INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

VIDEO-BASED APPROACH FOR TRACKING EYE ORIENTATION IN THREE DIMENSIONS

by

Danjie Zhu

A dissertation submitted to the Graduate Faculty in Computer Science

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy, The City University of New York

1997

This manuscript has been read and accepted for the Graduate Faculty in Computer Science in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

4/25/97
date

Theodore Raphan
Chairman of Examining Committee

4/29/97
date

Executive Officer

Professor Michael Anshel

Professor Bernard Cohen

Professor Stanley Habib

Professor Theodore Raphan (Advisor)

Professor George Wolberg

Dr. Steven Moore

Supervisory Committee

The City University of New York

ABSTRACT

VIDEO-BASED APPROACH FOR TRACKING EYE ORIENTATION IN THREE DIMENSIONS

by

Danjie Zhu

Advisor: Professor Theodore Raphan

Accurate eye movement recording and monitoring has been growing in importance in the differential diagnosis of diseases associated with vertigo and other neurological diseases. Tracking of eye movements in three dimensions is important for modelling and studying saccadic eye movements and those generated by the vestibular system. The scleral coil system is currently a standard technique for measuring three dimensional eye movements in both animals and humans. However this technique is invasive. To bypass the invasive nature of the coil system in measuring eye movements in three dimensions, there have been a number of investigations on the use of cameras and video-based techniques. The methodology developed by Moore et al. (1996) uses a three dimensional rotation model to find the eye orientations in three dimensions. It is based on finding the orientation of the optic axis by computing the center of mass of the pupil and then computing the rotation about the optic axis utilizing variation in intensity along an iral arc. This methodology is dependent on maintaining a clear and small pupil image so that the correct center position can be found. It is also dependent on maintaining a fixed pupil size in order not to have changes in the iral striation patterns which would cause errors

in computing torsional eye position. The purpose of this dissertation is to develop algorithms based on a more global model of how circular contours which inscribe the pupil and iris are mapped to elliptical contours under a rotation and orthographic projection mapping, i.e., affine transformations. We have implemented an edge detection algorithm that determines the contour at each given eye orientation. The orientation of the optic axis relative to the axis of the image plane is computed by finding the best fitting ellipse for this contour. The shape of the ellipse can then be related to the orientation of the optic axis. Torsion about the optic axis is computed by using the a minimum distance pattern matching algorithm on an annulus of the iris. The algorithm allows for the removal of the image noises due to the infra-red reflection from the annulus. In order to obtain a clear and uncovered iral signature, an adaptive algorithm has been implemented to choose the location of the annulus depending on the orientation of the eye. This approach to computing the torsional component of eye orientation tends to be more immune to artifacts which arise when the eye moves into positions where there is partial lid occlusion. The proposed methodology should be important for developing a robust system for tracking eye movements.

ACKNOWLEDGEMENTS

I wish to express my extreme gratitude to my mentor, Professor Theodore Raphan, for his guidance and encouragement throughout the course of this research. The many hours of technical discussions and many valuable criticisms and suggestions are most appreciated. The idea of developing a video-based three dimensional eye movements measurement system dated back five and a half years ago when I first met Professor Raphan in his lab. I was immediately attracted by the idea of tracking 3D eye position using image processing technique.

I would like to thank all the members of my doctoral guidance and examination committee, Professors Michael Anshel (City College), Bernard Cohen (Mt. Sinai School of Medicine), George Wolberg (City College), Stanley Habib (City College) and Dr. Steven Moore (Mt. Sinai School of Medicine) for the time and effort each has taken to read and constructively criticize this dissertation.

I would especially thank Professor Bernard Cohen for his support, encouragement and suggestions.

I would like to thank Dr. Steven Moore for his valuable suggestions and criticisms. The detailed discussions with him during the course of this work was a great inspiration and encouragement to me.

I would also thank Mr. Louis Pundis for his help in building the artificial eye and gimbal, and modifying the equipment again and again.

I would like to thank my wife, Chun Ning, for her encouragement through the course of this work.

I would like to thank the faculty of Computer and Information Science Department at Brooklyn College and the faculty of Computer Science at CUNY Graduate School for their support during the course of this study.

Finally I would like to acknowledge the financial support of The National Aeronautics and Space Administration (NASA) under grant number NAS 9-17720, the National Eye Institute under grant EY04148, CUNY-PSC grant 664230, and CUNY Collaborative Grant 991-962.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1

INTRODUCTION

Accurate eye movement recording and monitoring has been growing in importance in the differential diagnosis of diseases associated with vertigo and other neurological diseases (Leigh and Zee 1983). Tracking of eye movements in three dimensions is important for modelling and studying saccadic eye movements (Hepp and Henn 1983; Tweed and Vilis 1987; Schnabolk and Raphan 1994) and those generated by the vestibular system (Raphan et al. 1979; Raphan and Cohen 1995; Yakushin et al. 1995). However, many current techniques are either not suited for three dimensional recording or are invasive. Video-based techniques are suited for three dimensions and are non-invasive, but much work needs to be done both technically and algorithmically to improve the performance.

The most common eye movement measurement techniques used both in animals and humans are electrooculography (EOG) and the scleral search coil systems. Chronically implanted coils are almost standard in animal related work. One major shortcoming of EOG is that it can only measure eye movements in two dimensions and has associated drift and blink artifacts (Young and Sheena 1975). The coil system (Robinson 1963; Judge et al. 1980) can measure three dimensional eye rotations in animals (Hess 1990; Tweed et al. 1990; Hess et al. 1992; Dai et al. 1994) and humans (Furman et al. 1987; Furman et al. 1987; Collewijn et al. 1988; Tweed et al. 1990). It has been used successfully in a wide range of studies of the vestibular and oculomotor systems, including some clinical investigations (Collewijn et al. 1985; Straumann et al.

1996). One major shortcoming of the coil technique is that it is invasive. This has restricted its use in general clinical settings and has caused many investigators to use Electrooculography (EOG).

To bypass the invasive nature of the coil system and the difficulties associated with EOG, there has been some investigation on the use of video-based techniques (VOG) in determining eye position (Hatamian and Anderson 1983; Parker et al. 1985; Vieville and Masse 1987; Clarke et al. 1991; Moore et al. 1996). Recently, low cost miniature cameras have been utilized to develop systems to accurately measure eye orientation in three dimensions. These achievements have opened the possibility that video-based techniques could be developed as a standard for measuring eye orientation. To accomplish this, there is a need to develop both technical and algorithmic approaches which will be robust under a wide range of lighting conditions and ocular occlusions by the eyelids.

Most existing VOG methodologies for determining the orientation of the optic axis or direction of gaze are based on computing the center of the pupil using centroid or other algorithms (Hatamian and Anderson 1983; Parker et al. 1985; Vieville and Masse 1987; Clarke et al. 1991; Moore et al. 1996). Torsional eye position is based on intensity variations along a small arc of the iris centered on the center of the pupil with corrections made for eye geometric displacement (Moore et al. 1996). The algorithm for describing eye orientation is therefore intimately related to finding the exact pupil center and maintaining fixed pupil size. It also depends on a small pupil size, since a large pupil size causes errors in finding the exact pupil center due to the optic properties of the cornea (Haslwanter and Moore 1995). The degree to which occlusion of the iris and

pupil by the lid affects the computation of eye orientation is also not clear. In fact, in the latest studies by Moore et al. (1996), pilocarpine was used in the eye to maintain a fixed and small pupil size. There are also problems with the existing algorithms in distinguishing eye translation and eye rotation due to non-ideal "ball and socket" behavior (Clarke et al. 1991; Moore et al. 1996). Moreover, reflections due to varying lighting conditions cause errors in the computation of eye orientation. Another problem with existing algorithms is calibrating the camera offset angles. The calibration algorithm is sensitive to errors in computing the coordinates of the projection of the pupil center. In most systems, the offset angles of the camera are ignored by adjusting the camera so that the "line of sight" of the camera is aligned along the optic axis when the eye is looking straight ahead (Moore et al. 1996).

The purpose of this work is to develop an eye movement tracking scheme which will describe the eye orientation in a robust way using a model-based algorithm for determining the horizontal, vertical and torsional eye orientation. The algorithm which will be developed is based on modelling the eye as a sphere. The iris and pupil are modelled as the intersection of the surface of the sphere with a plane (Moore et al. 1996). The orthographic projection of boundaries between the sclera, iris and pupil onto an image plane will be ellipses for any arbitrary orientation of the eye. When the optic axis of the eye is orthogonal to the image plane, i.e., when the eye looks straight into the camera, the pupil and iral boundaries are circles. Computation of eye orientation is based on sampling the boundary points between the iris and the sclera and/or between the iris and the pupil and finding the best fitting ellipses to the boundaries. Even with partial lid closure, the imaged boundary points can be used to find the best fitting ellipse.

Thus, the utilization of an elliptical model is a novel aspect of this development and is potentially a robust method for finding eye orientation when there is lid obstruction.

With this model, it is theoretically possible to distinguish between a translation of the rotational center of the eye and a rotation about this center. Rotations of the eye will cause changes in the eccentricity and the direction of the major and minor axis of the ellipse, while translation will not affect these parameters. Using pupil centroid approaches, it was impossible to distinguish between translation and rotation of the eye (Clarke et al. 1991; Moore et al. 1996).

Torsion about the optic axis is calculated by comparing an annulus of the current iral image with the corresponding annulus of the reference iral image using a minimum distance template matching technique. Most existing systems used a fixed arc as the iral signature for computing torsion (Vieville and Masse 1987; Clarke et al. 1991). But as the eye moves around or eyelid droops, the arcs may be covered and become unsuitable for use. To partially overcome this problem, Moore et al. (1996) used the signatures of four arcs. We have extended the capability of computing torsion by making the algorithms adaptive. The adaptive sampling annulus technique used in this study defines an algorithm which allows for the adjustment of the sampling annulus as a function of the amount of eyelid closure and eye orientation. Therefore even if part of the iris is covered by the eyelid, torsion can still be computed using the uncovered iral annulus. This is another contribution of this study.

A "City Block" metric, a rather fast algorithm, is defined for computing the minimum distance between the signatures. This algorithm allows for the removal of the infra-red light reflection from the signatures, and therefore is another novel aspect of this

work. Most existing algorithms compute torsion by cross correlating iral signatures using FFT technique, which breaks down when part of the signature contains infra-red light reflection (Moore et al. 1996). The "City Block" metric makes the signature matching algorithm more robust with regard to the light reflection.

Another important contribution of this work is the implementation of the calibration procedure. The elliptical model is more robust with regard to calibrating the camera offset angles. The camera offset angles are related to the center of the pupil and the eccentricity of the ellipse at different eye orientation. Using both criteria makes the calibration algorithm more accurate and less sensitive to noise.

CHAPTER 2

BACKGROUND

In order to better understand the video-based three dimensional eye movement measurement technique presented in this study, a review of the eye movement measurement techniques and their development is shown below.

2.1 ELECTRO-OCULOGRAPHY

One of the oldest methods used in measuring eye movements is Electro-oculography (EOG) (Young and Sheena 1975). EOG is based on voltages from cornea to retina. The cornea remains 0.40 to 1.0 mV positive with respect to the retina. The position of the eye can be measured by placing skin electrodes around the eye and recording the potential differences. As the eye rotates, the potential difference also varies. EOG has the largest range of any objective methods practical for human studies, since it does not require visualization of the eye. The method can be used to measure eye movements up to $\pm 70°$. However, the linearity becomes progressively worse at excursions greater than $\pm 30°$. The typical accuracy with surface electrodes is 1.5°-2°. Muscle artifacts, eyelid interferences, basic non-linearities in the technique, variation in the cornea-retinal potential attributable to light adaptation and diurnal variations all cause errors in measurement (Young and Sheena 1975).

2.2 CORNEAL REFLECTION

One early imaging technique that has been attempted in tracking eye movements is the corneal reflection method. Incident light from a source is reflected from the convex surface of the cornea in a pattern of diverging light and is imaged through a concave lens onto a film plate or other light detector (Buswell 1935; Carmichael and Dearborn 1947; Taylor 1971) (See (Young and Sheena 1975) for Review). The output of this eye movement monitoring technique was limited to graphical displays in two dimensions and to small angles of eye deviation ($\pm 12°$ to $\pm 15°$). The accuracy of this method is limited to 0.5°-1.0°. It was also cumbersome to calibrate and the technology was never developed as a viable means of ocular tracking in three dimensions which includes tracking torsional eye movement.

2.3 LIMBUS AND PUPIL TRACKING

Another technique that has been used for tracking eye movement is the limbus and pupil tracking technique. The sharp boundaries between the iris and sclera (the limbus) and the pupil and iris provide easily identifiable edges. These edges can be detected by two or more photodetectors or other means to find the center of the pupil, which coincides with the optic axis (Young and Sheena 1975). Although the limbus and pupil tracking technique yields high accuracy (0.1°-0.5°), it is limited to the range of $\pm 15°$. The tracking of vertical eye movement also causes a problem because of the eye lid coverage of the pupil and iris. This technique is also not suitable for detecting torsional eye movement.

2.4 THE PHOTOGRAPHIC TECHNIQUE

To solve the problem of measuring torsional movement of the eye, a photographic technique using the natural landmarks of the eye for extracting the angular position of the eye have been utilized (Miller and Graybiel 1972; Balliet and Nakayama 1978; Diamond et al. 1979). The eye is photographed before and after the counterrolling, the films are developed, and transparencies are made. Using a projector, the two transparencies are magnified by a large factor and superimposed by projection on a screen. One image is fixed, and the other is manually rotated with a fine mechanical device until a match between two images occurs. The angle of rotation can then be read directly from a scale available on the mechanical device. In spite of its good resolution (0.09°) and ability to measure large angles of eye rotations, this technique is very time-consuming and not suitable for eventual real-time application.

2.5 THE SCLERAL SEARCH COIL TECHNIQUE

The scleral search coil technique (Robinson 1963; Collewijn et al. 1985) uses a contact lens about which are wound two perpendicular coils which are fitted over the eye. The eye then moves in two head-fixed alternating magnetic fields in phase and space quadrature, vertical and horizontal. As the eye moves, the induced voltages in these coils are related to the projection of the magnetic fields onto the search coil axis. The induced voltages can then be used to represent eye rotation relative to the head (Tweed et al. 1990). Although the scleral coil technique has high resolution and is presently the standard method for measuring three dimensional eye movements, it is in some sense invasive. It also causes discomfort, image obstruction for the subject, and the time of

testing using the contact lens is limited. The possibility of coil slippage, particularly in the torsional direction, also remains a potential source of error.

2.6 VIDEO-OCULOGRAPHY

One of the latest developments in the area of eye movement measurement is the application of video-oculography (VOG), a video-based measurement technique (Nakayama 1974; Hatamian and Anderson 1983; Parker et al. 1985; Clarke et al. 1991; Groen et al. 1996; Moore et al. 1996; Mulligan 1997). In a VOG system, a video camera which is interfaced to a microcomputer is used to track the rotation of the eye. A frame grabber continuously grabs images which are stored in memory or on video tapes for future processing. Digital image processing techniques and other analytical algorithms are applied to the images.

Algorithms for determining eye orientation have been separated into two sub-algorithms. One sub-algorithm has concentrated on finding horizontal and vertical eye position, while another algorithm has focused on determining the torsion of the eye about the optic axis. The algorithm for finding the horizontal and vertical eye position has focused on tracking the pupil center and various algorithms have been developed which accomplish this task. Hatamian and Anderson (1983) determined horizontal and vertical rotations of the eye by averaging the coordinates of the edge points of the pupil. Other techniques have included fitting a circle to the pupil boundary (Sung and Anderson 1991) or using a "center of mass" computation (Vieville and Masse 1987; Moore et al. 1996). There are also commercial systems available (ISCAN) (See Mulligan (1997) for a

Description) that compute two-dimensional eye position using proprietary but probably similar types of algorithms.

A more difficult problem from an algorithmic perspective is the computation of the torsional angle about the optic axis. This problem has been approached by tracking a variety of landmarks of the eye (Nakayama 1974; Parker et al. 1985; Ott et al. 1990; Yamanobe et al. 1990; Groen et al. 1996; Mulligan 1997). Parker et al. (1985) have developed algorithms to measure ocular torsion from the translation of two landmarks on the eye using both rectilinear coordinate system and polar coordinate system. Accuracy of 0.13° was reported. Groen et al. (1996) have developed a torsion algorithm by tracking a set of 36 significant iris patterns using template matching technique. Each landmark results in a single estimate of the torsion angle. Ocular torsion is computed from the total set of individually determined torsion angles. Ott et al. (1990) have used a contact lens with a special pattern on it to provide the marker, but it has a similar problem to the scleral search coil, i.e. the contact lens is in some sense invasive. The landmark tracking technique is critically dependent on computer operator interaction in choosing the landmarks (Parker et al. 1985). Because of inter-subject variability, it also does not readily lend itself to the implementation of automatic recognition.

Another algorithmic approach has used cross correlation of iral patterns of circular arcs about the pupil center (Hatamian and Anderson 1983; Vieville and Masse 1987; Clarke et al. 1991; Moore et al. 1996). It relies on the fact that most of the variation of the iris image is in the angular direction (Hatamian and Anderson 1983). Because this method uses information about variations in intensity over an arc, it is more likely to be representative of a given torsional eye rotation without the need for operator

selection. Moreover, while gray level intensity may exist between different people, image enhancement techniques can be utilized to normalize the intensity variations for a given subject. Measurement of ocular torsion about the optic axis has been implemented by sampling the pixel intensity of the iris along a circular sampling window at a fixed radius from the pupil center (Hatamian and Anderson 1983; Vieville and Masse 1987; Clarke et al. 1991; Moore et al. 1996).

Vieville and Masse (1987) have developed a VOG system (EMIRAT) which was used in the First European Space-lab Flight (Vieville and Masse 1987). Moore et al. (1991) have developed the VTM system which measured the eye movements in a clinic environment (Moore et al. 1991). Resolution of 0.1° was reported. The VOG technique is therefore accurate and non-invasive and has the potential for being a standard for tracking eye movements.

The VOG technique is presently limited by its ability to grab frames at only 30 fps, although some systems have gone to 60 fps at the cost of reduced spatial resolution (Clarke et al. 1991). This limitation is being overcome by faster hardware acquisition (Clarke & Teiwis, 1995, Personal Communication). Other problems with video-based techniques are algorithmic. Present algorithms rely on knowledge of the complete image of the pupil and an arc of the iris (Hatamian and Anderson 1983; Vieville and Masse 1987; Clarke et al. 1991; Moore et al. 1996). Thus, partial occlusion of the iris and/or pupil poses difficulties in accurately tracking the eye. To overcome some of these difficulties, pilocarpine has been used to maintain a fixed and small pupil size so that it can be imaged completely (Moore et al. 1996). This is a physiologically invasive

procedure. The light reflection from the pupil and the iris also causes problems in finding the exact pupil center and obtaining a usable iral signature.

An additional problem with VOG technique is calibrating the camera orientation with regard to the head. A general algorithm for calibrating camera position for 3D eye position measurement has been described by Moore et al. (1996), but a practical application of this method was found to be sensitive to small errors in the pupil center estimate. This algorithm has been implemented using a non-linear optimization technique (Peterka and Merfeld 1996), which has overcome problems associated with sensitivity to the pupil center estimate. A simulation of the effect of camera offset on 3D eye position measurement has demonstrated that small horizontal and vertical camera offsets (less than 5°) have a negligible effect on eye position measurement, but torsional offset causes significant error and must be determined by calibration (Moore et al. 1996). Based on these results, the system of Moore et al. (1996) ignores the effects of horizontal and vertical camera offsets by accurate positioning of the camera, but takes the torsional offset of the camera into account.

In this study, eccentricity of the elliptical boundary surrounding the pupil or iris as well as the pupil or iris center are used in calibration. This can overcome the calibration problem. We will show that a model-based algorithm for tracking eye orientation is potentially robust with regard to partial occlusion of the pupil and/or iris and the light reflection, and can overcome the calibration problem.

CHAPTER 3

DEVELOPMENT OF EYE TRACKING ALGORITHMS

3.1 CONCEPTUAL BASIS OF ALGORITHM

Mathematically, the eye can be considered as a sphere (Figure 1). Light enters the eye through the transparent cornea, and then through an approximately circular opening in the iris called the pupil. As a first approximation, the iris and pupil can be modelled as a plane which intersects the sphere (Wolff 1940) (refer to Appendix C for notes on pure spherical eye model). The cornea can be modelled as a single thin lens which magnifies the image of the pupil and iris with little distortion (Gullstrand 1909; Bennett and Rabbetts 1984; Haslwanter and Moore 1995). Thus, images of the eye give two closed contours lying on the iral plane (Figure 1): the boundary between the pupil and the iris and the boundary between the iris and the sclera. When viewing the orthographic projection of these closed contours onto an image plane, the curves change shape as the eye rotates about its center relative to the image plane (Figure 2). Because of the symmetry of the normal eye, the contours are circular when the optic axis is directed perpendicular to the image plane and transform into ellipses when the eye rotates (Figure 2). Thus, the model-base for the image analysis of eye orientation that is studied here is the affine transformation of a closed curve lying on a plane which projects as a circular contour when it is facing the image plane and an ellipse when the sphere is rotated. Eye orientation can then be extracted from the orientation and eccentricity of the ellipses.

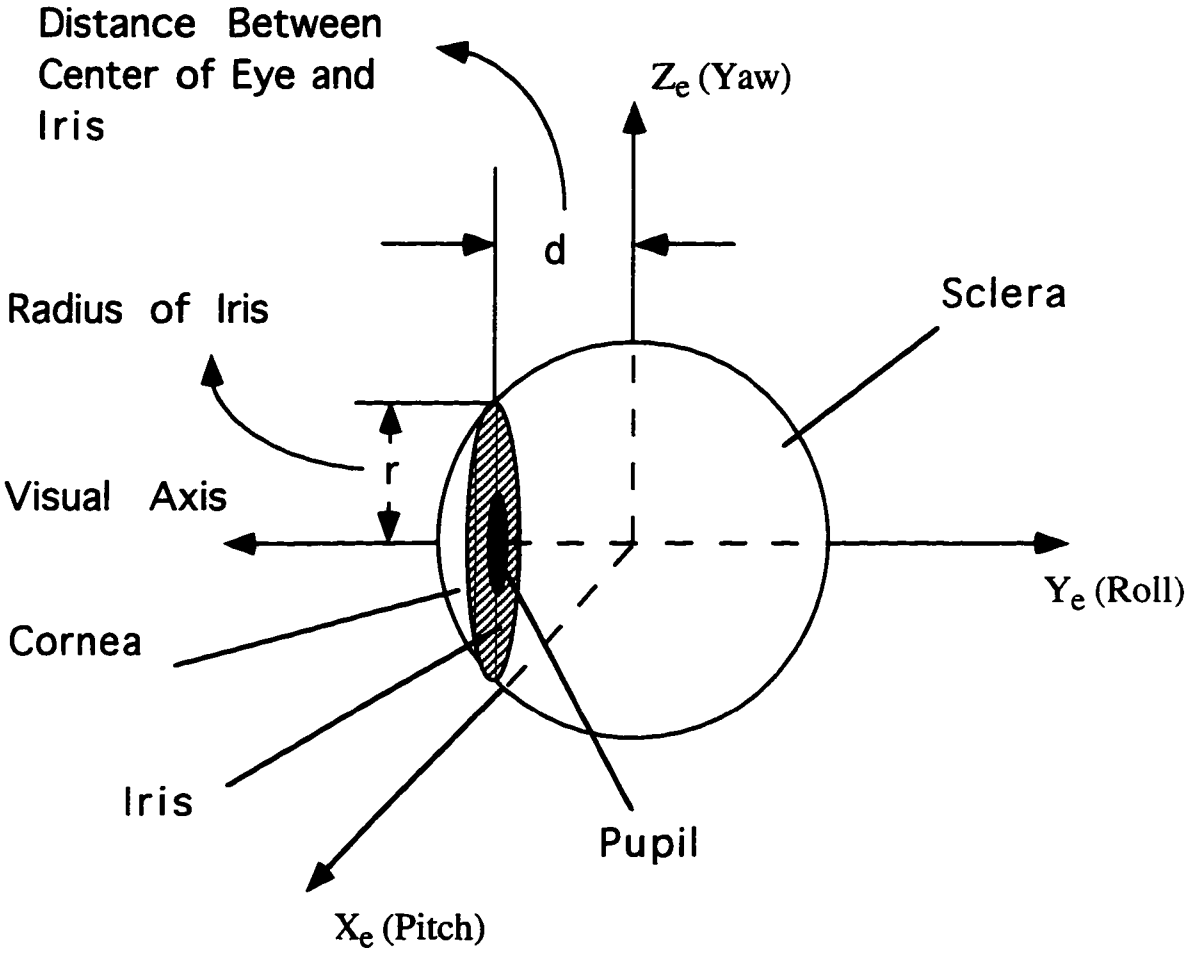


Figure 1: Modelling of Eye Landmarks for Tracking Eye Position

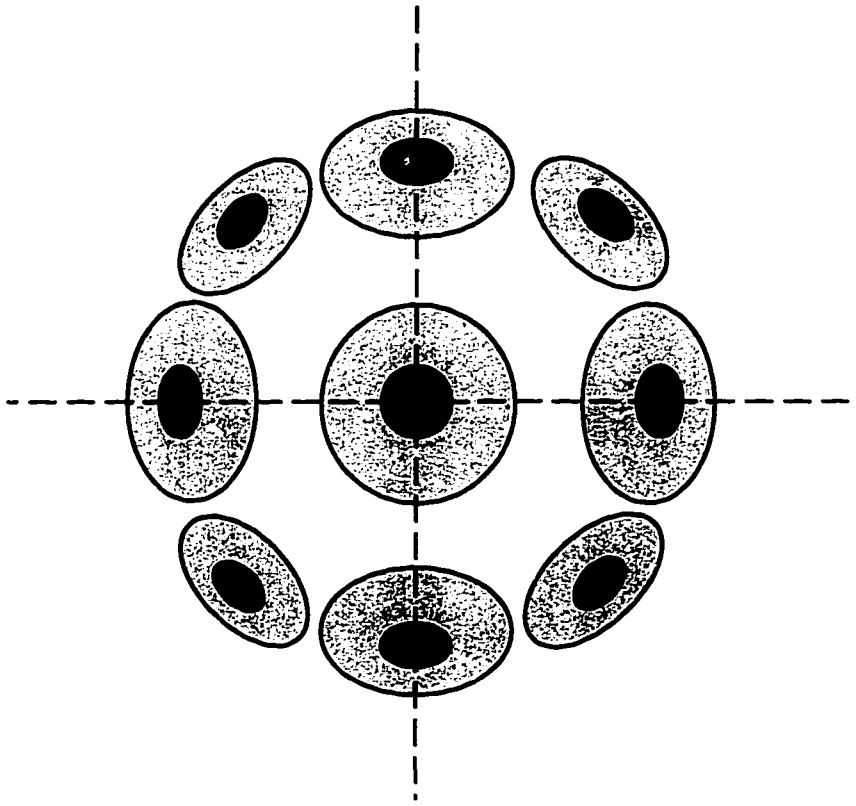


Figure 2: Elliptical Projection of Iris and Pupil as Eye Rotates

To derive model-based formulae for tracking eye movements, the following assumptions are made:

1. The eye is assumed to be a perfect sphere which rotates about its center.

2. The eye exhibits ideal ball and socket behavior, so that all eye movements are pure rotations around the center of the eye, with no translation of this center. This assumption leads to a simple algorithm for finding eye orientation. However we will examine the effects of eye translation and conclude that eye orientation can theoretically be computed from the eccentricity of the iris/pupil projection when the eye is translated.

3. The projections of the iral and pupillary boundaries are concentric. Studies indicate that the iral and pupil boundaries are not in general concentric (Wyatt 1995). However, the orientation of the eye is computed by fitting the iral and pupil boundaries by ellipses separately. In addition, the centers are separately computed and the system could be calibrated separately from the individual elliptical boundaries. Therefore, this assumption can be relaxed without altering the development or conclusions of this study.

4. The iris and pupil image distortion caused by the optic properties of the cornea can be ignored.

3.2 MATHEMATICAL BASIS OF ALGORITHM

3.2.1 COORDINATE FRAMES

Eye orientation is computed from images captured by a camera attached to the head. It is, therefore, convenient to define a head-fixed coordinate frame whose roll axis

is normal to the image plane. This will be referred to as the camera-coordinate-frame (X-Y-Z, Figure 3). Compatibility with other studies requires that eye movements be represented in a head-fixed coordinate frame which is defined as the pitch, roll, and yaw axes when the eye is looking straight ahead. This frame will be referred to as the reference-coordinate-frame (X'-Y'-Z', Figure 3). To fully specify the eye orientation, an eye-fixed coordinate frame is also defined as the pitch, roll and yaw axes of the eye (X_e-Y_e-Z_e, Figure 1). It coincides with the reference-coordinate-frame when the eye is looking straight ahead, and the camera-coordinate-frame when the eye is looking straight at the camera (Figure 3). The centers of the three above coordinate frames coincide with the center of the eye (Figure 1, Figure 3).

In general, the reference-coordinate-frame has a basis of unit vectors given by $\beta' = \{e'_1, e'_2, e'_3\}$ (Figure 3). The roll axis basis vector (e'_2), for example, is along the "line of sight" or optic axis of the eye, when the eye is looking straight ahead. The positive direction for the unit vector is defined as pointing toward the back of the head.

The camera-coordinate-frame has a basis of unit vectors given by $\beta = \{e_1, e_2, e_3\}$ (Figure 3). The roll axis basis vector (e_2), is aligned with the normal to the image plane, i.e., along the optic axis when the eye is looking straight into the camera. The pitch and yaw axes are defined as parallel to X_{im} and Z_{im} axes of the image plane (Figure 4). The point of intersection of the roll axis with the image plane is the origin of the two dimensional coordinate frame of the image plane. The X_{im} and Z_{im} axes are in the same direction as the pitch and yaw axes of the camera-coordinate-frame.

The rotation angle of the eye is assumed to be measured counter-clockwise when looking at the coordinate center from a point on either +X, +Y, or +Z axes. The Euler

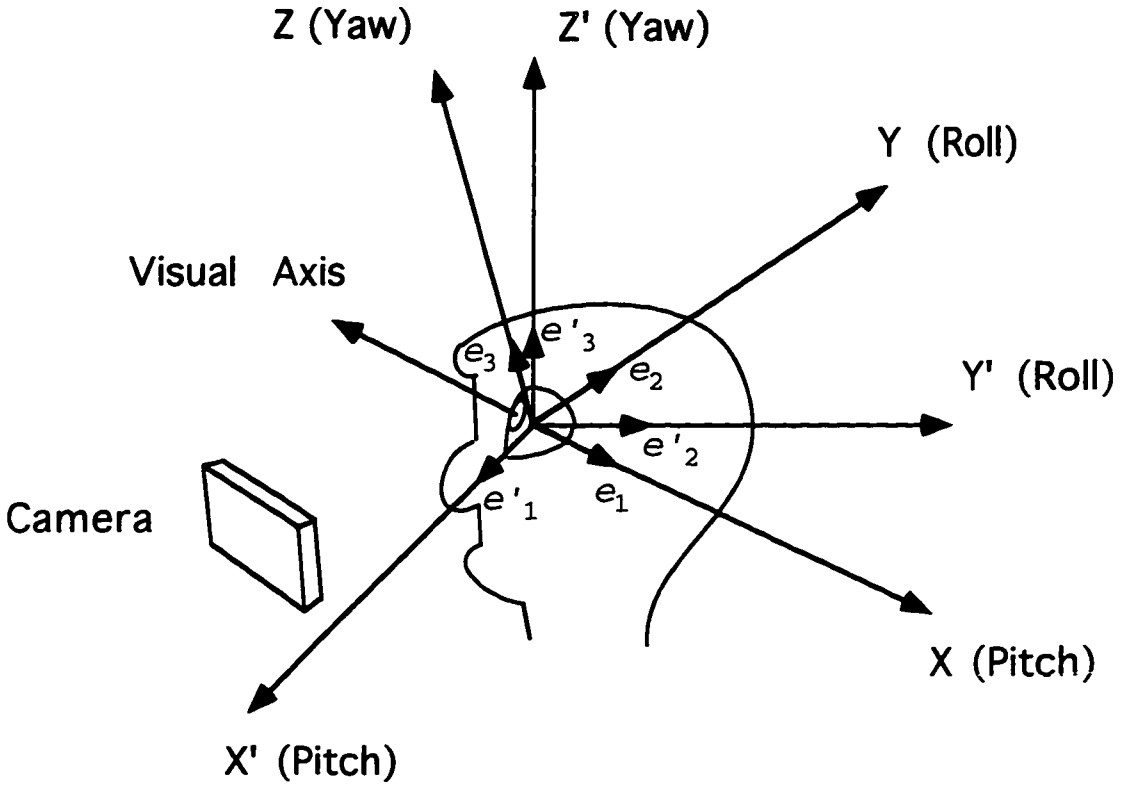


Figure 3: Relationship between Camera-Coordinate-Frame (X-Y-Z) and Reference-Coordinate-Frame (X'-Y'-Z')

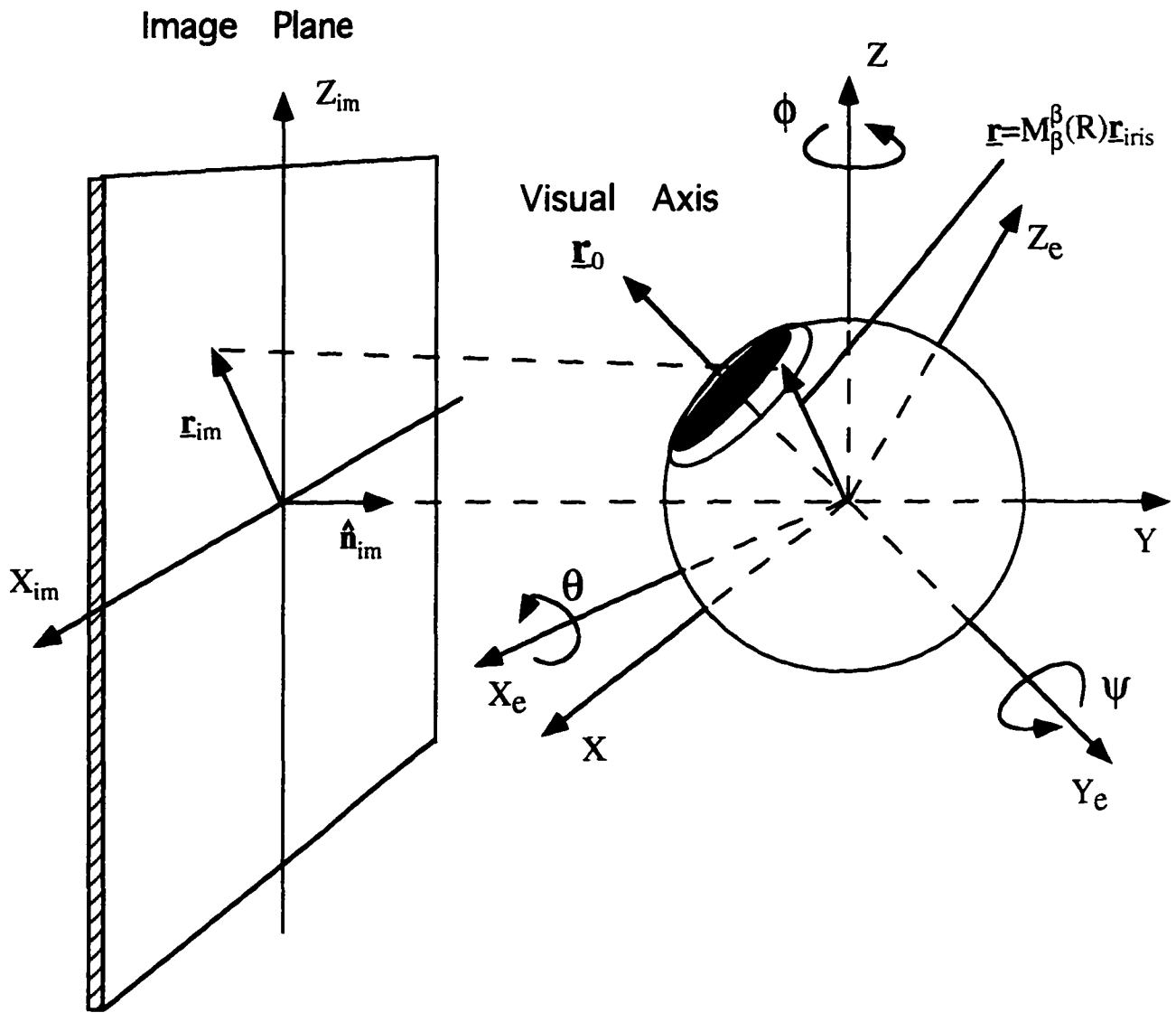


Figure 4: Eye Position in Camera-Coordinate-Frame as Seen by Image Plane

angles describing the rotation (Goldstein 1980) are first the angle ϕ about the +Z axis, then θ about $+X_e$ axis (line of nodes), followed by ψ about the $+Y_e$ axis or optic axis (Figure 4).

In general, the camera is tilted and there will be a relative rotation between the camera-coordinate-frame and reference-coordinate-frame. This offset is determined during calibration. In this dissertation, we will first consider computation of eye orientation when the camera has no offset and the camera and reference coordinate frames coincide. We will then consider the computation of eye orientation when they are offset with respect to eachother.

3.2.2 TRANSFORMATION MATRIX

Eye rotations in the head can be described by a rotation matrix in the camera-coordinate-frame given by $M_\beta^\beta(R)$ (Figure 4). If we let $\underline{r}_{iris}$ be the position vector of a point on the iris when the eye is in the reference position (looking straight ahead) (Figure 4), then any eye rotation will transform the vector $\underline{r}_{iris}$ to the vector $\underline{r}$ in the camera-coordinate-frame (Figure 4) given by:

$$\underline{r} = M_\beta^\beta(R)\underline{r}_{iris} \qquad (1)$$

The orthographic projection of the rotated position vector, $M_\beta^\beta(R)\underline{r}_{iris}$, onto the image plane is $\underline{r}_{im}$ (Figure 4) and can be given as:

$$\mathcal{L}_{im} = M_{\beta}^{\beta}(R)\mathcal{L}_{iris} - \langle M_{\beta}^{\beta}(R)\mathcal{L}_{iris}, \hat{n}_{im}\rangle \hat{n}_{im} \tag{2}$$

where $\hat{n}_{im}$ is a unit vector normal to the image plane, i.e., $\hat{n}_{im} = e_2$, $<,>$ is the inner product between two vectors. The matrix representing the rotation, $M_{\beta}^{\beta}(R)$, can be parameterized by Euler angles (Goldstein 1980), Axis-Angle (Goldstein 1980; Schnabolk and Raphan 1994), or Euler-Rodrigues parameters (Rodrigues 1840). For the current study, the parameterization in terms of Euler angles leads to a simple transformation because only two angles (ϕ and θ) need to be considered. Torsion about the optic axis can be computed separately. The rotation matrix, in terms of Euler angles can be given as a product of three sequential rotations given by:

$$R_{\phi} = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$R_{\theta} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \tag{3}$$

$$R_{\psi} = \begin{pmatrix} \cos\psi & 0 & \sin\psi \\ 0 & 1 & 0 \\ -\sin\psi & 0 & \cos\psi \end{pmatrix}$$

corresponding to rotations about the head-fixed yaw (Z) axis (R_ϕ), a rotated pitch (X_c) axis (R_θ), and the optic (Y_c) axis (R_ψ) (Figure 4). The composite rotation is then given by:

$$M_\beta^\beta(R) = R_\phi\, R_\theta\, R_\psi =$$

$$\begin{pmatrix} \cos\phi\cos\psi-\sin\phi\sin\theta\sin\psi & -\sin\phi\cos\theta & \cos\phi\sin\psi+\sin\phi\sin\theta\cos\psi \\ \sin\phi\cos\psi+\cos\phi\sin\theta\sin\psi & \cos\phi\cos\theta & \sin\phi\sin\psi-\cos\phi\sin\theta\cos\psi \\ -\cos\theta\sin\psi & \sin\theta & \cos\theta\cos\psi \end{pmatrix} \quad (4)$$

Using Eqs. (1) and (4), we can express the relationship between the vector $\underline{r}$ and $\underline{r}_{iris}$ as follows:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = M_\beta^\beta(R) \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \quad (5)$$

where

$$\underline{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (6)$$

and

$$\mathbf{r}_{iris} = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \tag{7}$$

We now derive an algorithm for determining the two Euler angles associated with a given eye orientation from the iral and pupillary contours.

3.3 ALGORITHM FOR DETERMINING HORIZONTAL AND VERTICAL EYE ROTATIONS

3.3.1 FINDING EQUATION FOR IRIS/PUPIL BOUNDARY

When the camera and reference coordinate frames coincide, and the eye is looking straight ahead, contours of the iris-scleral and pupil-iral boundary are circles when projected onto the image plane. An algorithm is derived for tracking eye position based on the iris-scleral boundary, but it can be equally applied to the pupil-iral boundary.

When the eye is in the reference position, let (x', y', z') be the coordinates of a boundary point in camera-coordinate-frame. The boundary point satisfies the equation of a circle and can be given by the following equation:

$$\begin{aligned} x'^2 + z'^2 &= r^2 \\ y' &= -d \end{aligned} \tag{8}$$

where r is the radius of the circle in the iral plane.

After the eye has rotated, the coordinates of the boundary points are transformed as given by Eq. (5). Since the torsional movement of the eye about the optic axis does not change the shape and eccentricity of the ellipse projected onto the image plane, R_ψ can be ignored for purposes of describing the ellipse. Therefore, the rotation matrix given by Eq. (4) can be simplified and given by:

$$M_\beta^\beta(R) = \begin{pmatrix} \cos\phi & -\sin\phi\cos\theta & \sin\phi\sin\theta \\ \sin\phi & \cos\phi\cos\theta & -\cos\phi\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \tag{9}$$

The boundary after the eye has rotated can be described by a matrix operation as follows:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi\cos\theta & \sin\phi\sin\theta \\ \sin\phi & \cos\phi\cos\theta & -\cos\phi\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x' \\ -d \\ z' \end{pmatrix} \tag{10}$$

or

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x'\cos\phi + d\sin\phi\cos\theta + z'\sin\phi\sin\theta \\ x'\sin\phi - d\cos\phi\cos\theta - z'\cos\phi\sin\theta \\ -d\sin\theta + z'\cos\theta \end{pmatrix} \tag{11}$$

We have defined X_{im} and Z_{im} axes of the image plane as orthographic projections of X and Z axes of the camera-coordinate-frame onto the image plane, and the center of the image plane as the projection of the center of the camera-coordinate-frame onto the image plane. Thus, the projection of the boundary onto the image plane is equivalent to

its projection onto the X-Z plane of the camera-coordinate-frame. Using Eq. (11), the projection of the boundary onto the image plane when the eye has rotated is given by:

$$\begin{aligned} x &= x'\cos\phi + d\sin\phi\cos\theta + z'\sin\phi\sin\theta \\ z &= -d\sin\theta + z'\cos\theta \end{aligned} \tag{12}$$

Solving for x' and z', we obtain:

$$\begin{aligned} z' &= \frac{z + d\sin\theta}{\cos\theta} \\ x' &= \frac{x - d\sin\phi\cos\theta - \dfrac{z + d\sin\theta}{\cos\theta}\sin\phi\sin\theta}{\cos\phi} \end{aligned} \tag{13}$$

The equation in the primed coordinates is a circle with radius r and is given by:

$$x'^2 + z'^2 = r^2 \tag{14}$$

Using Eq. (13), this can be expressed as:

$$\frac{\left(x - d\sin\phi\cos\theta - \dfrac{z + d\sin\theta}{\cos\theta}\sin\phi\sin\theta\right)^2}{\cos^2\phi} + \frac{(z + d\sin\theta)^2}{\cos^2\theta} = r^2 \tag{15}$$

The above equation is equivalent to the following:

$$x^2 = -\frac{\cos^2\phi+\sin^2\phi\sin^2\theta}{\cos^2\theta}z^2 + 2\sin\phi\tan\theta xz + \frac{2d\sin\phi}{\cos\theta}x$$

$$- \frac{2d\sin\theta}{\cos^2\theta}z + \left(r^2 - d^2\tan^2\theta - \frac{d^2\tan^2\phi}{\cos^2\theta}\right)\cos^2\phi \tag{16}$$

If we allow:

$$A = -\frac{\cos^2\phi+\sin^2\phi\sin^2\theta}{\cos^2\theta}$$

$$B = 2\sin\phi\tan\theta$$

$$C = \frac{2d\sin\phi}{\cos\theta} \tag{17}$$

$$D = -\frac{2d\sin\theta}{\cos^2\theta}$$

$$E = \left(r^2 - d^2\tan^2\theta - \frac{d^2\tan^2\phi}{\cos^2\theta}\right)\cos^2\phi$$

we obtain the following equation which describes the contour of the projection of the boundary onto the image plane:

$$x^2 = Az^2 + Bxz + Cx + Dz + E \tag{18}$$

3.3.2 DERIVING FORMULA FOR EULER ANGLES

The values for A, B, C, D, E in Eq. (18) can be found by fitting all the boundary points to Eq. (18). Thus, an important step in the algorithm is to find the optimal ellipse that fits the sampled boundary points in the mean square sense. We have used Singular Value Decomposition for this purpose. The parameter, d, which defines the distance between the center of the eye and the iral plane (Eq. (17)) can be found during the calibration process. Once A, B, C, D, E and d are known, θ and ϕ can be computed from Eq. (17) as follows:

$$\theta = \arcsin\left(\frac{d \pm \sqrt{d^2 + D^2}}{D}\right)$$
$$\phi = \arcsin\left(\frac{C\cos\theta}{2d}\right) \qquad (19)$$

Eq. (19) is dependent on the assumption that the eye exhibits ideal "ball and socket" behavior (assumption 2 in section 3.1). Since parameters C and D change when the eye translates as well as rotates, Eq. (19) can not distinguish the translation of the eye from the rotation of the eye. To overcome this problem, an equation for obtaining θ and ϕ was obtained from parameters A and B, which are related to the eccentricity of the ellipse. These parameters depend only on the rotation and are not affected by translation. The Euler angles, θ and ϕ, can be given by:

$$\theta = \pm \arccos\left(\sqrt{\frac{2(A-1)\pm 2\sqrt{(A+1)^2+B^2}}{B^2+4A}}\right) \qquad (20)$$

$$\phi = \pm \arcsin(\sqrt{A+1+\tan^2\theta})$$

3.3.3 EVALUATION OF ERRORS FOR ELLIPSE FITTING ALGORITHM

Experimentally, we have found that Eq. (20) is accurate when the eye is at eccentric positions, but accuracy deteriorates when the eye is close to the reference position where the boundary is a circle, i.e. ϕ and θ are small. The inaccuracy arises because the fits are less sensitive to changes of eccentricity when the boundary is close to a circle.

To better understand the nature of the inaccuracies, we have simulated rotations of a circle using a range of ϕ and θ, which produced elliptical boundaries when imaged on a head-fixed plane. Assume that the size of the eye images used is 512x512, and the ratio of r:d (Figure 1) is 2:3, which were the values for the frame grabber and the artificial eye used in this study. If the ratio of r:d is reduced when r is fixed, i.e. if d is increased, the sensitivity of the algorithm to the changes of the eccentricity will be increased. The points on the boundary of the circle are obtained using Eq. (8), where typical values r=200 and d=300 were used to produce 800 sampling points for the simulation. The rotated points on the boundary, given by Eq. (12), were then utilized to compute the five parameters of the ellipse (Eq. (18)), i.e., A, B, C, D and E. From these parameters the angles, ϕ and θ, were estimated using Eq. (20) and an error

between the known rotation and the estimated rotation from ellipse fitting was computed. Figure 5 shows the error as a function of ϕ for a range of values for θ, assuming that the image can be specified with a continuous resolution. That is, each coordinate of the boundary can be specified as a real number. Under this assumption, the computational error is largest for $\theta=0°$, and is less than approximately $1.5\text{x}10^{-4}$ degree for ϕ ranging over $\pm 10°$ (Figure 5C). The largest error occurs close to the reference position where the boundary is approximately circular (Figure 5C).

For a discrete image, where the coordinates of boundary points are specified as integer numbers, the error is larger. Figure 6 shows the error functions for the same simulation using integer coordinates. The error ranges from approximately -2.5° to +2.5° at the reference position. As the absolute value of ϕ and/or θ increase, the error decreases. When the absolute value of ϕ and θ are larger than 5°, the error is less than approximately 0.15°.

The error at eccentric eye positions can be further reduced by increasing the image resolution or equivalently increasing the value of r and d. Suppose the size of the images used is 1024x1024. Figure 7 shows the error functions using 2000 samples with r=500 and d=750 (Eq. (8)) at different eye orientations. Although the maximum error is approximately the same as that for 800 sampling points close to the reference position, there is significant reduction at the larger angles (Figure 7). The conclusion from these simulations is that increased resolution, which can be achieved by using a high resolution camera and frame grabber, can result in improved accuracy. However, this must be weighed against the increased cost in computation time.

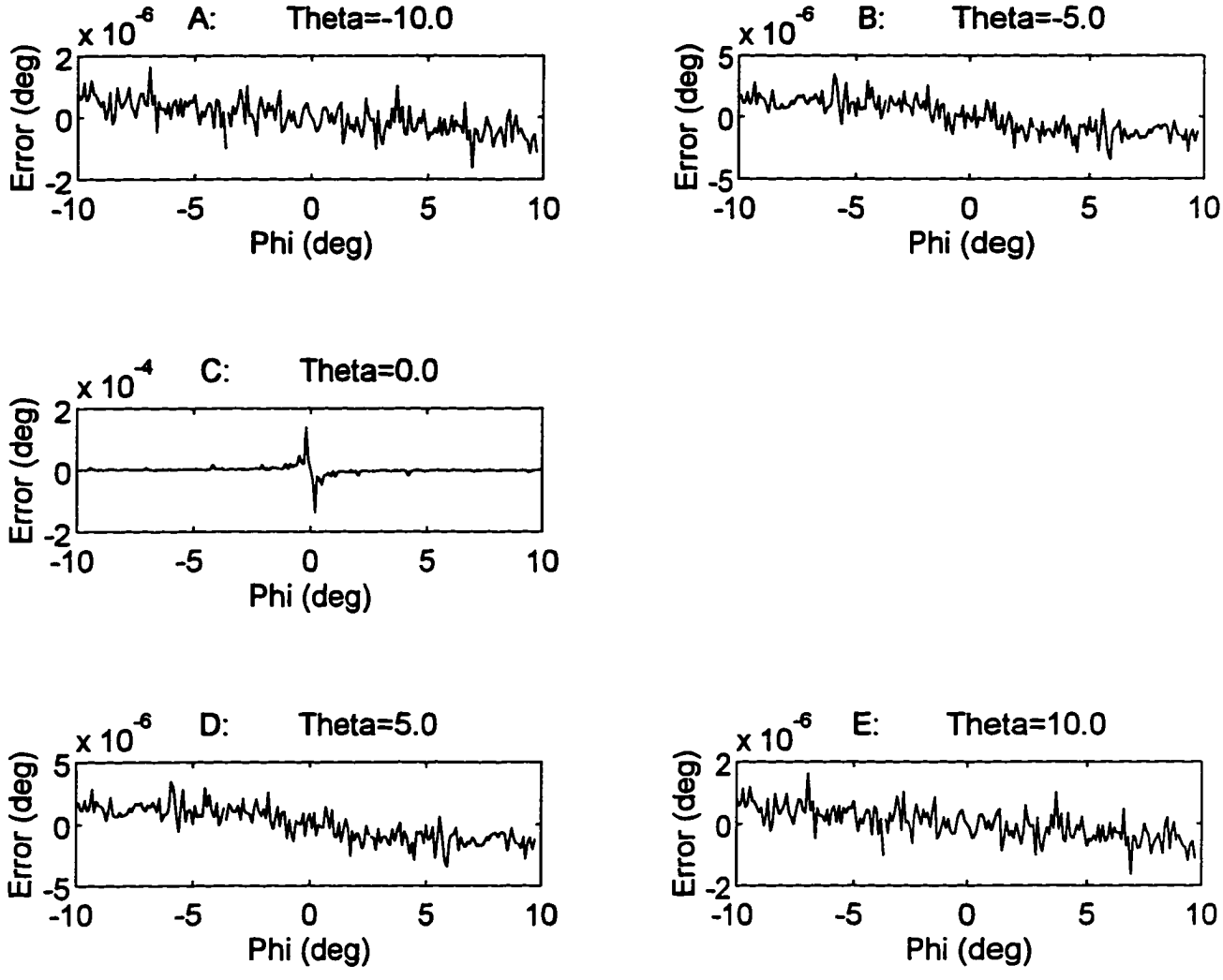


Figure 5: Computational Error of Ellipse Fitting as Function of Euler Angles Using Continuous Coordinates (800 Samples)

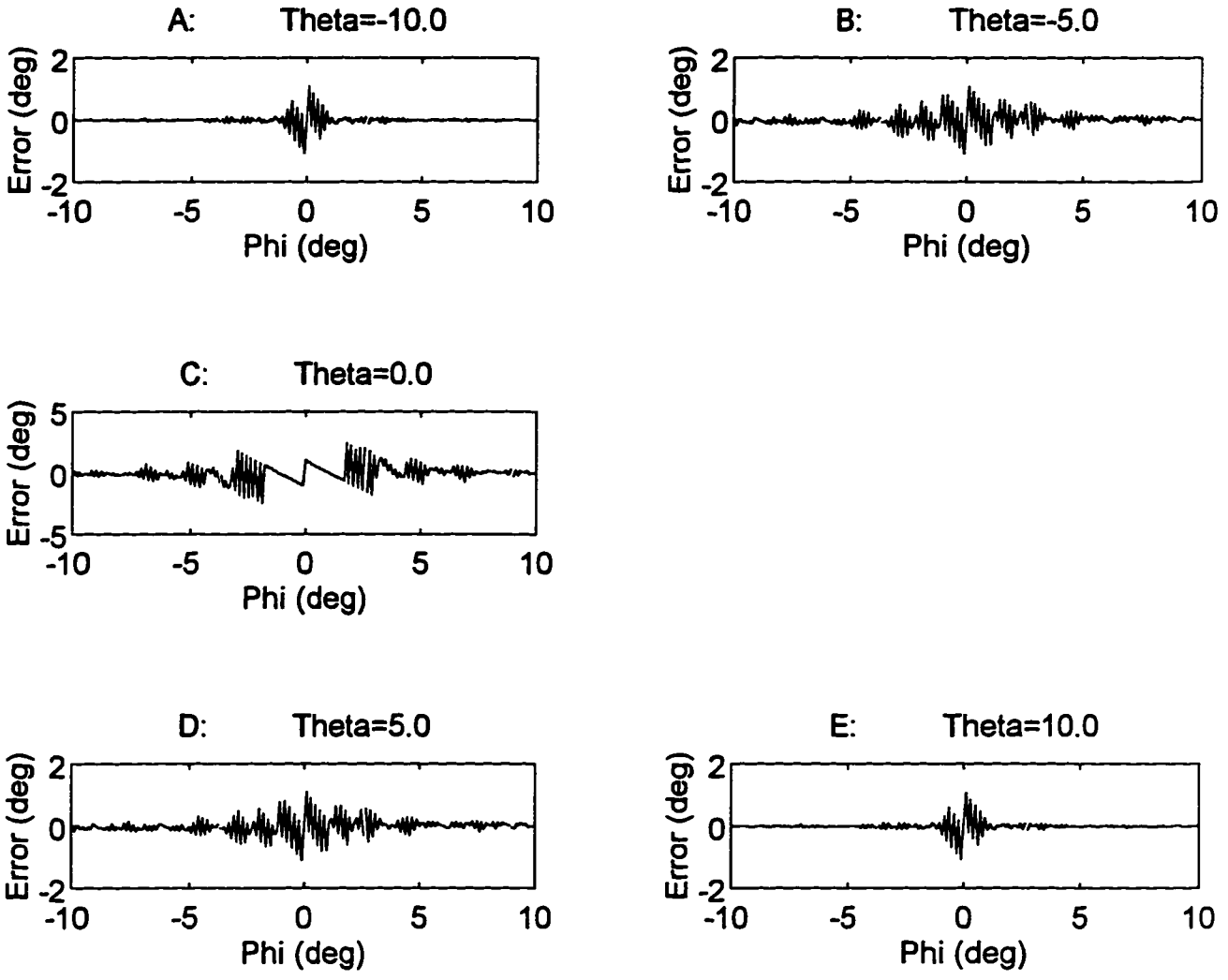


Figure 6: Computational Error of Ellipse Fitting as Function of Euler Angles Using Discrete Coordinates (800 Samples)

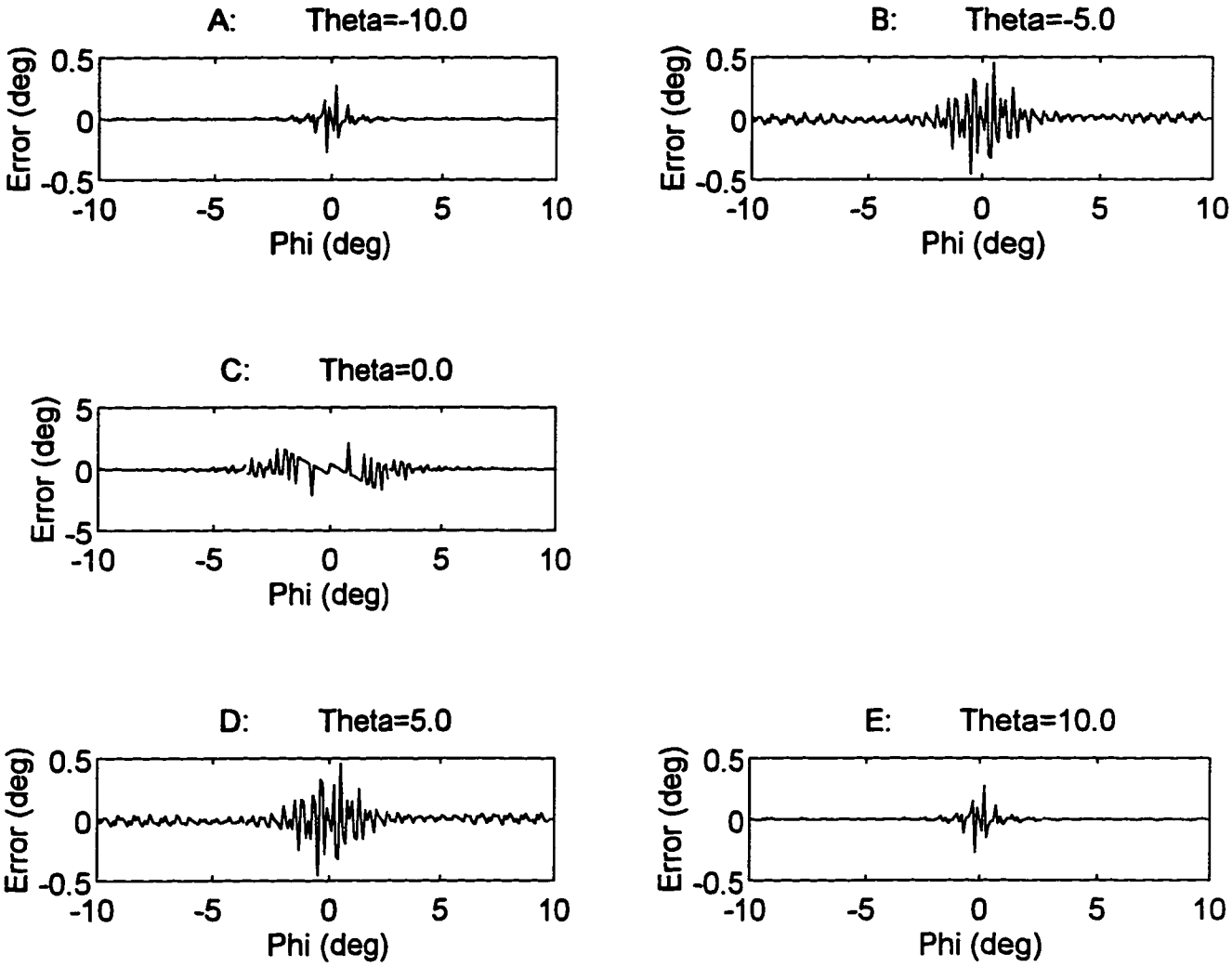


Figure 7: Computational Error of Ellipse Fitting as Function of Euler Angles Using Integer Coordinates for High Resolution Image (2000 Samples)

One way to keep a lower resolution image and partially overcome the problem of large computational error for eye orientation close to the reference position is to mount the camera at an angle relative to the optic axis of the eye when the eye is at the reference position. Under these circumstances, the projection of the boundary onto the image plane is an ellipse when the eye is looking straight ahead. If the eye position works over a range where the boundary is elliptical, the errors will be small and the accuracy of computing eye position will be high.

Another method for overcoming errors where the boundary is close to circular, is by computing the center of the ellipse from the fitted boundary. The center of the ellipse is dependent on all parameters of the fitted ellipse and is highly sensitive to changes in eye orientation at small angles. Therefore, a combined approach using ellipse eccentricity, readjusting of camera position and finding the center of the ellipse would add robustness to the algorithm.

3.3.4 ALGORITHM FOR CENTER FINDING

The projection of the pupil contour onto the image plane is an ellipse which is given by Eq. (18). The center of the ellipse projected onto the image plane corresponds to the intersection of the optic axis with the image plane. The projection of the iral boundary onto the image plane is an ellipse with center corresponding to the limbus center. Both can be computed directly from the parameters of the equation for the ellipse.

First the angle between the major or the minor axis of the ellipse and X or Z axis of the image plane is obtained from the parameters of its equation. Then a rotation

transformation is performed on the X-Z coordinate system of the image plane, so that the major and the minor axes are parallel to the rotated X and Z axes (Figure 8). This rotation transformation will simplify the equation of the ellipse by eliminating cross terms. The center of the ellipse is found from the simplified elliptical equation as follows:

The general equation for an ellipse is:

$$Ax^2 + Bxz + Cz^2 + Dx + Ez + F = 0 \tag{21}$$

Let α be the angle between the major axis of the ellipse and X axis of the image plane (Figure 8), we have (refer to Appendix A):

$$\alpha = \arctan\left(\frac{C - A \pm \sqrt{A^2 - 2AC + C^2 + B^2}}{B}\right) \tag{22}$$

Now we perform a rotation transformation on the X-Z coordinates of the image plane (Figure 8), such that the major and minor axes of the ellipse are parallel to the rotated X and Z axes, i.e. X' and Z' axes respectively (Figure 8). Mathematically, the above rotation transformation is described by the following equations:

$$\begin{aligned} x &= x'\cos\alpha - z'\sin\alpha \\ z &= x'\sin\alpha + z'\cos\alpha \end{aligned} \tag{23}$$

Therefore the equation for the ellipse in the rotated coordinate frame is obtained by transforming Eq. (21), using Eq. (23), which is shown below:

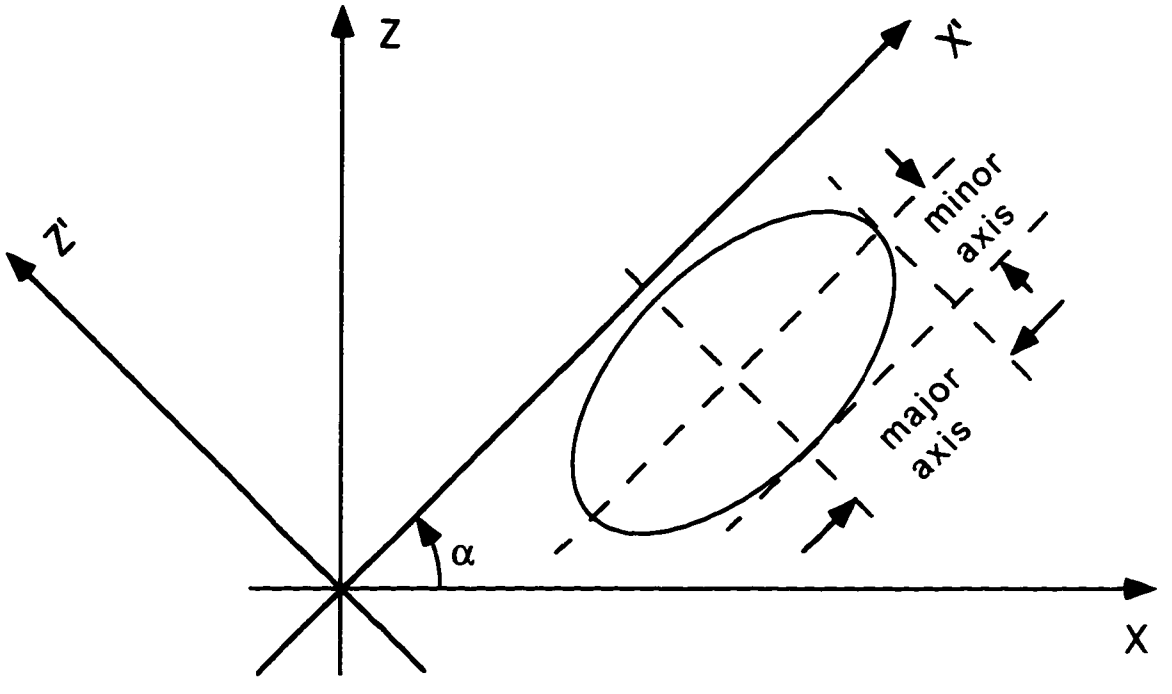


Figure 8: Rotational Transform of Two Dimensional Coordinate System

$$A'x'^2 + C'z'^2 + D'x' + E'z' + F' = 0 \tag{24}$$

where:

$$\begin{aligned} A' &= A\cos^2\alpha + B\cos\alpha\sin\alpha + C\sin^2\alpha \\ C' &= A\sin^2\alpha - B\cos\alpha\sin\alpha + C\cos^2\alpha \\ D' &= D\cos\alpha + E\sin\alpha \\ E' &= -D\sin\alpha + E\cos\alpha \\ F' &= F \end{aligned} \tag{25}$$

Eq. (24) is equivalent to the following:

$$\frac{\left(x' + \frac{D'}{2A'}\right)^2}{\frac{F''}{A'}} + \frac{\left(z' + \frac{E'}{2C'}\right)^2}{\frac{F''}{C'}} = 1 \tag{26}$$

where

$$F'' = \frac{D'^2}{4A'} + \frac{E'^2}{4C'} - F' \tag{27}$$

From Eq. (26), the center of the ellipse in the rotated image plane X'-Z' is: (-D'/(2A'), -E'/(2C')). Using Eq. (23) we can find the coordinates of the center of the ellipse in the original X-Z coordinate system of the image plane:

$$x_0 = -\frac{D'}{2A'}\cos\alpha + \frac{E'}{2C'}\sin\alpha$$
$$z_0 = -\frac{D'}{2A'}\sin\alpha - \frac{E'}{2C'}\cos\alpha \qquad (28)$$

3.3.5 RELATIONSHIP BETWEEN CENTER AND EYE ORIENTATION

Assume the distance between the iral plane and the center of the eye is d (Figure 1). When the eye is in the reference position, the center of the projected ellipse is given by (0, -d, 0) in camera-coordinate-frame. Torsional movement about an axis corresponding to the ellipse center does not change the center of the projected ellipse. Therefore, R_ψ can be ignored and $M_\beta^\beta(R)$ (Eq. (5)) is simplified as Eq. (9). The two Euler angles ϕ and θ can then be found from the center of the ellipse in the camera-coordinate-frame by mapping it by Eq. (9) as follows:

$$\begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi\cos\theta & \sin\phi\sin\theta \\ \sin\phi & \cos\phi\cos\theta & -\cos\phi\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 0 \\ -d \\ 0 \end{pmatrix} = \begin{pmatrix} d\sin\phi\cos\theta \\ -d\cos\phi\cos\theta \\ -d\sin\theta \end{pmatrix} \qquad (29)$$

where (x_0, y_0, z_0) is the coordinate vector of the rotated ellipse center.

The coordinates of the projection of the center onto the image plane are:

$$(x_0, z_0) = (d\sin\phi\cos\theta, -d\sin\theta) \tag{30}$$

Using Eq. (28) for finding the center of the ellipse and then using Eq. (30) to find the corresponding ϕ and θ, we obtain a formula for calculating the angles of the horizontal and the vertical eye rotations as follows:

$$\begin{aligned} \theta &= \arcsin\left(-\frac{z_0}{d}\right) \\ \phi &= \arcsin\left(\frac{x_0}{d\cos\theta}\right) \end{aligned} \tag{31}$$

By combining the ellipse fitting and center finding algorithms, a potentially robust estimate of pitch and yaw eye position over a wide range of angles has been obtained.

3.4 ALGORITHM FOR DETERMINING TORSION

3.4.1 CONCEPTUAL BASIS OF TORSION ALGORITHM

The algorithm for determining torsion of the eye is based on the gray level variations in images of the iris. Hatamian and Anderson (1983) found that the bandwidth in the angular direction was much higher than that in the radial direction. That is, most of the information is present along the angular direction. Using samples of the iral image on a circle about the optic axis can produce very good estimates of torsional rotation angles (Hatamian and Anderson 1983). One disadvantage of this technique is that portions of the circle may be blocked at specific eye positions. The gray level distribution

of a sampling annulus, or iral signature would give a more robust estimate of ocular torsion over a wider range of eye positions (Figure 9A).

We first consider the torsion determination problem assuming that the eye is in the reference position and there are no horizontal and vertical eye rotations. Under these conditions, the annulus can be mapped to a rectangle (Figure 9B), and a template matching algorithm can be utilized. In section **3.7**, we relax this restriction to show how to compute torsion about the optic axis when the eye has rotated about a pitch and/or yaw axis.

3.4.2 DETERMINATION OF PIXEL COORDINATES WITHIN SAMPLING ANNULUS

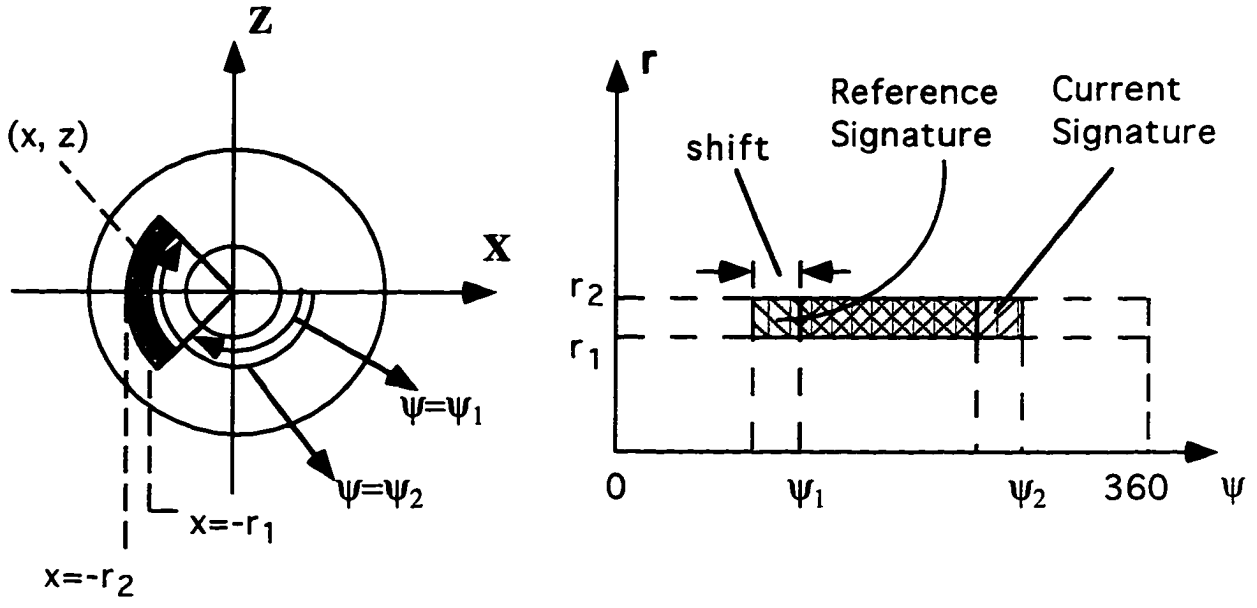
When the eye is in the reference position undergoing pure rotations about the optic axis, the annulus moves along a circle when projected onto the image plane. The center of the circle corresponds to the center of the pupil which in turn coincides with the center of the image plane. The pixel coordinates for obtaining image intensity are given in rectangular coordinates. Therefore, although the arrays for the iral signatures are maintained in polar coordinates, the information on intensity obtained from the frame grabber is given in rectangular coordinates. The relationship between the rectangular coordinates of the annulus image and its representation in the polar coordinate space is shown in Figure 9B. The rectangular coordinates of each pixel within the annulus (x, z) is obtained from the corresponding polar coordinates as:

$$\begin{aligned} x &= r\cos\psi \\ z &= -r\sin\psi \end{aligned} \tag{32}$$

A:

Iral Sampling Annulus Before and After Torsion

B:



Mapping from Annulus to Rectangle

Figure 9: Mapping Sampling Annulus to Rectangular Form

In the polar coordinate space, $g(\psi, r)$ is the gray level distribution function of the annulus, r_1 and r_2 are the radii of the inner and outer arcs respectively, and ψ_1 and ψ_2 are the starting and ending angles of the arc (Figure 9B). The parameters r_1, r_2, ψ_1 and ψ_2 are selected at calibration time. The gray level $g(\psi, r)$ at (ψ, r) corresponds to the intensity f(x,z) at pixel coordinates (x, z). The discrete form of the iral signature in polar coordinate space can then be given in terms of the discrete values at the pixel locations as:

$$\begin{aligned} g(r,\psi) &= f(x,z) = f(r\cos\psi, -r\sin\psi) \\ &= f((r_1+k*r_{inc})\cos(\psi_1+l*\psi_{inc}), -(r_1+k*r_{inc})\sin(\psi_1+l*\psi_{inc})) \end{aligned} \qquad (33)$$

$$(r_1 \le r \le r_2; \quad \psi_1 \le \psi \le \psi_2; \quad 0 \le k \le m; \quad 0 \le l \le n)$$

where r_{inc} and ψ_{inc} are the sampling intervals along the radial and angular directions. For simplicity, r_{inc} has been chosen equal to 1. The sampling interval ψ_{inc} approximately determines the torsional resolution of the algorithm. At each radius, it is approximately given by:

$$\psi_{inc} = \frac{360}{6r}(Degree) \qquad (34)$$

3.4.3 DETERMINATION OF TORSION

A template matching algorithm can be used to compute the gray level difference between the current signature and the reference signature in the polar coordinate space. This algorithm defines a "metric" or distance function between the two iral patterns. A

minimum value for this metric determines the shift of the iral signature $g(\psi, r)$ along ψ axis (Figure 9B). The rotation of the annulus is proportional to the shift of the rectangle which in turn is proportional to the torsion about the optic axis. The minimum distance is found by evaluating the metric at all positions along ψ axis using the following equation:

$$dist(np) = \frac{1}{ns} \sum_{k=0}^{m} \sum_{l=0}^{n} |g_1(r_1+k*r_{inc}, \psi_1+l*\psi_{inc}) - g_2(r_1+k*r_{inc}, \psi_1+(l+np)*\psi_{inc})| \quad (35)$$

$$(0 \le np \le n)$$

where np is the amount of shift in pixels and ns is the number of samples used for template matching (refer to section **4.11.1** for detail). If the two signatures contain no bright spot caused by the reflection of the infra-red LEDs, ns=(m+1)(n+1). Otherwise, the bright spots are removed, and ns<(m+1)(n+1). The minimum value for dist(np) determines the angular shift (np) between the current and reference iral signatures which is proportional to the torsion. The torsion is computed as the product of the shift in number of pixels (np) and the angular increment (ψ_{inc}), given by:

$$\psi = np*\psi_{inc} \quad (36)$$

Although the shift between the current and the reference iral signatures could be found by cross-correlating the two signatures using FFT (Moore et al. 1996), this technique operates on frequency domain and does not discriminate between gray level

noise caused by the infra-red reflection of the LEDs and the actual iral signature. The template matching algorithm (Eq. (35)) can effect a removal of bright spot caused by light reflection from the iral signatures. In this way, it can compute more robustly the shift between the current and reference iral signatures.

3.5 GENERALIZATION OF EYE ORIENTATION ALGORITHMS

The algorithms derived in sections **3.2, 3.3,** and **3.4** for computing the horizontal, vertical and torsional eye rotations have been based on the following two assumptions:

1) The camera-coordinate-frame coincides with the reference-coordinate-frame and the image plane of the camera is perpendicular to the optic axis of the eye when the eye is looking straight ahead.

2) The center of the image plane is the intersection of the roll axis (Y) of the camera-coordinate-frame with the image plane, i.e. it coincides with the projection of the center of the camera-coordinate-frame onto the image plane.

Practically, these assumptions must be relaxed because the camera may be tilted horizontally, vertically and torsionally relative to the reference-coordinate-frame which is coincident with the eye-fixed coordinate frame when the eye is in the reference position (Figure 10). This misalignment can be handled mathematically as a simple rotation of the camera-coordinate-frame relative to the reference-coordinate-frame. The center of the image plane may not coincide with the projection of the center of the camera-coordinate-frame onto the image plane. This offset can be treated as a translation of the image plane relative to the projection of the center of the camera-coordinate-frame onto the image plane.

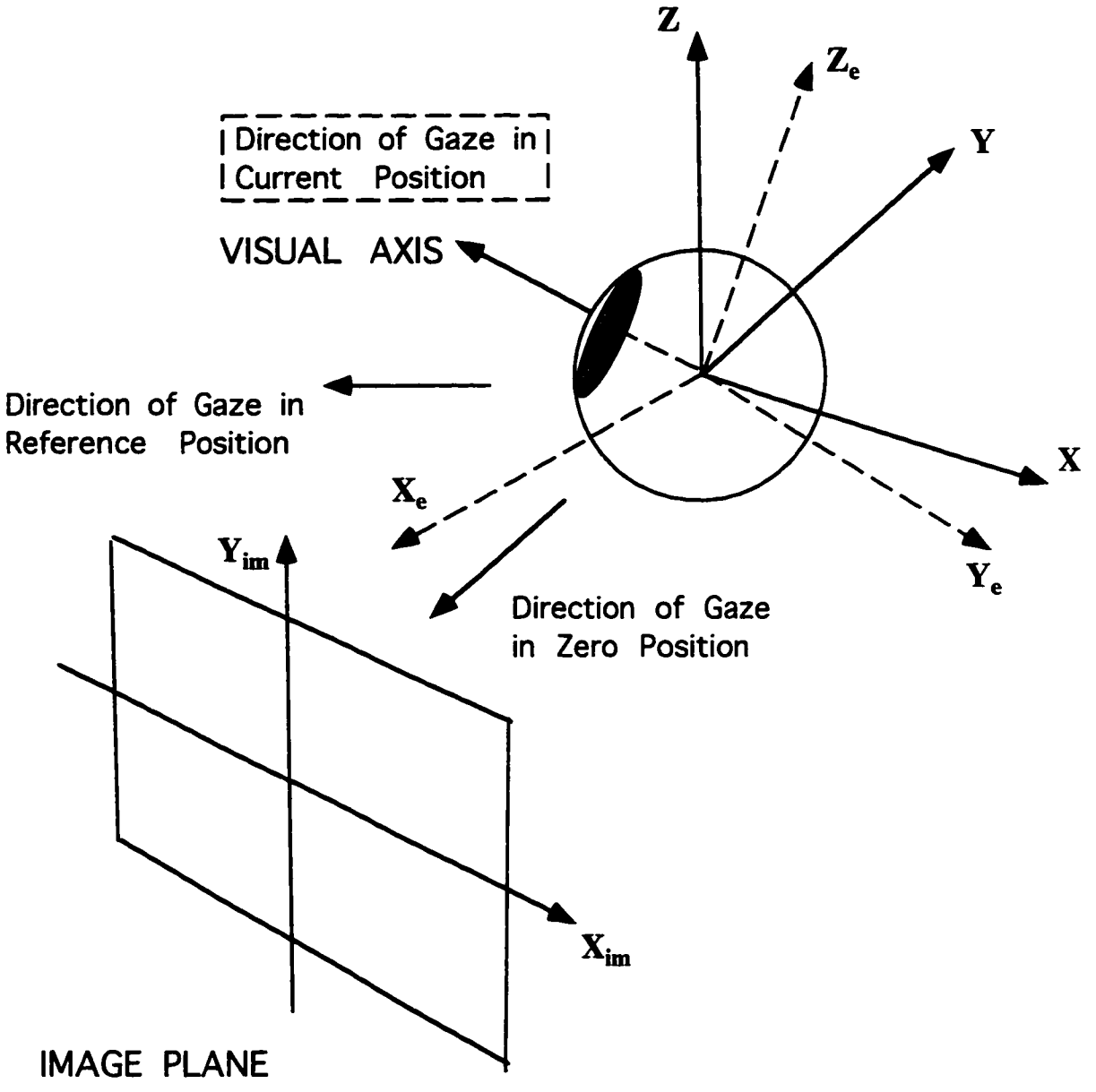


Figure 10: Eye Position in Head As Seen by Camera when Image Plane is Tilted And Translated

When the camera is tilted horizontally, vertically and torsionally relative to the reference-coordinate-frame, an eye rotation can be decomposed into two rotations: a first rotation from the zero position where the eye is looking straight at the camera to the reference position where eye is looking straight ahead without any torsion, and a second rotation from the reference position to the current position. When the eye is looking straight at the camera, Y_e axis (along the optic axis of the eye-fixed coordinate frame) is orthogonal to the image plane and X_e and Z_e axes of the eye-fixed coordinate system are parallel to X_{im} and Z_{im} axes of the image plane (Figure 10).

Let R_{eye} be the matrix representing eye rotation from the reference position to the current position, and R_{cal} be the matrix representing rotation from the zero position to the reference position. Let P be a point on the iris. The coordinates of P in the camera-coordinate-frame when the eye is at the zero, the reference and the current positions are (x", y", z"), (x', y', z') and (x, y, z) respectively. Then the relationship between positions of point P before and after the eye rotation is given by (compare with Eq. (5)):

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = R_{eye} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = R_{cal}\, R_{eye} \begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix} \tag{37}$$

where R_{eye} and R_{cal} are given by:

$$R_{eye} = R_{\phi}\, R_{\theta}\, R_{\psi} =$$

$$\begin{pmatrix} \cos\phi\cos\psi-\sin\phi\sin\theta\sin\psi & -\sin\phi\cos\theta & \cos\phi\sin\psi+\sin\phi\sin\theta\cos\psi \\ \sin\phi\cos\psi+\cos\phi\sin\theta\sin\psi & \cos\phi\cos\theta & \sin\phi\sin\psi-\cos\phi\sin\theta\cos\psi \\ -\cos\theta\sin\psi & \sin\theta & \cos\theta\cos\psi \end{pmatrix} \tag{38}$$

and

$$R_{cal} = \begin{pmatrix} \cos\phi_c & -\sin\phi_c & 0 \\ \sin\phi_c & \cos\phi_c & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_c & -\sin\theta_c \\ 0 & \sin\theta_c & \cos\theta_c \end{pmatrix} \begin{pmatrix} \cos\psi_c & 0 & \sin\psi_c \\ 0 & 1 & 0 \\ -\sin\psi_c & 0 & \cos\psi_c \end{pmatrix} =$$

$$\begin{pmatrix} \cos\phi_c\cos\psi_c-\sin\phi_c\sin\theta_c\sin\psi_c & -\sin\phi_c\cos\theta_c & \cos\phi_c\sin\psi_c+\sin\phi_c\sin\theta_c\cos\psi_c \\ \sin\phi_c\cos\psi_c+\cos\phi_c\sin\theta_c\sin\psi_c & \cos\phi_c\cos\theta_c & \sin\phi_c\sin\psi_c-\cos\phi_c\sin\theta_c\cos\psi_c \\ -\cos\theta_c\sin\psi_c & \sin\theta_c & \cos\theta_c\cos\psi_c \end{pmatrix} \tag{39}$$

If the camera is also translated relative to the camera-coordinate-frame, then the center of the camera-coordinate-frame will project to a point (x_c, z_c) in the image plane. Let (x_p, z_p) be the coordinates of the projection of point P in current position onto the image plane. Then the relationship between the projection of point P onto the image plane and point P in the camera-coordinate-frame is given below:

$$\begin{pmatrix} x_p \\ z_p \end{pmatrix} = \begin{pmatrix} x \\ z \end{pmatrix} + \begin{pmatrix} x_c \\ z_c \end{pmatrix} \tag{40}$$

R_{cal} and (x_c, z_c) have to be determined during calibration.

3.6 CALIBRATION PROCEDURE

3.6.1 CONCEPTUAL BASIS

To develop general algorithms to compute the horizontal, vertical and torsional rotations of the eye, seven parameters (ϕ_c, θ_c, ψ_c, x_c, z_c, **d**, and **r**) must be identified:

1) ϕ_c, θ_c and ψ_c are the Euler angles associated with the camera offset relative to the reference-coordinate-frame.

2) (x_c, z_c) are the coordinates of the projection of the center of the camera-coordinate-frame onto the image plane.

3) **d** is the distance between the center of the eye and the iral plane (Figure 1).

4) **r** is the radius of the iris.

The calibration algorithm is based on the idea that at different eye orientations, the optic axis which goes through the center of the pupil into different positions on the image plane as the eye rotates. The calibration parameters can then be computed using the projected pupil centers at different known eye orientations. Moore et al. (1997) have computed ψ_c, (x_c, z_c) and **d** by asking subjects to look at four different positions with known angles in the front. In general, the projections of the pupil center at these positions were found. Formulae were then applied to compute ϕ_c, θ_c, ψ_c, (x_c, z_c) and **d** using the coordinates of the projection of the pupil centers at different positions.

In order to obtain accurate estimates of the camera offset, we found experimentally that computation of the coordinates of the pupil centers be accurate to within 0.5 pixel. This was done choosing 25 calibration eye positions and introducing random noise of up to 1.5 pixels in center coordinates. The errors found in computing

ϕ_c, θ_c and ψ_c could be more than 3° and more than 5 pixels for (x_c, z_c) and **d**. Increasing the number of calibration positions would increase the accuracy, but it did not produce an acceptable result. For 121 eye positions, the errors could be more than 1° for ϕ_c, θ_c and ψ_c, and more than 2 pixels for (x_c, z_c) and **d**. For 961 eye positions, the result yielded 0.5° accuracy in ϕ_c, θ_c, ψ_c and less than 1 pixel accuracy for (x_c, z_c) and **d**. However, using such a large number of calibration positions is impractical. Thus, calibrating the system using only the coordinates of the pupil centers at different positions is not a practical calibration procedure. Most applications which utilize the pupil centers for calibration mount the camera on goggles such that the "line of sight" of the camera is parallel to the "line of sight" of the eye when the eye is looking straight ahead (Moore et al. 1996).

In our study of an artificial eye, we have calibrated the system from the pupil center projection at different eye orientations and from the projection of the contour ellipses. We have found this dual approach to be more robust with regard to image noise which causes errors in finding the center of pupil. We have simulated the ellipse algorithm using 4 eye positions with a noise level of up to 1.5 pixels in edge coordinates and up to 1.5° in eye orientations. The accuracy was 0.1° for ϕ_c and θ_c, and 0.2 pixel for (x_c, z_c) and **r**.

However the ellipse algorithm can only solve for five of the seven calibration parameters: ϕ_c, θ_c, x_c, z_c, and **r**. There are also two possible solutions for ϕ_c and θ_c. On the other hand, we have found that the center algorithm is robust in computing ψ_c and d despite of the errors in the coordinates of the pupil centers. Therefore the ellipse algorithm is used in conjunction with the center algorithm to compute the calibration

parameters: ϕ_c, θ_c, ψ_c, x_c, z_c, **d** and **r**. The approximate value for ϕ_c and θ_c are used by the ellipse algorithm to find the final answers for ϕ_c and θ_c.

In this study, images of the eye looking at 16 or more specified positions with known eye orientations were grabbed. The iral or pupil boundaries were detected by the contour following algorithm. Then the equations for the projected boundaries were found by fitting the edge points to the ellipses using the SVD algorithm. From these equations the pupil centers were computed. The center algorithm was then used to compute ψ_c and **d** from the coordinates of the centers and the corresponding eye orientations. Then the ellipse algorithm was used to compute the remaining five of the seven parameters (ϕ_c, θ_c, x_c, z_c, and **r**), using ψ_c, **d** and the approximate values for ϕ_c and θ_c computed by the center algorithm.

We now derive equations for the center algorithm which computes the camera offset angles and the other parameters from the coordinates of the projected pupil centers. Then we derive equations for the ellipse algorithm which computes the camera offset angles and the other parameters from the equations for the projected ellipses. The parameters: ϕ_c, θ_c, ψ_c, x_c, z_c, **d** and **r** completely determine the calibration so that any eye orientation can be determined relative to the reference position.

3.6.2 CALIBRATION USING PUPIL/IRIS CENTER

The torsional eye movement about the optic axis does not change the projection of the center of the pupil. Therefore, ψ has been arbitrarily set to zero (ψ=0). When the eye is looking straight at the camera, the coordinates of the pupil center is (0, -d, 0) in the camera-coordinate-frame (refer to Appendix C for notes on spherical eye model).

When the eye is rotated by ϕ, θ, and ψ, the coordinates of the pupil center in the camera-coordinate-frame is given by (Eq. (37)):

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = R_{eye} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = R_{cal}\, R_{eye} \begin{pmatrix} 0 \\ -d \\ 0 \end{pmatrix} \tag{41}$$

Substituting Eqs. (38) and (39) into Eq. (41), we obtain:

$$\begin{aligned} x &= d\sin\phi\cos\theta(\cos\phi_c\cos\psi_c - \sin\phi_c\sin\theta_c\sin\psi_c) \\ &\quad + d\cos\phi\cos\theta\sin\phi_c\cos\theta_c \\ &\quad - d\sin\theta(\cos\phi_c\sin\psi_c + \sin\phi_c\sin\theta_c\cos\psi_c) \\ y &= d\sin\phi\cos\theta(\sin\phi_c\cos\psi_c + \cos\phi_c\sin\theta_c\sin\psi_c) \\ &\quad - d\cos\phi\cos\theta\cos\phi_c\cos\theta_c \\ &\quad - d\sin\theta(\sin\phi_c\sin\psi_c - \cos\phi_c\sin\theta_c\cos\psi_c) \\ z &= -d\sin\phi\cos\theta\cos\theta_c\sin\psi_c \\ &\quad - d\cos\phi\cos\theta\sin\theta_c - d\sin\theta\cos\theta_c\cos\psi_c \end{aligned} \tag{42}$$

Using Eq. (40), the projection of the pupil center onto the image plane is given by:

$$
\begin{aligned}
x_p &= d\sin\phi\cos\theta(\cos\phi_c\cos\psi_c - \sin\phi_c\sin\theta_c\sin\psi_c) \\
&\quad + d\cos\phi\cos\theta\sin\phi_c\cos\theta_c \\
&\quad - d\sin\theta(\cos\phi_c\sin\psi_c + \sin\phi_c\sin\theta_c\cos\psi_c) + x_c \\
z_p &= -d\sin\phi\cos\theta\cos\theta_c\sin\psi_c - d\cos\phi\cos\theta\sin\theta_c \\
&\quad - d\sin\theta\cos\theta_c\cos\psi_c + z_c
\end{aligned}
\tag{43}
$$

If we set:

$$
\begin{aligned}
A &= d(\cos\phi_c\cos\psi_c - \sin\phi_c\sin\theta_c\sin\psi_c) \\
B &= d\sin\phi_c\cos\theta_c \\
C &= -d(\cos\phi_c\sin\psi_c + \sin\phi_c\sin\theta_c\cos\psi_c) \\
D &= x_c
\end{aligned}
\tag{44}
$$

and

$$
\begin{aligned}
A' &= -d\cos\theta_c\sin\psi_c \\
B' &= -d\sin\theta_c \\
C' &= -d\cos\theta_c\cos\psi_c \\
D' &= z_c
\end{aligned}
\tag{45}
$$

and

$$\begin{aligned} x &= \sin\phi\cos\theta \\ y &= \cos\phi\cos\theta \\ z &= \sin\theta \end{aligned} \tag{46}$$

we have:

$$\begin{aligned} x_p &= Ax + By + Cz + D \\ z_p &= A'x + B'y + C'z + D' \end{aligned} \tag{47}$$

In the above equation, (x_p, z_p) which are the coordinates of the pupil center projected onto the image plane can be computed using Eq. (28) which gives the center of an ellipse in terms of parameters which describe it. The Euler angles for each position, ϕ and θ, are known angles, so x, y and z are known. Parameters A, B, C, D, A', B', C' and D' can therefore be found by fitting the 16 or more pupil center coordinates and the corresponding known angles to Eq. (47). Once these 8 parameters are found, ψ_c, d are obtained using the following equation:

$$\begin{aligned} \psi_c &= \arctan\left(\frac{A'}{C'}\right) \\ d &= \sqrt{A^2 + B^2 + C^2} \\ &= \sqrt{A'^2 + B'^2 + C'^2} \end{aligned} \tag{48}$$

The values of ϕ_c and θ_c are obtained from the following equations:

$$\theta_c = \arcsin\left(-\frac{B'}{d}\right)$$
$$\phi_c = \arcsin\left(\frac{B}{d\cos\theta_c}\right) \quad (49)$$

As described in section **3.6.1**, there is considerable noise associated with computing ϕ_c and θ_c. Therefore, this computation is used in conjunction with an ellipse fitting algorithm to determine the final values of the calibrated camera offset.

3.6.3 CALIBRATION USING ELLIPTICAL PROJECTION OF IRIS/PUPIL CONTOUR

As described in section **3.6.2**, the camera offset will be considered as a rotation of the camera-coordinate-frame, R_{cal}, relative to the reference-coordinate-frame and given by:

$$R_{cal} = \begin{pmatrix} \cos\phi_c\cos\psi_c - \sin\phi_c\sin\theta_c\sin\psi_c & -\sin\phi_c\cos\theta_c & \cos\phi_c\sin\psi_c + \sin\phi_c\sin\theta_c\cos\psi_c \\ \sin\phi_c\cos\psi_c + \cos\phi_c\sin\theta_c\sin\psi_c & \cos\phi_c\cos\theta_c & \sin\phi_c\sin\psi_c - \cos\phi_c\sin\theta_c\cos\psi_c \\ -\cos\theta_c\sin\psi_c & \sin\theta_c & \cos\theta_c\cos\psi_c \end{pmatrix} \quad (50)$$

Since the torsional movement of the eye about the optic axis does not change the shape and the orientation of the major and the minor axes of the ellipse projected onto the image plane, the torsion is ignored when the transformation of the iral contour projected

onto the image plane is considered. We will therefore assume $\psi=0$ for the eye rotation relative to the camera-coordinate-frame. The simplified expression for R_{eye} is given as:

$$R_{eye} = \begin{pmatrix} \cos\phi & -\sin\phi\cos\theta & \sin\phi\sin\theta \\ \sin\phi & \cos\phi\cos\theta & -\cos\phi\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \tag{51}$$

Assume the coordinates of an iral contour point in the camera-coordinate-frame are (x", -d, z") when the eye is looking straight at the camera and (x, y, z) when the eye is rotated to the current position. Because the eye rotation is relative to the reference-coordinate-frame to which the camera is relatively rotated, the coordinates (x", -d, z") and (x, y, z) are related by Eq. (37) and is given by:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = R_{cal}\, R_{eye} \begin{pmatrix} x'' \\ -d \\ z'' \end{pmatrix} \tag{52}$$

Substitute Eq. (51) and Eq. (50) into the above equation, we have:

$$
\begin{aligned}
x &= (\cos\phi\cos\phi_c\cos\psi_c - \cos\phi\sin\phi_c\sin\theta_c\sin\psi_c - \sin\phi\sin\phi_c\cos\theta_c)x'' \\
&\quad + (\sin\phi\cos\theta\cos\phi_c\cos\psi_c - \sin\phi\cos\theta\sin\phi_c\sin\theta_c\sin\psi_c \\
&\quad + \cos\phi\cos\theta\sin\phi_c\cos\theta_c - \sin\theta\cos\phi_c\sin\psi_c \\
&\quad - \sin\theta\sin\phi_c\sin\theta_c\cos\psi_c)d + (\sin\phi\sin\theta\cos\phi_c\cos\psi_c \\
&\quad - \sin\phi\sin\theta\sin\phi_c\sin\theta_c\sin\psi_c + \cos\phi\sin\theta\sin\phi_c\cos\theta_c \\
&\quad + \cos\theta\cos\phi_c\sin\psi_c + \cos\theta\sin\phi_c\sin\theta_c\cos\psi_c)z'' \\
y &= (\cos\phi\sin\phi_c\cos\psi_c + \cos\phi\cos\phi_c\sin\theta_c\sin\psi_c + \sin\phi\cos\phi_c\cos\theta_c)x'' \\
&\quad + (\sin\phi\cos\theta\sin\phi_c\cos\psi_c + \sin\phi\cos\theta\cos\phi_c\sin\theta_c\sin\phi_c \\
&\quad - \cos\phi\cos\theta\cos\phi_c\cos\theta_c - \sin\theta\sin\phi_c\sin\psi_c \\
&\quad + \sin\theta\cos\phi_c\sin\theta_c\cos\psi_c)d + (\sin\phi\sin\theta\sin\phi_c\cos\psi_c \\
&\quad + \sin\phi\sin\theta\cos\phi_c\sin\theta_c\sin\psi_c - \cos\phi\sin\theta\cos\phi_c\cos\theta_c \\
&\quad + \cos\theta\sin\phi_c\sin\psi_c - \cos\theta\cos\phi_c\sin\theta_c\cos\psi_c)z'' \\
z &= (\sin\phi\sin\theta_c - \cos\phi\cos\theta_c\sin\psi_c)x'' - (\sin\phi\cos\theta\cos\theta_c\sin\psi_c \\
&\quad + \cos\phi\cos\theta\sin\theta_c + \sin\theta\cos\theta_c\cos\psi_c)d \\
&\quad + (\cos\theta\cos\theta_c\cos\psi_c - \sin\phi\sin\theta\cos\theta_c\sin\psi_c - \cos\phi\sin\theta\sin\theta_c)z''
\end{aligned}
\tag{53}
$$

Assume (x_p, z_p) is the projection of (x, y, z) onto the image plane and (x_c, z_c) is the projection of the center of the camera-coordinate-frame onto the image plane. Using Eq. (40) and Eq. (53), we have:

$$
\begin{aligned}
x_p &= (\cos\phi\cos\phi_c\cos\psi_c - \cos\phi\sin\phi_c\sin\theta_c\sin\psi_c - \sin\phi\sin\phi_c\cos\theta_c)x'' \\
&\quad + (\sin\phi\cos\theta\cos\phi_c\cos\psi_c - \sin\phi\cos\theta\sin\phi_c\sin\theta_c\sin\psi_c \\
&\quad + \cos\phi\cos\theta\sin\phi_c\cos\theta_c - \sin\theta\cos\phi_c\sin\psi_c \\
&\quad - \sin\theta\sin\phi_c\sin\theta_c\cos\psi_c)d + (\sin\phi\sin\theta\cos\phi_c\cos\psi_c \\
&\quad - \sin\phi\sin\theta\sin\phi_c\sin\theta_c\sin\psi_c + \cos\phi\sin\theta\sin\phi_c\cos\theta_c \\
&\quad + \cos\theta\cos\phi_c\sin\psi_c + \cos\theta\sin\phi_c\sin\theta_c\cos\psi_c)z'' + x_c \\
z_p &= (\sin\phi\sin\theta_c - \cos\phi\cos\theta_c\sin\psi_c)x'' - (\sin\phi\cos\theta\cos\theta_c\sin\psi_c \\
&\quad + \cos\phi\cos\theta\sin\theta_c + \sin\theta\cos\theta_c\cos\psi_c)d + (\cos\theta\cos\theta_c\cos\psi_c \\
&\quad - \sin\phi\sin\theta\cos\theta_c\sin\psi_c - \cos\phi\sin\theta\sin\theta_c)z'' + z_c
\end{aligned}
\tag{54}
$$

The above equation is equivalent to the following:

$$
\begin{aligned}
x_p &= a_1x'' + b_1d + c_1z'' + x_c \\
z_p &= a_2x'' + b_2d + c_2z'' + z_c
\end{aligned}
\tag{55}
$$

where

$$
\begin{aligned}
a_1 &= \cos\phi\cos\phi_c\cos\psi_c - \cos\phi\sin\phi_c\sin\theta_c\sin\psi_c - \sin\phi\sin\phi_c\cos\theta_c \\
b_1 &= \sin\phi\cos\theta\cos\phi_c\cos\psi_c - \sin\phi\cos\theta\sin\phi_c\sin\theta_c\sin\psi_c \\
&\quad + \cos\phi\cos\theta\sin\phi_c\cos\theta_c - \sin\theta\cos\phi_c\sin\psi_c \\
&\quad - \sin\theta\sin\phi_c\sin\theta_c\cos\psi_c \\
c_1 &= \sin\phi\sin\theta\cos\phi_c\cos\psi_c - \sin\phi\sin\theta\sin\phi_c\sin\theta_c\sin\psi_c \\
&\quad + \cos\phi\sin\theta\sin\phi_c\cos\theta_c + \cos\theta\cos\phi_c\sin\psi_c \\
&\quad + \cos\theta\sin\phi_c\sin\theta_c\cos\psi_c \\
a_2 &= \sin\phi\sin\theta_c - \cos\phi\cos\theta_c\sin\psi_c \\
b_2 &= -(\sin\phi\cos\theta\cos\theta_c\sin\psi_c + \cos\phi\cos\theta\sin\theta_c + \sin\theta\cos\theta_c\cos\psi_c) \\
c_2 &= \cos\theta\cos\theta_c\cos\psi_c - \sin\phi\sin\theta\cos\theta_c\sin\psi_c - \cos\phi\sin\theta\sin\theta_c
\end{aligned}
\tag{56}
$$

Solving for x" and z", we have:

$$
\begin{aligned}
x'' &= \frac{c_2x_p - c_1z_p + (b_2c_1 - b_1c_2)d - c_2x_c + c_1z_c}{a_1c_2 - a_2c_1} \\
z'' &= \frac{a_2x_p - a_1z_p + (a_1b_2 - a_2b_1)d + a_1z_c - a_2x_c}{a_2c_1 - a_1c_2}
\end{aligned}
\tag{57}
$$

Since the projection of the iral contour onto the image plane is a perfect circle when the eye is looking straight at the camera, we have:

$$
x''^2 + z''^2 = r^2 \tag{58}
$$

Substituting Eq. (57), (122) into Eq. (58), (123), we have:

$$\left(\frac{c_2x_p - c_1z_p + (b_2c_1 - b_1c_2)d - c_2x_c + c_1z_c}{a_1c_2 - a_2c_1}\right)^2 + \left(\frac{a_2x_p - a_1z_p + (a_1b_2 - a_2b_1)d + a_1z_c - a_2x_c}{a_2c_1 - a_1c_2}\right)^2 = r^2 \quad (59)$$

The above equation is equivalent to the following:

$$x_p^2 = Az_p^2 + Bx_pz_p + Cx_p + Dz_p + E \quad (60)$$

The parameters A, B, C, D and E define the ellipse when the eye is looking into some position in terms of the camera offset angles and eye orientation angles (See Appendix B for a complete derivation).

A general formula for obtaining the calibration parameters in terms of eye orientations is not solvable in closed form. However, the calibration parameters can be obtained in closed form if calibrated eye orientations are taken along the horizontal or vertical directions. If we allow $\phi=0$ and $\theta=\theta_0$, i.e the eye rotates only vertically, we have:

$$\theta_c = \pm\arcsin\left(\sqrt{\frac{2(A+1)+B^2\pm 2\sqrt{(A+1)^2+B^2}}{(4A+B^2)(1-\sin^2\theta_0\sin^2\psi_c)}}\right)-\arctan(\tan\theta_0\cos\psi_c)$$

$$\phi_c = \pm\arcsin$$

$$\left(\sqrt{\frac{2A(A+1)+B^2\pm 2A\sqrt{(A+1)^2+B^2}}{(4A+B^2)[(\cos\theta_0\cos\theta_c-\sin\theta_0\sin\theta_c\cos\psi_c)^2+(\sin\theta_0\sin\psi_c)^2]}}\right)$$

$$+\arctan\left(\frac{\sin\theta_0\sin\psi_c}{\cos\theta_0\cos\theta_c-\sin\theta_0\sin\theta_c\cos\psi_c}\right) \qquad (61)$$

$$x_c = \frac{2AC-DB}{4A+B^2}-b_1 d$$

$$z_c = -\frac{2D+BC}{4A+B^2}-b_2 d$$

$$r = \frac{\sqrt{(a_2^2+c_2^2)(Az_c^2+Bx_cz_c+Cx_c+Dz_c+E-x_c^2)+[(a_1b_2-a_2b_1)^2+(b_2c_1-b_1c_2)^2]d^2}}{|a_2c_1-a_1c_2|}$$

where

$$\begin{aligned}
a_1 &= \cos\phi_c\cos\psi_c-\sin\phi_c\sin\theta_c\sin\psi_c \\
b_1 &= \cos\theta_0\sin\phi_c\cos\theta_c-\sin\theta_0\cos\phi_c\sin\psi_c-\sin\theta_0\sin\phi_c\sin\theta_c\cos\psi_c \\
c_1 &= \sin\theta_0\sin\phi_c\cos\theta_c+\cos\theta_0\cos\phi_c\sin\psi_c+\cos\theta_0\sin\phi_c\sin\theta_c\cos\psi_c \\
a_2 &= -\cos\theta_c\sin\psi_c \\
b_2 &= -(\cos\theta_0\sin\theta_c+\sin\theta_0\cos\theta_c\cos\psi_c) \\
c_2 &= \cos\theta_0\cos\theta_c\cos\psi_c-\sin\theta_0\sin\theta_c
\end{aligned} \qquad (62)$$

If we allow $\phi=\phi_0$ and $\theta=0$, i.e. the eye rotates only horizontally, we have:

$$\theta_c = \pm\arcsin\left(\sqrt{\frac{2(A+1)+B^2\pm2\sqrt{(A+1)^2+B^2}}{(4A+B^2)(1-\sin^2\phi_0\cos^2\psi_c)}}\right)-\arctan(\tan\phi_0\sin\psi_c)$$

$$\phi_c = \pm\arcsin$$

$$\left(\sqrt{\frac{2A(A+1)+B^2\pm2A\sqrt{(A+1)^2+B^2}}{(4A+B^2)[(\cos\phi_0\cos\theta_c-\sin\phi_0\sin\theta_c\sin\psi_c)^2+(\sin\phi_0\cos\psi_c)^2]}}\right)$$

$$-\arctan\left(\frac{\sin\phi_0\cos\psi_c}{\cos\phi_0\cos\theta_c-\sin\phi_0\sin\theta_c\sin\psi_c}\right) \qquad (63)$$

$$x_c = \frac{2AC-BD}{4A+B^2}-b_1d$$

$$z_c = -\frac{BC+2D}{4A+B^2}-b_2d$$

$$r = \frac{\sqrt{(a_2^2+c_2^2)(Az_c^2+Bx_cz_c+Cx_c+Dz_c+E-x_c^2)+[(a_1b_2-a_2b_1)^2+(b_2c_1-b_1c_2)^2]d^2}}{|a_2c_1-a_1c_2|}$$

where

$$a_1 = \cos\phi_0\cos\phi_c\cos\psi_c-\cos\phi_0\sin\phi_c\sin\theta_c\sin\psi_c-\sin\phi_0\sin\phi_c\cos\theta_c$$

$$b_1 = \sin\phi_0\cos\phi_c\cos\psi_c-\sin\phi_0\sin\phi_c\sin\theta_c\sin\psi_c+\cos\phi_0\sin\phi_c\cos\theta_c$$

$$c_1 = \cos\phi_c\sin\psi_c+\sin\phi_c\sin\theta_c\cos\psi_c \qquad (64)$$

$$a_2 = \sin\phi_0\sin\theta_c-\cos\phi_0\cos\theta_c\sin\psi_c$$

$$b_2 = -(\sin\phi_0\cos\theta_c\sin\psi_c+\cos\phi_0\sin\theta_c)$$

$$c_2 = \cos\theta_c\cos\psi_c$$

By averaging calibration parameters for a range of horizontal and vertical eye positions, a complete set of calibration values are obtained.

3.7 GENERAL ALGORITHM FOR DETERMINING HORIZONTAL AND VERTICAL ROTATIONS

As described in section 3.3, the horizontal and vertical eye rotations can be found from either the projection of the iris/pupil center coordinates or the eccentricity of the projected iris/pupil boundary if the axis of the camera is aligned with the roll (Y) axis of the camera-coordinate-frame. For a camera offset relative to the reference-coordinate-frame, the equations describing the ellipse become unmanageable and it is difficult if not impossible to solve for the angles of the eye orientation from the eccentricity of the ellipse. Therefore, the projection of the iris/pupil center was used to compute the angles of the pitch and yaw eye orientation. Torsion was computed separately.

The coordinates of the center of the ellipse is related to the pitch and yaw components (ϕ, θ) of eye orientation by Eq. (41). Once the calibration parameters ϕ_c, θ_c, ψ_c, x_c, z_c and **d** are found, ϕ, θ can be determined. Two algorithms for computing ϕ, θ are discussed below.

ALGORITHM 1:

Using Eq. (41), we have:

$$R_{cal}^{-1}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = R_{eye}\begin{pmatrix} 0 \\ -d \\ 0 \end{pmatrix} \tag{65}$$

where (0, -d, 0) and (x, y, z) are the coordinates of the center of ellipse in the camera-coordinate-frame when the eye is looking straight at the camera and when it is rotated

to the current position (refer to Appendix C for notes on spherical eye model). Using Eq. (40), we have:

$$R_{cal}^{-1}\begin{pmatrix} x_p - x_c \\ y \\ z_p - z_c \end{pmatrix} = R_{eye}\begin{pmatrix} 0 \\ -d \\ 0 \end{pmatrix} \tag{66}$$

where (x_p, z_p) and (x_c, z_c) are the coordinates of the projection of the ellipse center and the projection of the center of the camera-coordinate-frame onto the image plane. Since a sphere satisfies the following equation:

$$x^2 + y^2 + z^2 = d^2 \tag{67}$$

Therefore we have:

$$R_{cal}^{-1}\begin{pmatrix} x_p - x_c \\ -\sqrt{d^2 - (x_p - x_c)^2 - (z_p - z_c)^2} \\ z_p - z_c \end{pmatrix} = R_{eye}\begin{pmatrix} 0 \\ -d \\ 0 \end{pmatrix} \tag{68}$$

where R_{eye} is given by Eq. (51) and R^{-1}_{cal} is given by the following equation:

$$R_{cal}^{-1} = (R_{cal}(\phi)\ R_{cal}(\theta)\ R_{cal}(\psi))^{-1} = R_{cal}^{-1}(\psi)\ R_{cal}^{-1}(\theta)\ R_{cal}^{-1}(\phi)$$

$$= \begin{pmatrix} \cos\psi_c & 0 & -\sin\psi_c \\ 0 & 1 & 0 \\ \sin\psi_c & 0 & \cos\psi_c \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_c & \sin\theta_c \\ 0 & -\sin\theta_c & \cos\theta_c \end{pmatrix} \begin{pmatrix} \cos\phi_c & \sin\phi_c & 0 \\ -\sin\phi_c & \cos\phi_c & 0 \\ 0 & 0 & 1 \end{pmatrix} = \tag{69}$$

$$\begin{pmatrix} \cos\phi_c\cos\psi_c - \sin\phi_c\sin\theta_c\sin\psi_c & \sin\phi_c\cos\psi_c + \cos\phi_c\sin\theta_c\sin\psi_c & -\cos\theta_c\sin\psi_c \\ -\sin\phi_c\cos\theta_c & \cos\phi_c\cos\theta_c & \sin\theta_c \\ \cos\phi_c\sin\psi_c + \sin\phi_c\sin\theta_c\cos\psi_c & \sin\phi_c\sin\psi_c - \cos\phi_c\sin\theta_c\cos\psi_c & \cos\theta_c\cos\psi_c \end{pmatrix}$$

Substitute Eqs. (51) and (69) into Eq. (68), we have:

$$\begin{aligned} A &= d\sin\phi\cos\theta \\ B &= -d\cos\phi\cos\theta \\ C &= -d\sin\theta \end{aligned} \tag{70}$$

where

$$\begin{aligned} A &= (\cos\phi_c\cos\psi_c - \sin\phi_c\sin\theta_c\sin\psi_c)(x_p - x_c) - \\ &\quad (\sin\phi_c\cos\psi_c + \cos\phi_c\sin\theta_c\sin\psi_c)\sqrt{d^2 - (x_p - x_c)^2 - (z_p - z_c)^2} - \\ &\quad \cos\theta_c\sin\psi_c(z_p - z_c) \\ B &= -\sin\phi_c\cos\theta_c(x_p - x_c) - \cos\phi_c\cos\theta_c\sqrt{d^2 - (x_p - x_c)^2 - (z_p - z_c)^2} \\ &\quad + \sin\theta_c(z_p - z_c) \\ C &= (\cos\phi_c\sin\psi_c + \sin\phi_c\sin\theta_c\cos\psi_c)(x_p - x_c) - \\ &\quad (\sin\phi_c\sin\psi_c - \cos\phi_c\sin\theta_c\cos\psi_c)\sqrt{d^2 - (x_p - x_c)^2 - (z_p - z_c)^2} + \\ &\quad \cos\theta_c\cos\psi_c(z_p - z_c) \end{aligned} \tag{71}$$

In Eq. (71), (x_p, z_p) are found by computing the center of ellipse using Eq. (28). ϕ_c, θ_c, ψ_c, x_c, z_c and **d** are determined during calibration. Therefore using Eqs. (70) and (71), ϕ and θ are given by:

$$\begin{aligned} \phi &= \arctan\left(-\frac{A}{B}\right) \\ \theta &= \arcsin\left(-\frac{C}{d}\right) \end{aligned} \tag{72}$$

It should be noted that Eq. (72) reduces to Eq. (31), if there is no camera offset ($\phi_c=0$, $\theta_c=0$, $\psi_c=0$, $x_c=0$, $z_c=0$).

ALGORITHM 2:

The horizontal and vertical eye rotations (ϕ and θ) can also be found using another algorithm. This algorithm is based on the idea that from the projected iral contour onto the image plane, the Euler angles (ϕ_h and θ_h) of the eye relative to the camera-coordinate-frame can be computed immediately. It is computed by fitting the iral boundary points to Eq. (18) and solving for ϕ_h and θ_h from either the coordinates of the ellipse center or the eccentricity of the ellipse, using Eqs. (19), (20) or (31). The computation of ϕ_h and θ_h is not affected by camera offset angles.

Let R_h be the matrix representing the eye rotation from the zero position to the current position. Let (0, -d, 0) and (x, y, z) be the coordinates of the ellipse center when the eye is at the zero position and is rotated to the current position. Then we have:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = R_h \begin{pmatrix} 0 \\ -d \\ 0 \end{pmatrix} \tag{73}$$

Using Eq. (41), we have:

$$R_{cal}^{-1} R_h \begin{pmatrix} 0 \\ -d \\ 0 \end{pmatrix} = R_{eye} \begin{pmatrix} 0 \\ -d \\ 0 \end{pmatrix} \tag{74}$$

R^{-1}_{cal} and R_{eye} are defined by Eqs. (69) and (51). R_h is defined below:

$$R_h = \begin{pmatrix} \cos\phi_h & -\sin\phi_h\cos\theta_h & \sin\phi_h\sin\theta_h \\ \sin\phi_h & \cos\phi_h\cos\theta_h & -\cos\phi_h\sin\theta_h \\ 0 & \sin\theta_h & \cos\theta_h \end{pmatrix} \tag{75}$$

From Eq. (74), we have the following:

$$\begin{aligned} A' &= -\sin\phi\cos\theta \\ B' &= \cos\phi\cos\theta \\ C' &= \sin\theta \end{aligned} \tag{76}$$

where

$$
\begin{aligned}
A' &= (\sin\phi_c \sin\theta_c \sin\psi_c - \cos\phi_c \cos\psi_c)\sin\phi_h \cos\theta_h + \\
&\quad (\sin\phi_c \cos\psi_c + \cos\phi_c \sin\theta_c \sin\psi_c)\cos\phi_h \cos\theta_h \\
&\quad - \cos\theta_c \sin\psi_c \sin\theta_h \\
B' &= \sin\phi_c \cos\theta_c \sin\phi_h \cos\theta_h \\
&\quad + \cos\phi_c \cos\theta_c \cos\phi_h \cos\theta_h + \sin\theta_c \sin\theta_h \\
C' &= -(\cos\phi_c \sin\psi_c + \sin\phi_c \sin\theta_c \cos\psi_c)\sin\phi_h \cos\theta_h + \\
&\quad (\sin\phi_c \sin\psi_c - \cos\phi_c \sin\theta_c \cos\psi_c)\cos\phi_h \cos\theta_h \\
&\quad + \cos\theta_c \cos\psi_c \sin\theta_h
\end{aligned}
\tag{77}
$$

Therefore:

$$
\begin{aligned}
\phi &= \arctan\left(-\frac{A'}{B'}\right) \\
\theta &= \arcsin(C')
\end{aligned}
\tag{78}
$$

While both algorithms utilize the ellipse fitting concept, they are conceptually different in that one utilizes the fit to find the center while the other allows for the computation of both eccentricity and center to determine pitch and yaw eye position.

3.8 GENERAL ALGORITHM FOR DETERMINING TORSION

The derivation of the torsion algorithm derived in section **3.4** assumed that (1) there were no horizontal and vertical eye rotations, (2) the image plane was perpendicular to the "line of sight" of the eye at the reference position, and (3) the center of the image

plane coincided with the projection of the center of the camera-coordinate-frame onto the image plane. Under these assumptions, the iral signatures were sampled along a circular arc in the image plane (Eq. (33)) whose center coincided with the center of the image plane. In general, however, the eye is rotated about the yaw (ϕ) and pitch (θ) axes, the camera is offset relative to the reference-coordinate-frame, and the center of the image plane is translated relative to the projected center of the camera-coordinate-frame. Under these conditions, the iral signature is sampled along an elliptical arc in the image plane. The sampling positions on the elliptical arc are related to a circular arc by an affine transformation. Once the affine transformation is computed, the algorithm for computing torsion by comparing the reference and current signatures remain as described in section 3.4.3.

Let R_{eye} be the matrix representing the rotation of the eye from the reference position (looking straight ahead) to the current position. Let R_{cal} be the matrix representing the rotation of the eye from the zero position (looking straight at the camera) to the reference position. The transformations then follow analogously to those describing the transformation of the ellipse center (Eq. (41)), but applied to any arbitrary point on the iris.

Assume the coordinates of any point P on the iris are (x", -d, z") when the eye is at the zero position and (x, y, z) when the eye is at the current position (refer to Appendix C for notes on spherical eye model). The relationship between (x", -d, z") and (x, y, z) is as follows:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = R_{cal}\, R_{eye} \begin{pmatrix} x'' \\ -d \\ z'' \end{pmatrix} \tag{79}$$

R_{cal} and R_{eye} are given by Eq. (51) and Eq. (39).

Let (x_p, z_p) be the coordinates of the projection of the point P on the iris onto the image plane when the eye is rotated to the current position and $(x"_p, z"_p)$ the coordinates of P when the eye is looking directly at the camera, i.e., the zero position. Let (x_c, z_c) be the coordinates of the projection of the center of the camera-coordinate-frame onto the image plane. Substituting Eqs. (51) and (39) into Eq. (79), then using Eq. (40), the relationship between (x_p, z_p) and $(x"_p, z"_p)$ can be given as follows:

$$\begin{aligned}
x_p = {} & (\cos\phi\cos\phi_c\cos\psi_c - \cos\phi\sin\phi_c\sin\theta_c\sin\psi_c - \sin\phi\sin\phi_c\cos\theta_c)(x''_p - x_c) \\
& + (\sin\phi\cos\theta\cos\phi_c\cos\psi_c - \sin\phi\cos\theta\sin\phi_c\sin\theta_c\sin\psi_c + \\
& \cos\phi\cos\theta\sin\phi_c\cos\theta_c - \sin\theta\cos\phi_c\sin\psi_c - \sin\theta\sin\phi_c\sin\theta_c\cos\psi_c)d + \\
& (\sin\phi\sin\theta\cos\phi_c\cos\psi_c - \sin\phi\sin\theta\sin\phi_c\sin\theta_c\sin\psi_c + \\
& \cos\phi\sin\theta\sin\phi_c\cos\theta_c + \cos\theta\cos\phi_c\sin\psi_c + \\
& \cos\theta\sin\phi_c\sin\theta_c\cos\psi_c)(z''_p - z_c) + x_c \\
z_p = {} & (\sin\phi\sin\theta_c - \cos\phi\cos\theta_c\sin\psi_c)(x''_p - x_c) - (\sin\phi\cos\theta\cos\theta_c\sin\psi_c + \\
& \cos\phi\cos\theta\sin\theta_c + \sin\theta\cos\theta_c\cos\psi_c)d + (\sin\phi\sin\theta\sin\theta_c - \\
& \sin\phi\sin\theta\cos\theta_c\sin\psi_c + \cos\theta\cos\theta_c\cos\psi_c)(z''_p - z_c) + z_c
\end{aligned} \tag{80}$$

If we let $(x"_p, z"_p)$ be the coordinates of a point within the sampling annulus when the eye is at the zero position (looking straight at the camera), Eq. (80) uniquely determines the coordinates of the sampling point when the eye is rotated by (ϕ, θ, ψ).

When the eye is at the zero position, the sampling annulus is circular. Let (r, ψ) be the coordinates for any point within the sampling annulus in the polar coordinate system centered at (x_c, z_c). Let (x, z) be the coordinates for the same point in the rectangular coordinate system (Figure 11). The coordinates of the sampling annulus are defined by the following equation (refer to section **3.4.2** for the definition of the parameters):

$$\begin{aligned} x &= r\cos\psi + x_c = (r_1 + k*r_{inc})\cos(\psi_1 + l*\psi_{inc}) + x_c \\ z &= -r\sin\psi + z_c = -(r_1 + k*r_{inc})\sin(\psi_1 + l*\psi_{inc}) + z_c \end{aligned} \qquad (81)$$

$$(r_1 \le r \le r_2;\ \ \psi_1 \le \psi \le \psi_2;\ \ 0 \le k \le m;\ \ 0 \le l \le n)$$

Thus, the coordinates of the sampling annulus, when the eye is rotated to the current position, are found by substituting (x, z) in Eq. (81) into (x''_p, z''_p) in Eq. (80). The iral signatures found from the remapped sampling annulus are then used to compute the torsional rotations of the eye about the optic axis using Eqs. (35) and (36) (refer to section **3.4.3** for the detailed description).

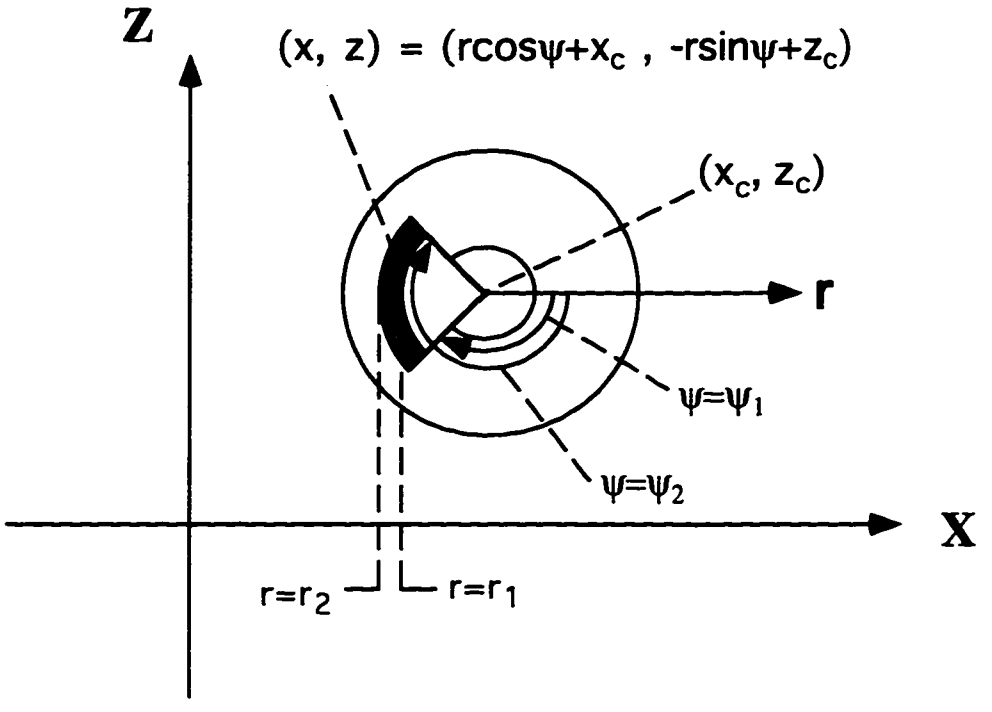


Figure 11: Representation of Sampling Annulus in Both Rectangular and Polar Coordinate Systems

CHAPTER 4

EYE MOVEMENT TRACKING IMPLEMENTATION METHODS

4.1 HARDWARE CONFIGURATION

A flow chart of the components of the system used in developing the eye tracking algorithms is shown in Figure 12. The system consists of a miniature CCD camera (V-1055, Marshall Electronics) which generates an image of the eye with 380 lines of resolution, a video frame grabber (Model 495, Colorado Video) which digitizes the picture into a 512x512 array with 8 bits gray level resolution, a host adapter (Model 745, Colorado Video) which connects the IBM-PC to the display memories in frame grabber, a monitor (HR2000, Dage-MTI) which displays the image stored in the frame grabber, and an IBM-PC which runs the software to determine eye orientation relative to a reference position.

An artificial eye (diameter 30 mm) of the dimension similar to that of a human eyeball was used for testing the accuracy of the algorithms. It was fixed to a gimbal whose center of rotation coincides with the center of the artificial eye (Figure 13). It is possible to rotate the artificial eye horizontally, vertically and torsionally around the center of the eye with an accuracy of 0.1°. During testing, the artificial eye was illuminated by a small flashlight from above. The image of the iris and pupil were reflected in a mirror and sensed by a CCD camera. The camera, the flashlight and the mirror were all mounted on the gimbal.

The algorithms were thoroughly tested on the artificial eye to calibrate the system. The algorithms were then tested on images obtained from a system developed by LETI

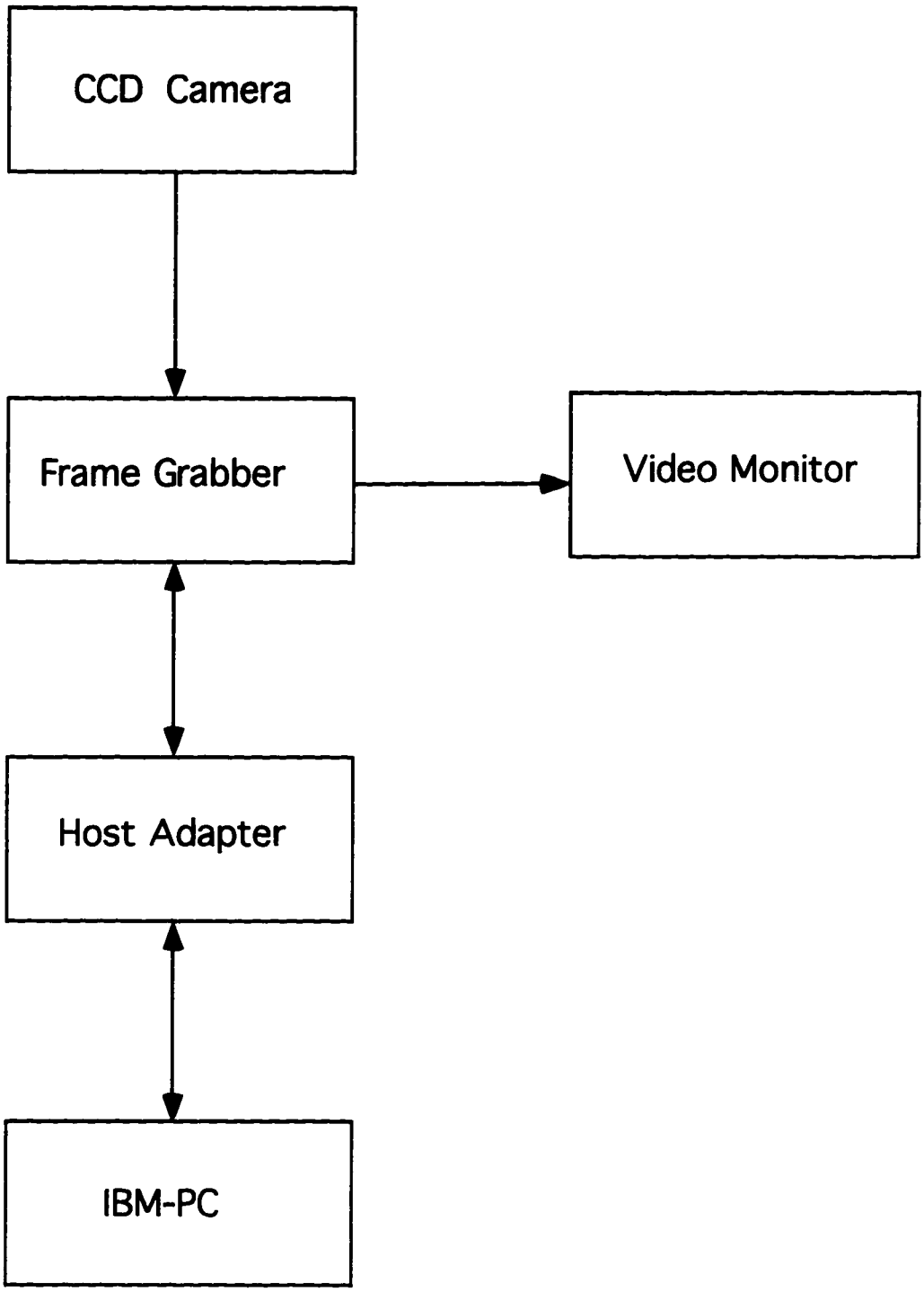


Figure 12: Hardware Configuration of System

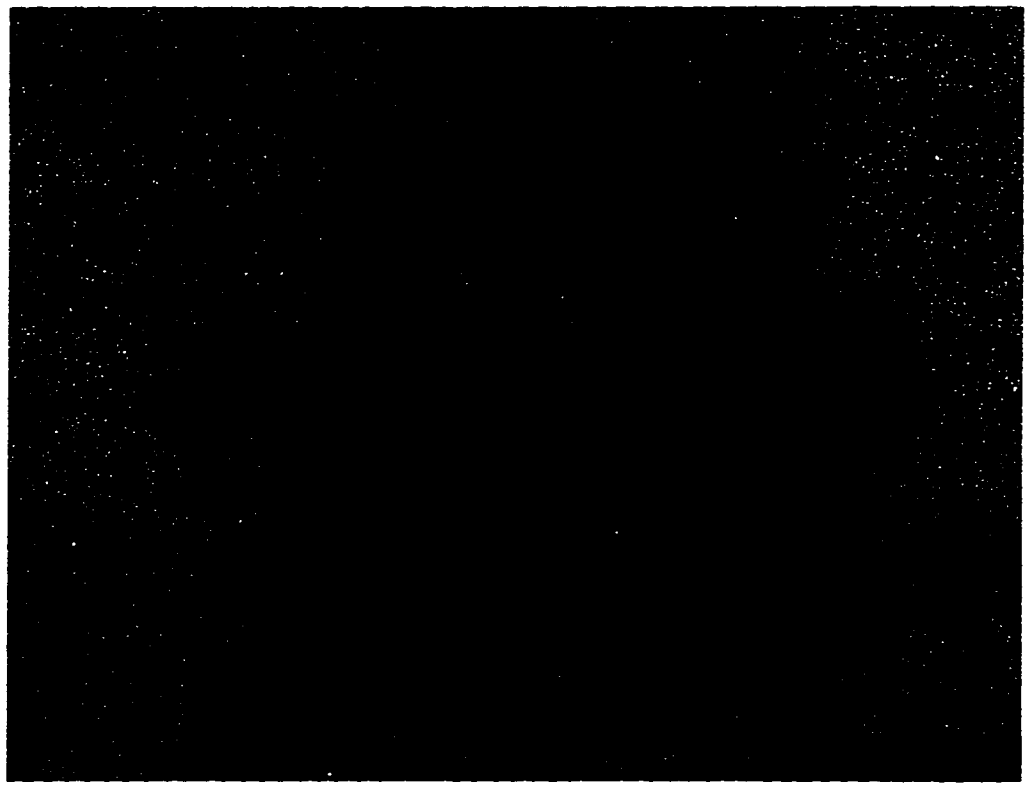


Figure 13: Gimbal Mounted with Artificial Eye and CCD Camera

for the Neurolab project. The eye was imaged by an infra-red CCD camera mounted on goggles worn by a subject. In that system, infra-red LEDs were attached to goggles to illuminate the eye. The infra-red sensitive mirror of the goggles which was transparent in the visible range reflected the image of the eye onto the CCD camera (Figure 14).

The flowchart for the algorithms to determine the horizontal, vertical and torsional eye movements is shown in Figure 15 and Figure 16. Each part of the algorithms is described below:

4.2 SAMPLING AND DIGITIZATION PROCESS

The digitization of the image is performed by a frame grabber that grabs images at 30 frames/sec and digitizes the image captured by the camera into a 512x512 digital array. Each pixel within the digital array has 8 bits and 256 gray levels. The digitization process is controlled by an IBM-PC which runs the program to freeze an image into the memory in the frame grabber and then to read the image into the IBM-PC.

4.3 THRESHOLDING

To speed up and simplify the algorithm, a thresholding algorithm was implemented to segment the eye image into three parts: the pupil, the iris and the remaining parts of the image representing the sclera, eyelid and blood vessels. Usually the pupil has the darkest gray level. The iris is darker than the sclera and the eye lid, although it is brighter than the pupil.

Thresholding is performed by analyzing the histogram of the gray-levels of the eye image. There are three dominant peaks in the histogram (Figure 17). They

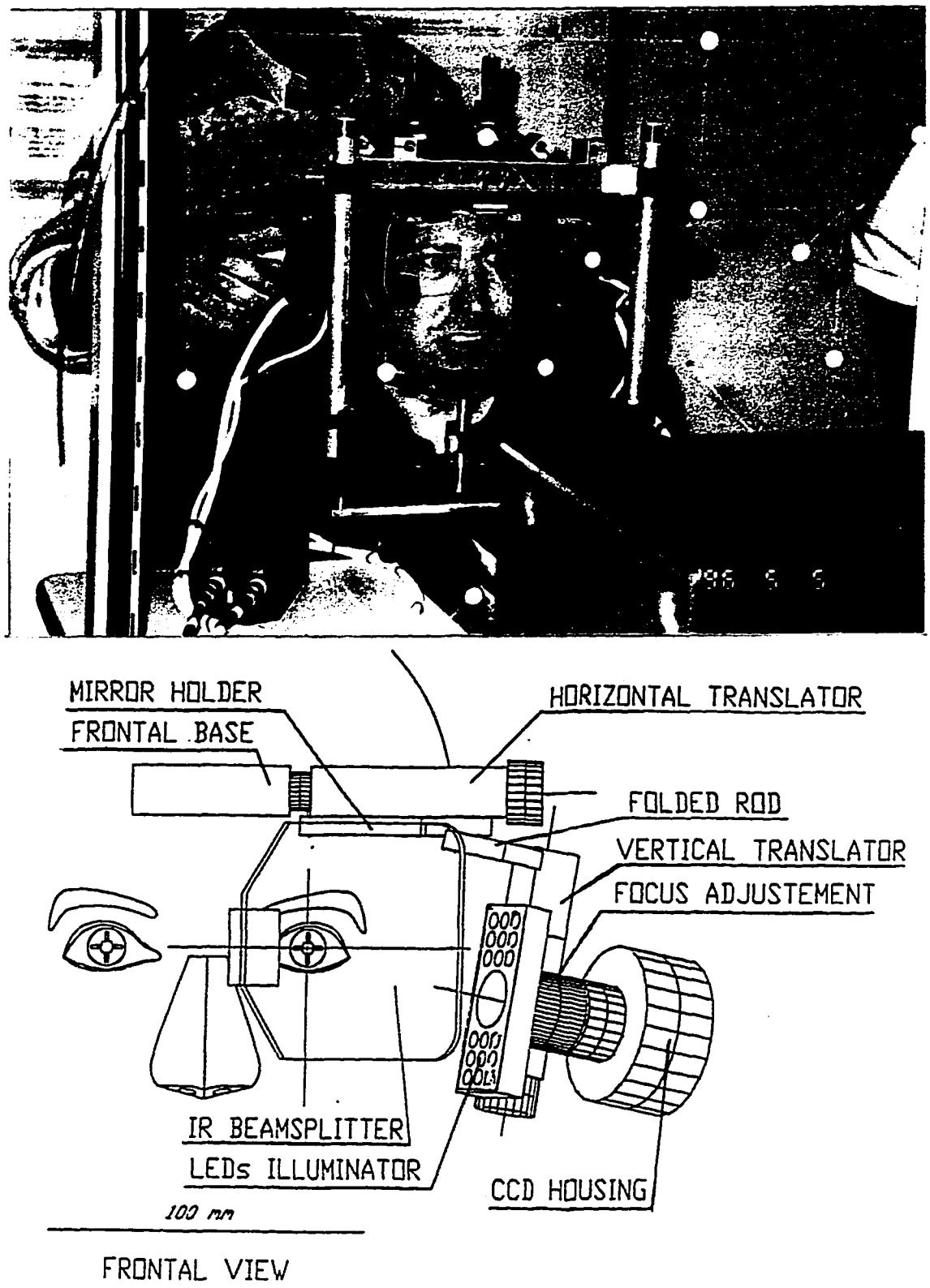


Figure 14: Goggles and Camera Set-up by LETI

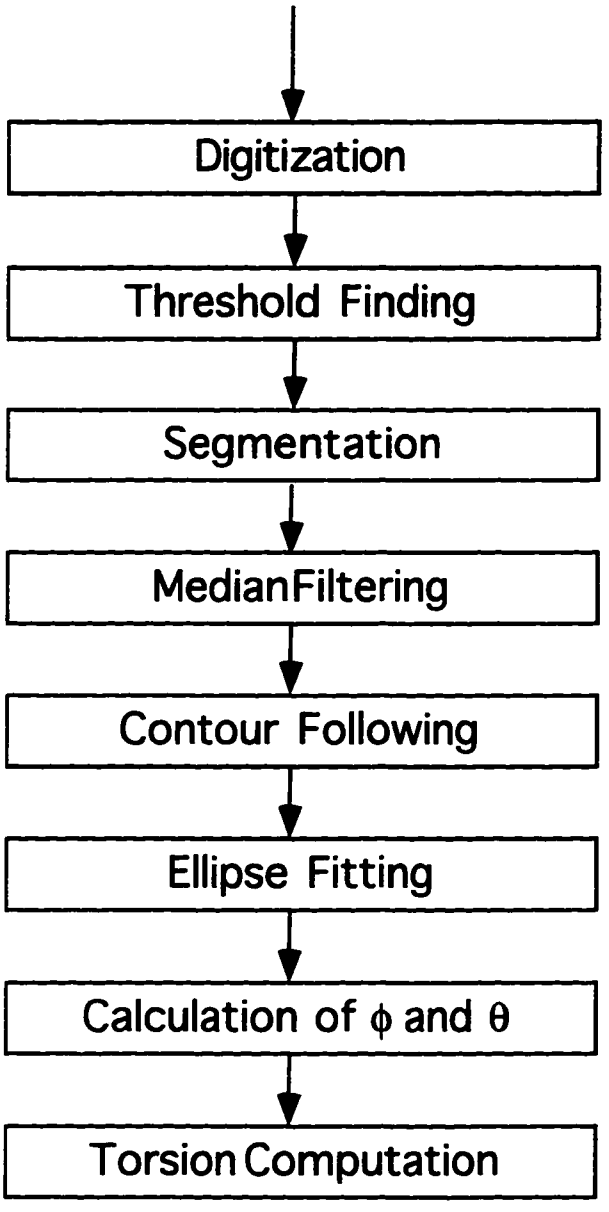


Figure 15: Flowchart of Algorithm for Computing Pitch and Yaw Angles

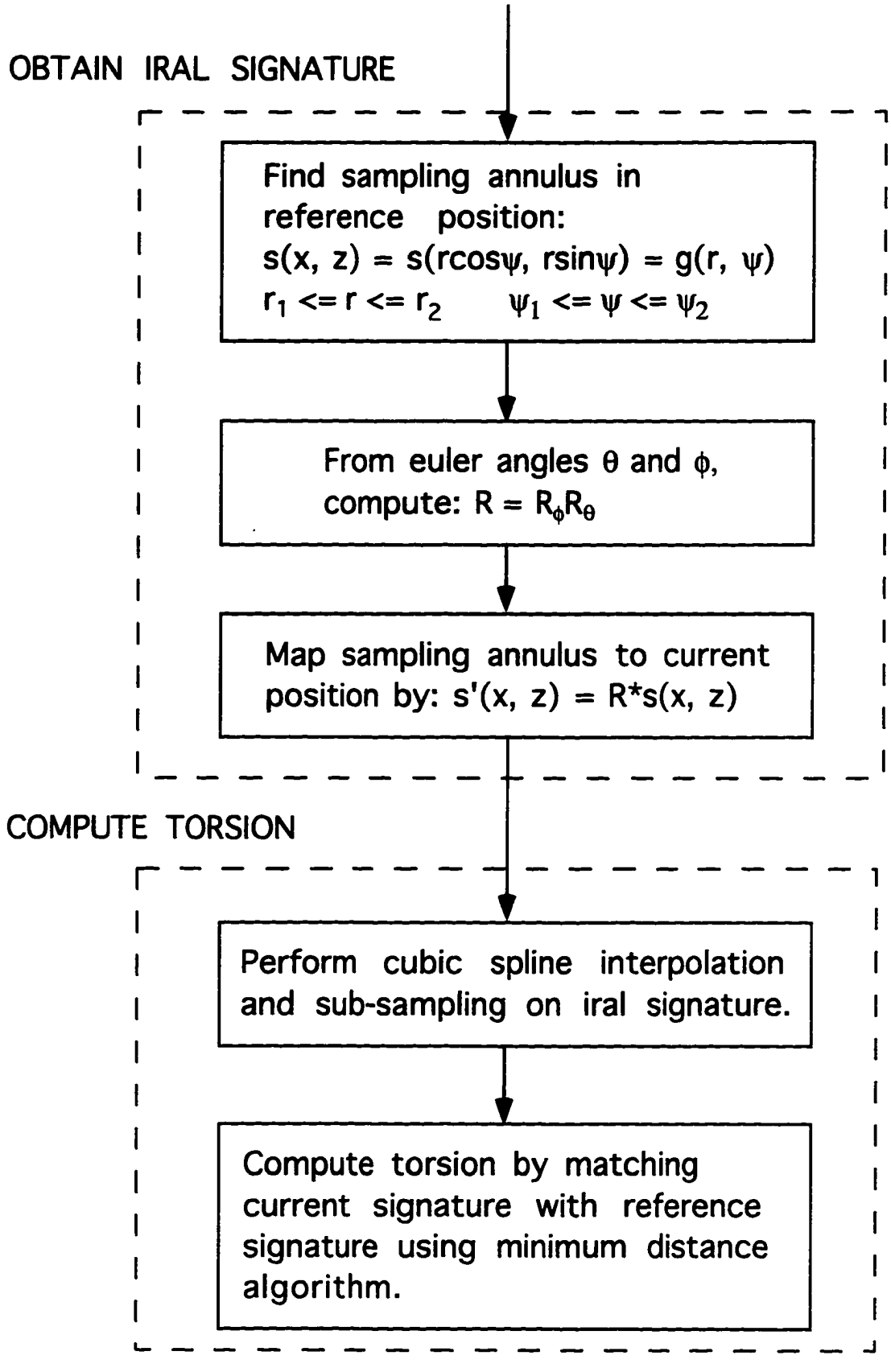


Figure 16: Flowchart of Algorithm for Computing Torsion

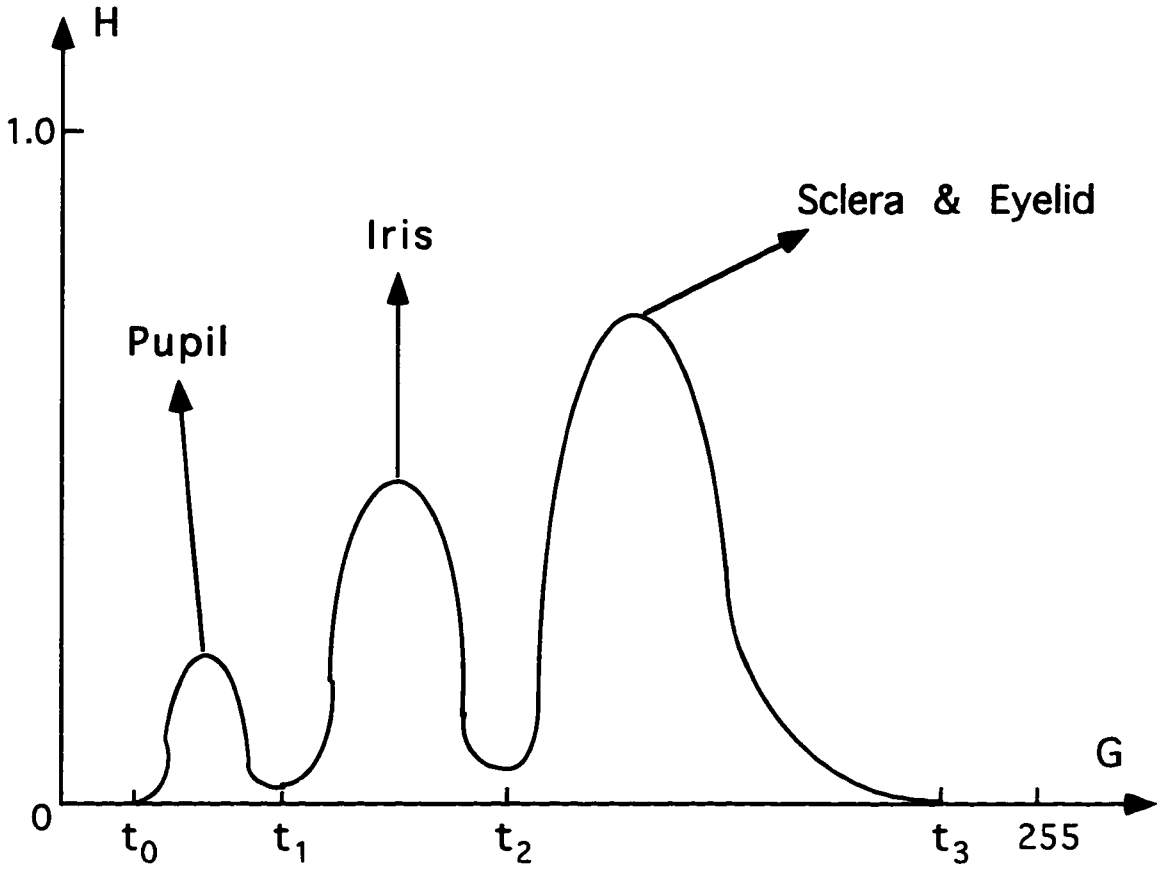


Figure 17: Gray Level Histogram of Eye Image

corresponds to the gray levels of the pupil, the iris and the rest of the image. If the proper gray level thresholds are selected, the iris and the pupil can be extracted from the rest of the eye image. The pixels associated with iris are determined by the following equation:

$$t_1 \leq f(x,z) \leq t_2 \tag{82}$$

where (x, z) are the coordinates of the pixel to be examined and f() is the gray level function (Figure 17). The thresholds for the histogram are determined during calibration.

4.4 SEGMENTATION

The segmentation program divides the eye image into three segments defined as the pupil, the iris and the rest of the image using the thresholds of the gray-level histogram of the eye image described above.

4.5 MEDIAN FILTERING

To reduce the video noise along the iral or pupillary contour, the boundary was smoothed with a median filter so that the correct iris-scleral or pupil-iral boundary points can be detected. Smoothing is performed only near the contours which are indicated by large variation in gray levels, from white (sclera) to dark (iris) or vice versa, across the boundaries to improve the speed. The smoothing algorithm is based on the fact that "noisy" pixels along the iral or pupillary contour are not on the smoothly changing curve of the iris. They are somewhat "isolated" and most of the points around the "noisy" pixels would be white. The smoothing is implemented by moving a square window (5x5)

on the iris image to locate the iral edge points. Once an iral edge point is found, it is checked to see if it is a "noisy" point. If the number of black pixels inside the window is less than the number of white pixels, or is below a certain number, the edge point which is at the center of the window is considered a "noisy" pixel and is removed.

4.6 CONTOUR FOLLOWING ALGORITHM

4.6.1 BASIS OF ALGORITHM

To locate all the edge points on the boundary of the iris, we implemented a contour following program (see Appendix D for a detailed flowchart). The program searches for a point on the boundary and extends the boundary by adding points in the contour direction. The starting point for searching for iral edges is an interior point of the iral boundary and is found during the course of segmenting the iris and the pupil images from the background. The interior point is obtained by dividing the sum of the coordinates of all the points associated with the iris and the pupil by the total number of points, i.e., computing the center of mass.

After the coordinates of the interior point is found, the contour following program scans the image to the left and to the right of the interior point until iral edge points which are on the boundary of the black iris and the white sclera are encountered. Once an initial boundary point is found, the contour is developed according to the following algorithm (Appendix D-3):

Suppose P is a pixel on the boundary of the iris with coordinates in the image plane given by (x, z). The eight connected neighbors of P are shown in Figure 18. Each

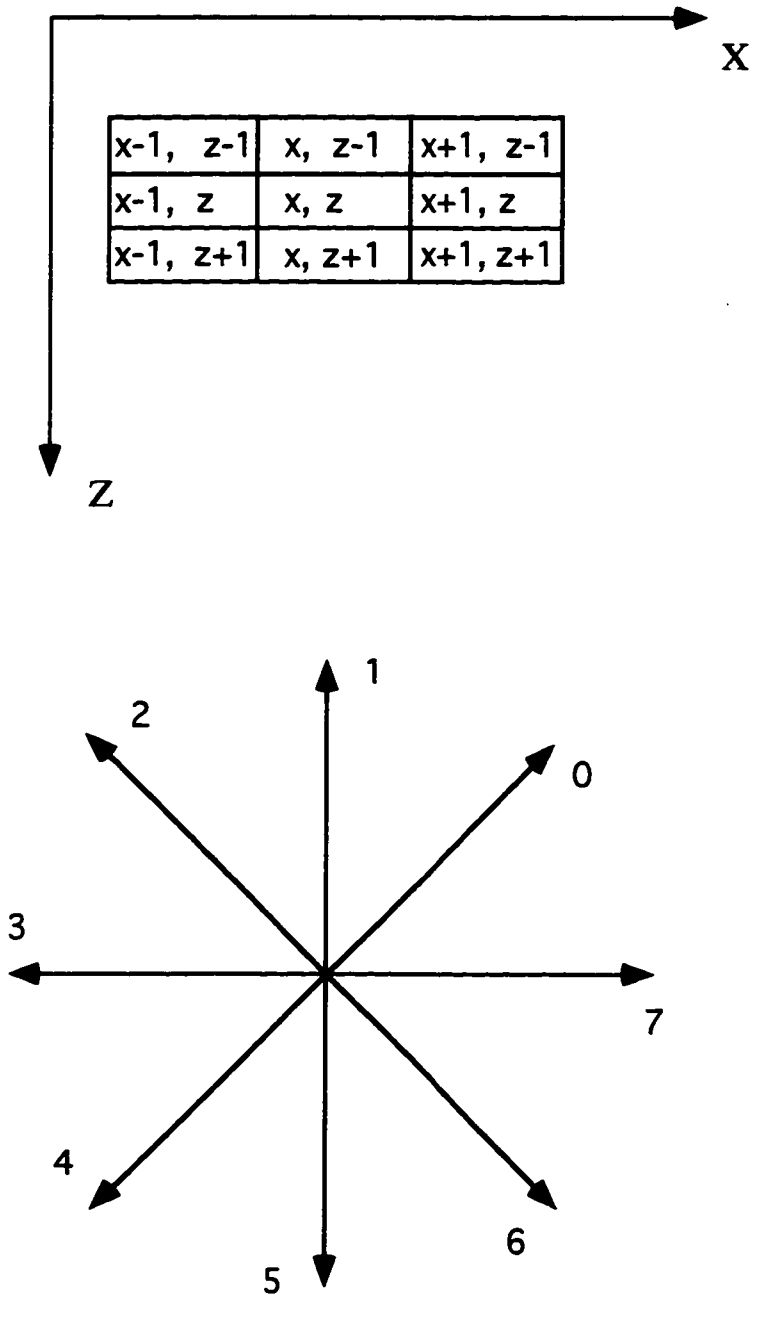


Figure 18: Eight Connected Neighbors of Point (x, z) and Corresponding Eight Search Directions

pixel on the iral boundary is assigned a direction value from "0" to "7" which represents the search direction of the iral contour at this point (Figure 18). The searching for the next edge point can be either clockwise or counter clockwise, which is determined by the contour following program. Suppose the searching for the next edge point is clockwise and the direction value at point P is "1". The contour following algorithm first checks point (x-1, z) to see whether it is a pixel related to the iris (black point). The points (x-1, z-1), (x, z-1), etc. are then examined in sequence. The first black point encountered is the next edge point after P. In this example, the searching is clockwise around point P. The searching starts from point (x-1, z) which has a direction value "3", which is before direction "1" in the clockwise order. If there is no discontinuity in curvature, i.e., no eye lid closure, the searching continues with one of P's eight neighbors which is before the other neighbors in either clockwise or counter clockwise order until the curve closes on itself.

If the iris is partially covered by the eyelid (Figure 19), the iral contour is not a complete circle or ellipse. To detect the uncovered iris or pupil edges, the contour following program computes the curvature at each edge point (Appendix D-5). The edge point associated with a curvature function value greater than a certain threshold is the point where the eyelid curve and the iris or the pupil curve intersect. Figure 20 illustrates the iral contour when 1) the iris is fully exposed (Figure 20A), 2) the iris is covered by the upper eyelid (Figure 20B), 3) the iris is covered by the lower eyelid (Figure 20C) and 4) the iris is covered by both the upper and the lower eyelids (Figure 20D).

The contour following program first locates two possible starting points for the iral or pupil boundary, P_0 and P_1. The program starts to search the edges at point P_0 in

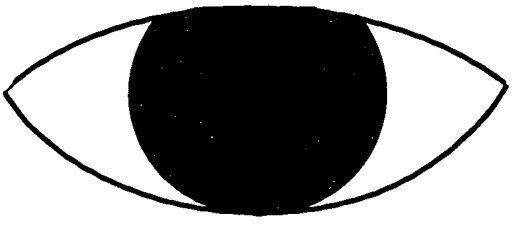


Figure 19: Image of Eye Partially Covered by Eyelid

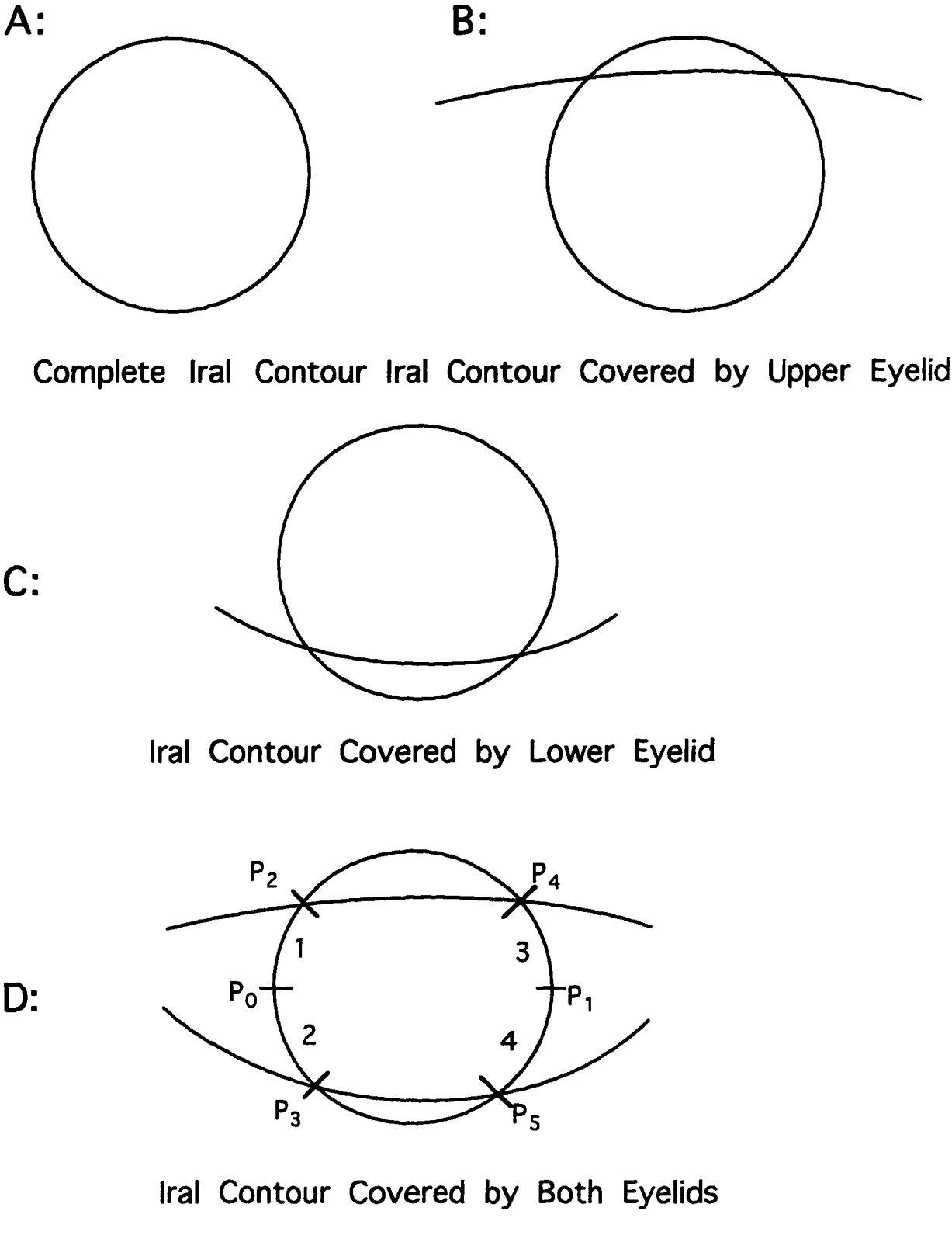


Figure 20: Segments of Iral Contour when Eye is covered by Eyelid

the clockwise direction. If the iris is not covered by the eyelid (case 1), the searching will finally end in point P_0. If the iris is covered by the upper eyelid (case 2), the searching will stop at point P_2 where the curvature function value is greater than the threshold. The program will then re-start the searching at point P_0 in the counter-clockwise direction until it reaches point P_4 where the curvature function value is again greater than the threshold. If the iris is covered by the lower eyelid (case 3), the searching will stop at point P_5. The program then re-starts the searching of the edges at point P_0 in the counterclockwise direction and ends at point P_3. If the iris is covered by both the upper and the lower eyelids (case 4), the searching will stop at point P_2 as it moves in the clockwise direction. Then it re-starts the searching at point P_0 moving counterclockwise until it reaches point P_3 where it stops again. The program then starts the searching of the edges at point P_1 in the counter-clockwise direction until it stops at point P_4. Finally the program re-starts the searching at point P_1 in the clockwise direction ending at point P_5.

With the elliptical eye model and the fitting method, it is sufficient to use only the uncovered iris or pupil edge points to find the equations of the projection of the iris or pupil boundary onto the image plane. During this study, we have implemented a simulation program which simulated the eye movements with the iris or pupil partially covered by the upper and lower eye lids, and computed the eye orientation using the uncovered edge points. The result showed that using only about 30% of the total boundary points would result in an acceptable estimate of the eye orientation.

4.6.2 CURVATURE CALCULATION ALGORITHM

The curvature κ of an arc is defined as the rate of change of the angle ϕ with respect to the arc length s, which is shown below:

$$\kappa = \frac{d\phi}{ds} \tag{83}$$

If the arc is given in the form y=f(x), then we have:

$$\kappa(x) = \frac{f''(x)}{[1+(f'(x))^2]^{3/2}} \tag{84}$$

The contour following program uses a discrete form of the curvature function to calculate the curvature (Appendix D-5). It uses a window of 30 to 40 edge points, which is moved along the boundary of the iris while searching for iral edges. The window is divided into two sub-windows: the first half and the second half. The directions for the edge points within the window are assigned one of the eight values (Figure 18). From the histogram distribution of the edge directions of the sub-windows, four directions with highest occurrences of within both the first sub-window and the second sub-window are obtained. The average of the four directions for the two sub-windows are computed. The difference between the average direction values of the first and second sub-windows is the curvature function value of the edge point associated with the point in the middle of the window.

During the course of testing the algorithm, the curvature function value of the edge points calculated by the contour following program usually ranged from 0.1 to 0.6.

The threshold used by the program ranged from 0.6 to 1.5. The actual threshold value is determined at calibration time.

4.6.3 BOUNDARY POINT LIST DATA STRUCTURE

A list of iral boundary points found by the contour following program are stored in a data structure shown in Figure 21. These boundary points are then used to solve for the parameters for the elliptical equation of the iral contour projected onto the image plane using SVD fitting algorithm.

4.7 ELLIPSE FITTING

The fitting program fits all the boundary points found by the contour following program to an ellipse which is the projection of the iral contour onto the image plane. The equation for the ellipse is defined by Eq. (18). The purpose of the fitting is to find all the coefficients (A, B, C, D and E) in Eq. (18). To measure the agreement between the data (the iral boundary points) and the model (Eq. (18)) with a particular choice of parameters (A, B, C, D and E), a merit function "chi-square" is designed which is shown below:

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{x_i^2 - (Az_i^2 + Bx_i z_i + Cx_i + Dz_i + E)}{\sigma} \right)^2 \qquad (85)$$

χ^2 is correspondingly a sum of n squares of quantities. σ is the standard deviation for each iral boundary points and is given by the following equation:

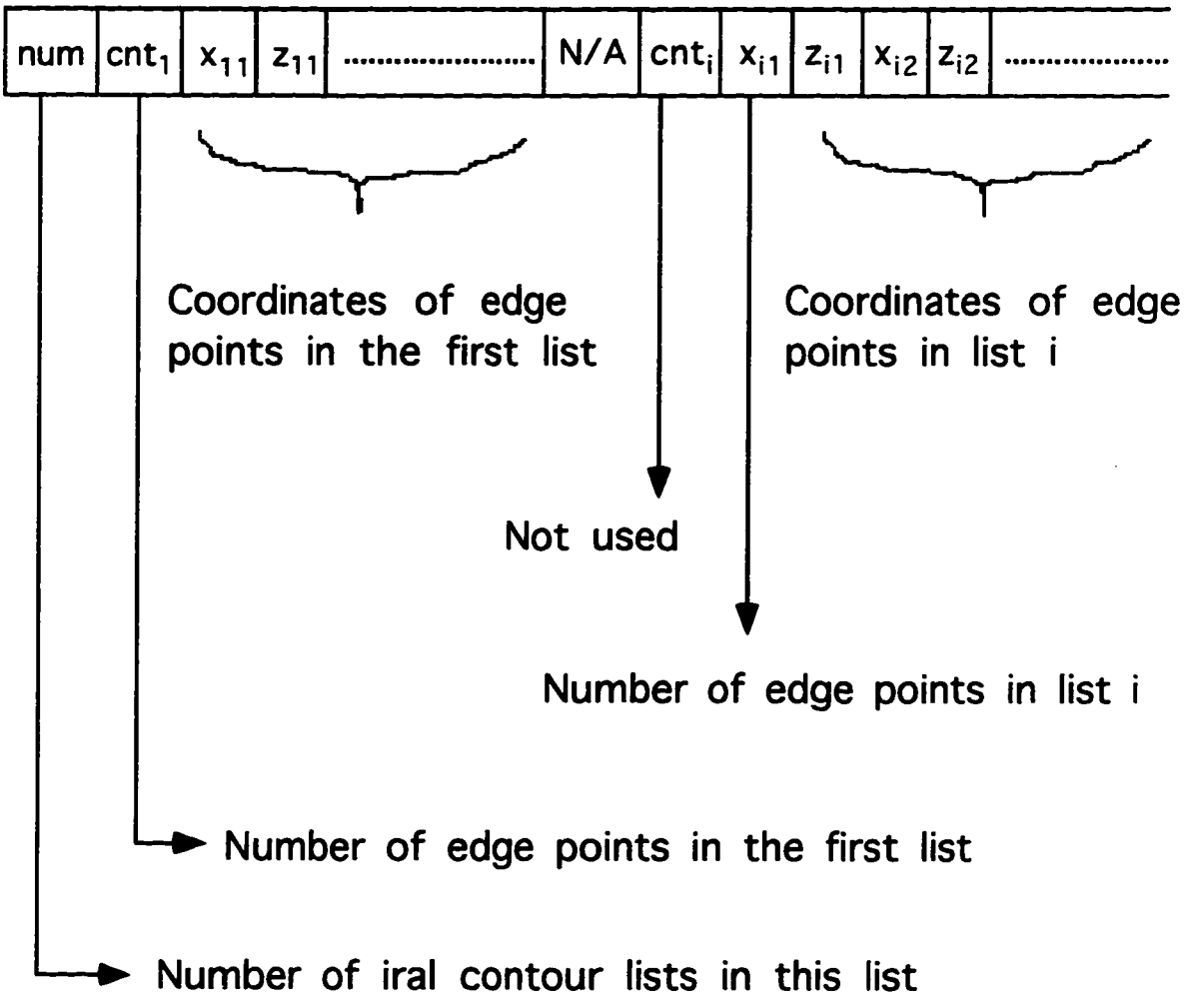


Figure 21: Data Structure of Iral Contour List

$$\sigma^2 = \frac{\sum_{i=1}^{n} [x_i^2 - (Az_i^2 + Bx_i z_i + Cx_i + Dz_i + E)]^2}{n-5} \tag{86}$$

Small values of χ^2 represent close agreement. A, B, C, D and E are then adjusted to achieve a minimum for χ^2, yielding the best-fit parameters. If we take the derivative of Eq. (85) with respect to A, B, C, D and E, we obtain the following equations that must hold at the chi-square minimum:

$$\begin{aligned}
0 &= \sum_{i=1}^{n} \left(\frac{x_i^2 - (Az_i^2 + Bx_i z_i + Cx_i + Dz_i + E)}{\sigma^2} \right) z_i^2 \\
0 &= \sum_{i=1}^{n} \left(\frac{x_i^2 - (Az_i^2 + Bx_i z_i + Cx_i + Dz_i + E)}{\sigma^2} \right) x_i z_i \\
0 &= \sum_{i=1}^{n} \left(\frac{x_i^2 - (Az_i^2 + Bx_i z_i + Cx_i + Dz_i + E)}{\sigma^2} \right) x_i \\
0 &= \sum_{i=1}^{n} \left(\frac{x_i^2 - (Az_i^2 + Bx_i z_i + Cx_i + Dz_i + E)}{\sigma^2} \right) z_i \\
0 &= \sum_{i=1}^{n} \left(\frac{x_i^2 - (Az_i^2 + Bx_i z_i + Cx_i + Dz_i + E)}{\sigma^2} \right)
\end{aligned} \tag{87}$$

The above equation is a set of five linear equations for the five parameters (A, B, C, D and E). The parameters of Eq. (87) were found using Singular Value Decomposition (SVD) (Press et al. 1992).

4.8 CALCULATION OF EULER ANGLES

The Euler angles (ϕ and θ) can be determined from the parameters A, B, C, D and E. There are two methods for computing the Euler angles. One is based on the relationship between the translation of the ellipse center projected onto the image plane and the horizontal and the vertical rotations of the eye (Eq. (19), Eq. (31), Eq. (72) and Eq. (78)). This method basically uses parameters C and D which are very sensitive to the translation of the ellipse center. The other method is based on the relationship between the change of the eccentricity of the iral contour projected onto the image plane and the horizontal and the vertical rotations of the eye (Eq. (20)). This method uses parameters A and B which are sensitive to the change of the eccentricity of the iral contour. The camera offset angles (ϕ_c and θ_c) are computed using this method (Eq. (132) and Eq. (137)). Since translation of the iral contour does not affect the eccentricity of the iral contour, this method can distinguish between the translation of the center of the eye and the rotation about the center of the eye. However the algorithm using the eccentricity of the projected iral contour has lower resolution than the algorithm using the projected iral center. This is because the change of the eccentricity of the iral contour is less sensitive to the eye rotations than the change of translation of the iral center. We also found that the change of the eccentricity of the iral contour projected onto the image plane is least sensitive to eye rotations when the eye is close to the zero position (looking straight at the camera) where the projection of the iral contour onto the image plane is a perfect circle. To accurately calculate the Euler angles of the eye using this method, the projection of iral contours should represent eccentric positions of the eye relative to the camera.

The resolution of the horizontal and vertical eye rotations depends on the value of d which is the distance from the center of the eye and the iral plane in pixels and is directly proportional to the size of the eye image. It is approximated by the following equation:

$$res_{hv} = \frac{180}{\pi} \arcsin\left(\frac{1}{d}\right)(degree) \tag{88}$$

4.9 SAMPLING OF IRAL SIGNATURE

4.9.1 DETERMINATION OF REFERENCE AND CURRENT SIGNATURES

The torsional movement of the eye about the optic axis can be determined by computing the angular shift between the current iral signature relative to the reference iral signature which is obtained during calibration. To find the degree of shift, a template matching algorithm is used that searches for a best match between the current and reference signatures by calculating the minimum distance. The metric used for computing distances between iral signatures was the "City Block" metric which is a relatively fast algorithm for computing distances between patterns (Ballard and Brown 1982). The torsional angle is proportional to the amount of shift between the current and reference signatures (Eq.(36)).

In this study, the whole uncovered iris has been used as the reference signature for computing the shift between the reference and current iral patterns. Storage of the complete visible iral pattern addresses an early problem with the VOG algorithm in

determining torsion caused by partial lid closures (Moore et al. 1996). When the eyelid droops, the sampling annulus of the iris may be covered by the eyelid, which makes sampling of the iral signature impossible. However, a methodology which utilizes the complete iral pattern as the current iral signature leads to computational intensiveness. Moore et al. (1996) have addressed this problem by dividing the iris into four quadrants and generating a sampling path in each quadrant. The approach taken in this dissertation was to generalize the approach taken by Moore et al. (1996) by using an "adaptive sampling annulus" of the current iral signature. The sampling annulus used in computing the angular shift between the signatures is shown in Figure 22C. Since we have stored the whole iris as the reference signature, the sampling annulus for the current iral signature can be adapted to use any part of the uncovered iris image (Figure 22A, B). This adaptive capability makes the torsion algorithm independent of eyelid closures and is an important contribution of this work.

4.9.2 DETERMINATION OF SAMPLING ANNULUS

The position of the sampling annulus for the current iral signature is determined by four parameters: the angular span of the sampling annulus β (Figure 22C), the inner and outer radii of the sampling annulus r_1 and r_2 (Figure 22C) and the offset angle of the sampling annulus α (Figure 22C). In this study, $\beta=90°$. The radii of the annulus, r_1 and r_2, are determined during calibration. The only parameter that has to be determined is α. It is selected dynamically according to the orientation of the eye at the current position (Figure 23). When the eye image translates to the right half of the image plane, the iral signature will be sampled from the left half of the iris (Figure 23A, D), and when the

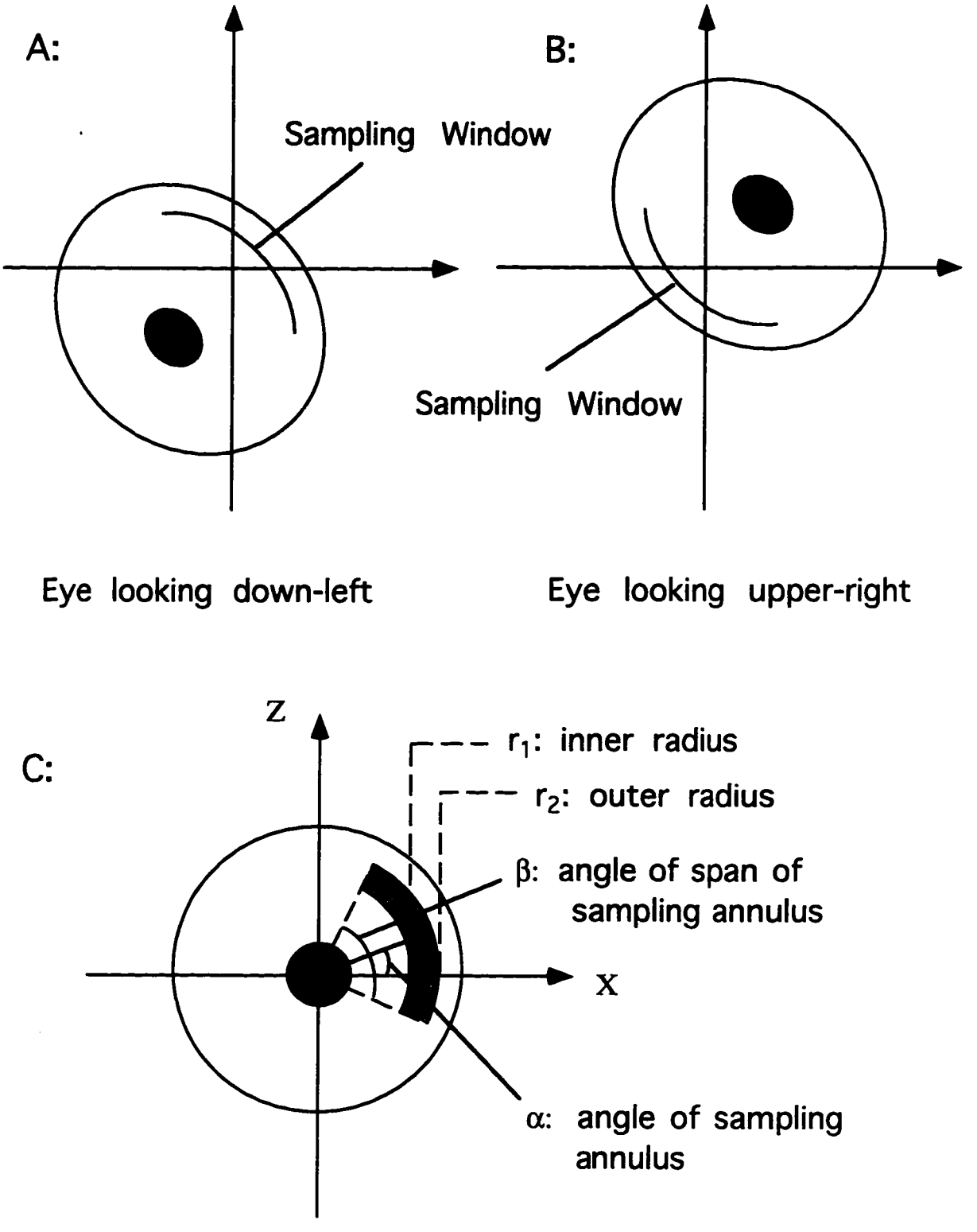


Figure 22: Parameters which Determine Sampling Annulus

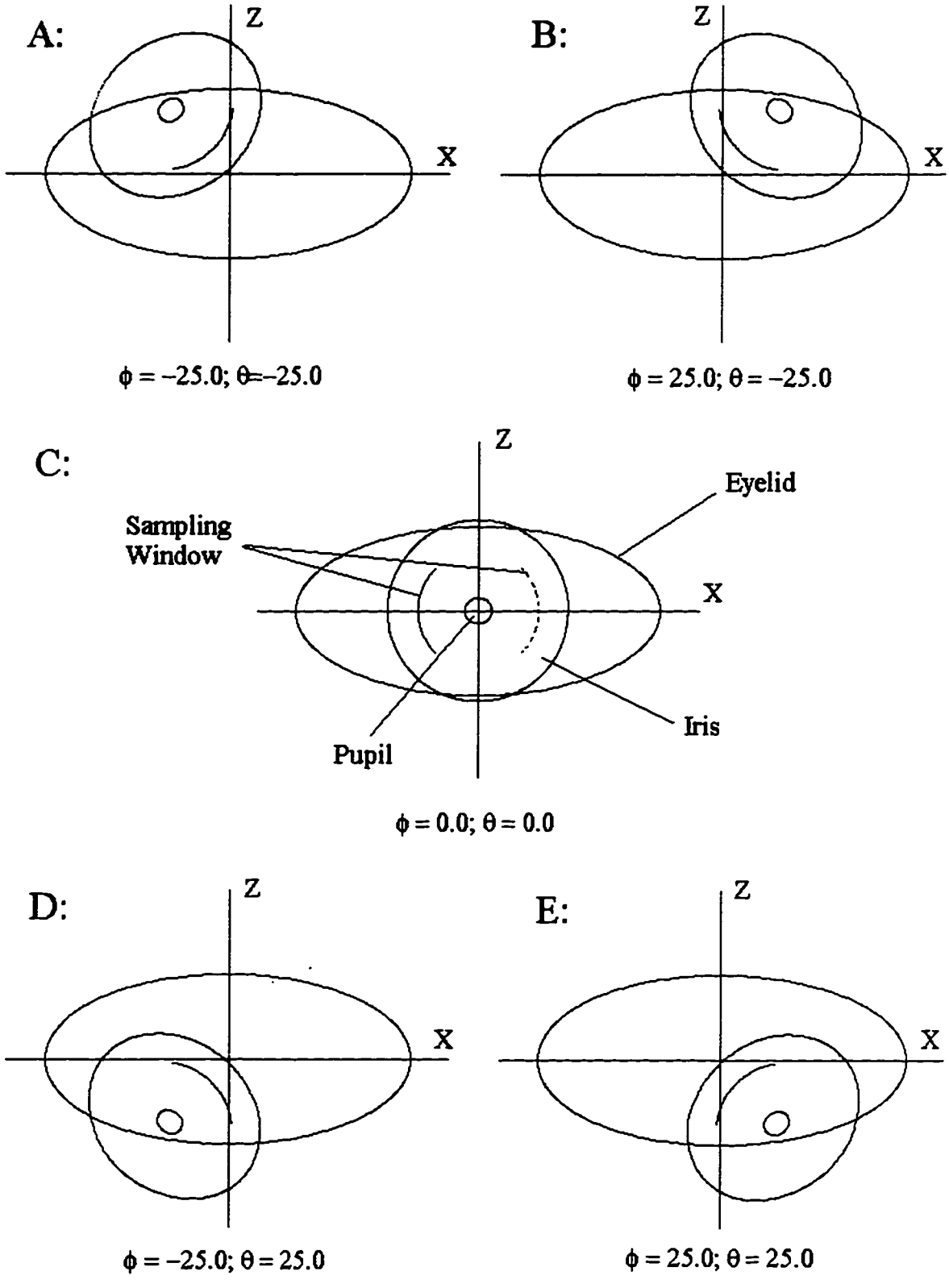


Figure 23: Simulation of How Sampling Window Adapts to Different Position for Different Eye Orientation

eye image translates to the left half of the image plane, the iral signature will be sampled from the right half of the iris (Figure 23B, E). When the eye moves upward, the sampling window moves down (Figure 23A, B) and when the eye moves downward, the sampling window moves up (Figure 23D, E). When the eye is directed straight ahead (Figure 23C), the sampling window can be chosen either to the right (Figure 23C, solid arc) or to the left (Figure 23C, dotted arc) of the pupil.

We now give the algorithm for choosing the sampling annulus as a function of α, ϕ and θ. A heuristic relationship has been used to shift the annulus according to eye orientation to a visible part of the iris by changing α as follows:

$$\alpha = \begin{cases} \pi - k(\theta+\theta_c) & \phi+\phi_c \geq 0 \\ k(\theta+\theta_c) & \phi+\phi_c < 0 \end{cases} \qquad (89)$$

$$(0<k<2)$$

where ϕ and θ are the Euler angles of the current eye image relative to the reference position, ϕ_c and θ_c are the horizontal and vertical offset angles of the camera, and k is a constant between the range 0 and 2 which is determined during calibration. Figure 23 simulates how the sampling window changes position for different ϕ and θ. The heuristic is based on the idea that when the eye looks to the left ($\phi > 0$) from the subject point of view, the annulus is best viewed on the right side of the pupil ($\alpha = \pi$) (Figure 23B, E). Similarly, when the eye looks to the right ($\phi < 0$), the annulus is best viewed on the left side of the pupil ($\alpha = 0$) (Figure 23A, D). The angle α is then adjusted as a linear function of pitch eye position. As the eye moves up when looking to the left, the annulus

is rotated clockwise (along the negative Y_e axis) (Figure 23B). When the eye moves down when looking to the left, the annulus is rotated counterclockwise (along the positive Y_e axis) (Figure 23E).

Selecting α, the angle of sampling annulus according to the Euler angles of the eye gives us a comparatively clear, uncovered and high quality iral signature.

4.9.3 DETERMINATION OF SAMPLING ANNULUS IN ROTATED EYE IMAGE

The position of the sampling annulus for the current eye image in the reference position is given by Eq. (89). The actual sampling annulus in the rotated eye image is not a circular annulus but it transforms into an elliptical annulus. The coordinates of each pixel within the sampling annulus in the rotated eye image can be found using a three step algorithm. First, determine α from ϕ and θ (Eq. (89)). Second, find the pixels associated with the circular annulus defined by Eq. (33) using the chosen β and determined α:

$$f(x,y) = f(r\cos\psi, -r\sin\psi) = g(r,\psi)$$
$$(r_1 \le r \le r_2; \qquad \alpha - \frac{\beta}{2} \le \psi \le \alpha + \frac{\beta}{2}) \tag{90}$$

Third, map each pixel given by Eq. (90) by a rotation and projection given by Eq. (80).

4.9.4 IRAL SIGNATURE DATA STRUCTURE

The gray levels of the annulus pixels are stored in a two dimensional array structure whose format is shown in Figure 24. Each arc of the annulus is an array

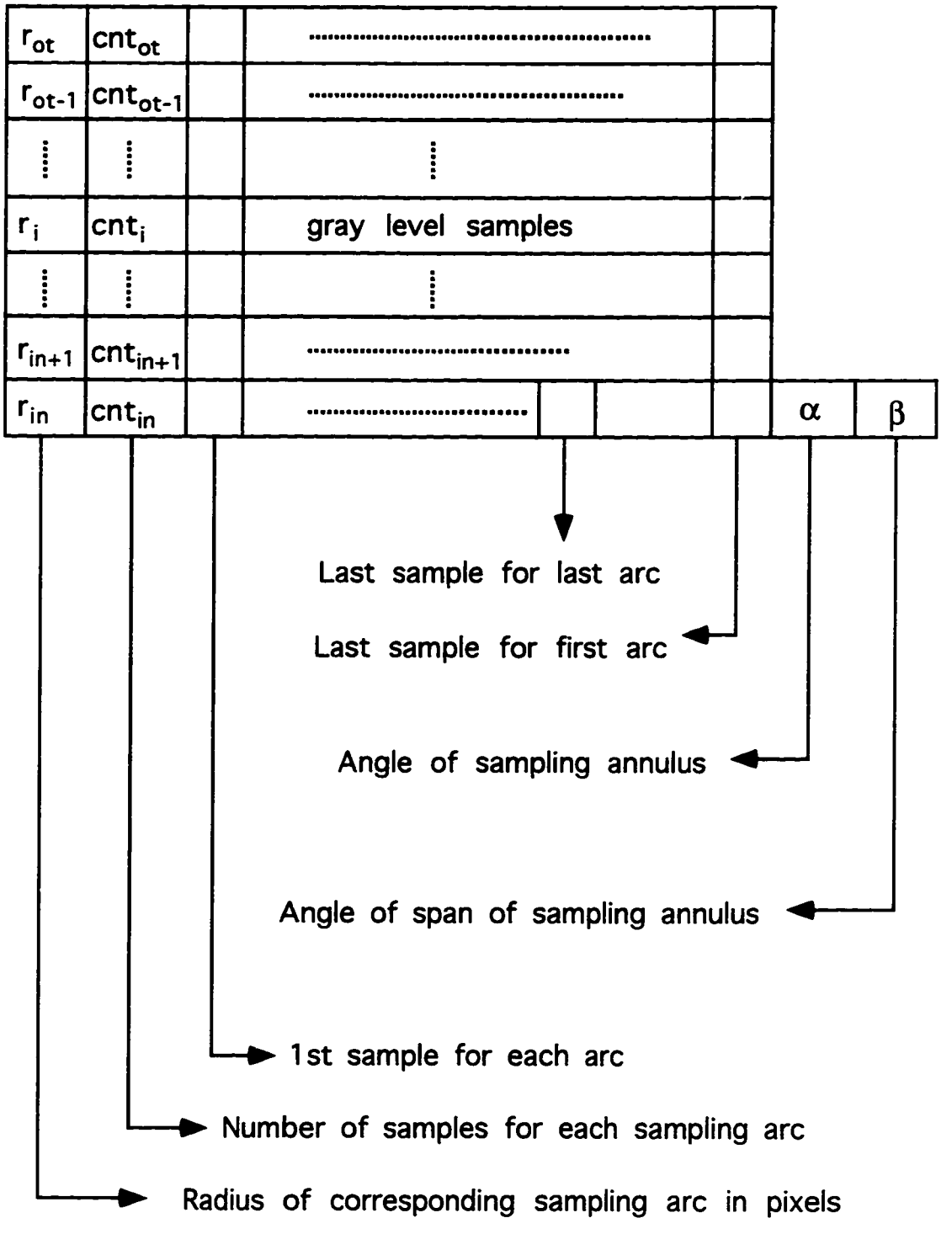


Figure 24: Data Structure for Iral Signature

containing the radius (r_i), the number of pixels sampled along the arc (cnt_i), and the gray level samples. The number of samples is dependent on the radius of the arc. The α and β are stored at the end of the two dimensional array.

4.9.5 TORSIONAL RESOLUTION OF IRAL SIGNATURE

The torsional resolution is defined as the average angular span per pixel of the signature and is dependent on the resolution of the iral signature. It is obtained by dividing the angular extent of the arc used for signature sampling by the number of pixels on it (deg/pixel). Thus, the smaller the deg/pixel, the higher the resolution.

The number of pixels on the arc within the sampling window, which has a fixed angular span, increases with the radius of the arc (Figure 25). Therefore, the resolution increases with the radius of the arc. The resolution can be computed by dividing 360° by the number of pixels on the circumference of a circle of the given radius. To do this, we will define a "digital circle" and what is meant by a pixel being on the circumference of a circle of radius, r.

DEFINITION:

For a given radius, r (given in pixels), the circumference of the "digital circle" centered at (x_c, z_c) is defined as the set of all pixel coordinates corresponding to maximal integer pairs (x, z), such that $(x-x_c)^2 + (z-z_c)^2 \leq r^2$.

The pixel coordinates of the digital circles with radii 4 and 8 with centers at (0, 0), according to the above definition, are shown in Figure 25. A simulation has generated

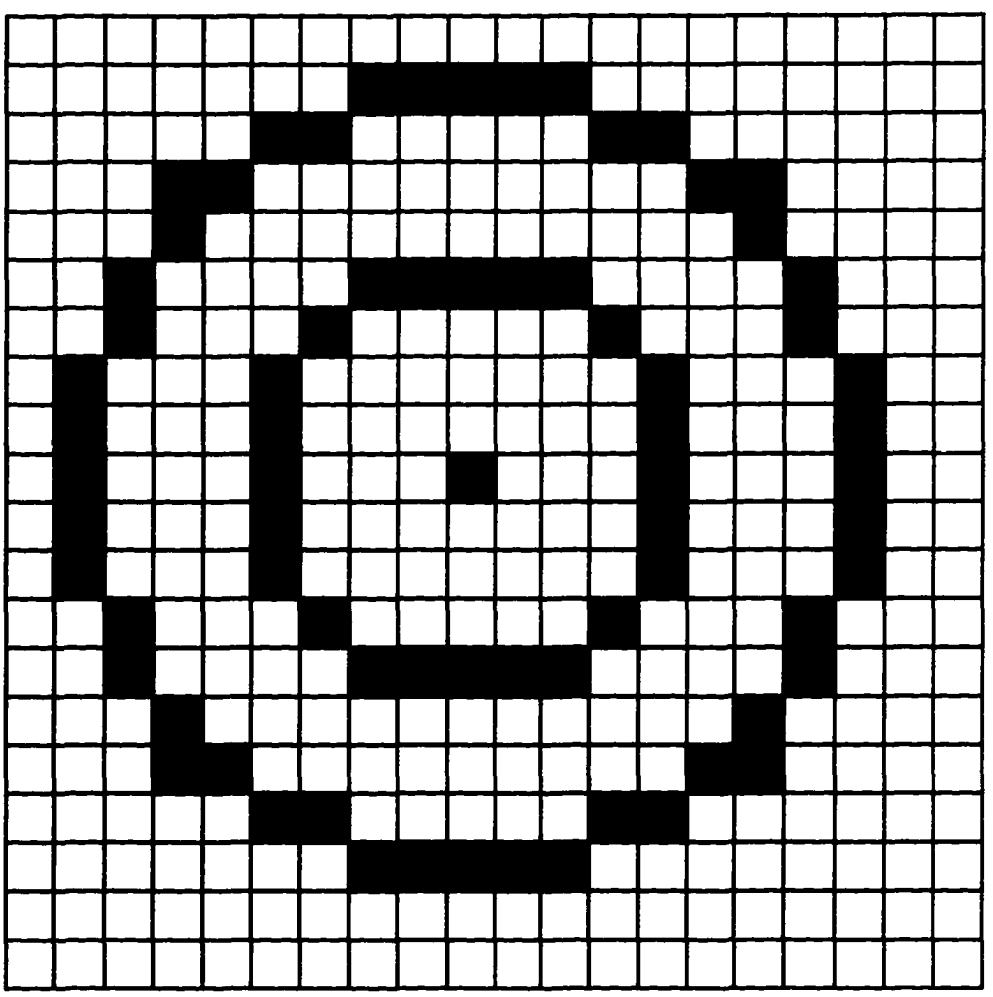


Figure 25: Digital Circles with Radii Being 4 and 8 Pixels

the coordinates on the circumference of the circle with radii ranging from 100 to 200 pixels, and computed the torsional resolution (Figure 26). The number of pixels increases approximately linearly with r (Figure 26A, Solid Line). Correspondingly, the resolution in deg/pixel becomes smaller with increased radius, corresponding to higher resolutions (Figure 26B, Solid Line). The resolution is bounded and approximately given by (Figure 26A, B, Dotted Line):

$$Res_t = \frac{360}{6r}(deg/pixel) \tag{91}$$

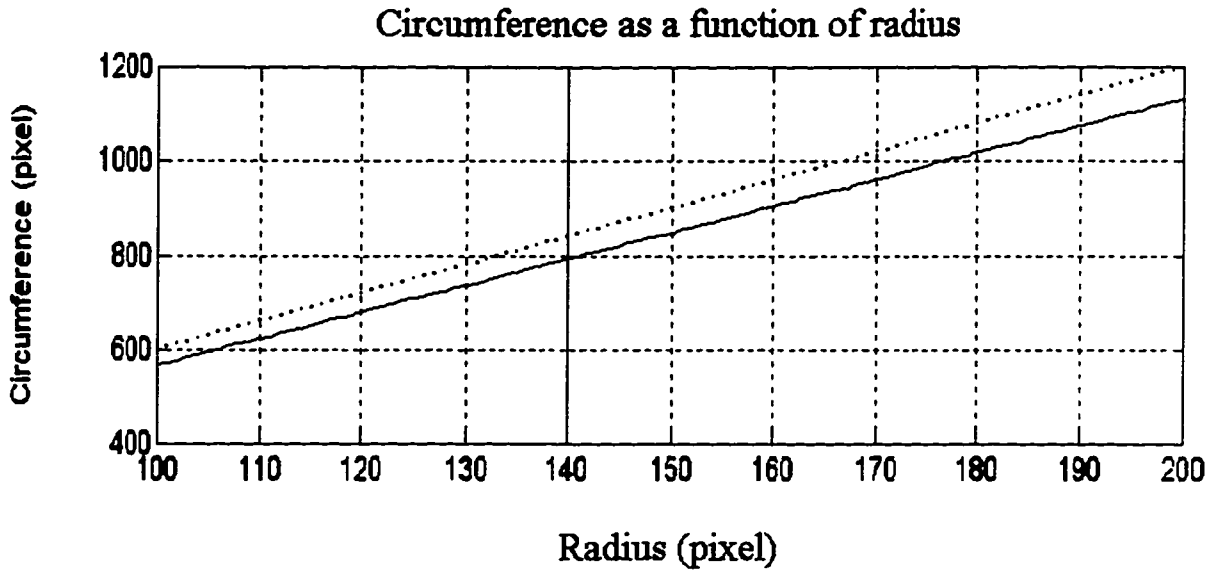
where r is the radius of the sampling arc.

4.10 INTERPOLATION

From Eq. (91), the torsional resolution is approximately inversely proportional to the radius of the sampled arc. Therefore, the value for Res_t can be used as the angular increment for signature sampling (Eq. (33)). If the angular increment for Eq. (33) is too large, not enough pixels are sampled, and the resolution is lost. If the increment is too small, the resolution is preserved, but there is redundant sampling.

To increase the torsional resolution of the iral signature, we have implemented a cubic spline interpolation function to resample the iral signature so that the interpolated iral signature has sub-pixel accuracy for the torsional movement of the eye. Since most of the pixel intensity changes of the iris are along the angular direction (Hatamian and Anderson 1983), only the arcs along a fixed radius are resampled. In this study, the resampled iral signature contains twice as many pixels in width as that of non-

A:



B:

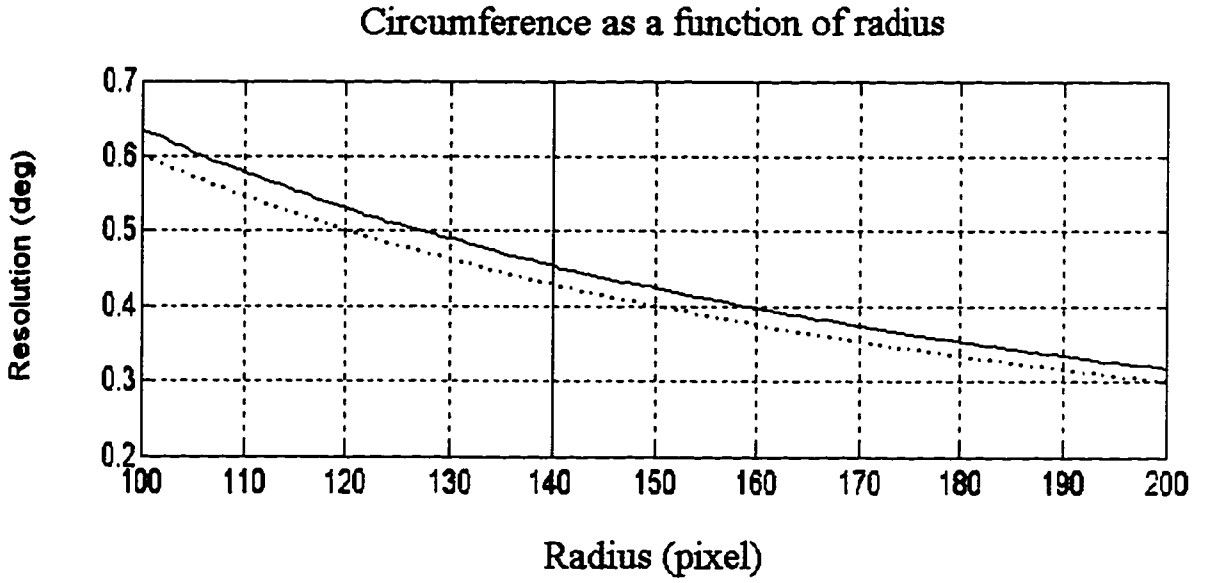


Figure 26: Resolution of Signature as Function of Radius

interpolated iral signature. Therefore after performing a cubic spline interpolation, the angular increment and hence the torsional resolution is given by the following equation:

$$Res_t = \frac{360}{12r}(degree) \tag{92}$$

The radius of the sampling arc r is approximately 150 pixels in this study. Therefore from Eq. (92), the theoretical torsional resolution using the interpolated iral signature is around 0.19°. Various interpolating methods are available (Press et al. 1992), but cubic spline interpolation provides the best resolution because of its high-frequency enhancement (Parker et al. 1983) and we give a short description of the methodology.

4.10.1 CUBIC SPLINE INTERPOLATION

Given a tabulated function $y_i=y(x_i)$, i=1...n, we can interpolate values between two sampling intervals, x_j and x_{j+1}, by a piecewise continuous third-order function which is given by (Press et al. 1992):

$$y = Ay_j + By_{j+1} + Cy_j'' + Dy_{j+1}'' \tag{93}$$

where

$$A = \frac{x_{j+1} - x}{x_{j+1} - xj}$$

$$B = \frac{x - x_j}{x_{j+1} - x_j} \qquad (94)$$

$$C = \frac{(A^3 - A)(x_{j+1} - x_j)^2}{6}$$

$$D = \frac{(B^3 - B)(x_{j+1} - x_j)^2}{6}$$

and y_j'' is the second derivative of the function $y=y(x)$ for $x=x_j$ (Eq. (93)) and x is a value between x_j and x_{j+1} whose interpolated function value is to be found. The above defined cubic spline function is smooth in the first derivative and continuous in the second derivative within any interval and at its boundaries and can be sampled between x_j and x_{j+1}.

4.11 TEMPLATE MATCHING

Torsion of the eye is computed by matching the current iral signature to the reference iral signature using minimum distance classifier to determine the amount of linear shift in pixels between the two signatures. Each arc of the current iral signature is compared as a template with the corresponding arc of the reference signature pixel by pixel. The sum of the gray level difference between each pixel of the current iral signature and the corresponding pixel in the reference iral signature is calculated. Let $X=(x_0, x_1 \ldots x_{n-1})$ be the current iral signature and $Y=(y_0, y_1 \ldots y_{m-1})$ be the reference

iral signature where $m \geq n$ and x_i and y_i are the gray levels for the ith pixels. The distance between X and Y at the ith pixel is given by the following equation:

$$dist(i) = \frac{1}{ns} \sum_{j=0}^{n-1} |x_j - y_{(i+j) \bmod(m)}| \qquad (95)$$

$$(0 \leq i \leq m-1)$$

where ns is the number of samples used for template matching. When X and Y contain no bright spot caused by the infra-red reflection of LEDs, ns=n. Otherwise the samples associated with the bright spots are removed and ns < n. The position where the minimum distance is found gives the amount of shift between X and Y. The average shift in pixels between the current iral signature and the reference iral signature is obtained by comparing all the arcs in the current signature with the corresponding arcs in the reference signature. In our study, the number of arcs used in the iral signature ranges from one to five. Increasing the number of arcs will increase the accuracy of the algorithm, but decreases its speed.

Assume m is the average shift in pixels between the current iral signature and the reference iral signature, the torsional angle of the eye ψ is computed by the following equation:

$$\psi = m * res_t \qquad (96)$$

4.11.1 REMOVING NOISE DUE TO LIGHT INTERFERENCE

One problem which has dominated torsion algorithms in the past is light interference. When monitoring the eye movement using a CCD camera, infra-red LEDs are mounted on the goggles to illuminate the eye. Thus, the image invariably contains a bright spot. This bright spot interferes with the template matching if it happens to be inside the sampling annulus. In our study, this problem is solved by ignoring all the bright pixels inside the iral signatures during template matching. The bright spot pixel has a gray level greater than the threshold for the iris (Figure 17). Therefore, before the gray level difference between a pixel in the current iral signature and its corresponding pixel in the reference iral signature is calculated, the gray levels of the two pixels are examined. If the gray levels of both pixels are less than the threshold, the difference is calculated and the result is added to the final distance. If the gray level of either pixel is greater than the threshold, the pixels are not used for the template matching.

Since a bright spot may occur within a sampling annulus, the actual pixels used for template matching varies from signature to signature. To solve this problem, the normalized distance has been used (Eqs. (35), (95)) and is refined below:

$$dist(i) = \frac{1}{ns} \sum_{j=0}^{n-1} |x_j - y_{(i+j) \bmod(m)}|$$

$$\forall \quad t_1 \le x_j \ \ and \ \ y_{(i+j) \bmod(m)} \le t_2 \qquad (97)$$

$$(0 \le i \le m-1)$$

where ns is the total number of pixels in the iral signature used in template matching and t_1 and t_2 are the lower and upper thresholds for the iris (Figure 17).

CHAPTER 5

RESULTS

5.1 INTRODUCTION

To validate the algorithm for measuring three dimensional eye movements, we have tested our programs on both artificial and real eye images. The artificial eye used for testing is mounted on a gimbal (Figure 13) and can be rotated horizontally, vertically and torsionally to known angles with an accuracy of 0.1°. Angles computed by the algorithm can then be compared with the known angles of rotation. Real eye images used in this study were obtained from LETI in connection with the Neurolab project (Figure 14) and stored on SVHS tapes. They are images of the eye looking at some fixed positions at known angles and sampled at 60 frames/sec.

5.2 VALIDATION OF ALGORITHM USING ARTIFICIAL EYE

5.2.1 TRACKING OF HORIZONTAL AND VERTICAL EYE MOVEMENTS

In this test, the camera was mounted on the gimbal and adjusted so that the optic axis of the artificial eye was orthogonal to the image plane in the reference position, i.e. there was no horizontal, vertical and torsional camera offset relative to the reference-coordinate-frame. In addition, the pupil and iris were completely visible.

First, an image of the artificial eye in the reference position was grabbed. The histogram of the gray-level distribution of the image was calculated to find the threshold

values that separate the pupil, the iris and the rest of the image (Figure 27). The whole iris image was stored and was used as the reference signature. Figure 27A shows the artificial eye image in the reference position. Figure 27B shows the computed gray-level histogram of the artificial eye image. The gray level thresholds that separate the pupil, the iris and the rest of the image are 90 and 140 respectively.

After the reference eye image was grabbed, the artificial eye was rotated horizontally and vertically to the position where $\phi=15°$ and $\theta=11°$ (Figure 28A). The program computed the eye position with high accuracy. The thresholds were used to segment the image into three parts (pupil, iris and the rest of the image). The contour following program then found all the edge points on the boundary of the iris (Figure 28B). The best fitting ellipse to the boundary points was obtained. The Euler angles ($\phi=14.81°$ and $\theta=11.04°$) were computed from the parameters of the best fitting ellipse.

In this test, images of eye at 11 different positions were grabbed. The Euler angles at each eye orientation were then computed. The mean and SD of the errors for the computed Euler angles were $0.341°\pm0.248°$ for the horizontal component, and $0.257°\pm0.197°$ for the vertical component (Table I).

5.2.2 TRACKING OF EYE WHEN IRIS IS COVERED BY EYELIDS

To test the accuracy of the fitting algorithm for the eye image with partial eyelid enclosure, the artificial eye was covered by two pieces of paper (Figure 29). The artificial eye was first set in the reference position (Figure 29A). Then it was rotated horizontally and vertically to the position where $\phi=15°$ and $\theta=15°$ (Figure 29C). The

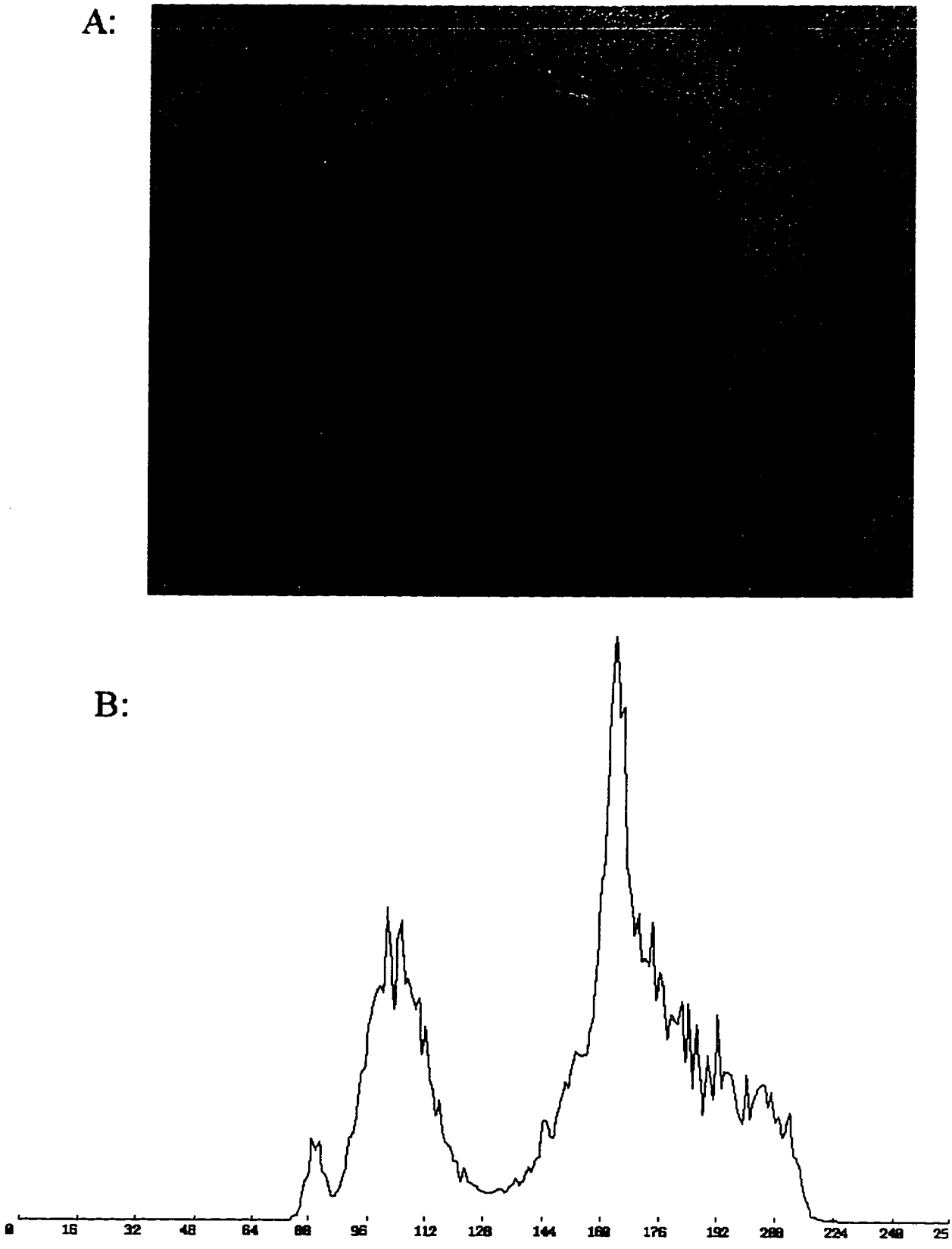


Figure 27: Artificial Eye In Reference Position (ϕ=0°, θ=0° and ψ=0°).

A:

B:

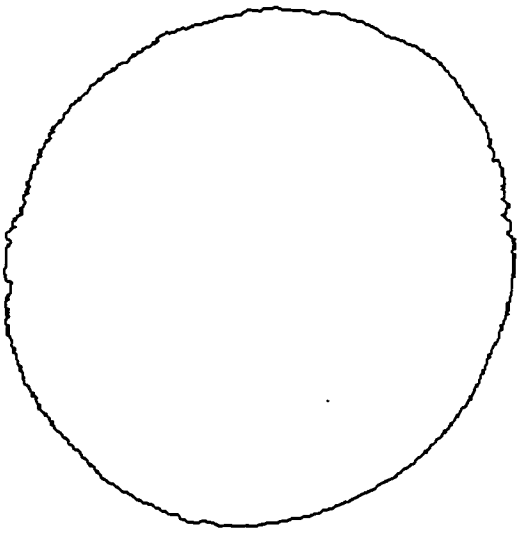
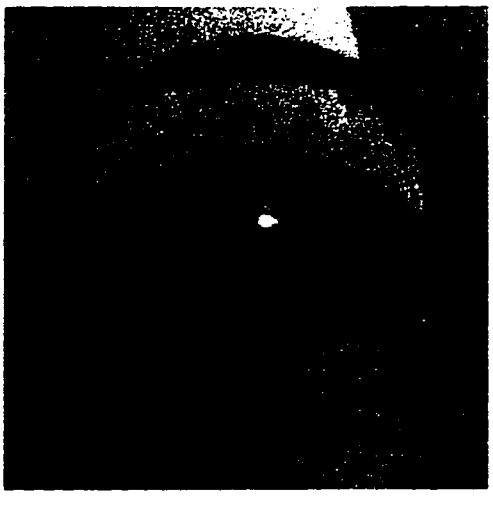


Figure 28: Eye Looking toward Lower Left (ϕ=15°, θ=11° and ψ=0°)

Table I: Comparison of actual and computed horizontal and vertical angles for 11 eye orientations using artificial eye images. Mean and standard deviation of errors are given in degrees.

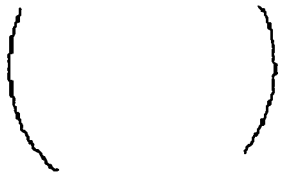
	Actual Angles (ϕ, θ)	Computed Angles (ϕ, θ)	Errors (ϕ, θ)
1	(0, 0)	(0.113, -0.283)	(0.113, 0.283)
2	(5, -5)	(4.798, -5.580)	(0.202, 0.580)
3	(5, 5)	(4.863, 5.073)	(0.137, 0.073)
4	(-5, 5)	(-4.716, 5.064)	(0.284, 0.064)
5	(-5, -5)	(-4.906, -5.266)	(0.094, 0.266)
6	(15, -10)	(14.516, -10.299)	(0.484, 0.299)
7	(15, 10)	(14.397, 10.511)	(0.603, 0.511)
8	(-15, 10)	(-14.213, 10.538)	(0.787, 0.538)
9	(-15, -10)	(-14.867, -10.105)	(0.133, 0.105)
10	(20, -10)	(19.280, -10.073)	(0.720, 0.073)
11	(15, 11)	(14.810, 11.040)	(0.190, 0.040)
Mean of Errors (ϕ, θ)			(0.341, 0.257)
Standard Deviation of Errors (ϕ, θ)			(0.248, 0.197)

A:



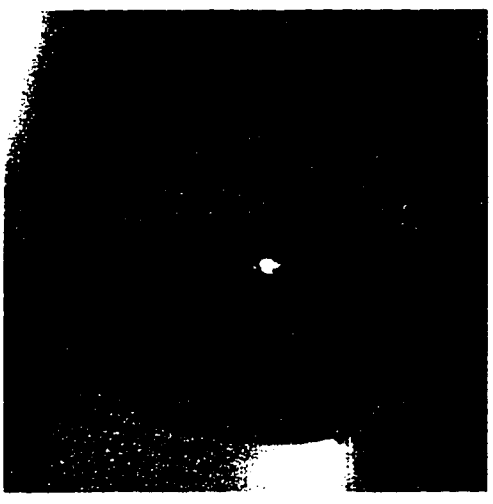
Eye Looking Straight Ahead
θ = 0.0
ϕ = 0.0

B:



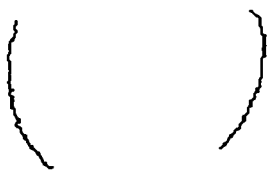
Contour of Uncovered Iris
Calculated Result:
θ = −0.140
ϕ = 0.340

C:



Eye Looking toward Lower Left
θ = 15.0
ϕ = 15.0

D:



Contour of Uncovered Iris
Calculated Result:
θ = 15.094
ϕ = 14.626

Figure 29: Artificial Eye Partially Covered by Eyelid

curvature values at each iris-scleral boundary point were computed by the contour following program (Figure 30). The four peaks in Figure 30A and B represent the four boundary positions in Figure 29A and C where iral contours intersect eyelid contours. In this test, a threshold curvature value of 0.65 was used to detect all the uncovered iris-scleral boundary points (Figure 29B, D). These points were then used to find the best fitting ellipses to the iris-sclera boundaries projected onto the image plane. The Euler angles (ϕ=-0.140°, θ=0.340°) and (ϕ=15.094°, θ=14.626°) were then computed from the best fitting ellipses.

Eye images at 11 different positions were used to test the accuracy of the algorithm in computing the eye orientation when the eye is partially covered by the eyelid. The overall discrepancies between the computed Euler angles and the actual angles were 0.352°±0.268° for the horizontal rotation (ϕ) and 0.203°±0.122° for the vertical rotation (θ) (Table II).

5.2.3 TRACKING OF TORSIONAL EYE MOVEMENT

To test the torsional capability of the system, the artificial eye was rotated to the position where ϕ=10°, θ=0° and ψ=5°. The yaw and pitch Euler angles computed from the projected iral contour were ϕ=9.759° and θ=0.258°. An adaptive sampling annulus was then located using Eq. (89) with ϕ=9.759°, θ=0.258°, k=1.5, ϕ_c=0 and θ_c=0 (Figure 31). The bright rectangle on the left side of the iris in Figure 31 simulated the infra-red light reflection of the LEDs. In this test, the inner and outer radii of the sampling annulus were chosen to be 141 pixels and 145 pixels respectively. The gray-

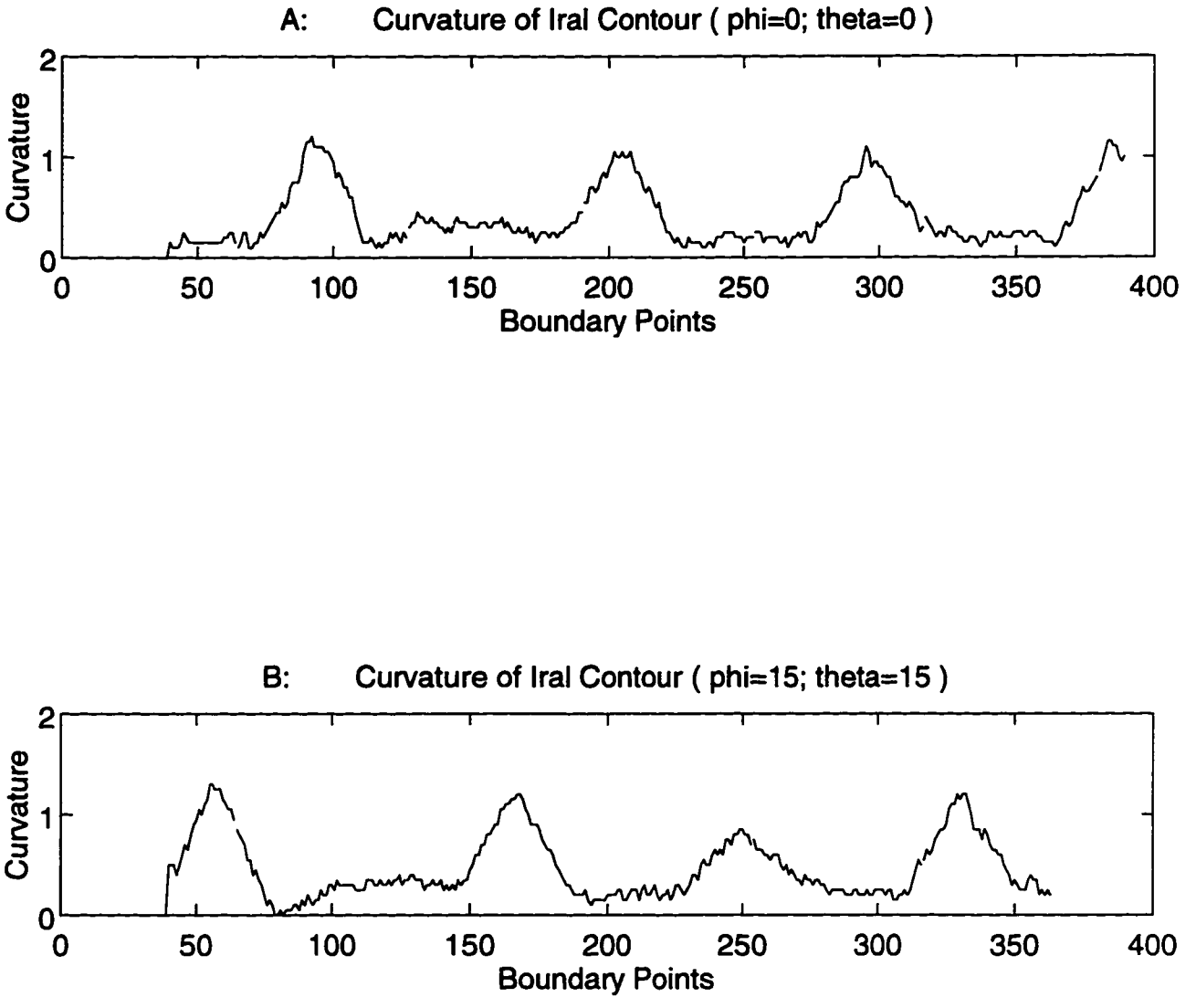


Figure 30: Curvature Value at Each Boundary Point

Table II: Comparison of actual and computed H and V angles for 11 eye orientations using artificial eye images with partially covered irises. Mean and standard deviation of errors are given in degrees.

	Actual Angles (ϕ, θ)	Errors (ϕ, θ)
1	(-5, 5)	(0.123, 0.123)
2	(-15, 0)	(0.390, 0.238)
3	(15, 0)	(0.775, 0.255)
4	(15, 10)	(0.696, 0.410)
5	(15, -10)	(0.672, 0.076)
6	(-5, -10)	(0.203, 0.169)
7	(-15, -10)	(0.034, 0.089)
8	(-15, 10)	(0.595, 0.047)
9	(-5, 10)	(0.145, 0.111)
10	(0, 0)	(0.140, 0.340)
11	(15, 15)	(0.094, 0.374)
Mean of Errors (ϕ, θ)		(0.352, 0.203)
Standard Deviation of Errors (ϕ, θ)		(0.268, 0.122)

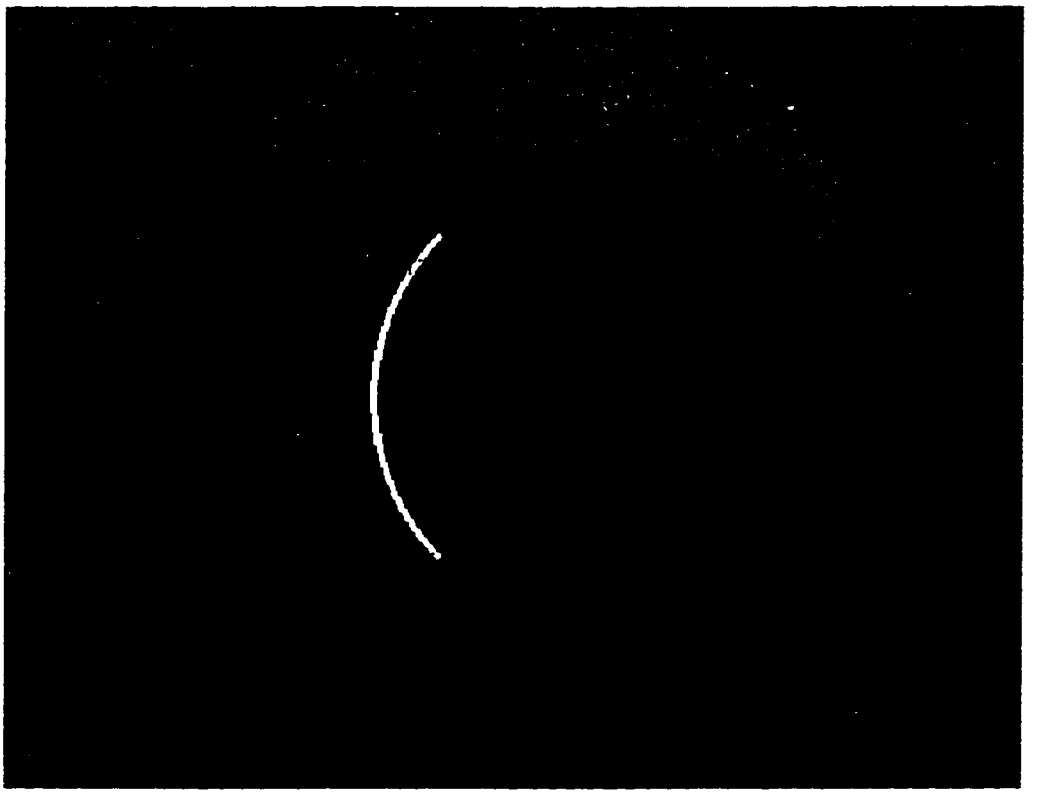


Figure 31: Sampling Annulus of Rotated Eye (ϕ=9.759° and θ=0.258°)

level distribution of one arc (radius=145) within the sampling annulus is shown in Figure 32A. The gray-level distribution of the corresponding arc in the reference signature is shown in Figure 32B. The gray-level distance between two arcs at each angular position was computed using Eq. (95) and is shown in Figure 33.

The computed angular shift between the arcs of the two signatures is $\psi=5.314°$, where the minimum distance (dist=4.847) is found. Each angular shift between arcs within the sampling annulus and the corresponding arcs in the reference signature was calculated. The average value of all the angular shifts gives the degree of torsion of the eye ($\psi=5.223°$).

We have also computed the torsion using non-interpolated iral signatures. The computed result using one arc (radius=145 pixels) within the sampling annulus was $\psi=5.411°$. The average torsion was $\psi=5.361°$. The accuracy of the computed torsion has been doubled by interpolating the iral signature to provide twice as many samples as the original iral signature.

The major error in computing torsion in this test was caused by the errors in computed yaw and pitch Euler angles. We have also computed the torsion using Euler angles $\phi=10°$ and $\theta=0°$. The result was $\psi=5.239°$ using one arc (radius=145 pixels) within sampling annulus. The average computed torsion over the entire annulus was $\psi=4.993°$. Thus, the more accurate the computed pitch and yaw Euler angles the more accurate the computed torsion.

In this test, images of eye at 11 different positions were used to verify the algorithm's capability in computing the torsional eye movement. The overall error for the computed torsion was $0.361° \pm 0.233°$ (Table III).

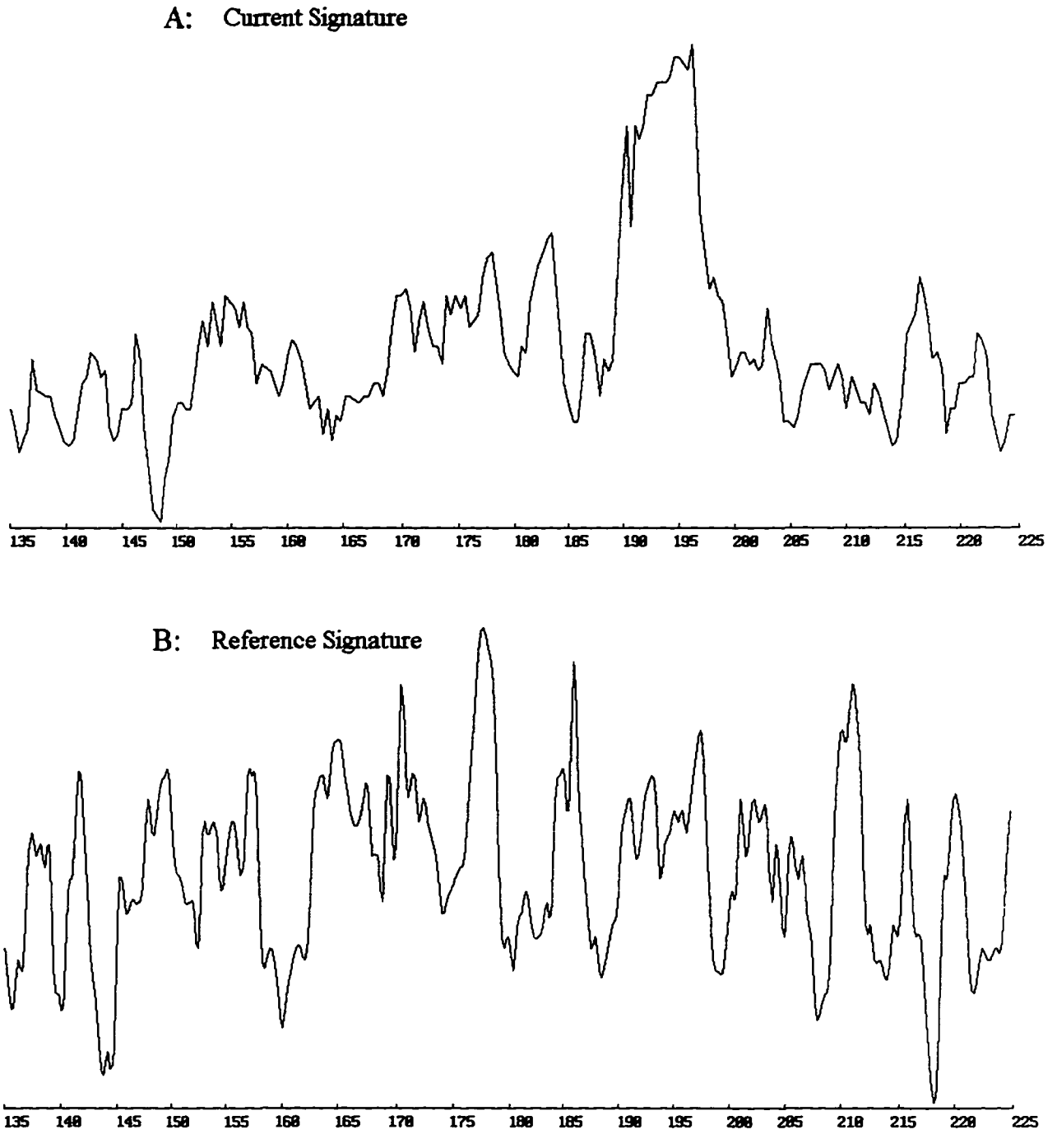


Figure 32: Gray-level Distribution of One Arc in Current Signature And Corresponding Arc in Reference Signature

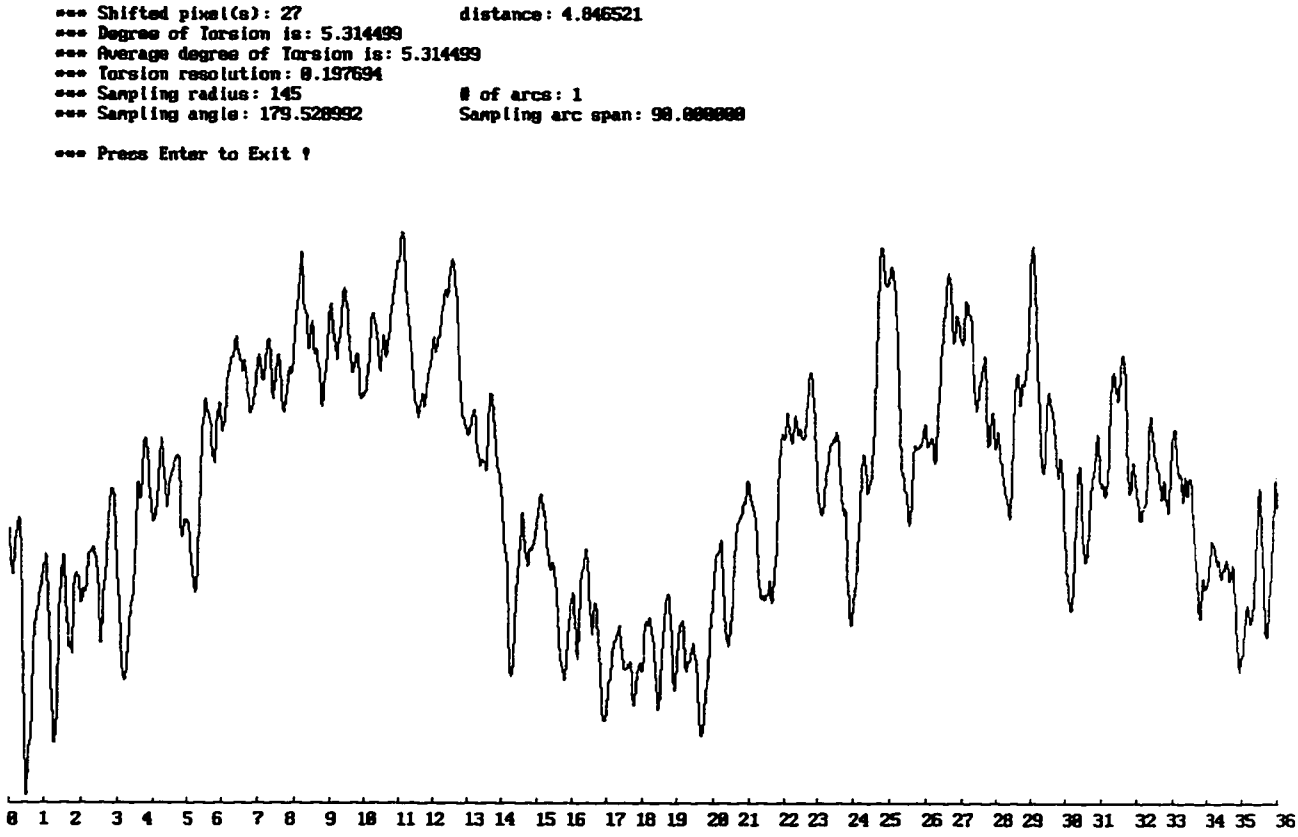


Figure 33: Distance as Function of Shift between One Arc in Current Signature and Corresponding Arc in Reference Signature

Table III: Comparison of actual and computed horizontal, vertical and torsional angles for 11 eye orientations using artificial eye images. Mean and standard deviation of errors are given in degrees.

	Actual Angles (ϕ, θ, ψ)	Computed Angles (ϕ, θ, ψ)	Errors (ϕ, θ, ψ)
1	(10, 0, 5)	(9.759, 0.258, 5.223)	(0.241, 0.258, 0.223)
2	(-10, 0, 5)	(-9.921, 0.104, 4.658)	(0.079, 0.104, 0.342)
3	(-5, 0, 10)	(-5.238, 0.248, 9.816)	(0.238, 0.248, 0.184)
4	(0, 0, -5)	(0.262, -0.289, -5.267)	(0.262 0.289, 0.267)
5	(5, 0, -10)	(5.196, -0.517, -9.942)	(0.196, 0.517, 0.058)
6	(5, 5, 0)	(4.811, 5.082, -0.340)	(0.189, 0.082, 0.340)
7	(4.120, 9.962, 5.077)	(4.049, 9.767, 4.281)	(0.071, 0.195, 0.796)
8	(-5.529, 9.986, 3.046)	(-5.064, 9.704, 3.893)	(0.465, 0.282, 0.847)
9	(-15.437, 4.981, 5.013)	(-15.353, 5.064, 4.718)	(0.083, 0.083, 0.301)
10	(-10.698, 4.951, 8.030)	(-10.543, 5.041, 7.801)	(0.155, 0.090, 0.229)
11	(4.537, -5.280, -5.021)	(4.807, -5.624, -4.641)	(0.270, 0.344, 0.380)
Mean of Errors (ϕ, θ, ψ)			(0.204, 0.227, 0.361)
Standard Deviation of Errors (ϕ, θ, ψ)			(0.108, 0.129, 0.233)

5.2.4 VALIDATION OF CALIBRATION ALGORITHM

As described in section **3.6**, seven parameters must be determined during calibration process. These parameters are ϕ_c, θ_c and ψ_c which are the three camera offset angles relative to the reference-coordinate-frame, (x_c, z_c) which are the coordinates of the center of the image plane, d which is the distance between the center of the eye and the iral plane, and r, the radius of the iral plane in the zero position (looking straight at the camera).

In this test, we adjusted the camera so that the approximate values of the offset angles were ϕ_c=-7°, θ_c=1° and ψ_c=0°. The image of the eye in approximately the zero position (looking straight at the camera) was grabbed, and the center of the iris projection and the radius were computed (x_c=357.43 , z_c=-246.96 and r=165) using the ellipse fitting routine (See section **3.3.4**). The distance between the center of the eye and the iral plane (d) was estimated as follows: First the eye was rotated relative to the "zero position" (optic axis aligned with camera axis) by a known angle. The center of the rotated iris projection was computed and the offset of the rotated iral center relative to the center of iris in the zero position was found. Finally the distance (d=221.79) was computed using Eq. (30).

To calibrate the system, 20 images were grabbed at 20 different eye orientations. Figure 34 shows the 20 eye orientations used for system calibration. The result calculated using Eqs. (61), (63), (48) and (49) matched the estimated result very closely and is shown below:

ϕ_c = -6.63°, θ_c = 0.53°, ψ_c = -0.46°,

x_c = 357.19, z_c = -247.45, d = 222.20, r = 164.55.

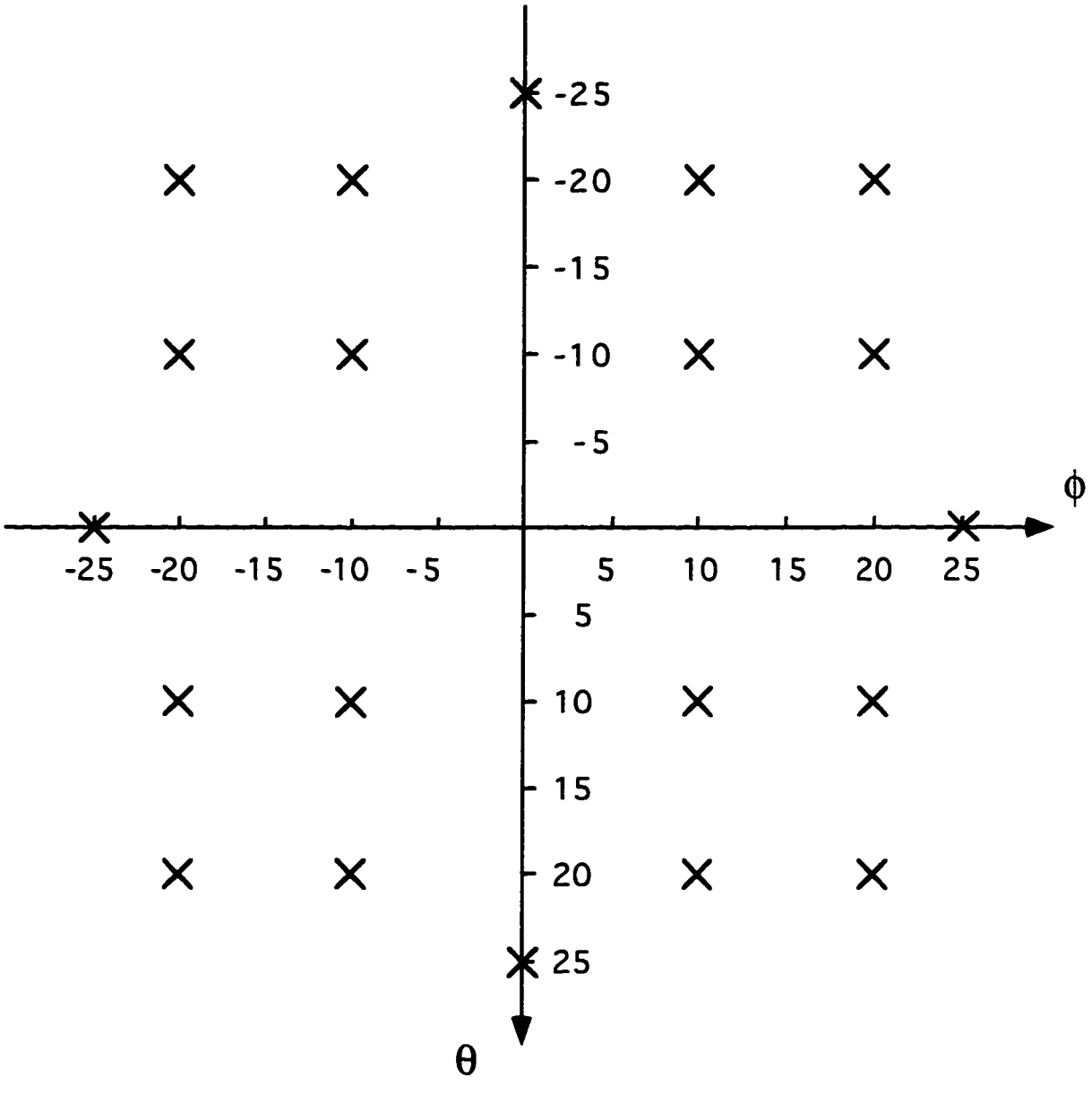


Figure 34: Twenty Eye Orientations for System Calibration

5.3 VALIDATION OF ALGORITHM USING VIDEO EYE IMAGES

We have obtained a video tape (Leti) on which were recorded eye movements obtained from their system. The eye looks successively at 34 positions with known orientations (Figure 35). The offset angles of the optic axis of the camera are: $\phi_c=0°$, $\Theta_c=-5°$ and $\psi_c=0°$ relative to the reference-coordinate-frame (Figure 35).

Sixteen (16) eye images, corresponding to positions 3, 9, 13, 19, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32 and 33 were used to calibrate the system (Figure 35). Since ϕ_c, Θ_c and ψ_c are already known, the parameters that needed to be computed are (x_c, z_c), the coordinates of the center of the image plane and d, the distance between the center of the eye and the iral plane. The computed result is shown below:

$x_c = 413.70$, $z_c = -248.58$ and $d = 237.79$

The calibrated parameters were then used to compute the eye orientations for other eye images.

Two images were grabbed from the video tape. Figure 36 shows an image of eye in the reference position ($\phi=0$ and $\theta=0$). The iris was partially covered by the upper eyelid. The contour following program was used to find all the uncovered iris-sclera boundary points (Figure 36). The computed Euler angles for this eye image, using the uncovered boundary points, were: $\phi=0.22°$ and $\theta=-0.21°$. Figure 37 shows an image of eye with Euler angles $\phi=10°$ and $\theta=10°$. The contour of the pupil was utilized to find the orientation of the eye. The computed Euler angles are: $\phi=9.62°$ and $\theta=10.39°$.

18 Images of the eye had been grabbed in this test. The Euler angles at each eye position were then computed. The overall accuracies were $0.671°\pm0.561°$ for the horizontal eye movement and $0.685°\pm0.551°$ for the vertical eye movement (Table IV).

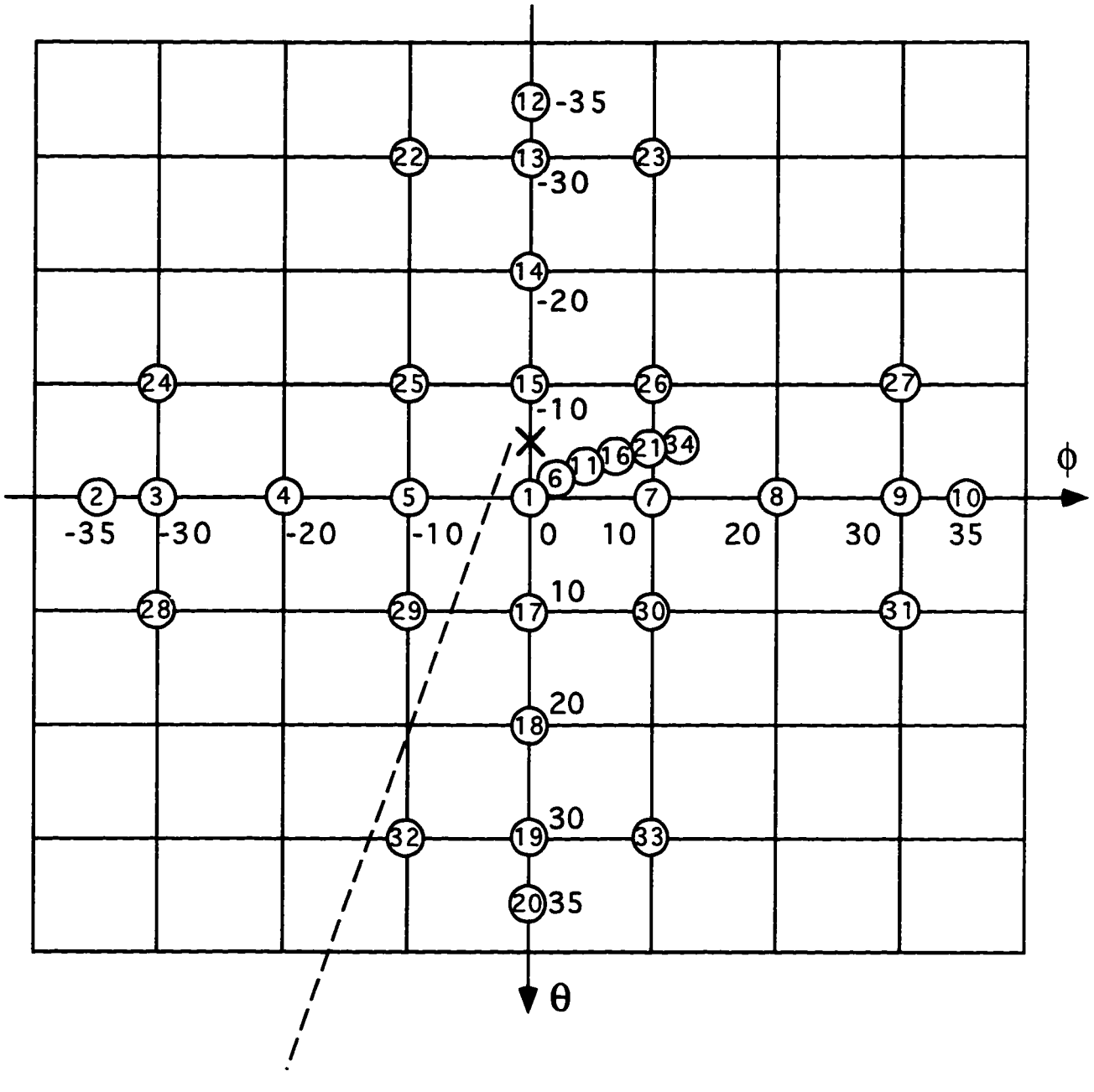
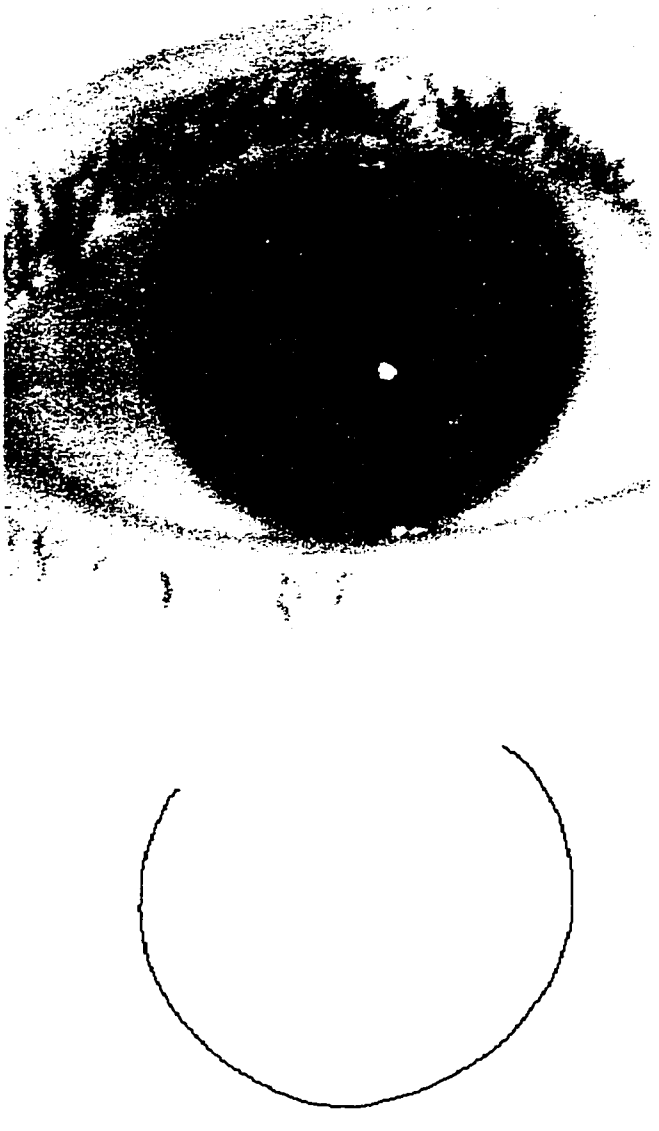
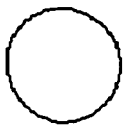


Figure 35: Scan Pattern of Eye Recorded on Video Tape and Angle of Camera Optical Axis



Direction of Gaze: $\theta = 0.0$; $\phi = 0.0$
Calculated Result: $\theta = -0.2147$ $\phi = 0.2230$

Figure 36: Video Image of Eye Partially Covered by Eyelid and Corresponding Iral Contour (ϕ=0°, θ=0°)



Direction of Gaze: $\theta = 10.000$ $\phi = 10.000$

Calculated Result: $\theta = 10.342$; $\phi = 10.032$

Figure 37: Video Image of Eye and Corresponding Pupillary Contour ($\phi=10°$, $\theta=10°$)

Table IV: Comparison of actual and computed horizontal and vertical angles for 18 eye orientations using real eye images from video tape (Leti). Mean and standard deviation of errors are given in degrees.

	Actual Angles (ϕ, θ)	Computed Angles (ϕ, θ)	Errors (ϕ, θ)
1	(0, 0)	(-0.052, 0.378)	(0.052, 0.378)
2	(10, 0)	(11.504, 1.215)	(1.504, 1.215)
3	(0, 0)	(0.640, 0.730)	(0.640, 0.730)
4	(-10, 0)	(-10.152, -0.201)	(0.152, 0.201)
5	(-20, 0)	(-21.450, -0.504)	(1.450, 0.504)
6	(-30, 0)	(-31.356, -1.353)	(1.356, 1.353)
7	(0, 0)	(0.372, 0.119)	(0.372, 0.119)
8	(0, -20)	(-0.029, -19.643)	(0.029, 0.361)
9	(0, -10)	(-0.313, -8.690)	(0.313, 1.310)
10	(0, 0)	(0.061, 0.315)	(0.061, 0.315)
11	(0, 10)	(0.281, 10.477)	(0.281, 0.477)
12	(0, 20)	(0.313, 19.603)	(0.313, 0.397)
13	(0, 30)	(0.986, 28.089)	(0.986, 1.911)
14	(10, -10)	(10.316, -8.293)	(0.316, 1.707)
15	(30, 10)	(31.313, 9.290)	(1.313, 0.710)
16	(10, 10)	(9.431, 10.406)	(0.569, 0.406)
17	(-10, 10)	(-11.809, 9.768)	(1.809, 0.232)
18	(0, 0)	(-0.560, 0.000)	(0.560, 0.000)
Mean of Errors (ϕ, θ)			(0.671, 0.685)
Standard Deviation of Errors (ϕ, θ)			(0.561, 0.551)

The major error was caused by the low resolution of the tape, image noises induced when frames were grabbed from the tape, and the relative movement of the camera with respect to the reference-coordinate-frame of the head.

CHAPTER 6

SUMMARY AND CONCLUSIONS

6.1 CONTRIBUTION OF THIS STUDY

In this study, we have developed new algorithms for computing three dimensional eye movements based on images of the eye. The novel aspect of the work was the utilization of a model-base for computing eye orientation. The model was based on the idea that the iris and pupil contours project as ellipses when the eye is in different orientation relative to the camera. Samples of the iris and pupil contours were fit by this model and algorithms were developed which could extract eye orientation from the shape and eccentricity of the projected ellipses.

Most existing image processing systems for eye movement measurement computed the pitch and yaw Euler angles of the eye by finding the center of the pupil using methods such as averaging the coordinates of the pupil edge points (Hatamian and Anderson 1983), fitting a circle to the pupil boundary (Sung and Anderson 1991) or calculating the "center of mass" (Moore et al. 1996). These algorithms rely on a clear and uncovered pupil image. We have shown in this thesis that the ellipse model used in our study is potentially more robust with regard to the eyelid closure and the infra-red light reflection of the pupil/iris image. In addition, by using the ellipse model, translation and rotation of the eye could theoretically be distinguished (Eq. (20)), especially for large translations. Such distinctions were not possible for the existing image processing systems (Clarke et al. 1991; Moore et al. 1996).

Another contribution of this study was to extend the capability of computing torsion by making the algorithms adaptive. Most existing systems used a fixed arc as the iral signature for computing torsion (Vieville and Masse 1987; Clarke et al. 1991). But as the eyelid droops or as tears build up in the lower eye, the arcs become unsuitable for use. To partially overcome this problem, Moore et al. (1996) used the signatures of four arcs. The adaptive sampling annulus technique used in this study defined an algorithm which allowed for the adjustment of the sampling annulus as a function of the amount of eyelid closure. This added greater flexibility and robustness to the algorithm.

Another novel aspect of this work was the removal of the image noise due to the infra-red light reflection from the iral signatures. This was accomplished by using a "city block" metric to compute the distance between iral signatures as opposed to commonly used cross correlation using FFT technique, which breaks down when part of the signature contains infra-red light reflection (Moore et al. 1996). We have shown that the infra-red light reflection can be overcome by removing the bright spots from the iral signatures and then using minimum distance function to compute the angular shift between the signatures. This makes the signature matching algorithm more robust with regard to the light reflection.

Another important contribution of this work was the implementation of the calibration procedure. In most systems, calibration to find the camera offset angles (ϕ_c, θ_c and ψ_c) has been a problem because the calibration algorithm is sensitive to errors in computing the coordinates of the projection of the pupil center. In fact, the offset angles of the camera were ignored in most systems by adjusting the camera so that the "line of sight" of the camera is aligned along the optic axis when the eye is looking straight

ahead. When the camera is not aligned accurately along the optic axis of the eye in the reference position, error in computing eye orientation occurs. On the other hand, it is sometime necessary to mount the camera with some offset angles, so that the eye can be viewed more clearly even if part of the iris is covered by the eyelid. The idea for this work was to relate the camera offset angles to the center of the pupil and the eccentricity of the ellipse at different eye orientation. Using both criteria made the algorithm less sensitive to noise and more accurate. Yet, it was easy to implement.

6.2 DIRECTION FOR FUTURE RESEARCH

Future research in this area could take a number of directions in developing a practical, accurate, reliable, and robust on-line eye movement tracking system. In this thesis, we have emphasized the robustness of the algorithms with regard to the eyelid enclosure and light reflection. We have also carefully considered the camera calibration problem. The resolution of the algorithm was about 0.3°. The processing speed was about 1.73 sec/frame (refer to Appendix E for detailed description). Further work should be done to enhance the speed, the resolution and the robustness of the system before it can be used in a wide range of applications.

SPEED:

In this work, we have used an external frame grabber (Model 495, Colorado Video Inc.) connected to IBM-PC through a DMA interface (Model 793, Colorado Video Inc.) to capture the images. The frame grabber can capture images at 30 frames/sec. However, the model 793 DMA interface is old and transfers a 512x512 frame in only

1.05 seconds. This configuration has limited the speed at which the image frame was processed. Faster image capturing and DMA transfer will speed up the algorithm by orders of magnitude. To speed up the algorithm, we can use a high speed internal image processor such as Matrox Pulsar (Matrox Electronics System Ltd.) to grab and process images. With this processor, it should be possible to read the image into PC's memory in real time (30 frames/sec). The processor also supports image processing with built-in commands such as histogram equalization, threshold finding, pattern matching, etc., which makes image processing programs run much faster.

In addition to the internal image processor, we can also use a more powerful computer such as a 200 MHz Pentium Pro to increase the processing speed. In this study we have used a 66 MHz 486-DX2 computer to run the programs.

Faster algorithms could also be developed for median filtering, ellipse fitting, template matching and other time consuming procedures. Some critical parts of the program can be re-written in assembly language. One approach which seems exceptionally promising is the development of parallel algorithms to implement the median filtering, ellipse fitting, template matching and other subroutines. This approach should help obtain processing speed close to real time.

RESOLUTION:

To enhance the resolution of the system, we can use more advanced imaging hardware. The high resolution CCD camera GP-MS112 (Panasonic Communication & System Company) has 500 TV lines compared to the 380 lines used in this study. However, higher resolution cameras will require higher resolution frame grabbers. The

Matrox Pulsar image processor contains a frame grabber with 768x576 resolution compared to the Colorado Video 512x512 frame grabber used in this study. The rapid development of the computer and video technology over the next few years will provide hardware with much higher resolution capability which will enhance the performance of the system.

Resolution enhancement can also be obtained through computational techniques. In this study, we have used the orthographic projection to model the projection of the eye onto the image plane. This method works well when the distance between the camera and the eye is much larger than the focal length of the camera. However, when the camera is close to the eye, there are distortions when the eye deviates significantly and resolution is compromised. The problem is exacerbated for large eye deviations. Using the pupil boundary as the model base for determining eye orientation, the resolution of the eye orientation is somewhat better than using the iris boundary. A higher resolution camera should enable the distance from the eye to the camera to be decreased while maintaining the accuracy. Another methodology to overcome this problem is to use the perspective projection to model the projection of the eye onto the image plane.

ROBUSTNESS:

One of the requirement for all existing VOG technique is that the pupil or iris image should be extracted using a single gray level threshold. This requires that the infra-red LEDs be adjusted to illuminate the eye uniformly, so that the gray-levels of the iris, the pupil and the rest of the eye are different, i.e. the gray-level histogram of the eye image consists of three separate dominant modes, representing the pupil, the iris and

the rest of eye (Gonzalez and Woods 1993). If the eye is not illuminated uniformly, the iris, the pupil and the sclera and the rest of the eye may overlap in gray levels. Thus, part of the sclera may be as dark as the iris. In this case, the algorithm is not able to detect the correct iris-sclera or pupil-iris boundaries. To solve this problem, an edge detection algorithm using gradient operators (Ballard and Brown 1982) can be used. Such an algorithm which relies on the gray level differences across the boundaries of the iris and the pupil, not on a certain gray level threshold, can make the edge detection more robust.

Another improvement can be made to the eye movement tracking algorithm by using a different curvature threshold, which is used by the contour following program to detect the iris-eyelid curve intersection, for different eye images. In the current study, the curvature threshold (between 0.6 to 1) was chosen by the operator during calibration. This value was then used for all iris/pupil images. However, the threshold value for each iris/pupil image is dependent on the eye position and the eyelid enclosure of this image. A more robust approach for choosing the curvature threshold for an iris/pupil image is to use the local maximas of the iris/pupil curvature function (Figure 30) of this image. The curvature threshold could be adaptively determined by a program. However, an adaptive algorithm would need to scan the iris/pupil boundary twice: computing the curvature values for all boundary points in the first scan, and then finding all the uncovered boundary points during the second scan.

The gray-level thresholds used to segment the eye image into three parts (the pupil, iris and the rest of the image) can also be determined automatically by a program. In this study, the threshold values were chosen by the operator from the gray-level

histogram of the reference eye image during calibration. It could be determined mathematically using the optimal thresholding technique (Gonzalez and Woods 1993). Optimal thresholding assumes that the image contains three principal brightness regions: the pupil, the iris and the rest of the image. The histogram of the image can then be considered as an estimate of the brightness probability density function, p(z), which is the sum of three unimodal densities given by the following equation:

$$p(z) = P_1 * p_1(z) + P_2 * p_2(z) + P_3 * p_3(z) \quad (98)$$

where P_1, P_2 and P_3 are the probabilities of the three brightness regions, and $p_1(z)$, $p_2(z)$ and $p_3(z)$ are the density functions for the three regions. If we assume the density functions to be Gausian, it is possible to determine the optimal thresholds from Eq. (98) (Gonzalez and Woods 1993).

While much work needs to be done to develop a practical, robust three dimensional video-based eye movement monitor, this thesis has made important contributions towards achieving that goal.

APPENDIX A

CALCULATION OF ANGLE FOR ROTATION TRANSFORM

To simplify the computation of the center of the ellipse, a rotation transform was applied to the ellipse in this study. The angle of rotation transform is derived below.

The general equation for the ellipse is as follows:

$$Ax^2 + Bxz + Cz^2 + Dx + Ez + F = 0 \tag{99}$$

To transform the above equation such that the major and the minor axes are parallel to X and Z axes respectively, we rotate X-Z plane about the center of X-Z plane, so that the rotated X and Z axes, i.e. X' and Z' are parallel to the major and the minor axes of the ellipse (Figure 8). This is done by substituting the following equation into Eq. (99).

$$\begin{aligned} x &= x'\cos\alpha - z'\sin\alpha \\ z &= x'\sin\alpha + z'\cos\alpha \end{aligned} \tag{100}$$

After rotation transform, we have:

$$
\begin{aligned}
&(A\cos^2\alpha + B\cos\alpha\sin\alpha + C\sin^2\alpha)x'^2 - \\
&(2A\sin\alpha\cos\alpha + B\sin^2\alpha - B\cos^2\alpha + 2C\sin\alpha\cos\alpha)x'y' + \\
&(A\sin^2\alpha - B\cos\alpha\sin\alpha + C\cos^2\alpha)z'^2 + \\
&(D\cos\alpha + E\sin\alpha)x' - (D\sin\alpha - E\cos\alpha)z' + F = 0
\end{aligned}
\tag{101}
$$

Since:

$$
-2A\sin\alpha\cos\alpha - B\sin^2\alpha + B\cos^2\alpha + 2C\sin\alpha\cos\alpha = 0 \tag{102}
$$

The above equation is equivalent to:

$$
B\sin^2\alpha + 2(A-C)\sin\alpha\cos\alpha - B\cos^2\alpha = 0 \tag{103}
$$

which is equivalent to:

$$
B\tan^2\alpha + 2(A-C)\tan\alpha - B = 0 \tag{104}
$$

therefore:

$$
\tan(\alpha) = \frac{C - A \pm \sqrt{A^2 - 2AC + C^2 + B^2}}{B}) \tag{105}
$$

therefore:

$$\alpha = \arctan\left(\frac{C-A\pm\sqrt{A^2-2AC+C^2+B^2}}{B}\right) \tag{106}$$

The same value for α can be derived with another method which computes the major and the minor axes first.

We can rewrite Eq. (99) in a quadratic form as follows (Lang 1987):

$$(x\ z)\begin{pmatrix} A & \frac{B}{2} \\ \frac{B}{2} & C \end{pmatrix}\begin{pmatrix} x \\ z \end{pmatrix} + (x\ z)\begin{pmatrix} D \\ E \end{pmatrix} + F = 0 \tag{107}$$

or

$$(x\ z)U\begin{pmatrix} x \\ z \end{pmatrix} + (x\ z)V + F = 0 \tag{108}$$

Since U is symmetric, the slopes of the major and the minor axes of the ellipse (Eq. (99)) can be found from the two orthogonal eigenvectors for matrix U (Lang 1987). Using the slopes obtained above, we can find the angles between the major or the minor axis of the ellipse and the X-axis or the Z-axis.

The eigenvalue λ for matrix U can be found from the following equation (Lang 1987):

$$\begin{vmatrix} \lambda - A & -\frac{B}{2} \\ -\frac{B}{2} & \lambda - C \end{vmatrix} = \lambda^2 - (A+C)\lambda + AC - \frac{B^2}{4} = 0 \tag{109}$$

Solving for λ we have:

$$\begin{aligned} \lambda_1 &= \frac{A+C+\sqrt{A^2-2AC+C^2+B^2}}{2} \\ \lambda_2 &= \frac{A+C-\sqrt{A^2-2AC+C^2+B^2}}{2} \end{aligned} \tag{110}$$

The eigenvectors can be found from the following equation (Lang 1987):

$$\begin{pmatrix} A & \frac{B}{2} \\ \frac{B}{2} & Cy \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix} \tag{111}$$

Let x=1, then y=2(λ-A)/B. We obtain the following eigenvector:

$$X(\lambda) = \begin{pmatrix} 1 \\ \frac{2(\lambda - A)}{B} \end{pmatrix} \tag{112}$$

There are two such eigenvectors for $\lambda=\lambda_1$ and $\lambda=\lambda_2$. From the two eigenvectors we can find the slopes for the major and the minor axes of the ellipse. Let α be the angle between the major axis of the ellipse and X axis of the image plane, we have:

$$\alpha = \arctan\left(\frac{2(\lambda-A)}{B}\right) = \arctan\left(\frac{C-A\pm\sqrt{A^2-2AC+C^2+B^2}}{B}\right) \tag{113}$$

APPENDIX B

DERIVATION OF CALIBRATION ALGORITHM

In this study, the seven calibration parameters: (ϕ_c, θ_c, ψ_c, x_c, z_c, **d**, and **r**) were computed using an algorithm which related the seven parameters to the eccentricity and the center of the projected iris/pupil contour at different eye orientation. We derive the algorithm for computing these seven parameters below:

Let (x", -d, z") and (x, y, z) be the coordinates of an iral edge point when eye is looking straight at the camera and when eye is rotated to the current position (Figure 10). Using Eq. (37), we have:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = R_{cal}\, R_{eye} \begin{pmatrix} x'' \\ -d \\ z'' \end{pmatrix} \tag{114}$$

Since the torsional movement about the optic axis does not change the eccentricity and the orientations of the major and the minor axes of the ellipse projected onto the image plane, the torsion is ignored when the transformation of the iral contour projected onto the image plane is considered. We will assume $\psi=0$. R_{cal} and the simplified R_{eye} are given below:

$$R_{cal} = \begin{pmatrix} \cos\phi_c\cos\psi_c - \sin\phi_c\sin\theta_c\sin\psi_c & -\sin\phi_c\cos\theta_c & \cos\phi_c\sin\psi_c + \sin\phi_c\sin\theta_c\cos\psi_c \\ \sin\phi_c\cos\psi_c + \cos\phi_c\sin\theta_c\sin\psi_c & \cos\phi_c\cos\theta_c & \sin\phi_c\sin\psi_c - \cos\phi_c\sin\theta_c\cos\psi_c \\ -\cos\theta_c\sin\psi_c & \sin\theta_c & \cos\theta_c\cos\psi_c \end{pmatrix} \quad (115)$$

$$R_{eye} = \begin{pmatrix} \cos\phi & -\sin\phi\cos\theta & \sin\phi\sin\theta \\ \sin\phi & \cos\phi\cos\theta & -\cos\phi\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \quad (116)$$

Substitute Eq. (116) and Eq. (114) into Eq. (114), we have:

$$
\begin{aligned}
x &= (\cos\phi\cos\phi_c\cos\psi_c - \cos\phi\sin\phi_c\sin\theta_c\sin\psi_c - \sin\phi\sin\phi_c\cos\theta_c)x'' \\
&\quad + (\sin\phi\cos\theta\cos\phi_c\cos\psi_c - \sin\phi\cos\theta\sin\phi_c\sin\theta_c\sin\psi_c \\
&\quad + \cos\phi\cos\theta\sin\phi_c\cos\theta_c - \sin\theta\cos\phi_c\sin\psi_c \\
&\quad - \sin\theta\sin\phi_c\sin\theta_c\cos\psi_c)d + (\sin\phi\sin\theta\cos\phi_c\cos\psi_c \\
&\quad - \sin\phi\sin\theta\sin\phi_c\sin\theta_c\sin\psi_c + \cos\phi\sin\theta\sin\phi_c\cos\theta_c \\
&\quad + \cos\theta\cos\phi_c\sin\psi_c + \cos\theta\sin\phi_c\sin\theta_c\cos\psi_c)z'' \\
y &= (\cos\phi\sin\phi_c\cos\psi_c + \cos\phi\cos\phi_c\sin\theta_c\sin\psi_c + \sin\phi\cos\phi_c\cos\theta_c)x'' \\
&\quad + (\sin\phi\cos\theta\sin\phi_c\cos\psi_c + \sin\phi\cos\theta\cos\phi_c\sin\theta_c\sin\phi_c \\
&\quad - \cos\phi\cos\theta\cos\phi_c\cos\theta_c - \sin\theta\sin\phi_c\sin\psi_c \\
&\quad + \sin\theta\cos\phi_c\sin\theta_c\cos\psi_c)d + (\sin\phi\sin\theta\sin\phi_c\cos\psi_c \\
&\quad + \sin\phi\sin\theta\cos\phi_c\sin\theta_c\sin\psi_c - \cos\phi\sin\theta\cos\phi_c\cos\theta_c \\
&\quad + \cos\theta\sin\phi_c\sin\psi_c - \cos\theta\cos\phi_c\sin\theta_c\cos\psi_c)z'' \\
z &= (\sin\phi\sin\theta_c - \cos\phi\cos\theta_c\sin\psi_c)x'' - (\sin\phi\cos\theta\cos\theta_c\sin\psi_c \\
&\quad + \cos\phi\cos\theta\sin\theta_c + \sin\theta\cos\theta_c\cos\psi_c)d \\
&\quad + (\cos\theta\cos\theta_c\cos\psi_c - \sin\phi\sin\theta\cos\theta_c\sin\psi_c - \cos\phi\sin\theta\sin\theta_c)z''
\end{aligned}
\tag{117}
$$

Assume (x_p, z_p) is the projection of (x, y, z) onto the image plane and (x_c, z_c) is the projection of the center of the camera-coordinate-frame onto the image plane. We have:

$$
\begin{pmatrix} x_p \\ z_p \end{pmatrix} = \begin{pmatrix} x \\ z \end{pmatrix} + \begin{pmatrix} x_c \\ z_c \end{pmatrix} \tag{118}
$$

Using Eq. (117), we have:

$$
\begin{aligned}
x_p &= (\cos\phi\cos\phi_c\cos\psi_c - \cos\phi\sin\phi_c\sin\theta_c\sin\psi_c - \sin\phi\sin\phi_c\cos\theta_c)x'' \\
&\quad + (\sin\phi\cos\theta\cos\phi_c\cos\psi_c - \sin\phi\cos\theta\sin\phi_c\sin\theta_c\sin\psi_c \\
&\quad + \cos\phi\cos\theta\sin\phi_c\cos\theta_c - \sin\theta\cos\phi_c\sin\psi_c \\
&\quad - \sin\theta\sin\phi_c\sin\theta_c\cos\psi_c)d + (\sin\phi\sin\theta\cos\phi_c\cos\psi_c \\
&\quad - \sin\phi\sin\theta\sin\phi_c\sin\theta_c\sin\psi_c + \cos\phi\sin\theta\sin\phi_c\cos\theta_c \\
&\quad + \cos\theta\cos\phi_c\sin\psi_c + \cos\theta\sin\phi_c\sin\theta_c\cos\psi_c)z'' + x_c \\
z_p &= (\sin\phi\sin\theta_c - \cos\phi\cos\theta_c\sin\psi_c)x'' - (\sin\phi\cos\theta\cos\theta_c\sin\psi_c \\
&\quad + \cos\phi\cos\theta\sin\theta_c + \sin\theta\cos\theta_c\cos\psi_c)d + (\cos\theta\cos\theta_c\cos\psi_c \\
&\quad - \sin\phi\sin\theta\cos\theta_c\sin\psi_c - \cos\phi\sin\theta\sin\theta_c)z'' + z_c
\end{aligned}
\tag{119}
$$

The above equation is equivalent to the following:

$$
\begin{aligned}
x_p &= a_1x'' + b_1d + c_1z'' + x_c \\
z_p &= a_2x'' + b_2d + c_2z'' + z_c
\end{aligned}
\tag{120}
$$

where

$$
\begin{aligned}
a_1 &= \cos\phi\cos\phi_c\cos\psi_c - \cos\phi\sin\phi_c\sin\theta_c\sin\psi_c - \sin\phi\sin\phi_c\cos\theta_c \\
b_1 &= \sin\phi\cos\theta\cos\phi_c\cos\psi_c - \sin\phi\cos\theta\sin\phi_c\sin\theta_c\sin\psi_c \\
&\quad + \cos\phi\cos\theta\sin\phi_c\cos\theta_c - \sin\theta\cos\phi_c\sin\psi_c \\
&\quad - \sin\theta\sin\phi_c\sin\theta_c\cos\psi_c \\
c_1 &= \sin\phi\sin\theta\cos\phi_c\cos\psi_c - \sin\phi\sin\theta\sin\phi_c\sin\theta_c\sin\psi_c \\
&\quad + \cos\phi\sin\theta\sin\phi_c\cos\theta_c + \cos\theta\cos\phi_c\sin\psi_c \\
&\quad + \cos\theta\sin\phi_c\sin\theta_c\cos\psi_c \\
a_2 &= \sin\phi\sin\theta_c - \cos\phi\cos\theta_c\sin\psi_c \\
b_2 &= -(\sin\phi\cos\theta\cos\theta_c\sin\psi_c + \cos\phi\cos\theta\sin\theta_c + \sin\theta\cos\theta_c\cos\psi_c) \\
c_2 &= \cos\theta\cos\theta_c\cos\psi_c - \sin\phi\sin\theta\cos\theta_c\sin\psi_c - \cos\phi\sin\theta\sin\theta_c
\end{aligned}
\tag{121}
$$

Solving for x" and z", we have:

$$
\begin{aligned}
x'' &= \frac{c_2x_p - c_1z_p + (b_2c_1 - b_1c_2)d - c_2x_c + c_1z_c}{a_1c_2 - a_2c_1} \\
z'' &= \frac{a_2x_p - a_1z_p + (a_1b_2 - a_2b_1)d + a_1z_c - a_2x_c}{a_2c_1 - a_1c_2}
\end{aligned}
\tag{122}
$$

Since the projection of the iral contour onto the image plane is a perfect circle when eye is looking straight at the camera, we have:

$$
x''^2 + z''^2 = r^2 \tag{123}
$$

Substituting Eq. (57), (122) into Eq. (58), (123), we have:

$$\left(\frac{c_2x_p - c_1z_p + (b_2c_1 - b_1c_2)d - c_2x_c + c_1z_c}{a_1c_2 - a_2c_1}\right)^2 +$$
$$\left(\frac{a_2x_p - a_1z_p + (a_1b_2 - a_2b_1)d + a_1z_c - a_2x_c}{a_2c_1 - a_1c_2}\right)^2 = r^2 \quad (124)$$

The above equation is equivalent to the following:

$$x_p^2 = Az_p^2 + Bx_pz_p + Cx_p + Dz_p + E \quad (125)$$

where

$$A = -\frac{a_1^2 + c_1^2}{a_2^2 + c_2^2}$$
$$B = \frac{2(a_1a_2 + c_1c_2)}{a_2^2 + c_2^2}$$
$$C = 2x_c - Bz_c + (2b_1 - Bb_2)d$$
$$D = -Bx_c - 2Az_c - (2Ab_2 + Bb_1)d \quad (126)$$
$$E = x_c^2 - Az_c^2 - Bx_cz_c - Cx_c - Dz_c + \frac{(a_2c_1 - a_1c_2)^2r^2}{a_2^2 + c_2^2}$$
$$- \frac{[(a_1b_2 - a_2b_1)^2 + (b_2c_1 - b_1c_2)^2]d^2}{a_2^2 + c_2^2}$$

Solving x_c, z_c and r from the above equation, we have:

$$x_c = \frac{2AC - BD}{4A + B^2} - b_1 d$$

$$z_c = -\frac{BC + 2D}{4A + B^2} - b_2 d \qquad (127)$$

$$r = \frac{\sqrt{(a_2^2 + c_2^2)(Az_c^2 + Bx_c z_c + Cx_c + Dz_c + E - x_c^2) + [(a_1 b_2 - a_2 b_1)^2 + (b_2 c_1 - b_1 c_2)^2]d^2}}{|a_2 c_1 - a_1 c_2|}$$

Using Eq. (121), we have:

$$
\begin{aligned}
a_2^2+c_2^2 = {} & 1-\sin^2\theta\cos^2\theta_c\cos^2\psi_c-\cos^2\phi\cos^2\theta\sin^2\theta_c \\
& -\sin^2\phi\cos^2\theta\cos^2\theta_c\sin^2\psi_c-2\sin\phi\cos\phi\sin\theta_c\cos\theta_c\sin\psi_c \\
& -2\sin\phi\sin\theta\cos\theta\cos^2\theta_c\sin\psi_c\cos\psi_c \\
& -2\cos\phi\sin\theta\cos\theta\sin\theta_c\cos\theta_c\cos\psi_c \\
& +2\sin\phi\cos\phi\sin^2\theta\sin\theta_c\cos\theta_c\sin\psi_c \\
a_1^2+c_1^2 = {} & 1-\sin^2\theta\cos^2\phi_c\sin^2\psi_c-\sin^2\theta\sin^2\phi_c\sin^2\theta_c\cos^2\psi_c \\
& -\sin^2\phi\cos^2\theta\cos^2\phi_c\cos^2\psi_c-\cos^2\phi\cos^2\theta\sin^2\phi_c\cos^2\theta_c \\
& -\sin^2\phi\cos^2\theta\sin^2\phi_c\sin^2\theta_c\sin\psi_c \\
& -2\sin^2\theta\sin\phi_c\cos\phi_c\sin\theta_c\sin\psi_c\cos\psi_c \\
& +2\sin^2\phi\cos^2\theta\sin\phi_c\cos\phi_c\sin\theta_c\sin\psi_c\cos\psi_c \\
& -2\sin\phi\cos\phi\cos^2\theta\sin\phi_c\cos\phi_c\cos\theta_c\cos\psi_c \\
& +2\sin\phi\cos\phi\cos^2\theta\sin^2\phi_c\sin\theta_c\cos\theta_c\sin\psi_c \\
& +2\sin\phi\sin\theta\cos\theta\cos^2\phi_c\sin\psi_c\cos\psi_c \\
& +2\sin\phi\sin\theta\cos\theta\sin^2\phi_c\cos\phi_c\sin\theta_c\cos^2\psi_c \\
& -2\sin\phi\sin\theta\cos\theta\sin\phi_c\cos\phi_c\sin\theta_c\sin^2\psi_c \\
& -2\sin\phi\sin\theta\cos\theta\sin^2\phi_c\sin^2\theta_c\sin\psi_c\cos\psi_c \\
& +2\cos\phi\sin\theta\cos\theta\sin\phi_c\cos\phi_c\cos\theta_c\sin\psi_c \\
& +2\cos\phi\sin\theta\cos\theta\sin^2\psi_c\sin\theta_c\cos\theta_c\cos\psi_c
\end{aligned}
\tag{128}
$$

$$
\begin{aligned}
a_1a_2+c_1c_2 = {} & \sin\phi\cos\phi\cos^2\theta\cos\phi_c\sin\theta_c\cos\psi_c \\
& -\sin\phi\cos\phi\cos^2\theta\sin\phi_c\sin^2\theta_c\sin\psi_c \\
& -\sin^2\theta\cos\phi_c\cos\theta_c\sin\psi_c\cos\psi_c \\
& +\sin\phi\sin\theta\cos\theta\cos\phi_c\cos\theta_c\cos^2\psi_c \\
& +\cos^2\phi\cos^2\theta\sin\phi_c\sin\theta_c\cos\theta_c \\
& +\sin^2\phi\cos^2\theta\cos\phi_c\cos\theta_c\sin\psi_c\cos\psi_c \\
& -\sin^2\phi\cos^2\theta\sin\phi_c\sin\theta_c\cos\theta_c\sin^2\psi_c \\
& +\sin\phi\cos\phi\cos^2\theta\sin\phi_c\cos^2\theta_c\sin\psi_c \\
& -\sin\phi\sin\theta\cos\theta\cos\phi_c\cos\theta_c\sin^2\psi_c \\
& -\cos\phi\sin\theta\cos\theta\cos\phi_c\sin\theta_c\sin\psi_c \\
& -\cos\phi\sin\theta\cos\theta\sin\phi_c\sin^2\theta_c\cos\psi_c \\
& -\sin^2\theta\sin\phi_c\sin\theta_c\cos\theta_c\cos^2\psi_c
\end{aligned}
\tag{129}
$$

To simplify the derivation for the formulas for computing calibration parameters, we will allow eye positions with only pure horizontal or vertical eye rotations. Let ϕ and θ be the angles of the horizontal and the vertical eye rotations relative to the reference position. Three eye positions are used to compute the calibration parameters: (ϕ=0, θ=0), (ϕ=0, $\theta=\theta_0$) and ($\phi=\phi_0$, θ=0).

When ϕ=0 and θ=0, we have:

$$
\begin{aligned}
A &= -\cos^2\phi_c - \tan^2\theta_c \\
B &= 2\sin\phi_c \tan\theta_c \\
C &= 2\left(x_c + \frac{d\sin\phi_c}{\cos\theta_c} - z_c \sin\phi_c \tan\theta_c\right) \\
D &= -2\left(\frac{d\tan\theta_c}{\cos\theta_c} + x_c \sin\phi_c \tan\theta_c - z_c(\cos^2\phi_c + \tan^2\theta_c)\right) \\
E &= -d^2(\sin^2\phi_c + \tan^2\theta_c) + \frac{2dz_c\tan\theta_c - 2dx_c\sin\phi_c}{\cos\theta_c} - \\
&\quad (x_c - z_c\sin\phi_c\tan\theta_c)^2 - \left(\frac{z_c\cos\phi_c}{\cos\theta_c}\right)^2 + r^2\cos^2\phi_c
\end{aligned}
\tag{130}
$$

Solving for ϕ_c, θ_c, x_c, z_c and r from Eq. (130), we have:

$$
\begin{aligned}
\phi_c &= \pm\arcsin\left(\frac{\sqrt{A+1\pm\sqrt{(A+1)^2+B^2}}}{2}\right) \\
\theta_c &= \pm\arctan\left(\sqrt{\frac{-A-1\pm\sqrt{(A+1)^2+B^2}}{2}}\right) \\
x_c &= \frac{2AC-BD}{4A+B^2} - d\sin\phi_c\cos\theta_c \\
z_c &= -\frac{BC+2D}{4A+B^2} + d\sin\theta_c \\
r &= \sqrt{\frac{\frac{C^2}{4}+E}{\cos^2\phi_c} + \frac{(d\sin\theta_c - z_c)^2}{\cos^2\theta_c}}
\end{aligned}
\tag{131}
$$

If we let $\phi=0$ and $\theta=\theta_0$, i.e the eye has only vertical movement, we have:

$$
\begin{aligned}
a_1^2+c_1^2 &= 1-(\sin\theta_0\cos\phi_c\sin\psi_c+\sin\theta_0\sin\phi_c\sin\theta_c\cos\psi_c-\cos\theta_0\sin\phi_c\cos\theta_c)^2 \\
a_2^2+c_2^2 &= 1-(\sin\theta_0\cos\theta_c\cos\psi_c+\cos\theta_0\sin\theta_c)^2 \\
a_1a_2+c_1c_2 &= -(\sin\theta_0\cos\theta_c\cos\psi_c+\cos\theta_0\sin\theta_c)* \\
&\quad (\sin\theta_0\cos\phi_c\sin\psi_c+\sin\theta_0\sin\phi_c\sin\theta_c\cos\psi_c-\cos\theta_0\sin\phi_c\cos\theta_c)
\end{aligned}
\tag{132}
$$

If we allow:

$$
\begin{aligned}
u &= \sin\theta_0\cos\theta_c\cos\psi_c+\cos\theta_0\sin\theta_c \\
&= \sqrt{\cos^2\theta_0+(\sin\theta_0\cos\psi_c)^2}\sin(\theta_c+\alpha) \\
v &= \sin\theta_0\cos\phi_c\sin\psi_c+\sin\theta_0\sin\phi_c\sin\theta_c\cos\psi_c-\cos\theta_0\sin\phi_c\cos\theta_c \\
&= -\sqrt{(\cos\theta_0\cos\theta_c-\sin\theta_0\sin\theta_c\cos\psi_c)^2+(\sin\theta_0\sin\psi_c)^2}\sin(\phi_c-\beta)
\end{aligned}
\tag{133}
$$

where

$$
\begin{aligned}
\alpha &= \arctan\left(\frac{\sin\theta_0\cos\psi_c}{\cos\theta_0}\right) = \arctan(\tan\theta_0\cos\psi_c) \\
\beta &= \arctan\left(\frac{\sin\theta_0\sin\psi_c}{\cos\theta_0\cos\theta_c-\sin\theta_0\sin\theta_c\cos\psi_c}\right)
\end{aligned}
\tag{134}
$$

then

$$\frac{v^2-1}{1-u^2} = A$$
$$\frac{2uv}{u^2-1} = B \tag{135}$$

Solving u and v from the above equation, we have:

$$u = \pm\sqrt{\frac{2(A+1)+B^2 \pm 2\sqrt{(A+1)^2+B^2}}{4A+B^2}}$$
$$v = \pm\sqrt{\frac{2A(A+1)+B^2 \pm 2A\sqrt{(A+1)^2+B^2}}{4A+B^2}} \tag{136}$$

Solving for ϕ_c, θ_c, x_c, z_c and r using Eq. (133) and Eq. (127), we have:

$$\theta_c = \pm\arcsin\left(\sqrt{\frac{2(A+1)+B^2\pm2\sqrt{(A+1)^2+B^2}}{(4A+B^2)(1-\sin^2\theta_0\sin^2\psi_c)}}\right)-\arctan(\tan\theta_0\cos\psi_c)$$

$$\phi_c = \pm\arcsin$$

$$\left(\sqrt{\frac{2A(A+1)+B^2\pm2A\sqrt{(A+1)^2+B^2}}{(4A+B^2)[(\cos\theta_0\cos\theta_c-\sin\theta_0\sin\theta_c\cos\psi_c)^2+(\sin\theta_0\sin\psi_c)^2]}}\right)$$

$$+\arctan\left(\frac{\sin\theta_0\sin\psi_c}{\cos\theta_0\cos\theta_c-\sin\theta_0\sin\theta_c\cos\psi_c}\right) \qquad (137)$$

$$x_c = \frac{2AC-DB}{4A+B^2}-b_1d$$

$$z_c = -\frac{2D+BC}{4A+B^2}-b_2d$$

$$r = \frac{\sqrt{(a_2^2+c_2^2)(Az_c^2+Bx_cz_c+Cx_c+Dz_c+E-x_c^2)+[(a_1b_2-a_2b_1)^2+(b_2c_1-b_1c_2)^2]d^2}}{|a_2c_1-a_1c_2|}$$

where

$$a_1 = \cos\phi_c\cos\psi_c-\sin\phi_c\sin\theta_c\sin\psi_c$$

$$b_1 = \cos\theta_0\sin\phi_c\cos\theta_c-\sin\theta_0\cos\phi_c\sin\psi_c-\sin\theta_0\sin\phi_c\sin\theta_c\cos\psi_c$$

$$c_1 = \sin\theta_0\sin\phi_c\cos\theta_c+\cos\theta_0\cos\phi_c\sin\psi_c+\cos\theta_0\sin\phi_c\sin\theta_c\cos\psi_c$$

$$a_2 = -\cos\theta_c\sin\psi_c \qquad (138)$$

$$b_2 = -(\cos\theta_0\sin\theta_c+\sin\theta_0\cos\theta_c\cos\psi_c)$$

$$c_2 = \cos\theta_0\cos\theta_c\cos\psi_c-\sin\theta_0\sin\theta_c$$

If we allow $\phi=\phi_0$ and $\theta=0$, i.e. the eye has only horizontal movement, we have:

$$
\begin{aligned}
a_1^2+c_1^2 &= 1-(\sin\phi_0\cos\phi_c\cos\psi_c+\cos\phi_0\sin\phi_c\cos\theta_c-\sin\phi_0\sin\phi_c\sin\theta_c\sin\psi_c)^2 \\
a_2^2+c_2^2 &= 1-(\cos\phi_0\sin\theta_c+\sin\phi_0\cos\theta_c\sin\psi_c)^2 \\
a_1a_2+c_1c_2 &= (\cos\phi_0\sin\theta_c+\sin\phi_0\cos\theta_c\sin\psi_c)* \\
&\quad (\sin\phi_0\cos\phi_c\cos\psi_c+\cos\phi_0\sin\phi_c\cos\theta_c-\sin\phi_0\sin\phi_c\sin\theta_c\sin\psi_c)
\end{aligned}
\tag{139}
$$

If we let:

$$
\begin{aligned}
u &= \cos\phi_0\sin\theta_c+\sin\phi_0\cos\theta_c\sin\psi_c \\
&= \sqrt{\cos^2\phi_0+\sin^2\phi_0\sin^2\psi_c}\sin(\theta_c+\alpha) \\
v &= \sin\phi_0\cos\phi_c\cos\psi_c+\cos\phi_0\sin\phi_c\cos\theta_c-\sin\phi_0\sin\phi_c\sin\theta_c\sin\psi_c \\
&= \sqrt{(\cos\phi_0\cos\theta_c-\sin\phi_0\sin\theta_c\sin\psi_c)^2+(\sin\phi_0\cos\psi_c)^2}\sin(\phi_c+\beta)
\end{aligned}
\tag{140}
$$

where

$$
\begin{aligned}
\alpha &= \arctan\left(\frac{\sin\phi_0\sin\psi_c}{\cos\phi_0}\right) = \arctan(\tan\phi_0\sin\psi_c) \\
\beta &= \arctan\left(\frac{\sin\phi_0\cos\psi_c}{\cos\phi_0\cos\theta_c-\sin\phi_0\sin\theta_c\sin\psi_c}\right)
\end{aligned}
\tag{141}
$$

Eq. (139) is equivalent to the following:

$$\frac{v^2-1}{1-u^2} = A$$
$$\frac{2uv}{1-u^2} = B \tag{142}$$

Solving u and v from the above equation, we have:

$$u = \pm\sqrt{\frac{2(A+1)+B^2 \pm 2\sqrt{(A+1)^2+B^2}}{4A+B^2}}$$
$$v = \pm\sqrt{\frac{2A(A+1)+B^2 \pm 2A\sqrt{(A+1)^2+B^2}}{4A+B^2}} \tag{143}$$

Solving ϕ_c, θ_c, x_c, z_c and r using Eq. (143), Eq. (140), Eq. (141) and Eq. (127), we have:

$$\theta_c = \pm\arcsin\left(\sqrt{\frac{2(A+1)+B^2\pm 2\sqrt{(A+1)^2+B^2}}{(4A+B^2)(1-\sin^2\phi_0\cos^2\psi_c)}}\right)-\arctan(\tan\phi_0\sin\psi_c)$$

$$\phi_c = \pm\arcsin$$

$$\left(\sqrt{\frac{2A(A+1)+B^2\pm 2A\sqrt{(A+1)^2+B^2}}{(4A+B^2)[(\cos\phi_0\cos\theta_c-\sin\phi_0\sin\theta_c\sin\psi_c)^2+(\sin\phi_0\cos\psi_c)^2]}}\right)$$

$$-\arctan\left(\frac{\sin\phi_0\cos\psi_c}{\cos\phi_0\cos\theta_c-\sin\phi_0\sin\theta_c\sin\psi_c}\right) \qquad (144)$$

$$x_c = \frac{2AC-BD}{4A+B^2}-b_1 d$$

$$z_c = -\frac{BC+2D}{4A+B^2}-b_2 d$$

$$r = \frac{\sqrt{(a_2^2+c_2^2)(Az_c^2+Bx_cz_c+Cx_c+Dz_c+E-x_c^2)+[(a_1b_2-a_2b_1)^2+(b_2c_1-b_1c_2)^2]d^2}}{|a_2c_1-a_1c_2|}$$

where

$$\begin{aligned}
a_1 &= \cos\phi_0\cos\phi_c\cos\psi_c-\cos\phi_0\sin\phi_c\sin\theta_c\sin\psi_c-\sin\phi_0\sin\phi_c\cos\theta_c\\
b_1 &= \sin\phi_0\cos\phi_c\cos\psi_c-\sin\phi_0\sin\phi_c\sin\theta_c\sin\psi_c+\cos\phi_0\sin\phi_c\cos\theta_c\\
c_1 &= \cos\phi_c\sin\psi_c+\sin\phi_c\sin\theta_c\cos\psi_c\\
a_2 &= \sin\phi_0\sin\theta_c-\cos\phi_0\cos\theta_c\sin\psi_c\\
b_2 &= -(\sin\phi_0\cos\theta_c\sin\psi_c+\cos\phi_0\sin\theta_c)\\
c_2 &= \cos\theta_c\cos\psi_c
\end{aligned} \qquad (145)$$

APPENDIX C

MODELLING IRIS AND PUPIL AS SPHERE

In this study, the iris/pupil were model as a plane which intersects the eye considered as a sphere (Figure 1). During the course of this work, we have also considered the spherical iris/pupil model (Figure 38). The effect of using the spherical iris/pupil model to the developed algorithm for computing the eye orientation is discussed below:

If the iris/pupil are modelled as a sphere, Y coordinate (-d') for any point (x, -d', z) inside the iris or pupil, when the eye is in the reference position, is not a constant (-d) (Figure 38). It is given by the following equation:

$$d' = \sqrt{r_{eye}^2 - x^2 - z^2} = \sqrt{r_{iris}^2 + d^2 - x^2 - z^2} \qquad (146)$$

where r_{eye} is the radius of the eye, r_{iris} is the radius of the iris and d is the distance between the center of eye and the iral plane (Figure 38). r_{iris} and d are determined during calibration.

If the iris and pupil are modelled as a sphere, the formula for torsion calculation using planar iris/pupil model (Eq. (79)) needs to be changed. This is done by replacing the coordinates (x, -d, z) in Eq.(79) by (x, -d', z). The formulae for computing the horizontal and vertical rotations are not affected by the spherical iris/pupil model, because the algorithm for computing the horizontal and vertical eye rotations used only the boundary points, where d equals d' (Figure 38).

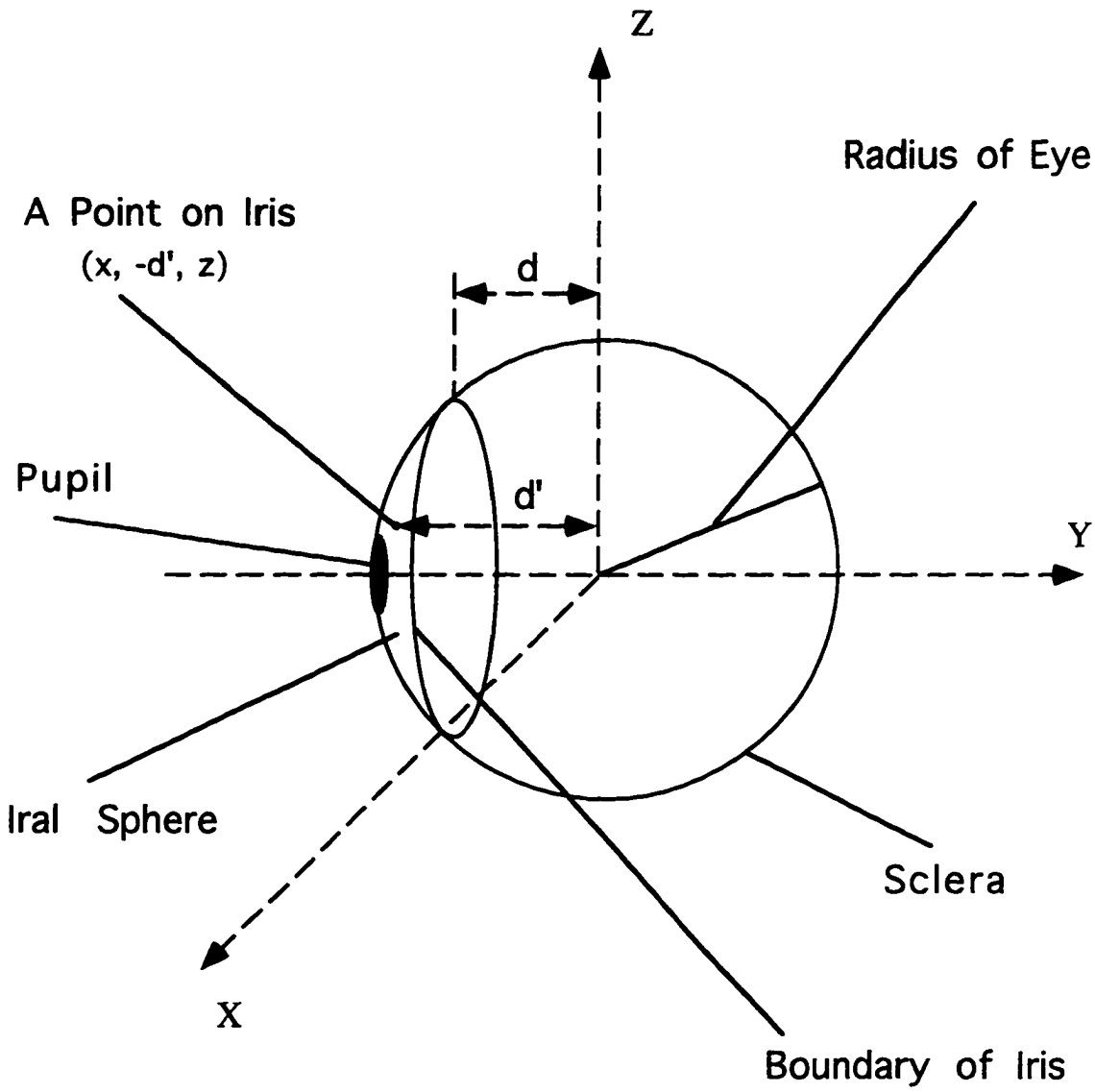
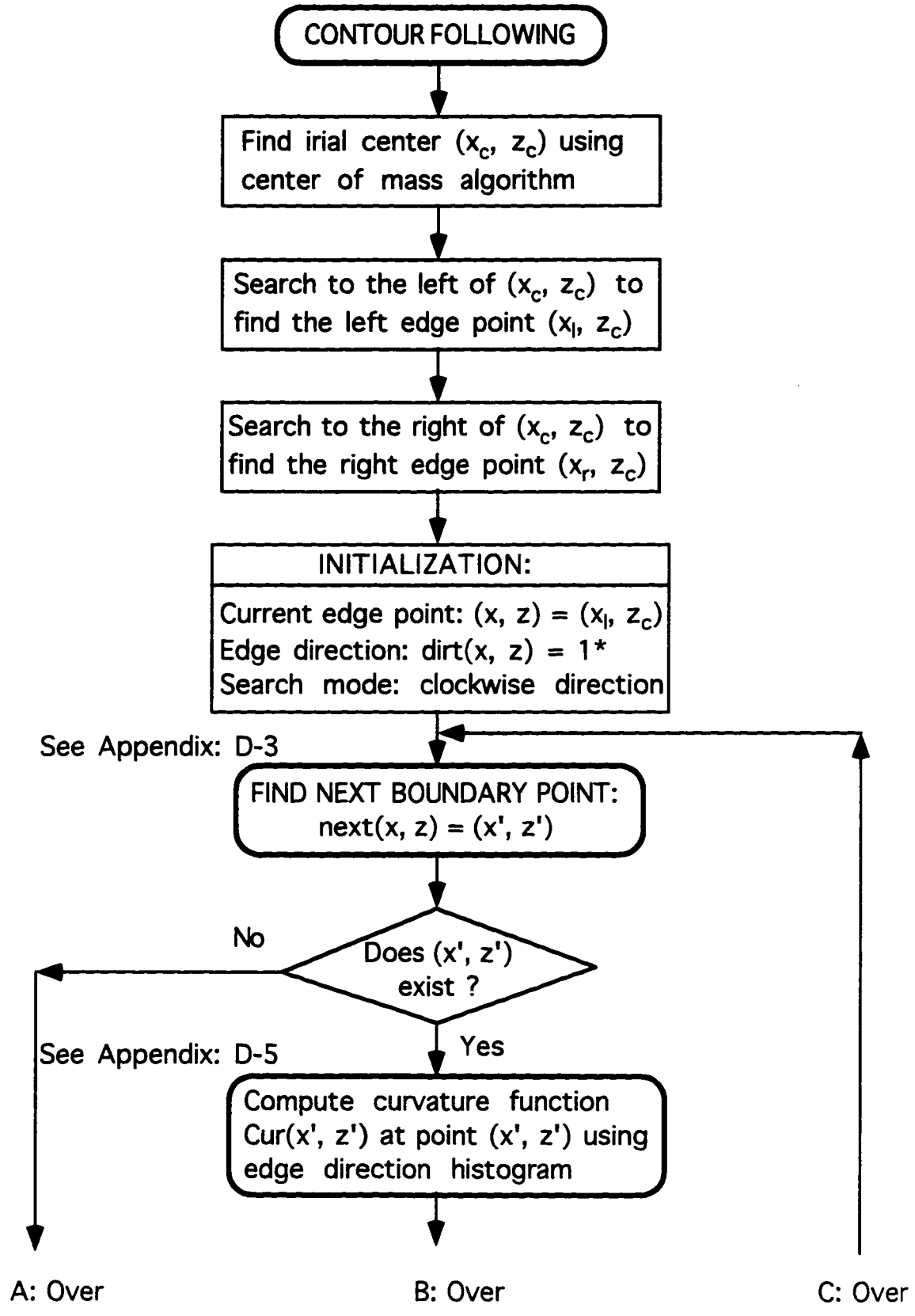


Figure 38: Spherical Iris Model

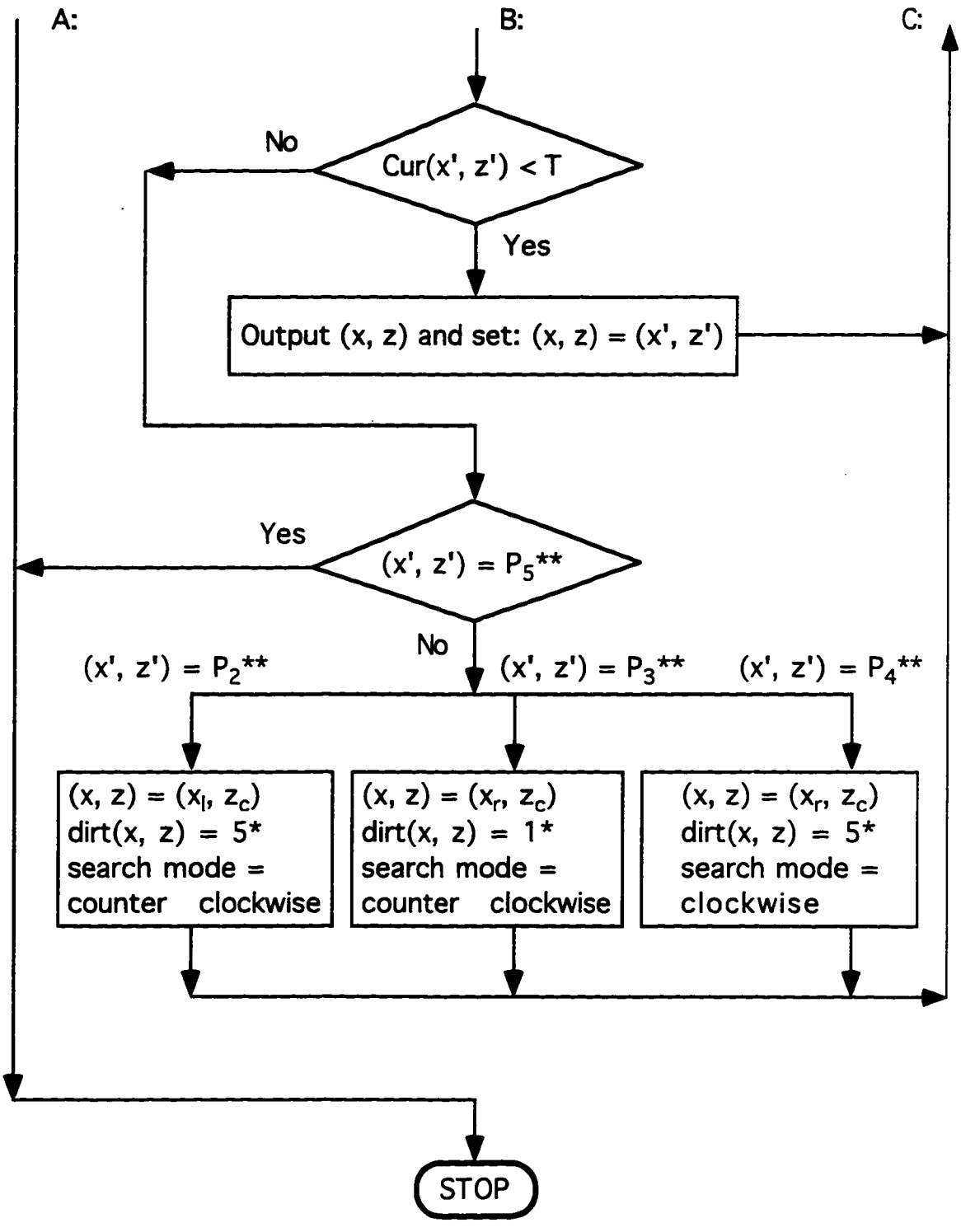
APPENDIX D

FLOWCHART FOR CONTOUR FOLLOWING ALGORITHM

The flowchart for the contour following algorithm is shown in the following pages:

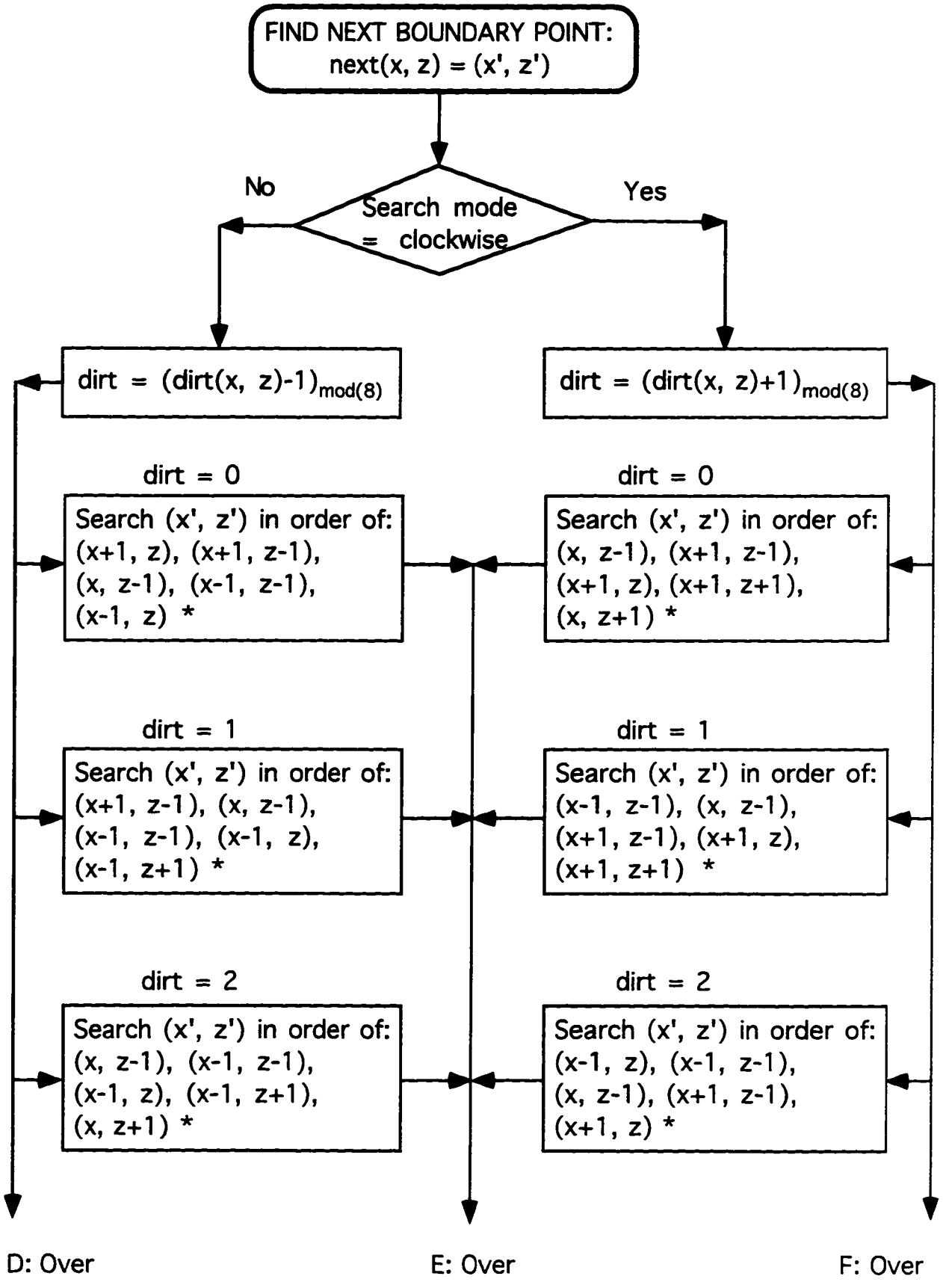


Appendix: D-1

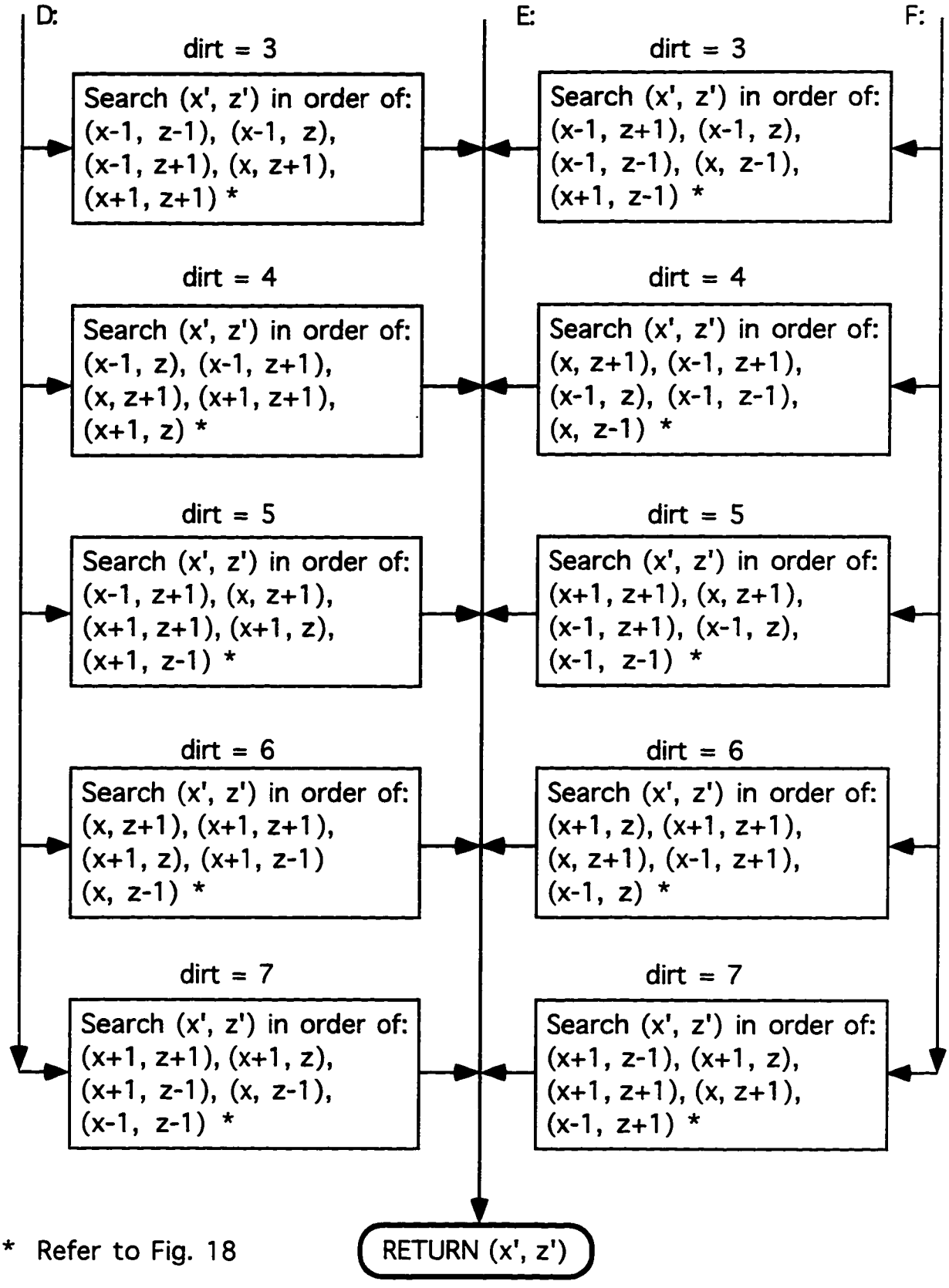


* Refer to Fig. 18
** Refer to Fig. 20

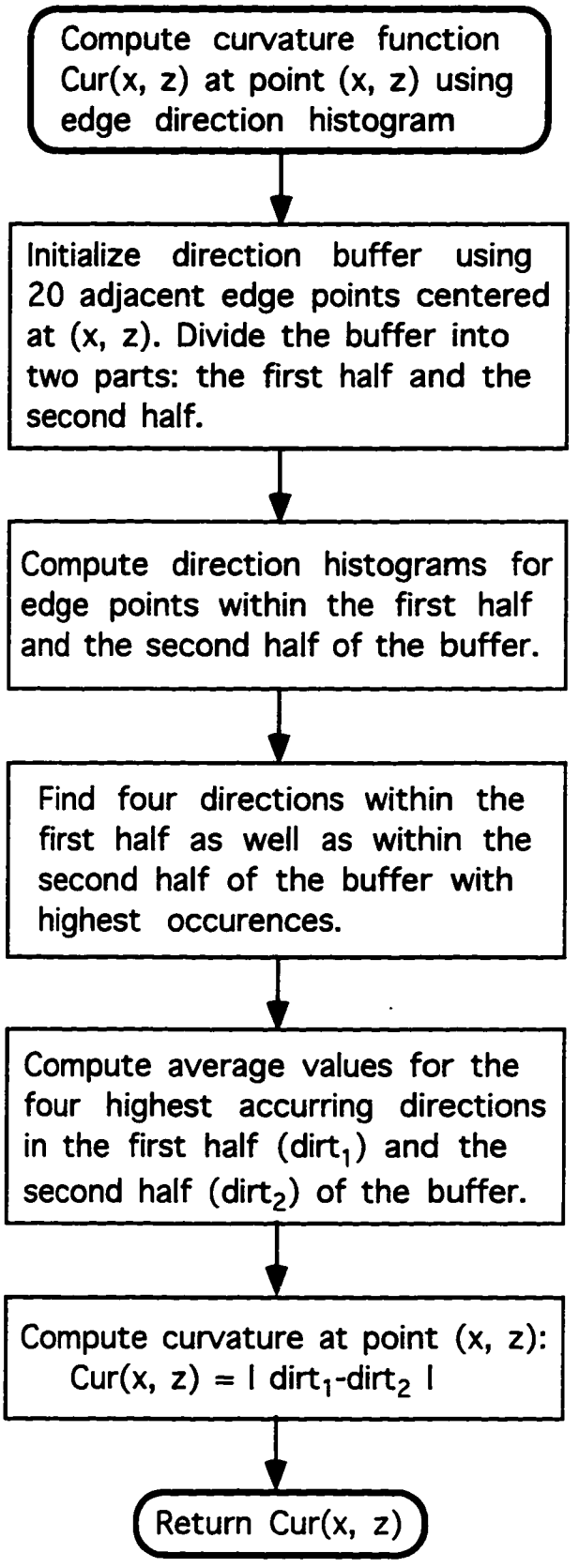
Appendix: D-2



Appendix: D-3



Appendix: D-4



Appendix: D-5

APPENDIX E

PROCESSING SPEED FOR EACH PROCEDURE

The execution time required to process one image for each of the procedures (Figure 15 and Figure 16) running on a 66 MHz 486-DX2 computer is listed below:

Digitization:	1.05 sec.
Segmentation:	0.11 sec.
median filtering:	1.48 sec.
Contour Following:	0.05 sec.
Ellipse Fitting:	0.22 sec.
Signature Sampling:	0.005 sec.
Interpolation:	0.005 sec.
Template Matching:	0.29 sec.

The total time needed to process an image is about 3.21 sec. The median filtering procedure can be ignored with little effect on the results. Thus the processing speed without median filtering is approximately 1.73 sec.

BIBLIOGRAPHY

Ballard, D. H. and C. M. Brown (1982). Computer Vision. Englewood Cliffs, New Jersey, Prentice-Hall, Inc.

Balliet, R. and K. Nakayama (1978). "Training of voluntary torsion." Invest. Ophthalmol. Visual Sci. **4**: 303-314.

Bennett, A. G. and R. B. Rabbetts (1984). Clinical Visual Optics. London, Butterworth.

Buswell, G. (1935). How people look at pictures. Chicago, University of Chicago Press.

Carmichael, L. and W. F. Dearborn (1947). Reading and visual fatigue. Boston, Houghton Mifflin.

Clarke, A. H., W. Teiwes and H. Scherer (1991). "Video-oculography - An Alternative Method for Measurement of Three-dimensional Eye Movements." Oculomotor Control and Cognitive Processes: 431-443.

Collewijn, H., L. Ferman and A. V. Van den Berg (1988). "The behavior of human gaze in three dimensions." Ann. N.Y. Acad. Sci. **545**: 105-127.

Collewijn, H. , J. Van der Steen, L. Ferman and T. C. Jansen (1985). "Human ocular counterroll: assessment of static and dynamic properties from electromagnetic scleral coil recordings." Exp. Brain Res. **59**: 185-196.

Dai, M., L. McGarvie, I. B. Kozlovskaya, T. Raphan and B. Cohen (1994). "Effects of spaceflight on ocular counterrolling and spatial orientaion of the vestibular system." Exp. Brain Res. **In Press.**

Diamond, S. G., C. H. Markham, N. E. Simpson and I. S. Curthoys (1979). "Binocular counterrolling in humans during dynamic rotation." Acta Otolaryngol **87**: 490-498.

Furman, J. M., H. Collewijn, T. C. Jansen and A. V. Van Den Berg (1987). "Human gaze stability in the horizontal, vertical and torsional direction during voluntary head movements, evaluated with a three-dimensional scleral induction coil technique." Vision Res. **27**: 811-828.

Furman, J. M., H. Collewijn and A. V. Van Den Berg (1987). "A direct test of Listing's law-I. Human ocular torsion measured in static tertiary positions." Vision Res. **27**: 929-938.

Goldstein, H. (1980). Classical Mechanics. Reading, MA., Addison-Wesley.

Gonzalez, R. C. and R. E. Woods (1993). Digital Image Processing, Addison Wesley.

Groen, E., J. E. Bos, P. F. M. Nacken and Bernd de Graaf (1996). "Determination of Ocular Torsion by Means of Automatic Pattern Recognition." IEEE Transactions on Biomedical Engineering **Vol. 43**(No. 5): 471-478.

Gullstrand, A. (1909). Appendix II.3. The optical system of the eye. Treatise on Physiological Optics. New York, Optical Society of America. **1:** 350-358.

Haslwanter, T. and S. T. Moore (1995). "A Theoretical Analysis of Three-Dimensional Eye Position Measurement Using Polar Cross-Correlation." IEEE Transaction On Biomedical Engineering **Vol. 42**(No.11): 1053-1061.

Hatamian, M. and D. J. Anderson (1983). "Design Considerations for a Real-Time Ocular Counterroll Instrument." IEEE Transactions on Biomedical Engineering **BME-30**(5): 278-288.

Hepp, K. and V. Henn (1983). "Spatio-temporal recording of rapid eye movement signals in the monkey paramedian pontine reticular formation (PPRF)." Exp. Brain Res. **52**: 105-120.

Hess, B. J. M. (1990). "Dual-search coil for measuring three-dimensional eye movements in experimental animals." Vis. Res **30**: 597-602.

Hess, B. J. M., A. J. Van Opstal, D. Straumann and K. Hepp (1992). "Calibration of three-dimensional eye position using search coil signals in the rhesus monkey." Vis. Res **32**: 1647-1654.

Judge, S. J. , B. J. Richmond and F. C. Chu (1980). "Implantation of magnetic search coils for measurement of eye position: an improved method." Vision Res **20**: 535-538.

Lang, S. (1987). Linear Algebra. New York, Springer-Verlag.

Leigh, J. R and D. S Zee (1983). The neurology of eye movements. Philadelpia, Davis.

Miller, E. F. and A. Graybiel (1972). "A Comparison of ocular counterrolling movements between normal persons and deaf subjects with bilateral labyrinthe defects." Ann. Otology, Rhinol. Laryngol. **72**: 885-893.

Moore, S. T., I. S. Curthoys and S. G. McCoy (1991). "VTM-an image-processing system for measuring ocular torsion." Computer Methods and Programs in Biomedicine **35**: 219-230.

Moore, S. T., T. Haslwanter, I. S. Curthoys and S. T. Smith (1996). "A Geometric Basis for Measurement of Three-dimensional Eye Using Image Processing." Vision Res. **Vol.36**(No.3): 445-459.

Mulligan, J. B. (1997). "Image Processing for Improved Eye Tracking Accuracy." Behavior Research Methods, Instruments and Computers **29**(1): 54-65.

Nakayama, K. (1974). "Photographic determination of the rotational state of the eye using matrices." American Journal of Optometry & Physiological Optics **51**: 736-742.

Ott, D., F. Gehle and R. Eckmiller (1990). "Video-oculographic measurement of 3-dimensional eye rotations." Journal of Neuroscience Methods **35**: 229-234.

Parker, J. A., R. V. Kenyon and D. E. Troxel (1983). "Comparison of Interpolating Methods for Image Resampling." IEEE Transaction on Medical Imaging **MI-2**(No.1): 31-39.

Parker, J. A., R. V. Kenyon and L. R. Young (1985). "Measurement of Torsion from Multitemporal Images of the Eye Using Digital Signal Processing Techniques." IEEE Transactions on Biomedical Engineering **BME-32**(1): 28-36.

Peterka, R. J. and D. M. Merfeld (1996). "Calibration Techniques for Video-oculography." Vestibular Research **6**(4S): S75.

Press, W. H., S. A. Teukolsky, W. T. Vetterling and B. P. Flannery (1992). Numerical Recipes in C: The Art of Scientific Computing. New York, New York, Cambridge University Press.

Raphan, T. and B. Cohen (1995). How the VOR works: Mechanisms for Gaze Compensation and Orientation. Handbook of Clinical Neuro-otology. R. W. B. G. M. Halmagyi. **I.**

Raphan, T., V. Matsuo and B. Cohen (1979). "Velocity storage in the vestibulo-ocular reflex arc (VOR)." Exp. Brain Research **35**: 229-248.

Robinson, D. A. (1963). "A method of measuring eye movement using a scleral search coil in a magnetic field." IEEE Trans Biomed Eng **10**: 137-145.

Rodrigues, O. (1840). "Des lois geometriques qui regissent les desplacements d'un systeme solide dans l'espace et de la variation des coordonnees provenant de deplacements consideres independamment des causes qui peuvent les produire." J. de Mathematiques Pures et Appliquees **5**: 380-440.

Schnabolk, C. and T. Raphan (1994). "Modeling three dimensional velocity-to-position transformation in oculomotor control." J. Neurophysiol. **71**(2): 623-638.

Straumann, D., D. S. Zee, D. Solomon and P. D. Kramer (1996). "Validity of Listing's law during fixations, saccades, smooth pursuit eye movements, and blinks." Exp Brain Res **112**: 135-146.

Sung, K. and D. J. Anderson (1991). "Analysis of two video eye tracking algorithms." Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society **13**: 1945-1950.

Taylor, S. E. (1971). The dynamic activity of reading: A model of the process. EDL Research and Information Bulletin. New York, McGraw-Hill: 9.

Tweed, D., W. Cadera and T. Vilis (1990). "Computing three-dimensional eye position quaternions and eye velocity from search coil signals." Vision Res. **30**(1): 97-110.

Tweed, D. and T. Vilis (1987). "Implications of rotational kinematics for the oculomotor system in three dimensions." J. Neurophysiol. **58**: 832-849.

Vieville, T. and D. Masse (1987). "Ocular Counter-rolling during Active Head Tilting in Humans." Acta Otolaryngol **103**: 280-290.

Wolff, E. (1940). The anatomy of the eye and orbit. London, H.K. Lewis & Co.

Wyatt, H. J. (1995). "The form of the human pupil." Vision Research **35**: 2021-2036.

Yakushin, S. B., M. J. Dai, J-I. Suzuki, T. Raphan and B. Cohen (1995). "Semicircular canal contribution to the three-dimensional vestibulo-ocular reflex: A model-based approach." J. Neurophysiol. **In Press**.

Yamanobe, S., S. Taira, T. Morizono, T. Yagi and T. Kamio (1990). "Eye Movement Analysis System Using Computerized Image Recognition." Arch Otolaryngol Head Neck Surg. **116**: 338-341.

Young, L. R. and D. Sheena (1975). "Methods & Designs: Survey of eye movement recording methods." Behavior Research Methods & Instrumentation **7**(5): 397-429.