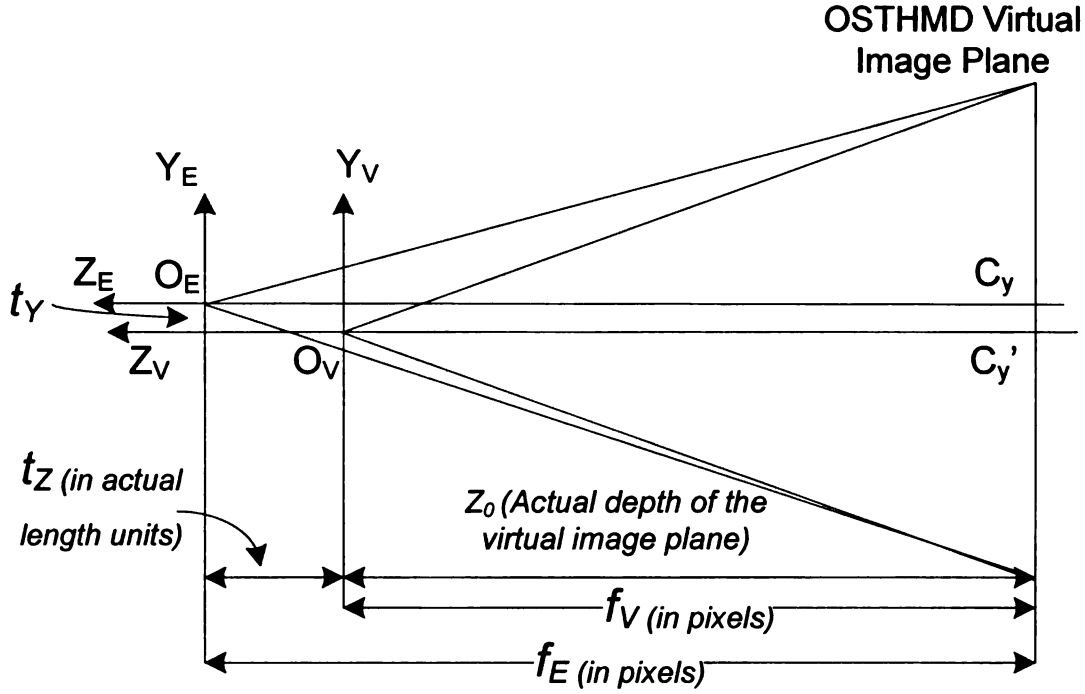
Figure 5.2 Computation of the updated effective focal length and image center

As illustrated in Figure 5.2, the displacement of virtual camera center of projection in the z direction introduces change of the effective focal length f, which is defined as the distance from the center of projection to the virtual image plane. Note that the translation vector $T_{VtoE} = (t_X, t_Y, t_Z)^T$ is measured in actual length units while f is computed in pixels as a relative value to accommodate the uncertain pixel distance of the OSTHMD virtual image plane. Based on the virtual image plane depth z_0 solved in the phase I calibration, the updated focal length f can be estimated by adjusting the z-direction offset t_Z scaled by z_0 as:

$$f_E = \frac{z_0 - t_Z}{z_0} f_V \tag{5.3}$$

The image center (C_{EX}, C_{EY}) of the phase II virtual camera can be computed

from the phase I virtual camera parameters as:

$$C_{EX} = C_{VX} - \frac{t_X}{z_0} f_V s_x$$

$$C_{EY} = C_{VY} - \frac{t_Y}{z_0} f_V$$

(5.4)

Disregarding distortion, a visible point (x_E, y_E, z_E) defined in the phase II virtual

camera coordinate system is related to its correspondent registered image point

(u_E, v_E) by:

$$s_{Ex} f_E \frac{x_E}{z_E} = u_E - C_{Ex}$$

$$f_E \frac{y_E}{z_E} = v_E - C_{Ey}$$

(5.5)

Combining (5.2) - (5.4) and assuming that the aspect ratio s_{Ex} is the same as s_{Vx},

we have:

$$s_{Vx} f_V \left(\frac{z_0 - t_z}{z_0} \right) \left(\frac{x_V + t_X}{z_V + t_z} \right) = u_E - \left(C_{Vx} - \frac{t_X s_x f_V}{z_0} \right)$$

$$f_V \left(\frac{z_0 - t_z}{z_0} \right) \left(\frac{y_V + t_Y}{z_V + t_z} \right) = v_E - \left(C_{Vy} - \frac{t_Y f_V}{z_0} \right)$$

(5.6)

where (x_V, y_V, z_V) is obtained by applying the rigid transformation T_{WtoV} to a point

(x_W, y_W, z_W) defined in the world coordinate system. T_{WtoV} is solved separately as the

tracker calibration as described in [22].

5.2   Options for Phase II calibration

The goal of the phase II calibration is to solve the user-dependent virtual camera model based on the phase I calibration results. The changed parameters from the phase I to the phase II include the focal length f_E, the translation vector T_{VtoE} defining the positional displacement of the center of projection, and the image center $\left(C_{Ex}, C_{Ey} \right)$.

The translation vector T_{VtoE} is the determinative factor from which the other parameters can be solved. These parameters are highly user-dependent due to the different eye position relative to the virtual image plane. To achieve precise registration for a specific user, certain user procedure is necessary to indicate the registration of the world points and the corresponding image points observed by the user.

Instead of providing a single fixed user procedure, multiple user procedures are provided in the phase II to adapt for different user efforts and registration accuracy requirements of various OSTHMD-based applications. The DRC phase II user procedures options include: 1) Use the phase I calibration directly; 2) Single-point alignment; 3) Manual adjustment and 4) Multiple-point alignment.

5.2.1   Use the DRC Phase I Calibration Directly

Although this option cannot reach accurate registration error without updating the center of projection displacement, it would be useful for applications such as OSTHMD-based AR games which need rough registration without any user procedure. However, it is necessary to understand how the eye position affects the registration

accuracy. A sensitivity analysis of the eye position on the registration errors will be presented in Section 6.2.

5.2.2 Single-Point Alignment

The center of projection of the phase II virtual camera is located at the user's eye position, which is very close to the center of projection of the digital camera used in the phase I calibration. The z-direction offset t_Z between the two projection centers is usually far less than the depth of the OSTHMD virtual image plane z_0 or the z-component of the calibration point z_V such that there are $z_0 \gg t_Z$ and $z_V \gg t_Z$. For example, the Sony Glasstron LDI-100 OSTHMD used in the user experiment has an estimated $z_0 = 1681mm$, and the z-offset of $5mm$ results in a focal length modification of less than 0.3%. Therefore, without significant loss of accuracy, we can assume $f_E = f_V$ and $z_E = z_V$ to simply (5.6) as:

$$
\begin{aligned}
s_{Vx} f_V \frac{x_V + t_X}{z_V} &= u_E - \left(C_{Vx} - \frac{t_X s_{Vx} f_V}{z_0} \right) \\
f_V \frac{y_V + t_Y}{z_V} &= v_E - \left(C_{Vy} - \frac{t_Y f_V}{z_0} \right)
\end{aligned}
\tag{5.7}
$$

Given a world point and its aligned image point on the OSTHMD virtual display observed by the user, the x and y components of the translation vector T_{VtoE} can be calculated by:

$$t_x = \frac{C_{Vx} - u_E}{\left(\dfrac{1}{z_V} + \dfrac{1}{z_0}\right) f_V s_{Vx}} - \frac{x_V}{z_V}$$

$$t_y = \frac{C_{Vy} - v_E}{\left(\dfrac{1}{z_V} + \dfrac{1}{z_0}\right) f_V} - \frac{y_V}{z_V} \qquad (5.8)$$

With the solved t_X, t_Y and the assumption $t_Z = 0$, the transformation T_{TtoE} from the tracker coordinate system to the phase II virtual camera coordinate system can be obtained by:

$$T_{TtoE} = T_{TtoV} T_{VtoE} \qquad (5.9)$$

where $T_{VtoE} = \left[R_{VtoE} \mid t_{VtoE} \right] = \left[I \mid t_{VtoE} \right]$.

As for the intrinsic parameters of the phase II virtual camera model, the focal length, and the aspect ratio can be estimated as the same values given in the phase I calibration, and the new image center can be estimated by (5.4).

The procedure of the single-point alignment option is similar to that of the SPAAM [88, 89], but the user only needs to do the alignment once for each eye. Usually the user indicates the alignment by clicking the mouse button when a predefined 2-D marker on the virtual display overlays on a known world point.

5.2.3   Manual Adjustment

One of the major advantages of the DRC method is that the degrees of freedom for the Phase II calibration are reduced to three for each eye. Meanwhile, the unknown parameters to be solved are independent of each other. Hence, a simple manual control

can be supplied that the user uses to adjust the registration directly. X and Y adjustments can be used to align edges and a Z adjustment used to correct for focal length. A simple mouse control has been implemented for the manual adjustment option. Shifting the mouse in the x and y direction directly modifies the X and Y location of the center of projection. A wheel-based mouse can be utilized for direct manipulation of the focal length using the scroll wheel. In practice, a simple rectangle pattern is displayed on the OSTHMD virtual image plane and the mouse is used to manually adjust the registration of that pattern and a real-world object. The advantage of this option is that it is easy to tweak the calibration if the display shifts on the head.

As shown in Figure 5.3, the tracker fiducial outlines are used as the alignment pattern for the manual adjustment option. The user uses the mouse to modify the observed fiducial outlines until they overlay onto the actual fiducials. Then the user clicks the mouse to indicate the completion of adjustment. The OSTHMD projection center position in x, y direction $\left(X_x, X_y\right)$ can be updated based on the lateral offsets of the fiducial outlines and the image center $\left(C_x, C_y\right)$ can be calculated using (5.4). The effective focal length f can be updated relative to the original value based on the movement of the mouse wheel.
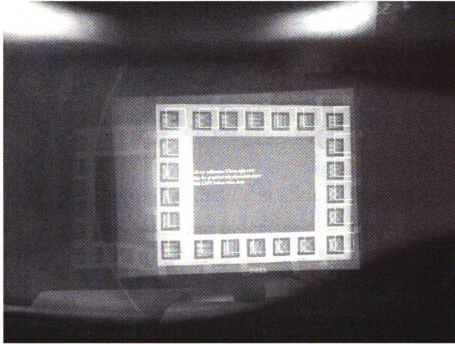
Figure 5.3 DRC phase II manual adjustment option

5.2.4   Multiple-Point Alignment

Given two or more image and world point correspondences, the focal length f_E and aspect ratio s_{Ex} ignored in the single-point alignment option can be solved.

From (5.8) there is:

$$
\frac{C_{Vx} - u_{E1}}{\left(\dfrac{1}{z_{V1}} + \dfrac{1}{z_0}\right) f_E s_{Ex}} - \frac{x_{V1}}{z_{V1}} = \frac{C_{Vx} - u_{E2}}{\left(\dfrac{1}{z_{V2}} + \dfrac{1}{z_0}\right) f_E s_{Ex}} - \frac{x_{V2}}{z_{V2}}
$$
$$
\frac{C_{Vy} - v_{E1}}{\left(\dfrac{1}{z_{V1}} + \dfrac{1}{z_0}\right) f_E} - \frac{y_{V1}}{z_{V1}} = \frac{C_{Vy} - v_{E2}}{\left(\dfrac{1}{z_{V2}} + \dfrac{1}{z_0}\right) f_E} - \frac{y_{V2}}{z_{V2}}
$$

(5.10)

Therefore, f_E and s_{Ex} can be solved by:

$$f_E = \frac{\dfrac{C_{Vx}-v_{E1}}{\left(\dfrac{1}{z_{V1}}+\dfrac{1}{z_0}\right)} - \dfrac{C_{Vx}-v_{E2}}{\left(\dfrac{1}{z_{V2}}+\dfrac{1}{z_0}\right)}}{\dfrac{y_{V1}}{\left(\dfrac{1}{z_{V1}}+\dfrac{1}{z_0}\right)z_{V1}} - \dfrac{y_{V2}}{\left(\dfrac{1}{z_{V2}}+\dfrac{1}{z_0}\right)z_{V2}}} \tag{5.11}$$

and

$$s_{Ex} = \frac{\dfrac{C_{Vx}-u_{E1}}{\left(\dfrac{1}{z_{V1}}+\dfrac{1}{z_0}\right)} - \dfrac{C_{Vx}-u_{E2}}{\left(\dfrac{1}{z_{V2}}+\dfrac{1}{z_0}\right)}}{\dfrac{x_{V1}}{\left(\dfrac{1}{z_{V1}}+\dfrac{1}{z_0}\right)z_{V1}} - \dfrac{x_{V2}}{\left(\dfrac{1}{z_{V2}}+\dfrac{1}{z_0}\right)z_{V2}}} \cdot \frac{1}{f_E} \tag{5.12}$$

The t_X, t_Y can be calculated by replacing f_V, s_{Vx} with f_E, s_{Ex} in (5.8) and the image parameter $\left(C_{Ex}, C_{Ey}\right)$ can be solved similarly by (5.4).

The user procedure for this option is similar as that of SPAAM [88, 89] but fewer alignments are required. The number of alignments can be determined based on the accuracy requirements and the user's performance of aligning virtual and physical points. In the user study described in Chapter 9, four-point alignment is used for this option, which achieves remarkable registration accuracy.

Furthermore, when a large number of point correspondences can be obtained, a nonlinear search can be used to optimize the calibration parameters including the rotation matrix R_{VtoE} and distortion factor k_1 to minimize the average distance of between the measured and expected image points on the virtual display. The initial guess of these parameters can be given by the computation of the above two-point alignment together

with $R_{VtoE} = I$ and $k_1 = 0$. This same strategy is used in the virtual camera calibration of the phase I.

5.3　Summary

Many computer vision algorithms including the Tsai camera calibration [98, 99] have applied the two-phase rough-to-refined approach in the DRC phase II, though usually as a computational method rather than a combination of computational and user interface method presented in the DRC phase II. For the OSTHMD calibration, there is no previous research presenting a successful two-phase algorithm like the DRC method. The design and implementation of the DRC method solve the following two challenges for OSTHMD calibration:

1)　The establishment of a computational model suitable for the two-phase approach

The DRC computational model separates the calibration parameters into system-invariant values and user-dependent variables. The DRC phase I uses a camera to simulate the user's eye for the OSTHMD imaging system and solves for the parameters of the OSTHMD virtual camera system. The phase II refines the user-dependent variables using multiple user options to reach better registration accuracy.

2)　The implementation of flexible interactive user procedures to improve the initial phase I calibration results

Except the DRC-None (Direct use of phase I results), the other three options of DRC phase II provide flexible user interaction to collect data required for improving the phase

I calibration results. The final calibration parameters are optimized for the specific user who completes one of the options.

Furthermore, the DRC phase II is extensible in that other options or processes could be applied as far as this process reduces the degrees of freedom or solves the required parameters. Such a process can still be a user procedure or some kind of measurements. For example, the approach used by Stereo SPAAM [89] could be applied as an option for the DRC phase II, with the important difference that only the eye position need be computed [91]. If the interpupillary distance for the user is already known or measured using pupilometer, the degrees of freedom for a stereo solution can be further decreased by one.

Chapter 6   Implementation and Sensitivity Analysis

This chapter introduces the implementation and the sensitivity analysis of the solved OSTHMD virtual camera model. OpenGL is used to reproduce the OSTHMD imaging system and to generate the expected imagery for AR registration. The implementations of the non-distorted model and the distorted one are different and both approaches are described. The sensitivity analysis discusses the effects of the center of projection on the registration accuracy, including a theoretical analysis based on the non-distorted virtual camera model and an experimental analysis based on calibration data solved with curved virtual image plane.

6.1   Implementation

In OpenGL, an object's 3-D world coordinates are modeled relative to a predefined world coordinate system and then converted to pixel coordinates through a series of viewing, projection and viewport transformations [55-57]. The viewing transformation establishes the position and orientation of the virtual camera and the projection transformation defines a viewing volume for projecting 3-D points into image plane and clipping. The viewport transformation determines the mapping from the normalized image coordinates to actual screen pixels.

Because the 2-D imagery is always rendered in full-screen on the miniature LCD

screen, the viewport transformation for the OSTHMD virtual camera can be simply set up by $glViewport(0,0,w,h)$ where w and h are the pixel numbers in the horizontal and vertical directions of the display given by the OSTHMD specification. The viewing and projection transformations can be implemented by the following code snippet:

```
glMatrixMode(GL _ PROJECTION);
glLoadIdentity();
glFrustum(-znear * cx /(sx * f), znear * (wid - cx)/(sx * f), znear * cy / f, znear * (hit - cy)/ f, znear, zfar);
glMatrix(GL _ MODEVIEW);
glLoadIdentity();
glMultMatrixd(m _ ttoe);
glMultMatrixd(m _ wtot);
```

The *glFrustum*() function defines a perspective projection viewing volume by specifying the coordinates of the lower-left and upper-right corners of the near clipping plane and the distances of the near and far clipping planes. These required parameters can be converted from the solved virtual camera intrinsic parameters. The *znear* and *zfar* are the distances from the center of projection to the near and far clipping planes, respectively. The m_ttoe and m_wtot are the transformation matrices from the tracker to eye coordinate system and from the world to the tracker coordinate system, respectively.

The above implementation does not incorporate the radial distortion or curved image plane. To implement the virtual camera with distortion, a warping transformation needs to be applied to the undistorted image coordinates to take into account the radial distortion. Although OpenGL does not support insertion of a user-defined transformation

68

to its rendering pipeline, indirect implementation such as using texture mapping can achieve the similar results. The basic idea is to render the undistorted 2-D image, then use its pixel data as a texture image whose texture mapping coordinates are established based on the radial distortion function. More details of such approaches can be found in the research work by Watson et al. [109] and Min et al. [110].

6.2   Sensitivity Analysis

The DRC phase I calibration estimates the OSTHMD virtual camera model based on the assumption that the center of projection of the camera is the same as that of the user's eye. In practice, this is unlikely to be the case because the actual center of projection of the OSTHMD virtual camera, i.e. the position of the user's eye relative to the display, always needs to be adjusted for different users. It is necessary to analyze the sensitivity of the estimation error of the center of projection on the registration accuracy. The sensitivity analysis can be used in estimation of the precision afforded by using phase I calibration directly and also the improvement brought about by applying the phase II calibration. Furthermore, the sensitivity analysis reveals the relationships among the registration errors, the eye offsets and the real object offsets. Should an OSTHMD shift on the user's head, the results of the sensitivity analysis can predict the errors that would be expected.

6.2.1 Theoretical Analysis

Assume C' is the center of projection of the camera in DRC phase I and $C = \left(x_e, y_e, z_e \right)^T$ is the actual user's eye position, as illustrated in Figure 6.1. The virtual camera coordinate system is centered at C' and its z axis is perpendicular to the virtual image plane and the x, y axes parallel to the horizontal and vertical sides of the OSTHMD virtual image plane, respectively. A point P of virtual camera coordinates $P = \left(x_p, y_p, z_p \right)^T$ should be projected to p but the calibration procedure based on the Phase I camera projects it to p'. The error vector $\delta_p = p - p'$ is coplanar with C, C' and P, and a measure of $\left\| \delta_p \right\|$ can be written as:

$$\left\| \delta_p \right\| = \left| \frac{z_p - z_0}{z_p} \sqrt{x_p^{\,2} + y_p^{\,2}} - \frac{z_p - z_0}{z_p - z_e} \sqrt{\left(x_p - x_e \right)^2 + \left(y_p - y_e \right)^2} \right| \qquad (6.1)$$

When $z_p \gg z_e$ (z_p is usually greater than $1m$ and z_e is less than 20mm), the above equation can be simplified as:

$$\left\| \delta_p \right\| = \left| \frac{z_p - z_0}{z_p} \right| \left(r_p - r_{pe} \right) \qquad (6.2)$$

where $r_p = \sqrt{x_p^{\,2} + y_p^{\,2}}$ and $r_{pe} = \sqrt{\left(x_p - x_e \right)^2 + \left(y_p - y_e \right)^2}$ are the radial distances from point P to the z-axis relative to the replacement camera coordinate and user's eye, respectively.
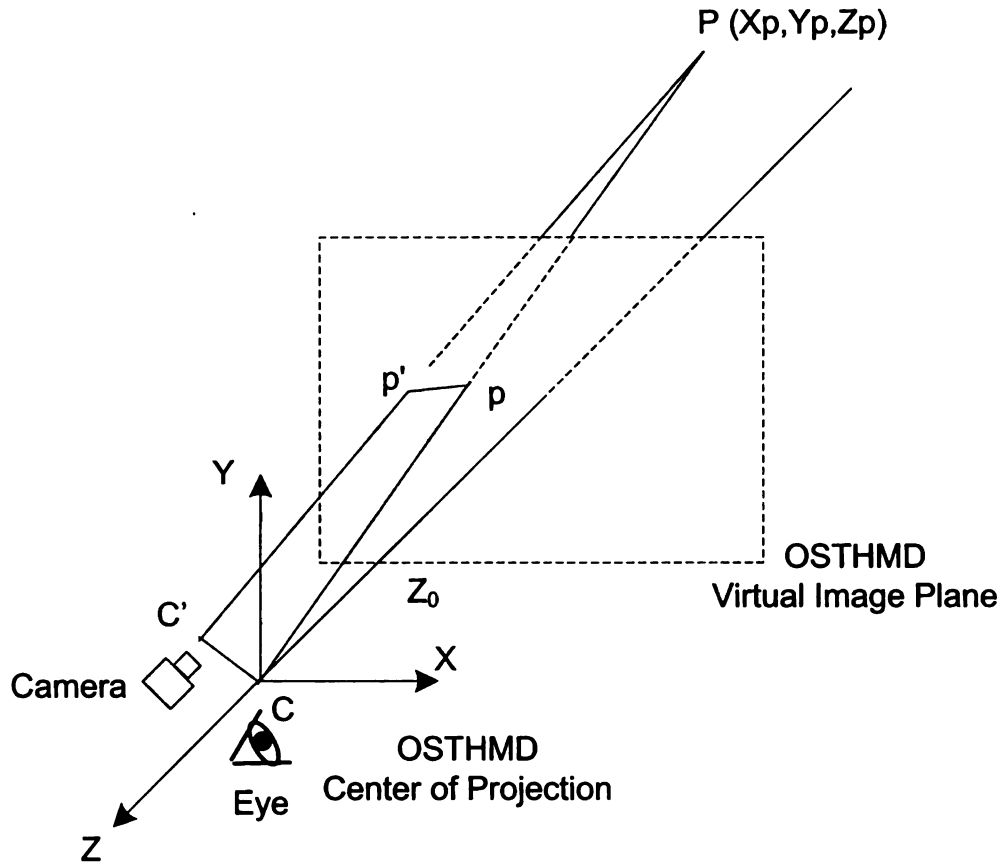
Figure 6.1 Effect of the OSTHMD virtual camera center of projection on the registration error

6.2.2 Experimental Analysis

The above theoretical analysis is based on a flat virtual image plane without taking into account optical distortion. It would be helpful to estimate the observed positional and angular error for different object depths relative to the eye offsets based on the actual data calculated with curved virtual image plane.

The OSTHMD under investigation is the Sony Glasstron LDI-100E display calibrated using the DRC phase I and the calculated parameters are $z_0 = -1681mm$,

$a_x = 1.56$, $a_y = 0.44$, and $K_2 = 19.5$. Figure 6.2 illustrates the pixel error for an object at an arm's length of 0.635 *meter*. The X axis plots the radial offset of the object from the center of the display and the Y axis plots the eye offset. Note that human user's interpupillary distance ranges from 53 to 73 *mm* and the computation is based on an interpupillary distance of 65 *mm*, approximately in the middle of that range. Hence, the maximum horizontal eye offset is 5mm. This results in an expected error in the center of the display of less than one pixel and less than 3 pixels at the edge of the display.
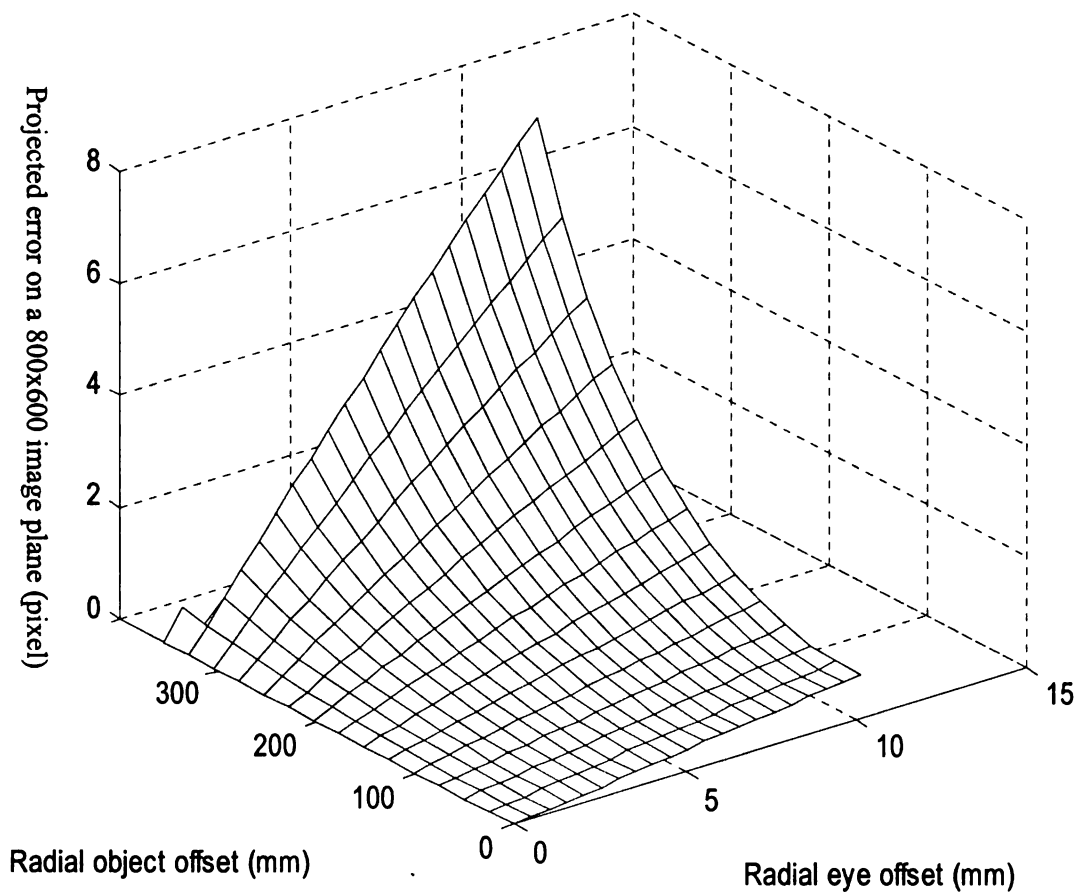
Figure 6.2 Registration errors due to eye offset (object distance z=0.635m)

Figure 6.3 illustrates that the pixel error for object at the distance of 2.5 *meters*. It is

shown that the error increases as the offset relative to the depth of the virtual display

increases. Hence, the DRC phase I calibration be used directly with little error when the

physical objects are at or near the virtual display depth. When working at larger distance,

the registration error will be more sensitive to the eye offset. This indicates that a phase

II calibration would not be necessary at all if the virtual image plane of the OSTHMD is

located at exactly the work depth.
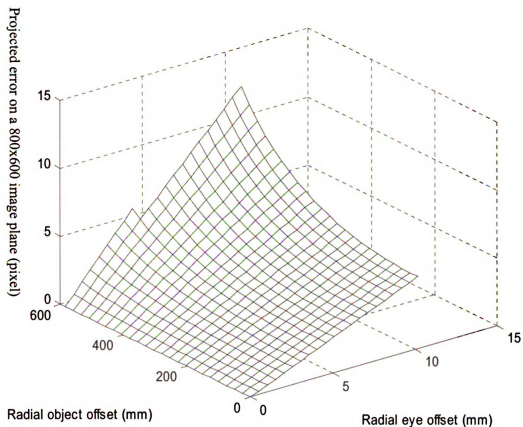
Figure 6.3 Registration errors due to eye offset (object distance z=2.5m)

Figure 6.4 illustrates the angular error for varying object depths relative to eye offset.

The maximum angular error of 0.8 degree in this figure is equivalent to a distance error

of 8.89 *mm* when the depth is 0.635m. This figure quantifies the improvements of any

Phase II strategy relative to tracking accuracy. As an example, the specification for the

Intersense IS-900 high-accuracy head tracker angular accuracy is 0.25 degrees. The

phase I calibration gives accuracy which is close to this value without any phase II

adjustments. Thus the phase II calibration would be of little use because the tracking

error would be more significant than the errors introduced by the calibration.
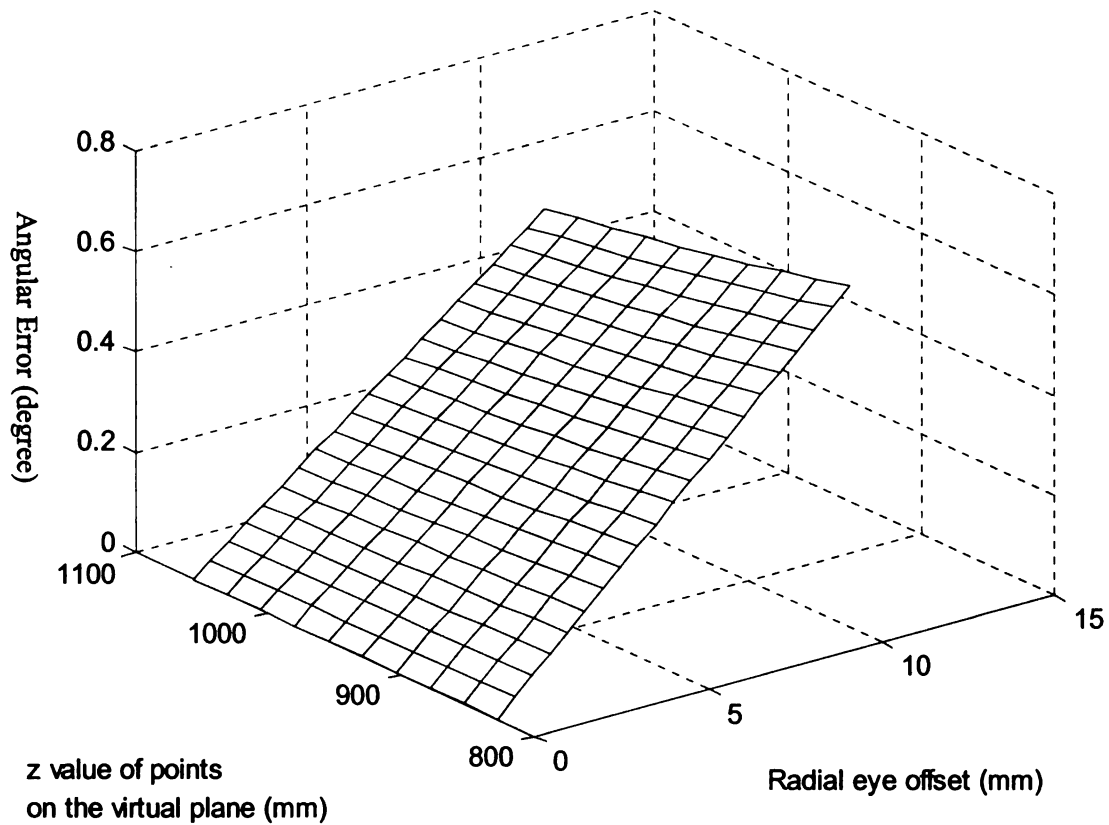
Figure 6.4 Angular errors due to eye offset for different depths

6.2.3 Summary

The sensitivity analysis indicates that the parameters from phase I calibration might be acceptable for AR applications which do not need very precise registration. Since the camera used in phase I is placed at the approximate position of the user's eye, the center of projections for camera-HMD and eye-HMD imaging systems can be very close to each other. Therefore, DRC phase I itself can be a calibration scheme for certain OSTHMD systems without the steps of adjusting the center of projection in the phase II. For those OSTHMD systems requiring accurate registration, phase II calibration is still

required for fine-tuning the projection center of the virtual camera model.

Chapter 7   Sources of Errors and Analysis

This chapter presents theoretical accuracy analysis and error estimation for the DRC method. It is necessary to determine the error range for an OSTHMD calibration method like DRC, although it is difficult to compute a quantitative accuracy assessment because there are no "ground truth" calibration parameters available as an absolute reference. The goal of the error analysis is to provide a theoretical error range composed of estimated errors from system components and the computation of the DRC method. Meanwhile, the error sources discussed in this analysis are related to the system used for the experimental evaluation described in Chapter 9 such that the theoretical analysis can be validated in the user experiment.

7.1   Related Work

Gomez et al. [96] analyze the expected accuracy level using different camera models and the stability of the calibration results, but this study mainly explains the effect of different camera models on AR calibration errors. Holloway presents a mathematical model of a medical OSTHMD system and analyzes the sensitivity of system registration error from different error sources [111, 112]. MacIntyre et al. describe a statistical model for estimating registration errors in AR systems [113]. Babbage evaluates the rendering error through optical modeling of the Display-Eye system of an OSTHMD [114].

7.2   Error Model

In camera calibration, the accuracy can be assessed in a variety of ways, the most common and generally useful being the calculation of the Root Mean Square (RMS) between the known 3-D coordinates of the test points and the re-projected 3-D points calculated using triangulation. The registration error model of an OSTHMD is different because OSTHMD users are relatively insensitive to depth information and only the lateral offsets between observed and actual points are distinguishable to the user. Meanwhile, it is difficult for the OSTHMD user to precisely locate the exact 3D position of an observed point on the OSTHMD virtual image plane.

As shown in Figure 7.1, the lateral error L between the expected point P_i and observed OSTHMD image point P_e provides a measurement of the OSTHMD registration offset perceived from the current eye position:

$$L = \left\| P_i - P_e \right\|$$ (7.1)
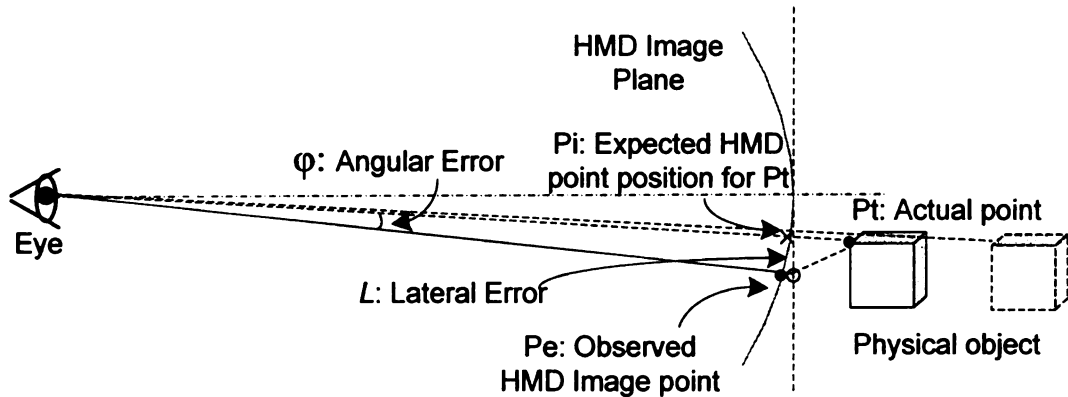
Figure 7.1 OSTHMD registration error model

78

Note that the lateral error metrics could be quite different, depending on the distance from the eye to the actual objects. Therefore, angular error φ is introduced as the visual angle between the ray from eye to the physical object and another ray from eye to observed OSTHMD image point. Usually the lateral error L is much less than the OSTHMD focal length (the distance between eye and image plane), and the angular error φ (in radians) can be calculated as:

$$\varphi = \arcsin\left(\frac{L}{D}\right) \tag{7.2}$$

where L is the lateral error and D is the distance between the user's eye and the OSTHMD virtual image plane.

7.3   Error Sources and Estimation

The OSTHMD registration errors come from two major sources, dynamic errors and static errors. Dynamic errors are caused by delays from tracker or graphics system when the user's position or viewing direction is changed. Although previous studies indicate that dynamic errors are the main contribution to the OSTHMD registration errors [87, 111, 112], the error analysis only discusses the static error sources because dynamic errors are independent of the calibration method.

7.3.1 Tracker Measurement

A webcam mounted on the OSTHMD is used as a vision-based tracker to provide the position and orientation of the camera relative to the world coordinate system defined by a set of fiducials. The fiducial scheme is based on the orthogonal DCT-based fiducial algorithm [48], and the tracking algorithm is based on a pose estimation algorithm using all corners of all visible fiducial images [115]. The webcam is calibrated beforehand and is compensated for radial distortion.

It is known that the image points extracted from fiducial program can reach sub-pixel accuracy, and the root mean square error e for the tracker calibration is 1.87 *pixels*. When the focal length of tracker camera f_{Trk} is 911.4 *pixels*, the reprojected lateral error d at $z = 0.635m.$ (*25in.*) can be roughly estimated as $d = \dfrac{e \cdot z}{f} = 1.3\ mm$ (*0.051 in.*) or 0.12 *degree* of equivalent angular error. This d can be considered as the theoretical upper bound for tracker measurement error without taking into account the delay or other dynamic error sources.

7.3.2 Camera Calibration

The Tsai camera calibration algorithm [98, 99] has been repeatedly used in the DRC phase I to calibrate both the measuring camera and the OSTHMD virtual camera model.

As explained earlier, the calibration of the measuring camera is only used for

computing the OSTHMD virtual camera model. The measuring camera parameters are used in the computation of the OSTHMD system registration. Therefore, any errors in this step would be reflected in the errors of virtual camera calibration.

In the OSTHMD virtual camera calibration, a non-linear minimization algorithm is used to search the parameters and normalized vectors are sent into the algorithm to compensate the unknown depth values. The computed root mean squared errors for left and right eye of the OSTHMD are 0.78 *pixel* and 0.82 *pixel*, respectively. The focal lengths of OSTHMD virtual camera are 1889.73 *pixels* and 1877 *pixels*, respectively. Therefore, the lateral error d at $z = 25$ *in.* can be estimated as: $d = \dfrac{e \cdot z}{f} = 0.254$ *mm* (0.01 *in.*) for left eye of the OSTHMD and 0.79 *mm* (0.011 *in.*) for right eye of the OSTHMD. Here the 0.01 *in.* lateral error is equivalent to 0.02 *degrees* of angular error.

7.3.3   Triangulation

In DRC phase I, Hartley's optimized triangulation algorithm is used to compute the depth of the OSTHMD image plane [103, 107]. Because the positions of the camera center during the DRC computation are very close to each other, the rays used for triangulation may be very close to parallel and bring certain uncertainty regions to the 3-D reconstruction. Blostein et al. present the following probability distribution function which gives the estimation of the maximum relative error based on the horizontal pixel disparity between the two images used for triangulation [116, 117].

Define the relative error in ranging a point as:

$$\varepsilon_z = \frac{\hat{z} - z}{z} \tag{7.3}$$

where z is the exact range of the point and $\hat{z}$ is the estimated range of the point. The probability of the range value being within a certain tolerance τ_z can be formulated as:

$$P_{|\varepsilon_z|}\left(|\varepsilon_z| < \tau_z\right) \equiv \begin{cases} 1 - (1 - \tau_z D)^2 & \tau_z < \dfrac{1}{D} \\ 1 & \tau_z \geq \dfrac{1}{D} \end{cases} \tag{7.4}$$

where D is the horizontal pixel difference between the left and right images.

In the triangulation of the DRC phase I, the image pairs with less than 40 pixels horizontal difference have been removed. Thus the minimum distance of the remaining image pairs is $D = 40$ *pixels*. From (7.4) the maximum relative error for reconstructed depth would be $\dfrac{1}{D} = 0.025$, which corresponds to an estimated depth difference of 1.06 *in* when the calculated depth is 42.26 *in* (1.07m). According to (5.3), such a depth change effects the focal length by 2.5%, which is trivial when the focal length of OSTHMD virtual camera model f is over 1880 *pixels*. The changes on lateral errors cause by such depth difference can be ignored.

7.3.4   User Alignment

The DRC method collects the required image and world point correspondences by requesting the user to align a number of image points to the corresponding predefined locations in the workspace. Previous studies [118] shows that the SPAAM method [88, 89] based on head motion alone is more subject to manual errors than the head-hand pointing method used in the Stylus-Marker calibration [75-77]. The head-hand style is more suitable for the human anatomy because the user can adjust the relative position and cooperation of the head and hand to find the most comfortable pose to complete the alignments.

The use of the head-hand pointing method as the input for the user alignment decreases the error introduced by user alignments, but the error from the point correspondences still contributes the largest portion to the calibration error and the overall registration error. However, it is impractical to separate the error introduced by the user alignment and the errors from other sources because the overall errors are combined by many errors sources including user alignment errors.

Furthermore, the user alignment error is dependent on the personal spatial sense and locating ability in the augmented environment. This increases the uncertainty of this error source. A previous study by McGarrity et al. shows registration errors of four users ranges from 0.11 *inch* to 0.43 *inch*, with a variance of 0.33 [119].

For the sake of simplicity, the user alignment error can be considered as the remaining error source when all other error sources are identified and quantified. Based on past study, the error introduced by the user alignments can be roughly estimated as up to 10.16 *mm* (0.4 *in.*).

7.4 Summary

Among the discussed error sources, the user alignment error contributes the largest portion to the overall error. The precision of tracker measurement and optical distortion also have certain effects on the registration accuracy. However, the errors from computational procedures in calibration algorithm such as camera calibration and triangulation are trivial. The analysis and estimation of different error sources for DRC leads to the following observations:

1)   The errors from the computation and the DRC algorithm itself are trivial compared to errors introduced by the user interaction. These errors can be minimized by computational optimization and adjusted procedures.

2)   As the most significant error source in the OSTHMD calibration, the user alignment procedure must be carefully designed to reduce the effects of the user performance on the calibration results. The existing user alignment procedures require excessive user efforts, which make the results inaccurate and unstable. A practical and reliable calibration method prefers simple and user-friendly user operation like the user

procedures proposed in the DRC.

3)   A flexible user interaction procedure like the multiple options of DRC is important to provide some choices for both OSTHMD system designer and users. A single calibration procedure with a fixed accuracy level cannot satisfy different types of applications and users. It is helpful to allow an offline calibration procedure with acceptable accuracy such that the OSTHMD systems can be quickly utilized without user interaction. Meanwhile, offline calibration would be helpful for user groups such as seniors, disabled persons and children. Offline methods would make things easier for OSTHMD-based game or entertainment applications where quick start-up and simple operation are preferred.

Chapter 8   Optimization

This chapter discusses the optimization strategies for the OSTHMD calibration, including the computational optimization and improving user procedure design. As described in Chapter 3, the OSTHMD computational model is established as a pin-hole virtual camera which converts a 3-D world point to a 2-D image point on the OSTHMD virtual image plane through perspective projection. Due to the special optical design of OSTHMD, the user procedure is necessary to collect data required for solving the OSTHMD virtual camera model. Both the computational process and user procedure design may affect the final calibration results.

8.1   Computational Optimization

The following computational optimization approaches can be applied for OSTHMD calibration:

1) Reduce the number of required point correspondences

Sufficient world and image point correspondences are required to solve the OSTHMD virtual camera model. The number of required point correspondences depends on the degree of freedom in the computational model. Fewer point correspondences can reduce the user efforts and the chance of making mistakes in the user procedure of the OSTHMD calibration. Currently, the SPAAM method [89] needs at least six point

correspondences and Fuhrmann's stylus-marker method [75] requires at least four. The DRC method reduces the number of free parameters to three, and thus only two points correspondences are required to obtain a solution to the OSTHMD calibration. Generally, each point correspondence will be subject to error induced by the inaccuracy in the user procedure. Although larger number of correspondences might be required in order to overcome the influence of the error points, starting with a lesser minimum requirement allows a smaller number of required points when errors are taken into account.

2) Detect false point correspondence (outlier)

RANSAC [120] can be used to detect the false point correspondence from the collected point correspondences for OSTHMD calibration, but the disadvantage of this method is that it only works well for a large amount of data. To find the false correspondences from a small data set for OSTHMD calibration is still an open problem, which can be a research topic for the future work.

8.2   Improving User Procedure Design

In OSTHMD calibration, the user procedure can be more effective if the following human factors are properly considered:

1)   Limitations of head-pointing alignment

Most existing OSTHMD calibration methods including the DRC method proposed in

this thesis use a head-pointing user task to align the virtual and real points, although previous research has demonstrated that a "head-mouse" is not precise enough even for a character pointing task for text editing [121]. In order to achieve pixel-level accuracy for OSTHMD calibration, the angular error of head rotation must be less than 0.05 degree. This is almost an impossible mission for simple head-pointing interaction like the method used in SPAAM [89]. Fuhrmann's stylus-marker approach [75] incorporates head-pointing with hand, wrist and finger motion to help the user stabilize his/her head with the body. Previous experiments shows that this approach improves the accuracy of head-pointing alignment but still cannot provide sufficient precision for OSTHMD calibration, as even holding the head still is a difficult task for most users [118].

2)   Difficulties with holding the head still while aligning

Another problem with head-pointing calibration is that it is very difficult to hold the head still while aligning the virtual and real correspondences. Even when the intrinsic movements of the neck and head can be stable, the extrinsic movements of the user's body still cause slight swinging relative to the calibration points. If this body jitter is unavoidable, some measures should be taken to avoid or eliminate its effect on the OSTHMD calibration.

8.3   Summary

Note that the above considerations of computational optimization and user procedure

design are general guidelines to the OSTHMD calibration. The utilization of these approaches is based on the computational model and algorithm design of the OSTHMD calibration method. The two-phase DRC method proposed in this thesis uses an offline computation to estimate the OSTHMD virtual camera model and refines the parameters using a minimal number of point correspondences. Thus the only concern of the DRC method is to improve the accuracy of the user alignment used in the phase II. Considering the limitations of SPAAM or Stylus-Maker calibration, a better user alignment procedure is yet to be established.

Chapter 9 Experimental Evaluation

This chapter describes the design, procedure and results of the user experiment which evaluates the performance of the DRC method. Compared with offline non-user testing, the online user experiment provides direct and practical results which are more effective for the evaluating the accuracy and usability of an OSTHMD calibration method.

This chapter describes measuring and evaluation of accuracy and usability for the OSTHMD system at first. Then the experiment design, system setup and the procedure are introduced. The experimental results are analyzed in both descriptive and quantitative statistical methods.

9.1 Related Work

It is essential to involve the user in the evaluation procedure because the user plays a key role in the OSTHMD viewing system and the perception of AR registration. A well-designed user study obtains the actual registration accuracy observed by the user and reveals the real user experience of the operational usability. However, due to the complexity of conducting user experiments, most previous research work on OSTHMD calibration either performs an offline evaluation using a camera to simulate the user's eye [89] or conducts a limited user study with small number of subjects [88, 89]. Such approaches cannot provide sufficient information about the usability or accuracy of an OSTHMD calibration method including user procedure.

The experimental evaluation in this thesis is similar to a previous research [118], which comparatively evaluates the registration errors of four calibration methods using McGarrity's online evaluation scheme [119, 122, 123]. However, the past study does not include evaluation of the usability or validation of the results using statistical analysis. The experiment conducted by Tang et al. also provides useful references for the experimental design of this research [11, 12] . Other similar research work on user studies for OSTHMD include the comparison of Head-Mounted Display and Displays by Pausch et al. [124] and the user-centered design and evaluation of a battlefield virtual environment by Hix et al. [125].

9.2   Evaluation of Accuracy and Usability

To properly measure the registration errors indicated by the user, the user procedure and the measurement of the OSTHMD accuracy evaluation face similar challenges as those for the OSTHMD calibration. The measurement of the OSTHMD registration errors should reflect the observable differences between the perceived and the expected points. Meanwhile, the user procedure should support the measurement in a user-friendly manner. Furthermore, the OSTHMD evaluation method and measurement should work for the OSTHMD virtual image plane. For an OSTHMD system, the user is insensitive to the depth of a virtual point and it is hard for a user to indicate the exact depth value. The measurement of the OSTHMD registration error is different from that of the camera calibration because only the lateral difference on the OSTHMD virtual image plane is

observable and expressible to the user.

An effective OSTHMD evaluation technique is the online method proposed by McGarrity et al [119, 122, 123]. As shown in Figure 9.1, P_e is the projection of a virtual point P_o perceived on the plane of an evaluation board, and the registration error can be measured based on the lateral offset $\left| P_t - P_e \right|$ on the evaluation board between the perceived point P_e and the expected point P_t.
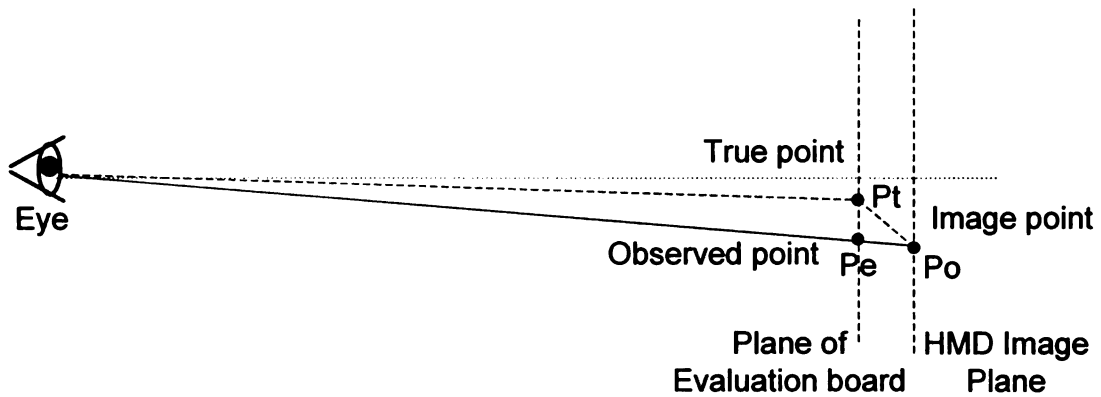
Figure 9.1 Planar measurement of the OSTHMD registration error

A potential problem for planar evaluation is that the error metric depends on the user's eye position relative to the evaluation board. As shown in Figure 9.2, the measured planar error $\left| Pt - Pe \right|$ from the eye position E is different from the error measurement of $\left| Pt - Pe' \right|$ viewed from E'. Due to this metric ambiguity, the planar errors between different subjects are not comparable directly. To solve this problem, the user is asked to align the OSTHMD image point with an observed point on the evaluation plane twice for each evaluation point, with the head at different positions. The

3-D position of this OSTHMD image point P_o can be determined by triangulating the two rays connecting the eye and the observed image point. Once the 3-D point is reconstructed, a fixed eye position can be chosen as a standard viewpoint for all subjects to recalculate the planar registration error. The recomputed planar error represents the perceived registration accuracy for a current subject, and the error value is comparable because of the same viewpoint.
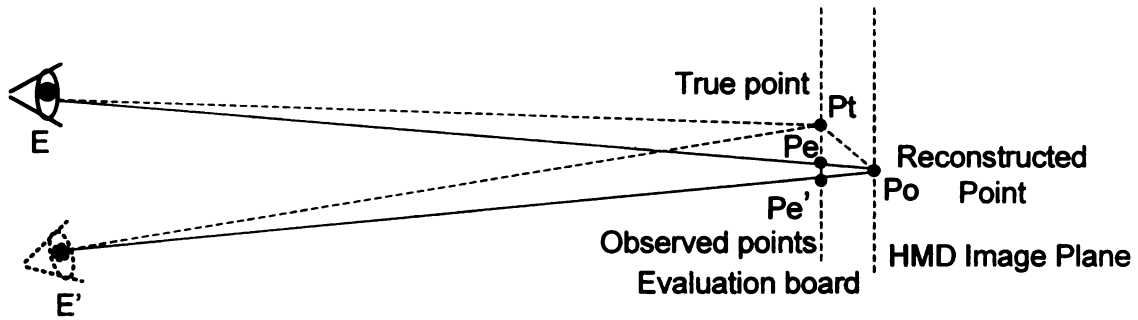
Figure 9.2 Calculating the rendered 3-D point using triangulation based on two observed image points

This experimental evaluation provides quantitative measurements of the registration error between the physical objects and the virtual objects as observed by the user. Analysis of the user data can verify the error estimation results presented in Chapter 7. Together with the previous theoretical error analysis, the errors as exhibited by the OSTHMD calibrated using the DRC method can be determined and experimentally validated. It is important to understand this error range because it determines whether or not this calibration method can be used for certain OSTHMD-based applications. Because of uncertain human performance, a calibration method involving user

interaction cannot always be guaranteed to achieve the required accuracy, even if its estimated error range satisfies the expectation. For such a situation, repeated calibrations with trained user may increase the accuracy to reach the estimated error range. The ability of a new user to achieve a usable calibration is shown in this chapter through the use of user trials that are shown statistical significance.

The registration accuracy and the user efforts required by different OSTHMD-based applications can be quite different. To fit for this diversification, the usability of an OSTHMD calibration method is an important consideration for both calibration algorithm developers and OSTHMD system designers. Therefore, this usability of the DRC method needs to be evaluated as well. There are many usability evaluation methods for virtual environments [126], but a quantitative workload assessment procedure such as NASA TLX [127] is accepted as an effective way to compare the usability of competing methods that require user concentration and effort. The workload assessment results indicate the mental workload and operation difficulty of different calibration methods.

Meanwhile, the completion time of each user calibration procedure is also a comparable factor of usability to be measured in the experiment. Although a lesser workload is usually correlated with a shorter completion time, it is hard to quantitatively connect the workload index with the completion time of a task. Similarly, it is not always true that a complicated and time-consuming calibration procedure provides accurate registration. Only a comprehensive analysis of accuracy, workload, and completion time

can reveal useful information for designing new calibration algorithms and choosing one that is suitable for a specific OSTHMD application. The analysis is also expected to determine whether or not the DRC method is a more effective and flexible calibration scheme than other approaches.

9.3   Experiment Design

The evaluation experiment uses a between-subjects experiment, in which each subject is assigned to only one factor level of the tested calibration methods. Compared with a within-subjects experiment, the between-subjects experiment is not affected by "learning effects" or "order effects", and the data are guaranteed to be independent between groups of different factor levels [128]. Another reason for choosing between-subjects approach is to make the evaluation data comparable with results from the previous user study [118] and the other research papers. Furthermore, a between-subjects test is simpler for the participants and easier to extend. More subjects can be easily added using the same experimental setup to increase the statistical significance of this user study. The disadvantage of a between-subjects test is the requirement of more subjects because of the greater variability. To ensure the statistical significance, the evaluation data needs to pass an analysis of variance (ANOVA) test [128-130].

This experiment has only one independent variable, the class of calibration methods (or user procedure options). There are six factor levels (treatments): the four options of

DRC phase II (DRC-None, DRC-Mouse, DRC-Single Point, and DRC-Multiple Points) and the two existing OSTHMD calibration methods, SPAAM [88, 89] and Stylus-Marker Calibration (SMC) [75-77]. The measurements include calibration completion time, registration error, and mental workload index (NASA TLX).

9.3.1 Participants and Choice of Sample Size

The participants of this experiment are undergraduate or graduate students without previous experience in OSTHMD or AR systems. Each of the 6 treatment conditions is assigned to a group of 6 participants. The determination of the sample size in this experiment is based on statistical power analysis of hypothesis test [131-134] . The specified values in the power analysis come from results of a past study [118] and the estimation of the registration errors presented in Chapter 7. The following is the approach to determine the sample size:

Among the three dependent variables of calibration completion time, registration error and mental workload index, the registration error is most important and the collected registration error data must be statistically significant. It is known from the previous chapters that the DRC method provides smaller registration error than other calibration methods do. A previous study [118] with a similar user experiment shows that the mean registration errors of Stylus-Marker Calibration (SMC) and several SPAAM variants range from 0.9 ~ 3 inches, and the SMC method has minimum mean registration

error of 0.9 inch measured on a planar evaluation board at a distance of $1.2m$. The estimated registration error of DRC method ranges from $0.1 - 0.5$ *inch*, and the DRC-None procedure has the upper bound of registration error of 0.5 *inch*. The experiment is expected to detect the minimum difference of 0.4 inch registration error between SMC and DRC-None. So a one-side test of hypothesis can be set up:

$H_0 : \mu_{DRC-None} = \mu_{SMC}$ versus $H_0 : \mu_{DRC-None} < \mu_{SMC}$, where $\mu_{DRC-None}$ is the (estimated) mean registration error of treatment DRC-None method and μ_{SMC} is the mean registration error of treatment SMC method. Since the parameter $\theta : \mu_{DRC-None} - \mu_{SMC}$ is actually the effect being tested, the above hypothesis can be written as $H_0 : \theta = 0$ and $H_1 : \theta < 0$.

The power of effect θ to be detected is specified to be 80% when the significance level is $\alpha = .05$. Data from [118] also suggest that the standard deviation of the SMC registration error data is $\sigma = 0.2$, which can be estimated as the standard deviation of the DRC-II options. Using a two-sample pooled t-test with equal numbers n of subjects in group, a sample size of $n = 6$ per group is calculated as the required sample size to detect the minimum registration error difference, and the actual power is 0.8764 in this situation. A program used for the above power analysis computation can be found at [135].

9.3.2   Comparative Treatments

In experimental design, the treatment refers to the factor levels of the independent variable. This user experiment includes the following calibration methods or user procedures:

1) Single Point Active Alignment Method (SPAAM) [88, 89]

The SPAAM method calibrates each eye of OSTHMD separately. For each eye, the subject aligns the OSTHMD image point with a single 3D point in the world coordinate system. 9 repeated alignments per eye are required for calculating calibration parameters and the whole calibration takes 18 alignments totally. During the calibration, the participant is encouraged to move the head around such that more systematic errors in the tracker measurement can be covered. The user interaction for SPAAM is comparatively simple but the computation might fail if coplanar data are collected or if the point alignment errors are too large. In the former case the algorithm fails to proceed due to internal singularities and in the later case the algorithm locates an incorrect local minimum.

2) Stylus-Marker Calibration (SMC) [75-77]

SMC is actually a variant of SPAAM calibration in that the alignment point in the world coordinate system is replaced by the reference point on a handheld tracked stylus.

The subject aligns the observed OSTHMD image point with a point on this stylus. Other than this, the alignment procedure, repeated times and the computation are the same as those of SPAAM. The cases of failure to compute in SPAAM apply equally to the SMC method.

But the difference on the user interaction does change the alignment difficulty and the calibration results. A common difficulty for both SPAAM and SMC is holding the head steady and clicking the mouse to collect alignment data at the same time. Usually there is a very small interval after the completion of alignment and before the data collection. During this interval the user needs to hold his/her head still to keep the "registered" status and to obtain the precisely aligned spatial information of the image and world points. At the moment of data collection, most users hardly realize a minor head movement and the corresponding imperfect alignment, which is one of the most common error sources in the alignment-based calibration approaches. Therefore, an easier or quicker alignment procedure involving fewer points usually results in more effective and accurate calibration even though less data is available for computation of the result. A past study [118] shows that the SMC method makes alignment easier because holding a stylus in the hand reduces extrinsic body movements. Head motion in SMC is limited to fixing the head relative to the body and any body motion is transferred to both the head and the body simultaneously. In SPAAM it is more difficult to hold the head still while aligning a fixed world point because the larger extrinsic body movement causes the head

swinging relative to the alignment target.

The configuration of SMC in this experiment is different from the scheme in the original paper to better fit the vision-based tracking system. As shown in Figure 9.3, a small square board painted with different fiducial set replaces the traditional stylus and a tracked point (crosshair) printed on this board is used as the alignment target. The vision-based tracking algorithm used in this experiment supports multiple tracking channels simultaneously, which means the position and orientation information of two different fiducial sets can be extracted from a single camera image at the same time. The use of four fiducials on the stylus card allows for computation of the pose of the card from the 16 corner points.
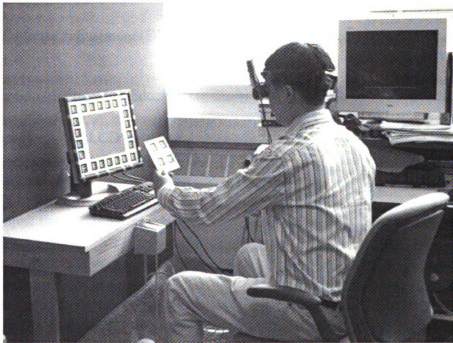
Figure 9.3 Stylus-Marker Calibration (SMC) used in the user experiment

During the experiment, the subject holds the tracked fiducial board in the hand during the calibration and moves it to align the observed OSTHMD image point with a reference point on the tracked board. Both must be in the view of the tracker camera. In the same way as SPAAM, 9 alignments per eye are collected.

3) DRC II-None (DRC-N)

The DRC-II-None procedure directly uses the offline DRC phase I calibration parameters for the OSTHMD system. There is no user interaction or alignment required. Thus it's a very simple and fast calibration option. The DRC phase I calibration is done before the experiment. Thus, the participants chosen for this method in this experiment only perform the evaluation procedure.

4) DRC Mouse Adjustment (DRC-M)

In this option, the tracking fiducials are displayed as wire-frame at the position of the actual fiducials on the LCD panel. As shown in Figure 5.3, the observed wire-frame fiducials may not exactly coincide with their counterparts. The subject uses a wireless mouse to adjust the horizontal and vertical position of the wire-frame fiducials (actually the position of the OSTHMD virtual camera projection center) and overlay them onto the physical fiducials. The user cannot rotate or scale the viewed fiducials. This adjustment needs to be done separately on each eye.

5) DRC Single-Point Alignment (DRC-SP)

This option takes only one image-world alignment for each eye, and updates the projection center of the OSTHMD virtual camera model based on a single point correspondence as described in Section 5.2.2. Like the DRC-Mouse, this option will only adjust the x, y positions of the OSTHMD projection center. The calibration procedure is fast and simple, but may be unstable due to errors during the one point correspondence. This experiment uses the head-hand pointing method in stylus-marker calibration for the point alignment because a past study [118] shows that this interactive method is more suitable for the human user and easier to reach high accuracy.

6) DRC Multiple-Point Alignment (DRC-MP)

Multiple point correspondences allow the DRC phase II calibration to adjust the OSTHMD projection center in the x, y, and z directions. From the math derivations presented in Section 5.2.4, at least two point correspondences for each eye are required for this option. In this experiment, 4 alignments are taken with each eye. For these 4 point correspondences, every combination of 3 is considered to form a group to compute the updated OSTHMD projection center. The group with least variance of the computed OSTHMD virtual camera projection centers is chosen as the effective group, and the (x, y, z) values of the projection centers in this group would be averaged to get the final projection center. This procedure has been chosen to make the calibration more reliable

by omitting one potential outlier. Similar to DRC-Single Point, this procedure also uses stylus-marker alignments to collect the point correspondences.

9.4   System Setup

The OSTHMD used in this study is a Sony LDI-100B Glasstron stereoscopic display. The product specification describes it as an SVGA resolution display device with resolution of 800 × 600, horizontal angle of view of 28°, and virtual image size of 1.18 *meters* (*diagonal*), and focal length of 1.2 *meters* [136]. Figure 9.4 shows a webcam with a 640 × 480 resolution rigidly mounted on the OSTHMD as a vision-based tracker.
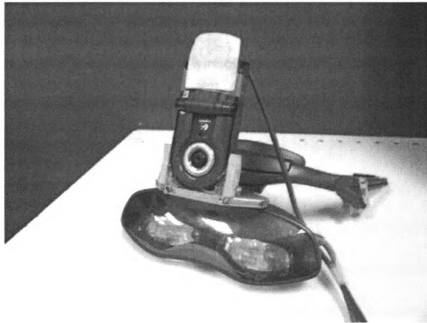
Figure 9.4 Webcam mounted on the OSTHMD as a vision-based tracker

During the calibration process a digital camera is used to capture images through the OSTHMD. These images will be processed to compute the Phase I calibration. A

professional camera is unnecessary but a camera with high resolution and minimal lens distortion is preferred to obtain precise calibration results. In this study, a Canon S70 consumer-grade camera (maximum image resolution of 3158×2342 *pixels*) is used to capture the required images, and all images are saved in uncompressed RAW format.

The evaluation method used in the experiment is based on the online planar evaluation scheme proposed by McGarrity et al [119, 122, 123]. As shown in Figure 9.5, the experiment system is implemented on two desktop computers connected using 100Mbps Ethernet. One of the computers works as a visualization workstation, which tracks the subject's head position and renders 3D reference points onto the OSTHMD for the evaluation procedure. The second computer works as a planar evaluation platform, which uses the LCD monitor panel and a wireless mouse as the evaluation board and the alignment device, respectively. The DRC phase I procedure is implemented using the ImageTclAR toolkit [137].

Figure 9.5 Evaluation system used in the user experiment

A pre-calibrated camera mounted on the OSTHMD providing real-time vision-based tracking information relative to the fiducial coordinate system by analyzing the current captured picture of fiducial images placed around the LCD monitor panel [115]. As shown in Figure 9.6, these fiducials are black-border square patterns with interior shaded images constructed from orthogonal Discrete Cosine Transform (DCT) to support orientation determination and unique identification [48]. The OSTHMD image is updated in real-time upon any changes of the subject's head position and direction. The fiducial images around the work area ensure reliable tracking and high accuracy.
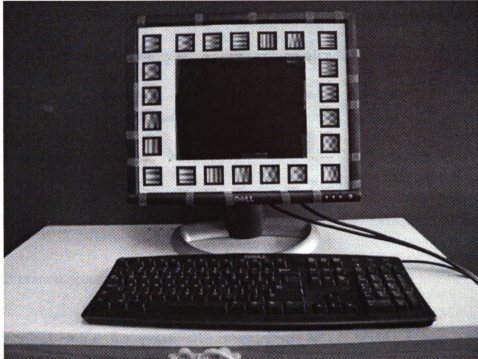
Figure 9.6 LCD monitor panel with fiducials used as the evaluation board

The central area of the LCD monitor panel is used as the planar evaluation board. The LCD resolution is set to 1280 by 1024 during the experiment, and the physical distances from the screen pixels to the origin of fiducial coordinate system are measured and matched to the physical distance. During the evaluation procedure, the subject moves the crosshair cursor displayed on the LCD screen and aligns it with the 3D point observed from the OSTHMD. The planar distance between the perceived point on the evaluation board (LCD monitor panel) and the expected position is collected as registration error data. This differs from the McGarrity et al. method in that they utilized a tracked stylus using a Polhemus Isotrack magnetic tracking system. However, the magnetic tracking system exhibited errors itself on orders of magnitude similar to those expected for the

DRC method and was deemed unacceptable. The mouse pointing method assumes hand-eye coordination, but is a familiar activity to most all adult users.

The calibration methods and evaluation procedure in this experiment are implemented using ImageTclAR toolkit [137].

This experiment measures the three types of measurements for each participant:

1) Calibration accuracy

As described earlier, the quantitative measurement of calibration accuracy used in this experiment is based on planar evaluation method. This method actually measures the registration errors of the whole OSTHMD system instead of the errors introduced by the calibration procedure itself. Because it is impractical to separate errors from difference sources, the data analysis of this experiment treats the planar registration error as a valid measurement for evaluating and comparing accuracy of different calibration methods or procedures. The same approach has proved to be effective in the past study [118]. Figure 9.7 shows the reconstructed and the actual points viewed from the OSTHMD after an evaluation test.
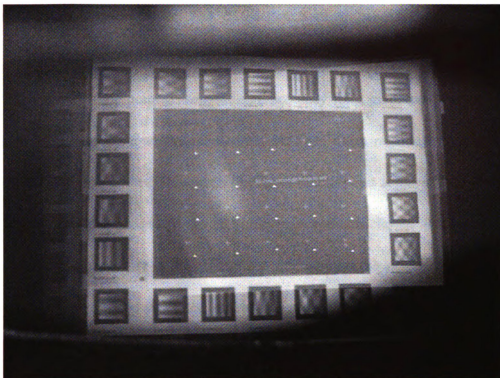
Figure 9.7 Reconstructed and actual points of an evaluation test

Another issue in the evaluation test is the choice of sample size, i.e. the number of evaluation points to be tested for each participant. In order to average out manual alignment errors, the subject is required to repeat the evaluation on certain number of points and the mean value of the collected data is selected as the position. Considering the fatigue effect of repeating alignment procedure, minimizing the number of samples while preserving effective evaluation results is important. A simple pilot test was conducted to investigate the relationship between the evaluation points and registration error. Two people participated in this study, and each performed the evaluation experiment with different numbers of evaluation points from 5 to 40. The order for tests with different numbers of evaluation points was randomized and each subject was given

a 5-minute rest between tests. The calibration method used for this pilot test is

DRC-Mouse because this procedure requires no user point correspondences during

calibration and thus the point correspondence accuracy for a given user does not affect

the results. Figure 9.8 shows the relationship between mean registration error and the

number of evaluation points. This graph illustrates that this test achieved better accuracy

when evaluation point size is 20. So 20 points each eye was chosen as the sample size

for the evaluation procedure of the experiment. These points are evenly distributed in the

planar evaluation area, and presented in a random order during the evaluation process.
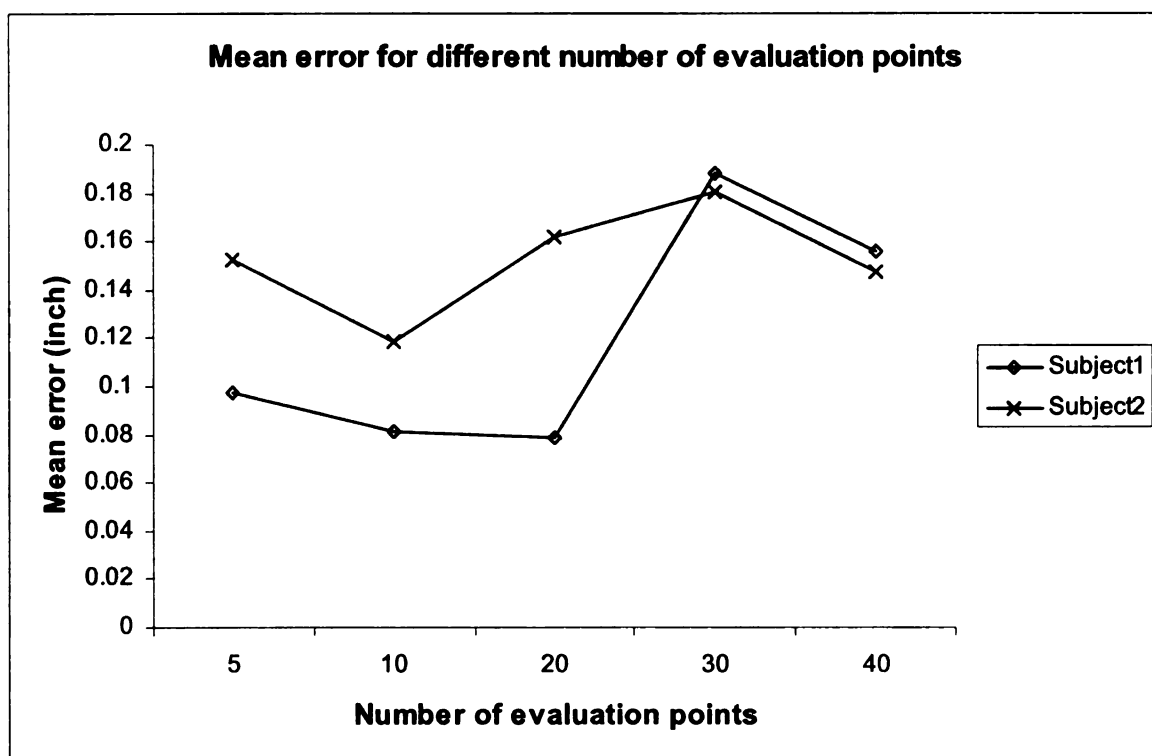
Figure 9.8 Relationship between the mean registration error and the number of
evaluation points

2) Mental workload

Mental workload perceived by subjects is measured using NASA Task Load Index (TLX) [127]. The NASA TLX program requires a subject to rate his/her experience on a task from 6 aspects, including mental demand, physical demand, temporal demand, effort, performance and frustration level. The NASA TLX score is a weighted average score of ratings on the above categories to reflect an overall workload perceived by the subject. Figure 9.9 shows the user interface of the part one of the NASA TLX program.

Figure 9.9 User interface of the NASA TLX program

3) Completion time

Completion time measures the time (in seconds) for a participant to finish the assigned calibration procedure. The participants are kept unaware of this measurement to prevent unintended speedup of the process. The completion time of the DRC-None option is not measured because there is no user interaction.

9.5   Procedure

The procedure of the user experiment includes the follow steps:

1) Pretest

Most subjects in this experiment had no previous experience in using OSTHMD or AR environment. So a warm-up procedure is conducted to make subjects comfortable with the OSTHMD and familiar with the alignment of virtual and physical objects. Subjects in this experiment are first briefed about the use of an OSTHMD and the experimental procedure. Then the experiment administrator would instruct the subject to run a 2-point alignment test using the OSTHMD. Errors made in the alignments are explained to the subjects to make sure that they understand the basic skills of the alignments. Participants can either repeat the pretest once or proceed to the two parts of the main test: calibration and evaluation.

2) Calibration

The participants are randomly assigned to one of the treatment conditions (calibration methods). The procedures for all treatments in the calibration part are described in section 9.3.2. All the calibration methods are single-eye operation and subjects are told to close the other eye while working on one eye. As shown in Figure 5.3, instruction text is displayed on the OSTHMD image screen to help the participants complete the procedure step by step. The participants assigned to the DRC-None treatment also run an unused 2-point alignment on each eye to balance the fatigue and learning effect with other groups.

3) Evaluation

The evaluation is a stereoscopic operation, and the participants are required to keep both eyes open during the procedure. Participants use a wireless mouse to move a crosshair displayed on the LCD screen and superpose it on the evaluation point (a small square) observed from the OSTHMD, and then repeat this step from another view point. Totally 40 alignments (2 alignments for each of 20 evaluation points) are required for the evaluation procedure.

9.6   Analysis of Results

36 undergraduate and graduate students at Michigan State University volunteered to participate in this experiment. Each of the 6 treatment conditions was randomly assigned to a group of 6 subjects. Most of them have no previous experience in OSTHMD system

or any AR application. 21 (58.3%) of the participants are female, and 15 (41.7%) are male. Female and male participants were evenly distributed to treatment groups to avoid possible gender-correlative effects [138]. All the collected experiment measurements are used for data analysis.

9.6.1 Descriptive Statistics

1) Calibration Accuracy

In this experiment, calibration accuracy is measured by the average planar registration error of the 20 evaluation points. As mentioned earlier, accuracy is one of the most critical factors for evaluating the performance of an OSTHMD calibration method. Smaller registration error means more precise registration and thus higher accuracy.

Figure 9.10 illustrates the experimental results for calibration accuracy. The values on the vertical bar are the maximum, mean, and minimum registration errors of each calibration method. These measurements reveal the accuracy and stability of these different calibration methods. This plot indicates that the DRC-Multiple Points method is the most accurate and stable calibration method with a mean registration error of 0.09 *inch* and a very small variance. The DRC-None method had the largest mean registration error of 0.48 *inch*. The mean registration errors of the DRC-Mouse and DRC-Single Point options are 0.15 *inch* and 0.13 *inch*, respectively, but the DRC-Single Point option is not as stable as DRC-Mouse as indicated by the larger variance. Both the SPAAM and

the Stylus-Maker calibration suffer from large variances of registration errors, but Stylus-Marker demonstrates better mean registration error of 0.20 *inch* compared to SPAAM's 0.25 *inch* mean registration error. The SPAAM and Stylus-Marker results are similar to past reported results [118]. The improvement over the past results for these methods is predominantly due to the much higher accuracy of the vision-based tracking method.

The experimental results clearly show that the 3 interactive DRC options (Mouse, Single Point and Multiple Point) provide better and more stable calibration accuracy than SPAAM or Stylus-Marker methods.
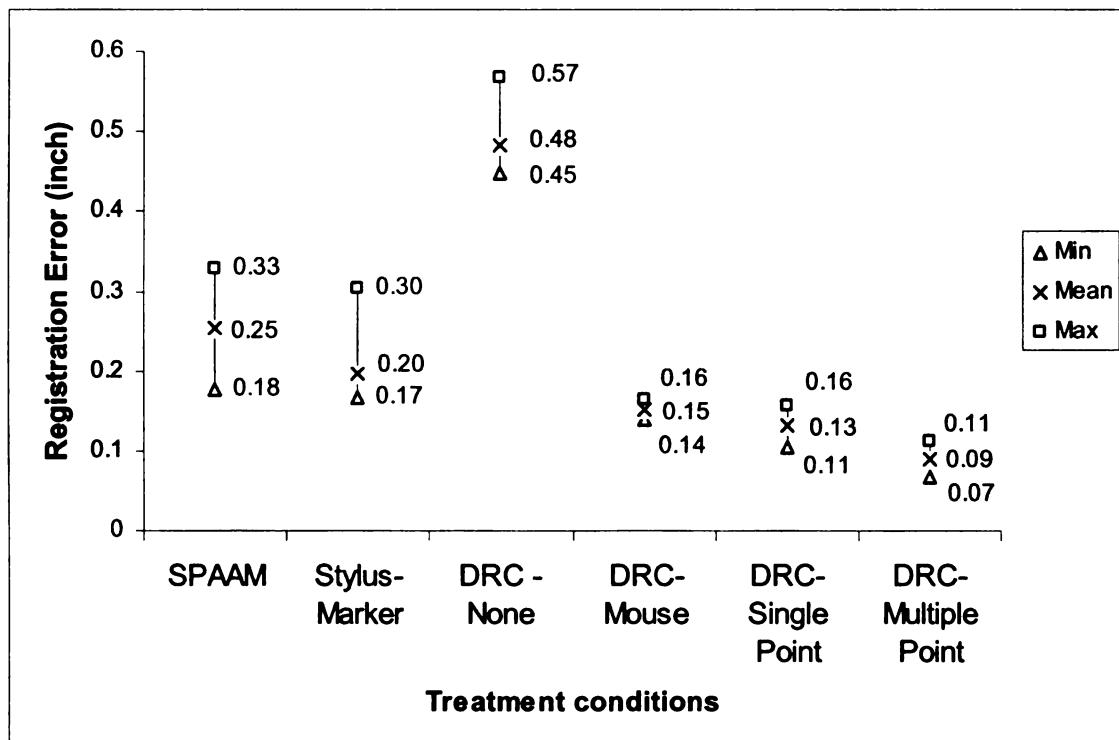
Figure 9.10 Registration errors of the tested calibration methods

2) Mental Workload

In this experiment, the NASA TLX score is used as a measurement of a participant's perceived mental workload on different calibration methods. This score is a quantitative direct evaluation on the usability of the tested procedure. Figure 9.11 illustrates the NASA TLX scores of each treatment. The DRC-None option was not rated because it is an offline calibration requiring no user interaction. It is shown that subjects running the DRC-Mouse procedure had the lowest mental workload, which indicates that it is the easiest and most straightforward calibration among all the tested methods. The TLX score for the DRC-Single Point method is also very small and only a little higher than the DRC-Mouse method. The mental workload of DRC-Multiple Point method is higher than other DRC options, but still less than SPAAM or Stylus-Marker calibration. This is not surprising, given the similarity between these procedures from a user standpoint.
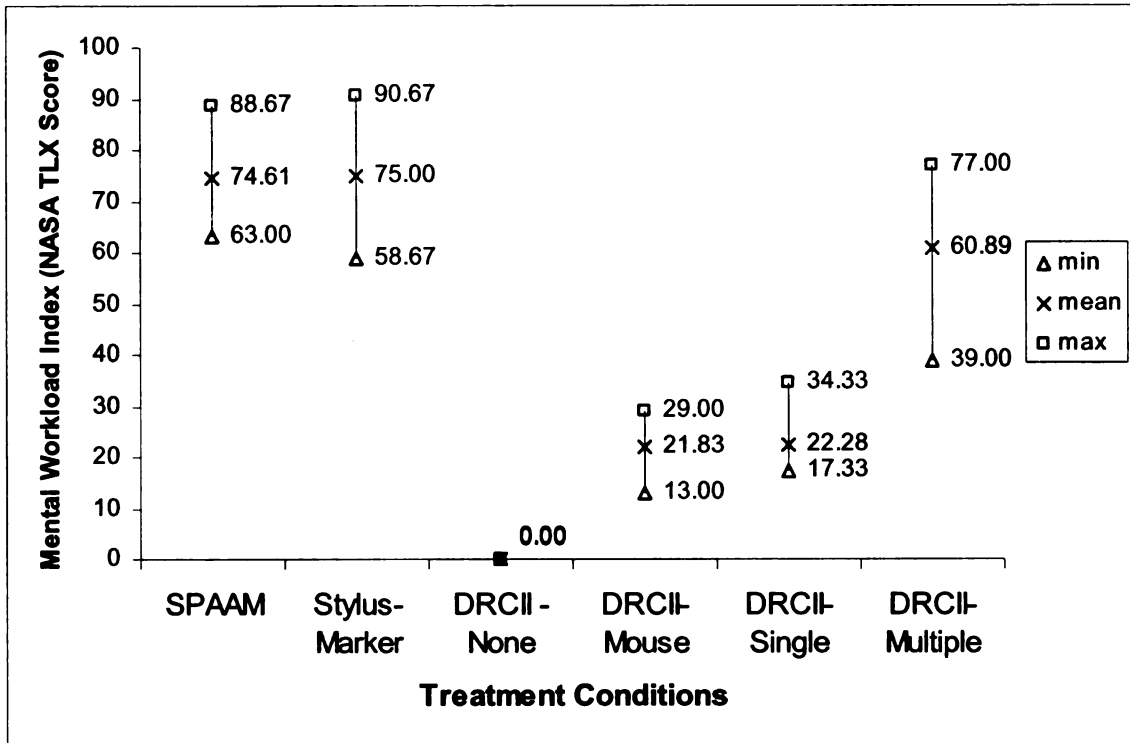
Figure 9.11 Mental Workload (NASA TLX Score) of the tested calibration methods

3) Completion Time s

The completion time of a calibration procedure measures the time a subject requires to finish the calibration process (not the evaluation). This value can be considered as an indirect measurement of the usability of this procedure. Figure 9.12 illustrates the completion time of each treatment except for the DRC-None option which exhibits a completion time of zero. Among the other 5 calibration methods, the DRC-Single Point method is the fastest calibration procedure requiring only about 10 seconds to complete. The DRC-Mouse method is the second fast calibration with average completion time of 12 seconds. The average completion time of DRC Multiple-Point option is around 34

116

seconds, which is longer than that of DRC-SP and DRC-Mouse but shorter than the two slowest methods, SPAAM and Stylus-Maker calibration. All of the DRC methods exhibit significantly faster completion times than previously reported methods.
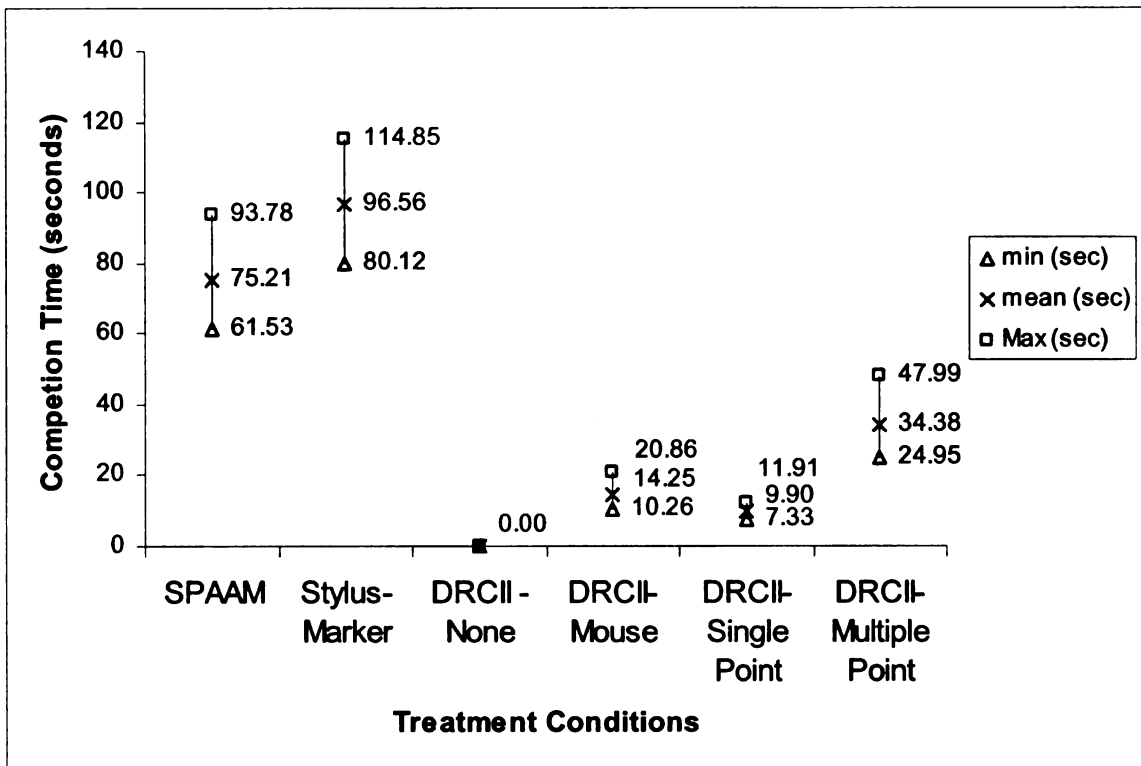
Figure 9.12 Completion time of the tested calibration methods

9.6.2 Statistical Analysis

The comparative observations seem to indicate relationships and differences of the means and variances of the tested calibration methods. However, due to variances, the actual mean differences may not be significant as indicated by the values themselves. Both individual differences and treatment effects could contribute to the mean value of a measurement. So it is necessary to run Analysis of Variance (ANOVA) tests to justify the

statistical significance of the collected measurements. ANOVA is a statistical testing procedure used to evaluate differences between two or more groups (variability) [130] [139]. ANOVA separates or partitions the variation observable in a response variable into between-subjects variance and within treatment variance, and explains the major sources of differences. One-way ANOVA deals with a single independent variable (the calibration method in this experiment) with two or more levels (the 6 different calibration methods). ANOVA is used to statistically compare different treatments by measuring the significance of the mean differences. An alpha value of 0.05 was used for all ANOVA tests in this experiment.

1) Effect of Calibration Methods on Registration Error

A one-way ANOVA is conducted to evaluate the significance of different calibration methods on the registration error. The results shown in Table 9.1 indicate there is statistical significance on the effect of calibration methods on the registration error, with $F_0 = 89.30 > F_{(0.05, 5, 30)} = 2.53$ and $p \ll 0.01$.

Anova: Single Factor

SUMMARY

Groups	Count	Sum	Average	Variance
Stylus-Marker	6	1.521814	0.253636	0.002555
SPAAM	6	1.174082	0.195680	0.002935
DRC-None	6	2.901866	0.483644	0.001870
DRC-Mouse	6	0.906369	0.151062	0.000120
DRC-Single Point	6	0.785089	0.130848	0.000320
DRC-Multiple Point	6	0.535767	0.089294	0.000342

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	0.605843	5	0.121169	89.300727	4.35223E-17	2.533555
Within Groups	0.040706	30	0.001357			
Total	0.646549	35				

Table 9.1 ANOVA test for registration errors of different calibration methods

Note that this test can only be used to reject the null hypodissertation that there is no

treatment effect or the means of the treatments are not all the same ($H_0 : \mu_1 = \mu_2 = ...\mu_6$),

but does not indicate a statistical test on pair-wise mean differences. *Post hoc* ANOVA

tests were further conducted to obtain all possible pair-wise comparisons, and the results

(*p*-values) are shown in Table 9.2. The results indicates that the measured registration

errors of the SPAAM and SMC trend to statistically significant ($p = 0.084$), but it cannot

be stated that SMC exhibits a statistically significant smaller error. However, that is not

a serious issue here, as the comparison is SMC and SPAAM is not a concern of this

dissertation. The mean registration error of SPAAM is significantly different from the 4

DRC options with all p-values $<<$ 0.01. The registration error of Stylus-Marker

calibration method (SMC) is significantly different between 3 of the 4 DRC options (DRC-None, DRC-Single Point, DRC-Multiple Point), but there is near significance between registration errors of SMC and DRC-Mouse (p = 0.076). Table 9.2 also indicates that the mean registration errors of any two of the 4 DRC options are statistically significant.

	SPAAM	SMC	DRC-None	DRC-Mouse	DRC-SP	DRC-MP
SPAAM	-	0.084	<<0.01	0.00066	0.00022	<<0.01
SMC	-	-	<<0.01	0.076	0.019	0.001
DRC-None	-	-	-	<<0.01	<<0.01	<<0.01
DRC-Mouse	-	-	-	-	0.014	<<0.01
DRC-SP	-	-	-	-	-	0.0027
DRC-MP	-	-	-	-	-	-

Table 9.2 P-values of pairwise ANOVA for registration errors of different calibration methods

2) Effects of Calibration Methods on Completion Time and Mental Workload

One-way ANOVA tests were conducted on the effects of different calibration methods on completion time and mental workload. The DRC-None option was not tested and compared because no user interaction is involved in the procedure.

A one-way ANOVA is conducted on the effect of calibration method on completion time, and the result is statistically significant, with F_0 = 102.01 > $F_{(0.05, 4, 25)}$ = 2.87 and p << 0.01. Table 9.3 lists the *post hoc* pairwise ANOVA tests results for completion time of different calibration methods. The test results indicate that all of the pairwise comparison of calibration completion time are statistically significant with most having very small

p-values. The difference between the completion time of DRC-SP and DRC-Mouse is still statistically significant, with comparatively higher p-value ($p = 0.039$) than that of other comparisons.

	SPAAM	SMC	DRC-None	DRC-Mouse	DRC-SP	DRC-MP
SPAAM	-	0.017	-	<<0.01	<<0.01	<<0.01
SMC	-	-	-	<<0.01	<<0.01	<<0.01
DRC-None	-	-	-	-	-	-
DRC-Mouse	-	-	-	-	0.039	0.0002
DRC-SP	-	-	-	-	-	0.0003
DRC-MP	-	-	-	-	-	-

Table 9.3 P-values of pairwise ANOVA for completion time of different calibration methods

A one-way ANOVA is conducted to test the effect of different calibration methods on the mental workload (NASA TLX ratings). The effect was statistically significant because $F_0 = 43.10 > F_{(0.05, 4, 25)} = 2.87$ and $p << 0.01$.

9.6.3 Discussion

The above ANOVA tests verify the statistical significance of the experimental findings described in Section 9.6.1. This section explores the possible reasons for these results and provides further considerations of the accuracy and usability of the different DRC options and the other OSTHMD calibration methods.

1) Accuracy

Except for DRC-None, the other DRC options demonstrate very good accuracy in

practice. The experiment results also verify the estimated error range of the previous

sensitivity analysis of the DRC-None option. The DRC-Multiple Point option provides

better accuracy and more reliable performance than any other options and the traditional

calibration methods. This validates the dissertation statement that the DRC method is

able to provide very precise calibration for high-accuracy applications. The registration

errors of DRC-Single Point and DRC-Mouse are close to those of SPAAM or SMC

methods, but the DRC options are more stable because of the approximate calibration

from phase I. The DRC-None option uses the offline phase I results directly, but its

registration error is still acceptable for some low-accuracy applications.

2) Usability

In this experiment, the usability of a calibration method is evaluated by the

completion time and the mental workload. It is not easy to provide a comprehensive

usability evaluation because there is no established method to combine the effects of

completion time and NASA TLX score. Interestingly, no calibration method has shortest

completion time and least mental workload simultaneously. The DRC-Single Point

method needs the shortest time to complete because of only one assignment for each eye,

but its mental workload is slightly more than the DRC-Mouse option, which requires no

point alignment at all. It seems that most experiment participants prefer traditional

mouse-based interaction rather than image point alignment. The DRC Multiple Point

option is less user-friendly than other DRC options, but it is still quicker and easier than

SPAAM or SMC because of fewer alignments.

9.7   Summary

The thesis statement about the accuracy and usability performance of DRC has been verified by the experimental evaluation process described in this chapter. The results are consistent with the theoretical analysis and the previous studies. The DRC method demonstrates its advantages over other calibration approaches by providing multiple options to satisfy the requirements of accuracy and usability.

The analysis of the experimental results demonstrates the following features of the different DRC user options:

1)   Directly Using DRC I calibration is an approximate offline calibration for low-accuracy OSTHMD systems.

2)   DRC Manual Adjustment using mouse is the easiest calibration method for most users. This procedure provides decent registration accuracy using simple mouse-based interaction.

3)   DRC Single Point Alignment is the quickest calibration procedure. Depending on the performance of the single point alignment, the calibration results might not be consistently stable, but very high accuracy can be achieved.

4)   DRC Multiple Point Alignment provides the highest accuracy with acceptable user efforts and more reliable computation. Precise alignments of image points with

physical markers are still required to achieve the best accuracy. This option is an

effective calibration procedure to replace the SPAAM or Stylus-Maker calibration

method in the professional OSTHMD-based AR systems which require high accuracy for

registration.

Chapter 10 Conclusions and Future Work

This chapter summarizes the outcome of this thesis and also proposes research topics for future work. This thesis is the first in-depth and comprehensive research on OSTHMD calibration for AR. The outcome of this thesis includes a new computational model, an effective and flexible procedure, and the relevant theoretical and experimental verifications. The Display Relative Calibration (DRC) proposed in this dissertation has been shown to be an accurate and flexible calibration method for OSTHMD.

10.1 Conclusions

The complexity and limited accuracy of existing calibration methods has limited the utilization and development of OSTHMD systems and applications. The OSTHMD calibration problem is different from that of other imaging systems because of the virtual image plane and the user interaction requirements. Therefore, calibration of an OSTHMD requires a comprehensive analysis and design instead of a mathematical solution.

This dissertation describes the computation, procedures, error analysis and experimental evaluation of the DRC method for OSTHMD-based AR systems. DRC is an innovative two-phase OSTHMD calibration method with proven high accuracy and flexible usability. This method demonstrated better performance than traditional calibration methods in practice, and its multiple options can be adapted for various

OSTHMD-based systems and applications.

In summary, this dissertation presents research results that lead to the following conclusions:

1) The DRC computational model determines the role of user's eyes in the OSTHMD imaging system and viewing process and establishes a virtual camera model with systematic parameters and user-dependent values. The theoretical analysis and experimental evaluation in this research have shown that this strategy is more accurate and flexible than existing approaches, which considers OSTHMD imaging system and user's eyes as a whole virtual camera model. These traditional approaches have been shown to be inaccurate and unstable compared with the DRC method because the solution to the calibration parameters is based on large number of user alignments of virtual and world points.

2) In the DRC method, the OSTHMD virtual camera parameters are solved in a two-phase rough-to-refined approach. This method combines offline image-based computation and online user procedures to provide reliable and flexible calibration results.

3) The DRC method takes into account the optical aberration and determines the attributes of the curved OSTHMD image plane using triangulation. The information about the image plane depth and the curve equation is important to improve calibration accuracy of DRC and to analyze the registration errors.

4) Analysis and estimation of the OSTHMD error sources explains that the calibration error contributes a considerable part to the total registration error. The high accuracy of the DRC method owes to the specially designed algorithm and computational optimization targeting to decrease these error sources.

5) This research demonstrates that there is a tradeoff between accuracy and usability for a calibration procedure. Therefore it would be better to design or choose a calibration procedure based on the application requirements and target user group instead of designing a "fit-all" best calibration algorithm. The DRC has advantages over other traditional methods because it has multiple options with different levels of accuracy and usability for different applications.

10.2 Future work

The following research topics can be further investigated in the future work on the OSTHMD calibration.

1) Correction of curved image plane

Both the sensitivity analysis and experimental evaluation have indicted that the registration error introduced by the curved image plane is very small, especially when the image plane is comparatively far from the user's eye. But a correction of this curved plane is still expected to reduce the error caused by the offsets of the user's eyes. Most previous studies use texture mapping to correct optical distortion such as the method

proposed by Watson et al. [109] and polar-based texture mapping meshes by Bax [140]. These techniques suffer from slow performance and A real-time, precise correction scheme for OSTHMD curved image plane requires further investigation.

2) Design of user alignment method

In the DRC and other existing calibration methods, the user alignment is done by moving and overlaying an observed graphical mark (usually a crosshair) in the OSTHMD view onto a physical target. The user needs to start the data collection procedure at the exact moment when the alignment is reached. This is not an easy task for most untrained normal users because most of them cannot hold the head still without hardware support. A new alignment method or a computational optimization needs to be established to improve the performance of user alignment. In that way, both the registration accuracy and usability can be significantly improved.

BIBLIOGRAPHY

[1]   S. K. Feiner, "Augmented reality: A new way of seeing," *Scientific American*, vol. 286, no. 4, pp. 48-56, 2002.

[2]   S. Feiner, B. Macintyre, and D. Seligmann, "Knowledge-based Augmented Reality," *Communications of ACM*, vol. 36, no. 7, pp. 53-62, 1993.

[3]   R. Azuma, "A Survey of Augmented Reality," *Presence: Teleoperators and Virtual Environments*, vol. 6, no. 4, pp. 355-385, 1997.

[4]   P. Milgram, H. Takemura, A. Utsumi, and F. Kishino, "Augmented Reality: A Class of Displays on the Reality-Virtuality Continuum," *Proceedings of SPIE Conference on Telemanipulator and Telepresence Technologies (Vol. 2351)*, pp. 282-292, Boston, MA, USA, 1994.

[5]   D. Schmalstieg, A. Fuhrmann, G. Hesina, Z. Szalavari, L. M. Encarnacao, M. Gervautz, and W. Purgathofer, "The Studierstube Augmented Reality Project," *Presence: Teleoperators and Virtual Environments*, vol. 11, no. 1, pp. 33 - 54, 2002.

[6]   J. Rozier, K. Karahalios, and J. Donath, "Hear&There: An Augmented Reality System of Linked Audio," *Proceedings of 2000 International Conference on Auditory Display (ICAD '00)*, pp., Atlanta, GA, USA, 2000.

[7]   K. Lyons, M. Gandy, and T. Starner, "Guided by Voices: An Audio Augmented Reality System," *Proceedings of 2000 International Conference on Auditory Display (ICAD '00)*, pp., Atlanta, GA, USA, 2000.

[8]   H. Regenbrecht, G. Baratoff, and M. Wagner, "A tangible AR desktop environment " *Computers & Graphics*, vol. 25, no. 5, pp. 755-763, 2001.

[9]   E. Gelenbe, K. Hussainb, and V. Kaptana, "Simulating Autonomous Agents in Augmented Reality " *Journal of Systems and Software*, vol. 74, no. 3, pp. 255-268 2004.

[10]   D. C. Foyle, A. D. Andre, and B. L. Hooey, "Situation Awareness in an Augmented Reality Cockpit: Design, Viewpoints and Cognitive Glue,"

Proceedings of 11th International Conference on Human Computer Interaction (HCI '05), pp., Las Vegas, NV, USA, 2005.

[11]   A. Tang, C. Owen, F. Biocca, and W. Mou, "Comparative Effectiveness of Augmented Reality in Object Assembly," *Proceedings of 2003 Conference on Human Factors and Computing Systems (CHI '03)*, pp. 73-80, Ft. Lauderdale, Florida, USA, 2003.

[12]   K. H. Tang, "Spatial Information Display Framework for MobileAugmented Reality Interfaces (Ph.D. Dissertation)," Department of Computer Science and Engineering, Michigan State University, East Lansing, MI, 2005

[13]   J. Lanier, "Virtually There: Three-dimensinal tele-immersion may eventually bring the world to your desk," *Scientific American*, vol. 284, no. 4, pp. 66-72, 2001.

[14]   T. Starner, S. Mann, B. Rhodes, J. Levine, J. Healey, D. Kirsch, R. W. Picard, and A. Pentland, "Augmented Reality Through Wearable Computing," *Presence: Special Issue on Augmented Reality*, no., 1997.

[15]   R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre, "Recent Advances in Augmented Reality," *IEEE Computer Graphics and Applications*, vol. 21, no. 6, pp. 34-47, 2001.

[16]   M. Figl, C. Ede, W. Birkfellner, J. Hummel, R. Hanel, and H. Bergmann, "Design and Automatic Calibration of a Head Mounted Operating Binocular for Augmented Reality Applications in Computer Aided Surgery," *Proceedings of SPIE Volume 5744 (Medical Imaging 2005: Visualization, Image-Guided Procedures, and Display)*, pp. 726-730, San Jose, CA, USA, 2005.

[17]   M. Figl, C. Ede, W. Birkfellner, J. Hummel, R. Seemann, and H. Bergmann, "Automatic Calibration of an Optical See Through Head Mounted Display for Augmented Reality Applications in Computer Assisted Interventions," *Proceedings of SPIE Vol. 5367 (Medical Imaging 2004: Visualization, Image-Guided Procedures, and Display)*, pp. 59-64, San Diego, CA, USA, 2004.

[18]   M. Figl, C. Ede, W. Birkfellner, J. Hummel, R. Seemann, and H. Bergmann, "Automatic Calibration of an Head Mounted Display for Augmented Reality Applications in Computer Assisted Surgery," *Proceedings of International Workshop on Augmented environments for Medical Imaging and*

Computer-aided Surgery (AMI-ARCS '04), pp. 49-56, Rennes, France, 2004.

[19]   T. P. Caudell and D. W. Mizell, "Augmented Reality: An Application of Heads-Up Display Technology to Manual Manufacturing Processes," *Proceedings of 25th Hawaii International Conference on System Sciences*, pp. 659-669, Kauai, HI, USA, 1992.

[20]   A. L. Janin, D. W. Mizell, and T. P. Caudell, "Calibration of Head-Mounted Displays for Augmented Reality Applications," *Proceedings of Virutal Reality Annual International Symposium (VRAIS '93)*, pp. 246-255, Seattle, WA, USA, 1993.

[21]   D. Sims, "New Realities in Aircraft Design and Manufacture," *IEEE Computer Graphics and Application*, vol. 14, no. 2, pp. 91, 1994.

[22]   M. Tuceryan, D. S. Greer, R. T. Whitaker, D. E. Breen, C. Crampton, E. Rose, and K. H. Ahlers, "Calibration Requirements and Procedures for a Monitor-Based Augmented Reality System," *IEEE Transactions on Visualization and Computer Graphics*, vol. 1, no. 3, pp. 255-273, 1995.

[23]   H. Regenbrecht, G. Baratoff, and W. Wilke, "Augmented Reality Projects in the Automotive and Aerospace Industries " *IEEE Computer Graphics and Application*, vol. 25, no. 6, pp. 48-56, 2005.

[24]   N. Navab, "Industrial Augmented Reality(IAR): Challenges in Design and Commercialization of Killer Apps," *Proceedings of 2nd IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR '03)*, pp. 2-6, Tokyo, Japan, 2003.

[25]   S. Julier, Y. Baillot, M. Lanzagorta, D. Brown, and L. Rosenblum, "BARS: Battlefield Augmented Reality System," *Proceedings of NATO Symposium on Information Processing Techniques for Military Systems*, pp., Istanbul, Turkey, 2000.

[26]   S. Julier, R. King, B. Colbert, J. Durbin, and L. Rosenblum, "The Software Architecture of a Real-Time Battlefield Visualization Virtual Environment," *Proceedings of IEEE Virtual Reality Conference 1999 (VR 1999)*, pp. 29-36, Houston, Texas, USA, 1999.

[27]   S. Julier, M. Lanzagorta, Y. Baillot, and D. Brown, "Information Filtering for

Mobile Augmented Reality," *Proceedings of IEEE and ACM International Symposium on Augmented Reality, 2000. (ISAR 2000)*, pp. 3-11, München, Germany, 2000.

[28]   C. E. Rash, *Helmet-Mounted Displays: Design Issues for Rotary-Wing Aircraft*: SPIE PRESS, 2001.

[29]   F. Mulholland, "Helmet-Mounted Display Accuracy in the Aircraft Cockpit," *Proceedings of SPIE Conference on Helmet- and Head-Mounted Displays VII (Vol. 4711)*, pp. 145-151, Orlando, FL, USA, 2002.

[30]   D. Wilkins, "Helmet Mounted Displays in Long Range Target Visual Acquisition," *Proceedings of SPIE Conference on Helmet- and Head-Mounted Displays IV (Vol. 3689)*, pp. 174-177, Orlando, FL, USA, 1999.

[31]   L. T. Christian, L. Hart, and S. D. Fechtig, "Helmet Mounted Display as a Primary Flight Reference," *Proceedings of SPIE Conference on Helmet- and Head-Mounted Displays VII (Vol. 4711)*, pp. 389-394, Orlando, FL, USA, 2002.

[32]   G. K. Edgar, J. C. D. Pope, and I. R. Craig, "Visual Accommodation Problems with Head-up and Helmet-mounted Displays," *Displays*, vol. 15, no. 2, pp. 68-75, 1994.

[33]   R. Cavallaro, "The FoxTrax Hockey Puck Tracking System," *IEEE Computer Graphics and Applications*, vol. 17, no. 2, pp. 6-12, 1997.

[34]   C. B. Owen, J. Zhou, K. H. Tang, and F. Xiao, "Augmented Imagery for Digital Video Applications," *Handbook of Video Database: Design and Applications*, B. Furht and O. Marques, Eds. Boca Raton, FL, USA: CRC Press, 2004, pp. 319-348.

[35]   T. Ohshima, K. Satoh, H. Yamamoto, and H. Tamura, "RV-Border Guards: A Multi-Player Mixed Reality Entertainment," *Transaction of Virtual Reality Society of Japn*, vol. 4, no. 4, pp. 699-705, 1999.

[36]   B. Thomas, B. Close, J. Donoghue, J. Squires, P. d. Bondi, M. Morris, and W. Piekarski, "ARQuake: An Outdoor/Indoor Augmented Reality First Person Application," *Proceedings of 4th International Symposium on Wearable Computers (ISWC'00)*, pp. 139-146, Atlanta, Georgia, 2000.

[37] R. Azuma, "Tracking Requirements for Augmented Reality," *Communications of the ACM*, vol. 36, no. 7, pp. 50-51, 1993.

[38] G. Roberts, A. Evans, and A. Dodson, "Integrating GPS, INS and Augmented Reality for sub-surface Visualization," *Proceedings of 13th International Technical Meeting of the Satellite Division of the Institute of Navigation*, pp., Salt Lake City, UT, USA, 2001.

[39] R. Behringer, "Registration for an augmented reality system enhancing the situational awareness in an outdoor scenario," *Proceedings of SPIE Conference on Enhanced and Synthetic Vision 1999*, pp. 231-242, Orlando, FL, USA, 1999.

[40] D. K. Bhatnagar, "Position trackers for Head Mounted Display Systems: A Survey," Department of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA, Technical Report TR93-010, 1993.

[41] V. V. Kindratenko, "A Survey of Electromagnetic Position Tracker Calibration Techniques," *Virtual Reality: Research, Development, and Applications (The VRS Journal)*, vol. 5, no. 3, pp. 169-182, 2000.

[42] J. P. Rolland, Y. Baillot, and A. A. Goon, "A Survery of Tracking Technology for Virtual Environments," *Fundamentals of Wearable Computers and Augmented Reality*, W. Barfield and T. Caudell, Eds. Mahwah, NJ, USA: Lawrence Erlbaum, 2001, pp. 67-112.

[43] D. Koller, G. Kinker, E. Rose, D. Breen, R. Whitaker, and M. Tuceryan, "Real-time Vision-Based Camera Tracking for Augmented Reality Applications," *Proceedings of Symposium on Virtual Reality Software and Technology (VRST-97)*, pp., Lausanne, Switzerland, 1997.

[44] J. Park, B. Jiang, and U. Neumann, "Vision-based Pose Computation: Robust and Accurate Augmented Reality Tracking," *Proceedings of IEEE and ACM Interntional Workshop on Augmented Reality (IWAR '99)*, pp. 3-12, San Francisco, California, 1999.

[45] Y. Genc, M. Tuceryan, A. Khamene, and N. Navab, "Optical See-through Calibration with Vision-Based Trackers: Propagation of Projection Matrices," *Proceedings of IEEE and ACM International Symposium on Augmented Reality (ISAR '01)*, pp., New York, NY, USA, 2001.

[46] S. You, U. Neumann, and R. Azuma, "Hybrid Inertial and Vision Tracking for Augmented Reality Registration," *Proceedings of IEEE Virtual Reality Conference 1999 (VR 1999)*, pp. 260-269, Houston, Texas, 1999.

[47] X. Zhang, S. Fronz, and N. Navab, "Visual Marker Detection and Decoding in AR Systems: A Comparative Study," *Proceedings of 1st IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'02)*, pp., Darmstadt, Germany, 2002.

[48] C. B. Owen, F. Xiao, and P. Middlin, "What is the Best Fiducial," *Proceedings of IEEE International Augmented Reality Toolkit Workshop*, pp. 98-105, Darmstadt, Germany, 2002.

[49] H. Kato and M. Billinghurst, "Marker Tracking and HMD Calibration for a Video-Based Augmented Reality Conferencing System," *Proceedings of IEEE and ACM Interntional Workshop on Augmented Reality (IWAR '99)*, pp., San Francisco, California, 1999.

[50] G. Simon, V. Lepetit, and M.-O. Berger, "Computer Vision Methods for Registration: Mixing 3D Knowledge and 2D Correspondences for Accurate Image Composition," *Proceedings of IEEE International Workshop on Augmented Reality (IWAR '98)*, pp., San Francisco, CA, USA, 1998.

[51] R. Azuma, J. W. Lee, B. Jiang, J. Park, S. You, and U. Neumann, "Tracking in Unprepared Environments for Augmented Reality Systems," *Computers and Graphics*, vol. 23, no. 6, pp. 787-793, 1999.

[52] R. T. Azuma, "The Challenge of Making Augmented Reality Work Outdoors," *Mixed Reality: Merging Real and Virtual Worlds (Proceedings of 1st International Symposium of Mixed Reality (ISMR '99))*, Y. Ohta and H. Tamura, Eds. Yokohama, Japan: Ohmsha, Ltd. & Springer-Verlag, 1999, pp. 379-390.

[53] U. Neumann and S. You, "Natural Feature Tracking for Augmented Reality," *IEEE Transactions on Multimedia*, vol. 1, no. 1, pp. 53-64, 1999.

[54] U. Neumann, S. You, Y. Cho, J. Lee, and J. Park, "Augmented Reality Tracking in Natural Environments," *Proceedings of International Symposium on Mixed Realities*, pp., Tokyo, Japan, 1999.

[55] M. Segal and K. Akeley, "The OpenGL Graphics System: A Specification

(Version 1.5)." OpenGL ARB, 2003.

[56] J. Neider and T. Davis, *OpenGL Programming Guide: The Official Guide to Learning OpenGL*, 2 ed. Boston, MA, USA: Addison-Wesley Publishing Company, 1997.

[57] F. S. Hill Jr., *Computer Graphics Using Open GL*, 2 ed: Prentice Hall, 2000.

[58] J. D. Foley, A. v. Dam, S. K. Finer, and J. F. Hughes, *Computer Graphics: Principles and Practice*, 2nd in C ed. Reading, MA: Addison-Wesley, 1996.

[59] H. Fuchs and J. Ackerman, "Displays for Augmented Reality: Historical Remarks and Future Prospects," *Mixed Reality: Merging Real and Virtual Worlds (Proceedings of 1st International Symposium of Mixed Reality - ISMR '99)*, Y. Ohta and H. Tamura, Eds. Yokohama, Japan: Ohmsha, Ltd. & Springer-Verlag, 1999, pp. 31-40.

[60] M. Deering, "High Resolution Virtual Reality," *Proceedings of 19th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '92)*, pp. 195-202, Chicago, Illinois, United States, 1992.

[61] R. Raskar, G. Welch, and H. Fuchs, "Spatially Augmented Reality," *Proceedings of 1st International Workshop on Augmented Reality*, pp., San Francisco, CA, USA, 1998.

[62] R. Raskar, G. Welch, M. Cutts, A. Lake, L. Stesin, and H. Fuchs, "The Office of the Future: A Unified Approach to Image-Based Modeling and Spatially Immersive Displays," *Proceedings of 25th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '98)*, pp. 179-188, Orlando, Florida, USA, 1998.

[63] R. Raskar, G. Welch, and W.-C. Chen, "Table-Top Spatially-Augmented Reality: Bringing Physical Models to Life with Projected Imagery," *Proceedings of 2nd IEEE and ACM International Workshop on Augmented Reality*, pp. 64-71, San Francisco, California, 1999.

[64] O. Bimber and R. Raskar, *Spatail Augmented Reality*. Wellesley, MA, USA, 2005.

[65] E. Patrick, D. Cosgrove, A. Slavkovic, J. A. Rode, T. Verratti, and G. Chiselko,

"Using a Large Projection Screen as an Alternative to Head-Mounted Displays for Virtual Environments," *Proceedings of SIGCHI conference on Human factors in computing systems*, pp. 478-485, The Hague, The Netherlands, 2000.

[66]   I. E. Sutherland, "A Head-Mounted Three Dimensional Display," *Proceedings of Fall Joint Computer Conference*, pp. 757-764, San Fransisco, CA, USA, 1968.

[67]   M. Bajura, "Camera Calibration for Video See-Through Head-Mounted Display," Department of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA, Technical Report TR93-048, July 1993.

[68]   E. K. Edwards, J. P. Rolland, and K. P. Keller, "Video see-through design for merging of real and virtual environments," *Proceedings of IEEE 1993 Virtual Reality Annual International Symposium (VRAIS '93)*, pp. 223-233, Seattle, WA , USA, 1993.

[69]   N. Yokoya, H. Takemura, T. Okuma, and M. Kanbara, "Stereo Vision Based Video See-through Mixed Reality," *Mixed Reality: Merging Real and Virtual Worlds (Proceedings of 1st International Symposium of Mixed Reality (ISMR '99))*, Y. Ohta and H. Tamura, Eds. Yokohama, Japan: Ohmsha, Ltd. & Springer-Verlag, 1999, pp. 131-145.

[70]   J. P. Rolland and H. Fuchs, "Optical Versus Video See-Through Head-Mounted Displays in Medical Visualization," *Presence: Teleoperators and Virtual Environments*, vol. 9, no. 3, pp. 287 - 309, 2000.

[71]   J. P. Rolland, "Wide-angle, off-axis, see-through head-mounted display," *Optical Engineering*, vol. 39, no. 7, pp. 1760-1767, 2000.

[72]   P. K. Manhart, R. J. Malcolm, and J. G. Frazee, ""Augeye" a compact, solid Schmidt optical relay for helmet mounted displays," *Proceedings of IEEE 1993 Virtual Reality Annual International Symposium (VRAIS '93)*, pp. 234-245, Seattle, WA , USA, 1993.

[73]   K. Kiyokawa, Y. Kurata, and H. Ohno, "An Optical See-through Display for Mutual Occlusion of Real and Virtual Environments," *Proceedings of IEEE and ACM International Symposium on Augmented Reality (ISAR '00)*, pp. 60-67, Munich, Germany, 2000.

[74]   E. McGarrity and M. Tuceryan, "A Method for Calibrating See-through

Head-mounted Displays for AR," *Proceedings of IEEE and ACM Interntional Workshop on Augmented Reality (IWAR '99)*, pp. 75-84, San Francisco, CA, USA, 1999.

[75] A. Fuhrmann, D. Schmalstieg, and W. Purgathofer, "Fast Calibration for Augmented Reality," *Proceedings of ACM symposium on Virtual Reality Software and Technology*, pp., London, UK, 1999.

[76] A. Fuhrmann, D. Schmalstieg, and W. Purgathofer, "Practical Calibration Procedures for Augmented Reality," *Proceedings of Eurographics Workshop on Virtual Environments (EGVE00)*, pp., Amsterdam, The Netherlands, 2000.

[77] A. L. Fuhrmann, R. Splechtna, and J. Poikryl, "Comprehensive Calibration and Registration Procedures for Augmented Reality," *Proceedings of 7th Eurographics Workshop on Virtual Environments(EG VE '01)*, pp., Stuttgart, Germany, 2001.

[78] M. Figl, W. Birkfellner, J. Hummel, C. Ede, R. A. Hanel, and H. Bergmann, "Calibration of an optical see-through head-mounted display with variable zoom and focus for applications in computer-assisted interventions," *Proceedings of SPIE Medical Imaging 2003: Visualization, Image-Guided Procedures, and Display (SPIE Vol. 5029)*, pp. 618-623, San Diego, CA, USA, 2003.

[79] T. Oishi and S. Tachi, "Calibration Method of Visual Parameters for See-through Head-Mounted Display," *Proceedings of IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI '94)*, pp. 447-454, Las Vegas, NV, USA, 1994.

[80] C. Gao, H. Hua, and N. Ahuja, "Easy Calibration of a Head-Mounted Projective Display for Augmented Reality Systems," *Proceedings of IEEE Virtual Reality Conference 2003 (VR 2003)*, pp. 53-60, Los Angeles, CA, 2003.

[81] V. A. Summers, K. S. Booth, T. Calvert, E. Graham, and C. L. MacKenzie, "Calibration for Augmented Reality Experimental Testbeds," *Proceedings of 1999 ACM Symposium on Interactive 3D Graphics*, pp. 155-162, Atlanta, Georgia, United States, 1999.

[82] G. Klinker, D. Stricker, and D. Reiners, "Augmented Reality: A Balancing Act Between High Quality and Real-Time Constraints," *Mixed Reality: Merging Real and Virtual Worlds (Proceedings of 1st International Symposium on Mixed*

Reality (ISMR '99)), Y. Ohta and H. Tamura, Eds. Yokohama, Japan: Ohmsha, Ltd. & Springer-Verlag, 1999, pp. 325-346.

[83]   T. Oishi and S. Tachi, "Methods to Calibrate Projection Transformation Parameters for See-Through Head-Mounted Displays," *Presence: Teleoperators and Virtual Environments*, vol. 5, no. 1, pp. 122 - 135, 1996.

[84]   H. Hua, C. Gao, L. D. Brown, N. Ahuja, and J. P. Rolland, "A Testbed for Precise Registration, Natural Occlusion and Interaction in an Augmented Environment Using a Head-Mounted Projective Display (HMPD)," *Proceedings of IEEE Virtual Reality Conference 2002 (VR 2002)*, pp. 81-89, Orlando, Florida, USA, 2002.

[85]   H. Hua, C. Gao, and N. Ahuja, "Calibration of a Head-Mounted Projective Display for Augmented Reality Systems," *Proceedings of 1st IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'02)*, pp. 176-85, Darmstadt, Germany, 2002.

[86]   R. Grasset, X. Decoret, and J.-D. Gascuel, "Augmented Reality Collaborative Environment : Calibration and Interactive Scene Editing," *Proceedings of 3th Virtual Reality International Conference (Laval Virtual 2001)*, pp., Laval, France, 2001.

[87]   R. Azuma and G. Bishop, "Improving Static and Dynamic Registration in an Optical See-through HMD," *Proceedings of 21st Annual Conference on Computer Graphics and Interactive techniques (SIGGRAPH '94)*, pp. 197-204, Orlando, FL, 1994.

[88]   M. Tuceryan and N. Navab, "Single Point Active Alignment Method (SPAAM) for Optical See-through HMD Calibration for AR," *Proceedings of IEEE and ACM International Symposium on Augmented Reality (ISAR '00)*, pp. 149-158, Munich, Germany, 2000.

[89]   M. Tuceryan, Y. Genc, and N. Navab, "Single-Point Active Alignment Method (SPAAM) for Optical See-Through HMD Calibration for Augmented Reality," *Presence: Teleoperators and Virtual Environments*, vol. 11, no. 3, pp. 259 - 276, 2002.

[90]   Y. Genc, F. Sauer, F. Wenzel, M. Tuceryan, and N. Navab, "Optical See-through HMD Calibration: A Stereo Method Validated with a Video See-Through

System," *Proceedings of IEEE and ACM International Symposium on Augmented Reality (ISAR '00)*, pp. 165-174,, Munich, Germany, 2000.

[91]   Y. Genc, M. Tuceryan, and N. Navab, "Practical Solutions for Calibration of Optical See-Through Devices," *Proceedings of 1st IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'02)*, pp. 169-175, Darmstadt, Germany, 2002.

[92]   Y. Baillot, S. J. Julier, D. Brown, and M. A. Livingston, "A Tracker Alignment Framework for Augmented Reality," *Proceedings of 2nd IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR '03)*, pp. 142-150, Tokyo, Japan, 2003.

[93]   W. Robinett and J. P. Rolland, "A Computational Model for the Stereoscopic Optics of a Head-Mounted Display," *Presence: Teleoperators and Virtual Environments*, vol. 1, no. 1, pp. 45-62, 1992.

[94]   W. Robinett and R. Holloway, "The Visual Display Transformation for Virtual Reality," *Presence: Teleoperators and Virtual Environments*, vol. 4, no. 1, pp. 1 - 23, 1995.

[95]   D. J. Schroeder, *Astronomical optics*. San Diego, CA: Academic Press, 2000.

[96]   J.-F. V. Gomez, G. Simon, and M.-O. Berger, "Calibration Errors in Augmented Reality: A Practical Study," *Proceedings of Fourth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'05)*, pp. 154-163, Vienna, Austria, 2005.

[97]   C. B. Owen, J. Zhou, A. Tang, and F. Xiao, "Display-Relative Calibration for Optical See-Through Head-Mounted Displays," *Proceedings of Third IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'04)*, pp. 70-78, Arlington, VA, USA, 2004.

[98]   R. Y. Tsai, "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses," *IEEE Journal of Robotics and Automation*, vol. 3, no. 4, pp. 323-344, 1987.

[99]   R. Y. Tsai, "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology using Off-the-shelf TV Cameras and Lenses," IBM T. J. Watson Research Center, Yorktown Heights, NY, USA RC 11413 (#51342),

Septemeber 1985.

[100] J. Weng, P. Cohen, and M. Herniou, "Camera Calibration with Distortion Models and Accuracy Evaluation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 10, pp. 965-980, 1992.

[101] S. Ganapathy, "Decomposition of Transformation Matrices for Robot Vision," *Proceedings of 1st IEEE International Conference on Robotics and Automation*, pp. 130-139, Atlanta, Georgia, USA, 1984.

[102] O. D. Faugeras and G. Toscani, "The Calibration Problem for Stereo," *Proceedings of IEEE Conference on Computer Vision and Patter Recognition (CVPR '86)*, pp. 15-20, Miami, FL, USA, 1986.

[103] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2 ed. Cambridge, UK: Cambridge University Press, 2004.

[104] J. Weng, T. S. Huang, and N. Ahuja, "Motion and Structure From Two Perspective Views: Algorithms, Error Analysis, and Error Estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 5, pp. 451-476, 1989.

[105] J. E. Dennis, "Non-Linear Least Squares and Equations," *Proceedings of Conference on the State of the Art in Numerical Analysis: The state of the art in numerical analysis*, pp. 269-312, University of York, York, UK, 1976.

[106] H. C. Longuet-Higgins, "A Computer Algorithm for Reconstructing a Scene from Two Projections," *Nature*, vol. 293, no. 10, pp. 133-135, 1981.

[107] R. I. Hartley and P. Sturm, "Triangulation," *Computer Vision and Image Understanding*, vol. 68, no. 2, pp. 146-157, 1997.

[108] J. Ma, J. M. Hollerbach, and I. W. Hunter, "Optical Design for a Head-Mounted Display," *Presence: Teleoperators and Virtual Environments*, vol. 2, no. 3, pp. 185 - 202, 1993.

[109] B. A. Watson and L. F. Hodges, "Using texture maps to correct for optical distortion in head-mounted displays," *Proceedings of Virtual Reality Annual International Symposium (VRAIS'95)*, pp. 172-178, Research Triangle Park, North Carolina, USa, 1995.

[110] P. Min and H. Jense, "Interactive Stereoscopy Optimization for Head-Mounted Displays," *Proceedings of SPIE Conference on Stereoscopic Displays and Virtual Reality Systems (Vol. 2177)*, pp. 306-316, San Jose, CA, USA, 1994.

[111] R. L. Holloway, "Registration Error Analysis for Augmented Reality," *Presence: Teleoperators and Virtual Environments*, vol. 6, no. 4, pp. 413-432, 1997.

[112] R. L. Holloway, "Registration Errors in Augmented Reality Systems (Ph.D. Thesis)," Department of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, NC, 1995

[113] B. MacIntyre and S. J. Julier, "Estimating and Adapting to Registration Errors in Augmented Reality Systems," *Proceedings of IEEE Virtual Reality Conference 2002 (VR 2002)*, pp. 73-80, Orlando, Florida, 2002.

[114] J. P. Babbage, "Evaluation of Rendering Error through Optical Modeling of the Display Eye System (MS Thesis)," Department of Computer Science and Engineering, Michigan State University, East Lansing, MI, 2003

[115] C.-P. Lu, G. D. Hager, and E. Mjolsness, "Fast and Globally Convergent Pose Estimation from Video Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 6, pp. 610-622, 2000.

[116] S. D. Blostein and T. S. Huang, "Correction to "Error Analysis in Stereo Determination of 3-D Point Positions"," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 5, pp. 765, 1988.

[117] S. D. Blostein and T. S. Huang, "Error Analysis in Stereo Determination of 3-D Point Positions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, no. 6, pp. 752-765, 1987.

[118] A. Tang, J. Zhou, and C. Owen, "Evaluation of Calibration Procedures for Optical See-Through Head-Mounted Displays," *Proceedings of 2nd IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR '03)*, pp. 161-168, Tokyo, Japan, 2003.

[119] E. McGarrity, Y. Genc, M. Tuceryan, C. Owen, and N. Navab, "A New System for Online Quantitative Evaluation of Optical See-through Augmentation," *Proceedings of IEEE and ACM International Symposium on Augmented Reality (ISAR '01)*, pp., New York, NY, 2001.

[120] M. A. Fischler and R. C. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Apphcatlons to Image Analysis and Automated Cartography," *Communications of ACM*, vol. 24, no. 6, pp. 381-395, 1981.

[121] S. K. Card, J. D. Mackinlay, and G. G. Robertson, "The Design Space of Input Devices," *Proceedings of 1990 ACM SIGCHI Conference on Human Factors in Computing Systems (CHI '90)*, pp. 117-124, Seattle, WA, USA, 1990.

[122] E. McGarrity, M. Tuceryan, C. Owen, Y. Genc, and N. Navab, "Evaluation of Optical See-through Systems," *Proceedings of Euroimage International Conference on Augmented, Virtual Environments and Three-Dimensional Imaging (ICAV3D '01)*, pp., Mykonos, Greece, 2001.

[123] E. S. McGarrity, "Evaluation of Calibration for Optical See-through Augmented Reality Systems (Master's Thesis)," Department of Computer Science and Engineering, Michigan State University, East Lansing, MI, 2001

[124] R. Pausch, M. A. Shackelford, and D. Proffitt, "A User Study Comparing Head-Mounted and Stationary Displays," in *IEEE Symposium on Research Frontiers in Virtual Reality*. San Jose, CA, 1993.

[125] D. Hix, J. E. S. II, J. L. Gabbard, M. McGee, J. Durbin, and T. King, "User-Centered Design and Evaluation of a Real-Time Battlefield Visualization Virtual Environment," *Proceedings of 1999 IEEE Virtual Reality Conference (VR '99)*, pp., Houston, TX, USA, 1999.

[126] D. A. Bowman, J. L. Gabbard, and D. Hix, "A Survey of Usability Evaluation in Virtual Environments: Classification and Comparison of Methods," *Presence: Teleoperators and Virtual Environments*, vol. 11, no. 4, pp. 404 - 424, 2002.

[127] S. G. Hart and L. E. Staveland, "Development of NASA-TLX(Task Load Index): Results of Empirical and Theoeretical Research," *Human Mental Workload*, P. A. Hancocks and N. Meshkati, Eds.: Elsevier Science Publisher B. V. (North-Holland), 1988, pp. 139-183.

[128] D. C. Montgomery, *Design and Analysis of Experiments*, 5 ed. New York, NY: John Wiley & Sons, Inc., 2001.

[129] J. Jaccard and M. A. Becker, *Statistics for the Behaviroal Sciences*, 4 ed. Belmont, CA: Wadsworth/Thomson Learning, 2002.

[130] F. J. Gravetter and L. B. Wallnau, *Statistics for the Behavioral Sciences*, 5 ed. Stamford, CT, USA: Wadsworth/Thomsan Learning, 2000.

[131] R. V. Lenth, "Some Practical Guidelines for Effective Sample-Size Determination," *The American Statistician*, vol. 55, no. 3, pp. 187-193, 2001.

[132] J. Cohen, *Statistical power analysis for the behavioral sciences*, 2 ed. Hillsdale, NJ, USA: L. Erlbaum Associates, 1988.

[133] D. Gardner-Bonneau, "How many Subjects do I need for a Statistically Valid Survey?," *Usability Interface*, vol. 5, no. 1, pp. 1-2, 1998.

[134] K. R. Murphy and B. Myors, *Statistical Power Analysis: A Simple and General Model for Traditional and Modern Hypothesis Test*, 2 ed. Mahwah, NJ, USA: Lawrence Erlbaum Associates, Publishers, 2004.

[135] R. V. Lenth, "Java applets for power and sample size (http://www.stat.uiowa.edu/~rlenth/Power/)." Iowa City, Iowa, 2006.

[136] C. G. S. D. Corporation, "Sony HMD LDI-50 and LDI-100 Features (http://cgsd.com/OldSonyHMD/Comparison.html)," 1999.

[137] C. Owen, A. Tang, and F. Xiao, "ImageTclAR: A Blended Script and Compiled Code Development System for Augmented Reality," *Proceedings of International Workshop on Software Technology for Augmented Reality Systems (STAR 2003)*, pp., Tokyo, Japan, 2003.

[138] C. M. Jones, V. A. Braithwaite, and S. D. Healy, "The Evolution of Sex Differences in Spatial Ability," *Behavioral Neuroscience*, vol. 117, no. 3, pp. 403-411, 2003.

[139] S. E. Maxwell and H. D. Delaney, *Designing Experiments and Analyzing Data: A Model Comparison Perspective*, 2 ed. Mahwah, NJ, USA: Lawrence ErlBaum Associates, Publishers, 2004.

[140] M. R. Bax, "Real-Time Lens Distortion Correction: 3D Video Graphics Cards Are Good for More than Games," *Stanford Electrical Engineering and Computer Science Research Journal*, no. 1, pp. 9-13, 2004.