AAMRL-TR-89-019

HELMET MOUNTED EYE TRACKING FOR VIRTUAL PANORAMIC DISPLAY SYSTEMS — VOLUME I: REVIEW OF CURRENT EYE MOVEMENT MEASUREMENT TECHNOLOGY (U)

Joshua Borah

APPLIED SCIENCE LABORATORIES
DIVISION OF APPLIED SCIENCE GROUP
WALTHAM, MA 02154

AUGUST 1989

PERIOD OF PERFORMANCE: AUGUST 1987 TO APRIL 1988

Approved for public release; distribution is unlimited.

HARRY G. ARMSTRONG AEROSPACE MEDICAL RESEARCH LABORATORY
HUMAN SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433-6573

NOTICES

When US Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Please do not request copies of this report from Armstrong Aerospace Medical Research Laboratory. Additional copies may be purchased from:

National Technical Information Service
5285 Port Royal Road
Springfield, Virginia 22161

Federal Government agencies and their contractors registered with Defense Technical Information Center should direct requests for copies of this report to:

Defense Technical Information Center
Cameron Station
Alexandria, Virginia 22314

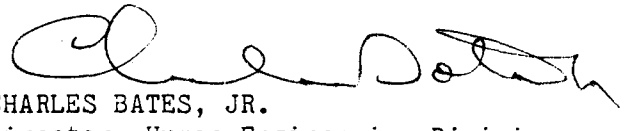
TECHNICAL REVIEW AND APPROVAL

AAMRL-TR-89-019

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER



CHARLES BATES, JR.
Director, Human Engineering Division
Armstrong Aerospace Medical Research Laboratory

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for Public Release; Distribution is Unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) AAMRL-TR-89-019
6a. NAME OF PERFORMING ORGANIZATION Applied Science Laboratories Div. of Applied Science Group 6b. OFFICE SYMBOL (*If applicable*)	7a. NAME OF MONITORING ORGANIZATION Armstrong Aerospace Medical Research Laboratory (AAMRL/HEA)
6c. ADDRESS (*City, State, and ZIP Code*) 335 Bear Hill Road Waltham, MA 02154	7b. ADDRESS (*City, State, and ZIP Code*) Wright-Patterson AFB OH 45433-6573
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (*If applicable*)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F33615-87-C-0542
8c. ADDRESS (*City, State, and ZIP Code*)	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 62202F; PROJECT NO. 7184; TASK NO. 26; WORK UNIT ACCESSION NO. 13

11. TITLE (*Include Security Classification*)
HELMET MOUNTED EYE TRACKING FOR VIRTUAL PANORAMIC DISPLAYS
VOLUME I: REVIEW OF CURRENT EYE MOVEMENT MEASUREMENT TECHNOLOGY (U)

12. PERSONAL AUTHOR(S)
BORAH, J.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (*Year, Month, Day*)	15. PAGE COUNT
FINAL	FROM Aug 87 TO Apr 88	1989, August	35

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (*Continue on reverse if necessary and identify by block number*)
FIELD	GROUP	SUB-GROUP	Eye Movements, Oculometers, Man-Computer Interface, Helmet-Mounted Displays
05	08		
12	09		

19. ABSTRACT (*Continue on reverse if necessary and identify by block number*)

The virtual cockpit concept being developed by the Air Force will require a helmet mounted eye tracker to be integrated with a helmet-mounted virtual panoramic display (VPD). Eye tracker measurements will be used with prototype systems to assist in candidate display evaluation. Operationally, eye tracking will be used for eye controlled switch selection, cueing, eye-slaved aiming, and pilot state monitoring.

Current eye tracking technology is reviewed in Volume I of this report. Relevant physiological considerations and the performance requirements implied by each of the above VPD tasks are thoroughly reviewed in Volume II.

A pupil center-to-corneal reflex technique is proposed as the most suitable technique for a VPD eye tracker. The need for robustness and dependability in the virtual cockpit

(SEE REVERSE)

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL GLORIA L. CALHOUN	22b. TELEPHONE (*Include Area Code*) (513) 255-7595 22c. OFFICE SYMBOL AAMRL/HEA

DD Form 1473, JUN 86 *Previous editions are obsolete.*

SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

19. Abstract (Continued):

application can best be met by using a full two-dimensional solid state array detector and a system that makes the complete image available to a digital processor.

Performance goals have been proposed that are feasible and will satisfy the virtual cockpit task requirements. An eye tracker design approach and prototype development plan have been outlined to meet these goals, including as examples, an analysis of possible optical paths for integration with the off-aperture and dual-mirror VPD designs.

PREFACE

The goal of this work was to explore integration of an eye line-of-gaze measurement system (eye tracker) with a helmet-mounted, virtual panoramic display (VPD). The report is presented in two volumes. Volume I is a review of current eye tracking technology, while volume II directly addresses the VPD application requirements.

The work reported on herein was performed under Contract Number F33615-87-C-0542 for a Phase I, FY87 effort in the Small Business Innovation Research Program. Ms Gloria Calhoun of the Armstrong Aerospace Medical Research Laboratory (AAMRL), Human Engineering Division, Visual Display Systems Branch, served as technical monitor. The work was done primarily at Applied Science Laboratories (ASL) in Waltham, Massachusetts, but also includes a brief set of tests performed at AAMRL Wright-Patterson Air Force Base, Ohio.

Important contributions were made by Professor Laurence R. Young, principal consultant to ASL and director of the MIT Man-Vehicle Laboratory; Professor Steven Benton, consultant to ASL and Director of the Spatial Imaging Group at the MIT Media Laboratory; and Jose Velez, ASL research scientist.

A great deal of support was also provided by several individuals at AAMRL who supplied information about Air Force goals and related display technology, participated in tests at AAMRL, generated feedback and suggestions, and served as editors. These people include Ms Gloria Calhoun, Lt German Valencia, Dr Brian Tsou, Mr Dean Kocian, Dr Herschel Self and Dr Wayne Martin.

The author is also grateful to Dr Hewitt D. Crane, of SRI International and Mr Paul Weisman, of Farrand Optical, for their suggestions and helpful information.

CONTENTS

LIST OF FIGURES

1.0 INTRODUCTION

In order to meet future operational needs the Air Force is currently developing technology for a revolutionary crew station design, often referred to as "Super Cockpit." Central to this program is the "virtual cockpit" concept in which information from many different sources is used to create a "virtual world" around the pilot.

To function as envisioned, the virtual cockpit must include eye movement and eye line-of-gaze measurement. As a first step towards the design of an appropriate eye tracking system for interface with the virtual cockpit, current eye tracking methods and devices are thoroughly reviewed in volume I of this report. Volume II details the eye tracker specifications implied by the virtual cockpit requirements and outlines a proposed design approach.

Eye tracking techniques now in use for humans can be divided into three categories: electro-oculography, scleral coil contact lens, and optical techniques. Optical techniques can, in turn, be divided into methods that detect single features or a landmark on the eyeball, methods that detect the differential motion of two features or landmarks, detection of eyelid motion, and a laser doppler velocity measurement technique. The recent technological developments of most importance to eye tracking have probably been advances in solid state optical sensors and in digital processing capabilities.

2.0 ELECTRO-OCULOGRAPHY

The human eye maintains a 0.4 to 1.0 mV electrical potential between the cornea and retina because of the higher metabolic rate of the retina which is at a negative potential. This dipole is approximately aligned with the optical axis of the eye. When the eye rotates, the dipole rotates with it and there is a corresponding variation of potential in the plane normal to the axis of rotation. The change in potential corresponding to eye rotation with respect to the head can be measured with surface skin electrodes (ref. 1, 2, and 3).

Electro-oculographic (EOG) measurement systems must overcome several difficulties. Signal levels are in the microvolt range, the conductive media is nonhomogeneous, skin resistance varies over time and the corneal-retinal potential itself varies with light adaptation, alertness and diurnal cycle. Muscle action potentials or external electrical activity can easily produce interference. These factors result in very nonlinear output functions and in significant dc drift.

EOG measurements are analog and allow eye movements to be measured with very high bandwidth and very low transport delay. The measurement range is virtually the entire physiological range of eye movements. The basic EOG technique was first used in the 1920's and '30's. Advances in the techniques since its inception have been the development of better electrodes, better preamplifiers, refinements in electrode positioning, and most recently, the capability for

sophisticated digital on-line processing of the data. Modern silver chloride skin electrodes are convenient to use and do not produce significant discomfort.

A variety of appropriate electrodes, as well as high impedance preamps designed for EOG use, are readily available on the commercial market. Turnkey systems for EOG measurements are available from a small number of sources, generally as part of a test system for neurological testing.

Even with high common mode rejection amplifiers, the EOG technique is highly sensitive to the ambient electromagnetic fields, including those at the line frequency. Although some noise can be eliminated by synchronous sampling or filtering, at the expense of higher frequency signal components, the effective resolution of EOG is essentially limited by noise to approximately 1 degree.

Modern computers are quite adequate to linearize the measurements while maintaining high sample rates. Linearization, however, is effective only to the extent that the input/output function remains stable over time. Although improvements in electrodes and preamplifiers have lessened the drift inherent in EOG measurements, such drift remains a major obstacle to accurate measurement of absolute eye position with EOG. The dc offset can easily drift by an amount representing several degrees over a period of several minutes. The sensitivity and shape of the input/output curve (eye position versus measured potential) may vary over time as well. EOG drift is typically slow enough that velocity measurements are not a problem, even for slow tracking eye movements.

As a result, current EOG technology provides an excellent means of measuring eye velocity and acceleration profiles, detecting saccades, measuring nystagmus, and supporting other applications that require good high frequency measurement, but do not require a highly accurate measure of absolute eye position. EOG is also very useful for making measurements when the eye is closed. EOG is not the best technique for measuring eye point-of-gaze on a scene or target.

3.0 SCLERAL COIL

The most accurate method for measuring eye position is the scleral coil technique. An induction coil is imbedded in a ring of flexible material which adheres to the eye sclera or limbus (boundary between cornea and sclera). A voltage is induced in the scleral coil by the uniform oscillating magnetic field generated by a large set of coils surrounding the subject's head. The induced voltage varies with the sine of the angle between the scleral coil and the magnetic field. If the scleral coil does not move with respect to the eye, it is possible to make a very precise measure of eye orientation with respect to the large coils surrounding the head (ref. 4 & 5).

Developments in recent years have led to systems that are fairly easy to apply, adhere to the limbus with little slippage, and cause

little discomfort for most users, at least over short periods of time up to about 20 minutes. The large electromagnetic coils are typically mounted on a cubical frame surrounding the subject's head. Orientation (azimuth and elevation) of the eye relative to this frame can be measured to within one minute of arc. Position of the eye with respect to the coil frame is not measured. The measurement bandwidth can be very high (e.g., 200 Hz). A second scleral coil and another radiating coil can also be incorporated to measure torsional motion.

The technique is slightly, but distinctly, invasive, requiring a drop of eye anesthetic and application of a contact lens. Very thin wire leads, typically positioned at the inner canthus (nasal corner of the eye), extend from the eye and are connected to an electronics unit.

Magnetic field distortions caused by externally generated fields or ferrous objects may cause measurement errors. It is possible to detect and correct for such effects by adding a stationary set of detector coils within the area enclosed by the induction coils. A complete scleral coil system is now commercially available from at least two sources (Skalar Instrumentation, Delft, The Netherlands; and C-N-C Engineering, Seattle, Washington).

The scleral coil technique is readily available, extremely dependable and accurate, and is impervious to subject differences and ambient light variation. If the technique were used for point-of-regard measurement in conjunction with a head position measurement system, almost no calibration would be required. Only one known data point would be needed to define the initial reference orientation of the eye line-of-sight axis. Changes from that orientation could be precisely measured without further mapping computations.

Unfortunately, the invasiveness associated with the technique, although minor in some research environments, probably rules out its use in an operational or training environment.

4.0 OPTICAL TECHNIQUES

All other practical eye movement measurement techniques involve tracking one or more features or reflections that can be optically detected on the eye. The features that have most often been used for this purpose include the limbus (iris sclera boundary), the pupil, the lower eyelid, the reflection of a light source from the cornea (1st Purkinje image or corneal reflex), and a similar reflection from the rear surface of the eye lens (4th Purkinje image). Devices exist which track several different combinations of these features with several different detection techniques.

Advances in these techniques in recent years are primarily due to advances in optical sensor technology and a virtual explosion in computer processing technology. Solid state cameras and linear detector arrays have become smaller and more sensitive, allowing better detection with less obtrusive equipment. Availability of increasingly small,

fast, and powerful computer processors allows more complex image information to be used and allows more complex nonlinearities to be handled effectively.

All of the optical techniques, except use of the lower eyelid and laser doppler velocimetry, rely to some extent on the following simple principles. If a landmark is fixed to a sphere, rotation of the sphere about its center will cause a translation of that landmark proportional to the sine of the rotation angle. The relation for a single axis is

$$d = r \sin\theta \tag{1}$$

where d is the landmark translation and r is the distance from the center of the sphere to the landmark (see figure 4.1). If translation of the landmark can be detected, and if rotation of the sphere is the only motion allowed with respect to the detector, then, still considering a single axis, the rotation angle is

$$\theta = \sin^{-1}(d/r) \tag{2}$$

Of course, if the entire sphere translates with respect to the sensor, the landmark will translate by the same amount. Use of equation 2 would result in an erroneous angle computation

$$\theta E = \sin^{-1}(d_T/r) \tag{3}$$

where d_T is translation of the entire sphere along the sensitive plane of the detector and θE is the erroneous rotation angle calculation.

Correct computation of θ in the presence of translation requires some means to distinguish landmark motion due to rotation and that due to translation. If translation of the entire sphere is measured by some independent means, this value can simply be subtracted from d before employing equation (2). Another approach is to detect the position of two landmarks fixed to the sphere, but located at different radii from its center. Two such landmarks will move together if the entire sphere translates, but will move differentially if the sphere rotates. From figure 4.2,

$$d_1 = d_T + r_1 \sin\theta \tag{4}$$

$$d_2 = d_T + r_2 \sin\theta \tag{5}$$

$$\Delta d = d_2 - d_1 = (r_2 - r_1)\sin\theta \tag{6}$$

where θ is sphere rotation, r_1 and r_2 are distances of landmarks 1 and 2 from the center of the sphere, d_T is translation of the entire sphere parallel to the sensitive plane of the detector, and d_1 and d_2 are the total displacements of landmarks 1 and 2 parallel to the plane of the detector. Note that, whereas d_1 and d_2 are functions of both the rotation and translation of the sphere, Δd (the relative motion of the two landmarks) is a function only of rotation. Furthermore, the sensitivity of the measurement is proportional to the difference in the distance of the two landmarks from the center of the sphere $(r_2 - r_1)$.

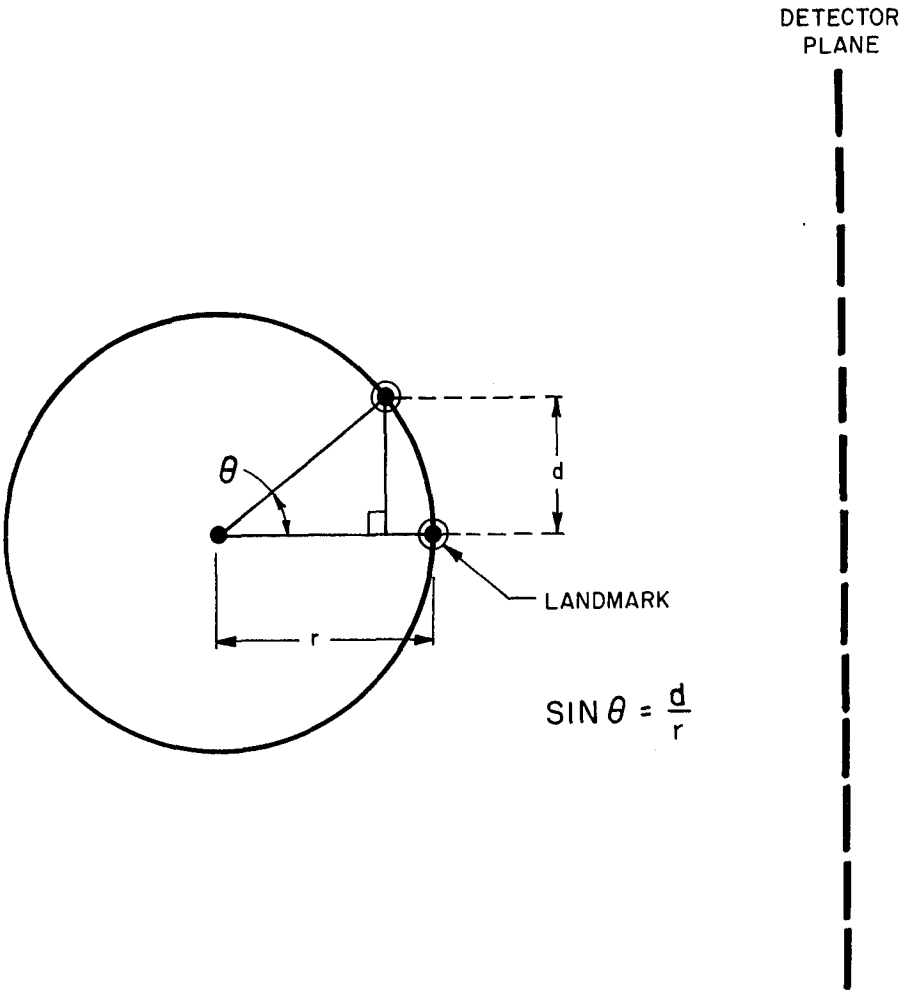


Figure 4.1 Cutaway view of sphere with surface landmark.

In the case of the eye, the limbus and the pupil center provide fixed features about 12.0 mm and 9.76 mm, respectively, from the center of rotation of the roughly-spherical eyeball; the corneal reflex provides a means of tracking the corneal center of curvature, which is about 5.6 mm from the eye center of rotation; and the 4th Purkinje image provides a means of tracking the lens rear surface equivalent mirror center of curvature, which is about 11.5 mm from the eye center of rotation.

Translation of the eyeball with respect to a detector can be caused either by eyeball translation within the eye socket or motion of the entire head with respect to the detector. The normal amount of eye translation within the eye socket is not precisely known, but is certainly very slight. The amount of head motion that occurs with respect to a detector depends on the specific apparatus being used. The eye dimensions shown in figure 4.3, computed from values in references 6 and 7, represent a nominal or "standard" eye and are the dimensions used in subsequent examples. These values do, however, vary somewhat between subjects.

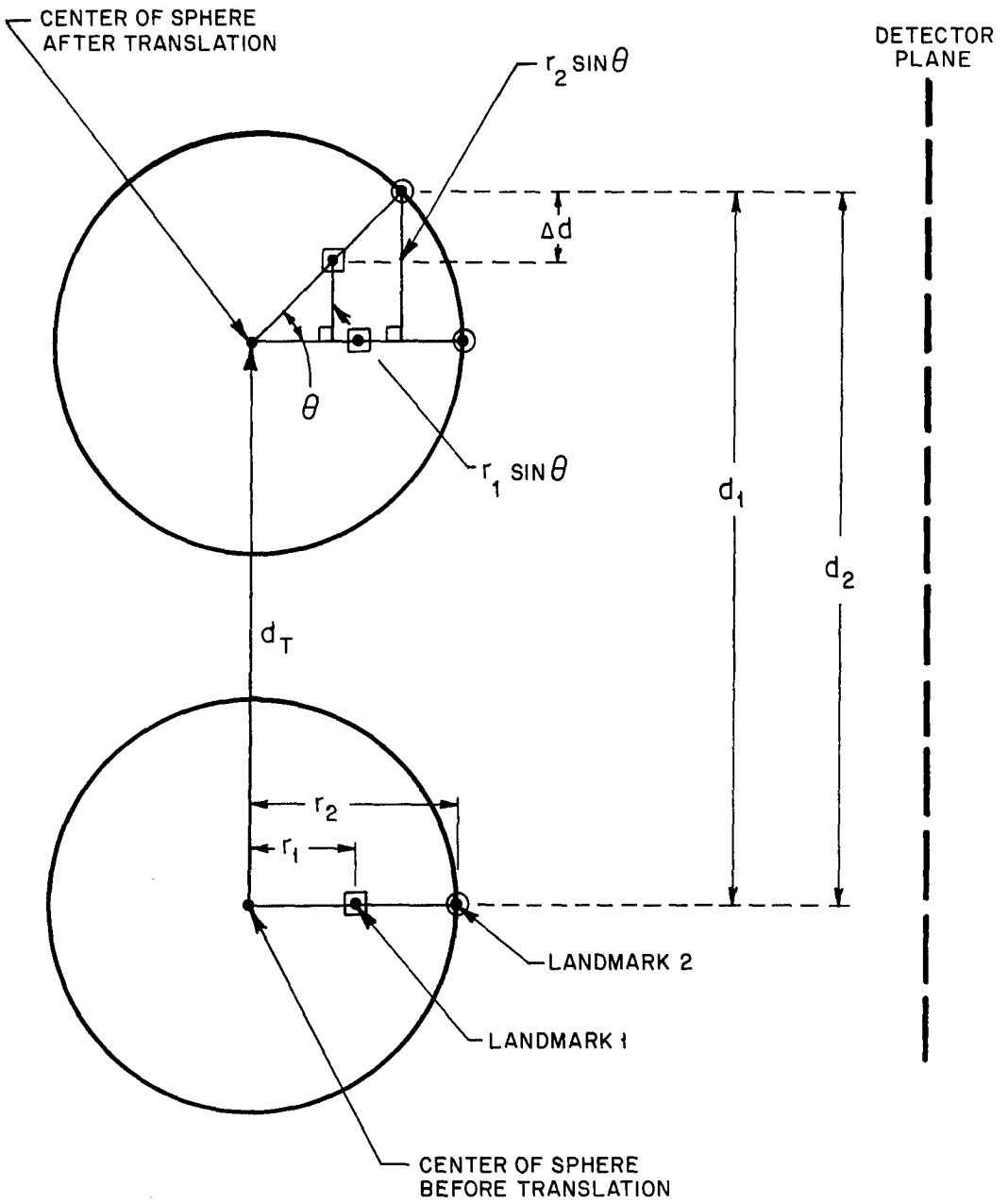


Figure 4.2 Cutaway view of sphere with two landmarks fixed to sphere at different radii from center.

4.1 Limbus

The limbus is the boundary between the iris and sclera. Since the limbus is fixed to the eyeball, a rotation of the eyeball will cause a translation of the limbus proportional to the sine of the rotation angle, as described by equations 1 and 2 in the previous section. The distance from the eye center-of-rotation to the limbus (r in equations 1 & 2) is about 12 mm.

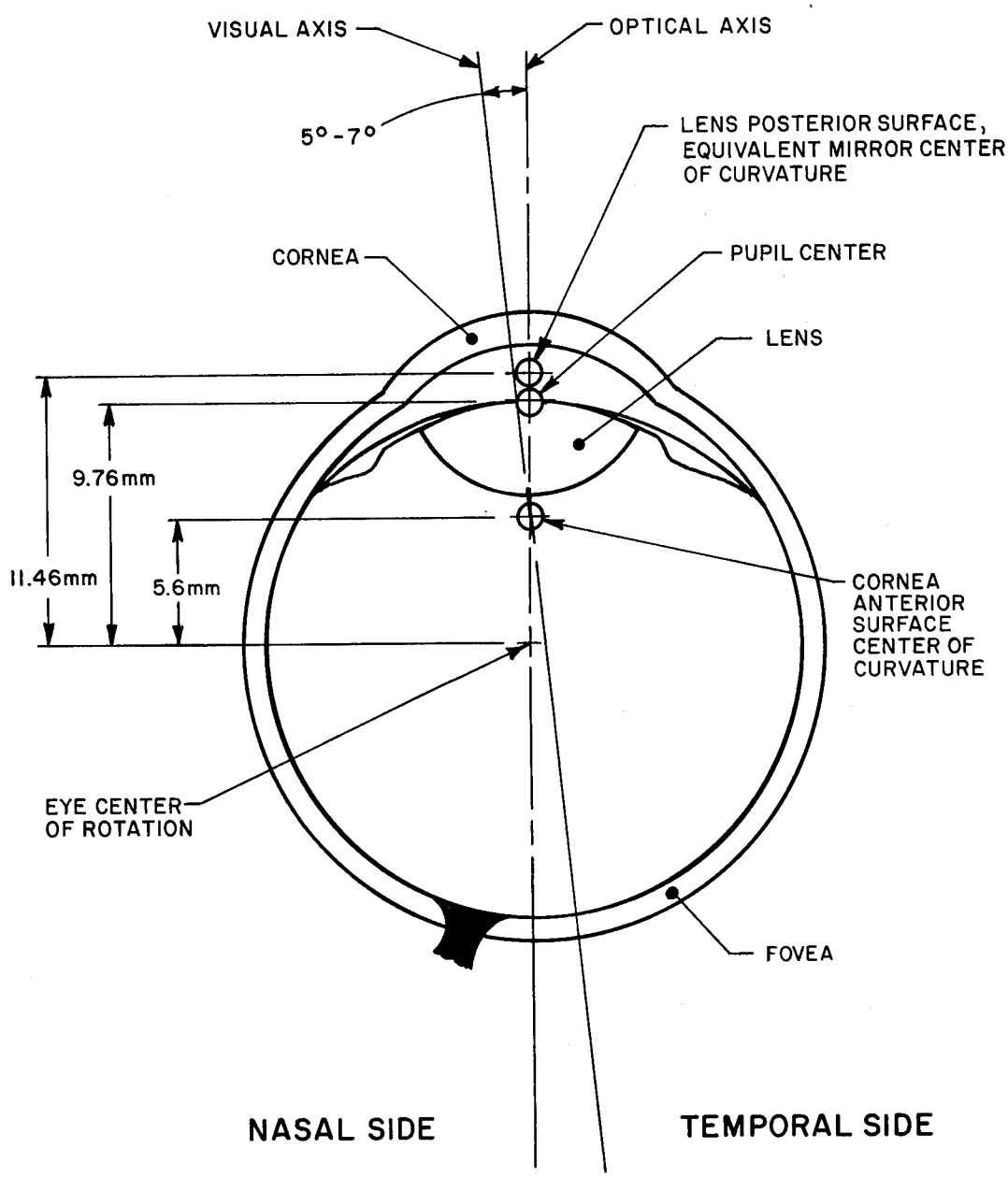


Figure 4.3 Eye showing features often used for eye tracking with nominal dimensions.

Note that an eye translation of 1 mm parallel to the sensitive plane of the sensor will be optically equivalent to a rotation of about 4.8 degrees.

If the entire limbus could be detected, motion due to eye rotation could be distinguished from translation by ellipticity considerations. In practice, ellipticity computation is difficult because the eyelids normally obscure a significant amount of the upper and lower portions of

the limbus. Ellipticity is also a very insensitive measure when the angle between the plane of the limbus and the plane of the detector is small.

If the entire limbus could be detected, motion due to eye rotation could be distinguished from translation by ellipticity considerations. In practice, ellipticity computation is difficult because the eyelids normally obscure a significant amount of the upper and lower portions of the limbus. Ellipticity is also a very insensitive measure when the angle between the plane of the limbus and the plane of the detector is small.

Eyelid obscuration hinders vertical eye position measurement using the limbus. Whereas horizontal position can be measured by tracking the left and right edges of the iris sclera boundary, the top and bottom edges, which would be the most sensitive indicators of vertical position, are usually occluded. Vertical position can be computed from the motion of the exposed sides of the limbus, but this is always a less sensitive measure, especially since such a large portion of the limbus is often obscured.

For these reasons, limbus tracking is most suitable for horizontal eye movement measurement. Since a good horizontal measurement can be made by tracking edge position alone, simple photocells can be used and high bandwidth is relatively easy to achieve.

4.2 Pupil

4.2.1 Eye Movement Measurement Using the Pupil

The pupil center corresponds closely to the optical axis of the eye and is almost fixed with respect to the eyeball. The center of the pupil can move with respect to the eyeball only to the small extent that the iris may open or close asymmetrically. Except for the small additional error introduced by asymmetrical iris motion, the geometrical relation between pupil center position and eye rotation angle is described by equations 1 and 2 (section 4.0) with r equal to about 9.76 mm. An eye translation of 1 mm parallel to the plane of the sensor will be the equivalent of about 5.8 degrees rotation.

Unlike the limbus, pupil diameter is not fixed, but varies with visual field luminance, fatigue, emotional arousal and other variables. It is, therefore, not possible to measure eye position by simply tracking a pupil edge. The position of the pupil center must be determined.

There are many possible algorithms for finding the pupil center. The most appropriate algorithm depends on the type of sensor being used, the desired measurement update rate, and the amount of computer processing power available.

The pupil has several advantages over the limbus for eye tracking. It is smaller than the limbus resulting in much reduced eyelid

occlusion; it is often possible to achieve greater optical contrast between pupil and iris than between iris and sclera; and the pupil edge is often sharper than the iris/sclera boundary. Vertical, as well as horizontal, position can usually be measured effectively, although eyelid occlusion may impose some limits on the vertical range.

It is not as easy to achieve high bandwidth measurements of pupil center as of limbus motion, since accurate centroid computation requires that more image information be processed.

4.2.2 Bright Versus Dark Pupil Image

The retina is highly reflective, but any light reflected back through the pupil will be directed towards its original source. In fact, if the eye is focused at the plane of the source, such retroreflected light will be imaged back at the source. Under normal viewing conditions, the pupil looks completely black because none of the reflected rays return to the observer. If, however, the observer is able to look along the axis of an illumination beam, then the observer will see the retinal reflection and the pupil will appear bright. A beam splitter can easily be used to align an illumination beam with the optical axis of a detector, thus creating a bright pupil image.

Under some conditions, there is substantially better contrast between a backlit "bright" pupil and the surrounding features (iris, sclera, eyelids, etc.) than between the normal dark pupil and the surrounding features. A bright pupil is most easily recognized when it is significantly brighter than its surround and a dark pupil is easily recognized when significantly darker than its surround.

Since the return from a bright pupil is a relatively narrow beam, while the return from surrounding features is diffusely scattered, or isotropic, bright pupil contrast increases as the detector to-the-eye distance increases.

If the detector aperture area is decreased, the radiation collected from the iris, sclera, and eyelids will decrease, but radiation collected from the bright pupil will not be affected so long as the detector aperture area is greater than the source area. (Remember that the source tends to be imaged back on to itself by the retinal reflection). Contrast, therefore, varies in inverse proportion to the source and detector aperture area if these apertures remain equal. Bright pupil contrast cannot be increased indefinitely by reducing the source and detector apertures, because such adjustments decrease the total signal strength. Detectors have limited sensitivity and safety considerations limit source strength.

In an idealized case, total radiant input to a detector from a bright pupil should be inversely proportional to the square of pupil area (fourth power of pupil diameter), since the light must pass through the pupil twice. The area of the image on the detector also increases with pupil area, so the apparent brightness of the image (radiant flux per unit area on the detector) will vary with the square of pupil

The brightness of the iris, sclera, and other surrounding features is not affected by pupil diameter and contrast between the pupil and surrounding features, therefore, increases with the square of pupil diameter.

Bright pupil contrast can be adversely affected by any illumination not coaxial with the detector. Such illumination may result in not only decreased pupil diameter, but also the return from this illumination may make surrounding features appear brighter to the detector without increasing the pupil signal. The reflectivity of the retina, the iris, the sclera, and the eyelids vary somewhat between subjects, and this will also affect contrast.

If we assume an equidistant illumination source and coaxial detector, assume that the eye is focused at the distance of the illuminator source, and further assume that the illuminator source is being imaged at the plane of the eye, then pupil-to-iris or sclera contrast can be described by the equation:

$$\text{contrast} = \begin{cases} \dfrac{r_d^2}{r_s^4}(r_p)^2 \dfrac{d_s^2}{d_e^2} \dfrac{R_r}{R_i} & ; r_s > r_d \\ \dfrac{1}{r_s^2}(r_p)^2 \dfrac{d_s^2}{d_e^2} \dfrac{R_r}{R_i} & ; r_d > r^s \end{cases} \qquad (7)$$

where d_s is distance from the source or detector to the eye, d_e is the distance from the pupil to the retina, r_p is the pupil radius, r_s is the source aperture radius, r_d is the detector aperture radius, R_r is retinal reflectance and R_i is iris or sclera reflectance.

Experience has shown that, if pupil diameter remains above a certain value, it is usually possible to create a bright pupil image that can be distinguished from surrounding features by little more than a simple threshold criterion. As pupil diameter decreases below this value, increasingly more sophisticated pattern recognition processing is required to distinguish the pupil from other features. The minimum pupil diameter value for which simple threshold recognition is possible depends on all of the contrast factors discussed above, but usually turns out to be between 3 and 4 mm.

A dark pupil image remains essentially black no matter what the characteristics of the noncoaxial illumination. Contrast is enhanced by increasing the radiant input to the detector from the iris and sclera. This can be done, for example, by decreasing the distance to the detector or increasing the illumination. Note that the most favorable conditions for recognizing a dark pupil image are usually the least favorable for bright pupil recognition and vice versa.

The presence of shadows caused by eyelids or the corneal bulge and the presence of corneal reflections that may be superimposed on parts of

the pupil usually require that more than a simple threshold criterion be used for dark pupil recognition. At Applied Science Laboratories the experience has been that, when the pupil diameter is above 3.5 mm, the pupil can clearly be more easily recognized with a bright, rather than dark, pupil technique. When pupil diameter is below 2.5 mm, it can be more easily recognized with a dark pupil technique. Between these values the choice is not clear and depends on all of the contrast factors previously discussed.

4.3 Corneal Reflex

The corneal reflex (CR), or first Purkinje image, is the reflection of a light source from the front surface of the cornea. The apparent position of the CR will be about half way between the corneal surface and corneal center of curvature along the ray that extends from the source through the corneal center of curvature. The apparent size of the CR will be proportional to the angle subtended by the source at the eye. Specifically, for sources of small angular subtense, the diameter of the reflection will be $0.5r\theta$, where r is the corneal radius of curvature and θ is the angular subtense of the source.

If the source is collimated, the apparent displacement of the CR will always be the same as the displacement of the corneal center of curvature. Even if the source is not collimated, the motion of the CR will be very close to that of the corneal center of curvature, as long as the source is at least several inches from the eye.

The CR generally appears brighter than any other return from a given source and is relatively easy to detect. It provides a landmark on the eye that can be used to compute eye rotation in the same way as for the pupil center or limbus. If the source is collimated (or far away compared to the corneal radius), the geometrical relation between the CR position and eye optical axis with respect to a detector is described by equation (1), but with r equal to the distance from the eye center of rotation to the corneal center of curvature (about 5.6 mm).

As with the pupil center and limbus tracking techniques, eye rotation cannot be distinguished from translation by measuring the CR position alone. Using equation (2) with r = 5.6 mm, a 1 mm translation parallel to the sensitive plane of the detector is equivalent to about 10.3 degrees of rotation. The CR technique is far more sensitive to translation than pupil or limbus tracking. Because of interference by scleral reflections, the range of CR detection is usually limited to about $\pm$25 degrees visual angle with respect to the source. The range may sometimes be further limited in the vertical direction by eye lid occlusion. Due to a slight flattening towards the outer edges of the cornea, the relation between eye rotation and CR displacement becomes more nonlinear at large angles. Nonlinearities may also be introduced by tear film and other conditions that distort the corneal surface.

The CR is an inherently more precise landmark than pupil center because its position is unaffected by pupil diameter changes and because, over a large range of eye motion, it remains completely

unoccluded by eyelids and other artifacts. Because of the potentially small diameter of the CR, however, higher resolution detectors are necessary to take advantage of the additional precision it can afford.

4.4 Fourth Purkinje Image

The reflections from the posterior corneal surface and the anterior and posterior lens surfaces (2nd, 3rd and 4th Purkinje images) are very dim compared to the CR (1st Purkinje image). Aside from the CR, the 4th Purkinje image is the easiest to detect. It makes no sense to measure 4th Purkinje image position alone, since equivalent information can be obtained much more easily from the CR, but it is sometimes used in combination with the CR. This will be discussed in section 4.5.

Light is refracted by the cornea and the anterior lens surface before reaching the posterior lens surface. Reflected light is also refracted by these surfaces as it returns from the posterior lens surface. We can account for these effects by considering an equivalent mirror, with about a 5.8 mm radius of curvature, for the posterior lens surface.

The 4th Purkinje image will appear along the ray extending from the source through the rear lens surface equivalent mirror center of curvature. The reflection will appear to be about half way between the equivalent mirror and the center of curvature. If the source is collimated, tracking the 4th Purkinje image is equivalent to tracking the equivalent mirror center of curvature. Once again, equation (1) applies, but this time with r equal to the distance from the center of eye rotation to the rear lens surface equivalent mirror center of curvature (about 11.5 mm).

The 4th Purkinje image is obscured if it falls behind the iris and, therefore, range of detection depends somewhat on pupil diameter.

4.5 Dual Feature Techniques

Tracking the position of a single landmark on the eye does not permit distinction between eye rotation and translation with respect to the detector. The ambiguity can be eliminated by tracking two features located at different radii from the eye center of rotation. As described in section 4.0, two such features will move together when the eye translates, but differentially when the eye rotates. The two sets of features that have been used most successfully are the pupil and CR (1st Purkinje image), and the 1st and 4th Purkinje images.

4.5.1. Pupil-to-CR Vector

The most commonly employed dual feature technique is measurement of the relative position of the pupil center and the CR. Assuming a collimated source, then from sections 4.2 and 4.3

$$d_p = (9.76\ mm)\sin\theta \quad (8)$$

$$d_{cr} = (5.6\ mm)\sin\theta \quad (9)$$

$$\Delta d = d_p - d_{cr} = (4.16\ mm)\sin\theta \quad (10)$$

where d_p is pupil displacement parallel to the sensitive plane of the detector, d_{cr} is CR displacement parallel to the detector plane, θ is the angle between the detector optical axis and the eye optical axis, and Δd is the relative displacement of the pupil center and CR. As shown in figure 4.4, if the detector optical axis and illumination beam are coaxial, then Δd is the absolute distance between the pupil and CR in the plane of the detector.

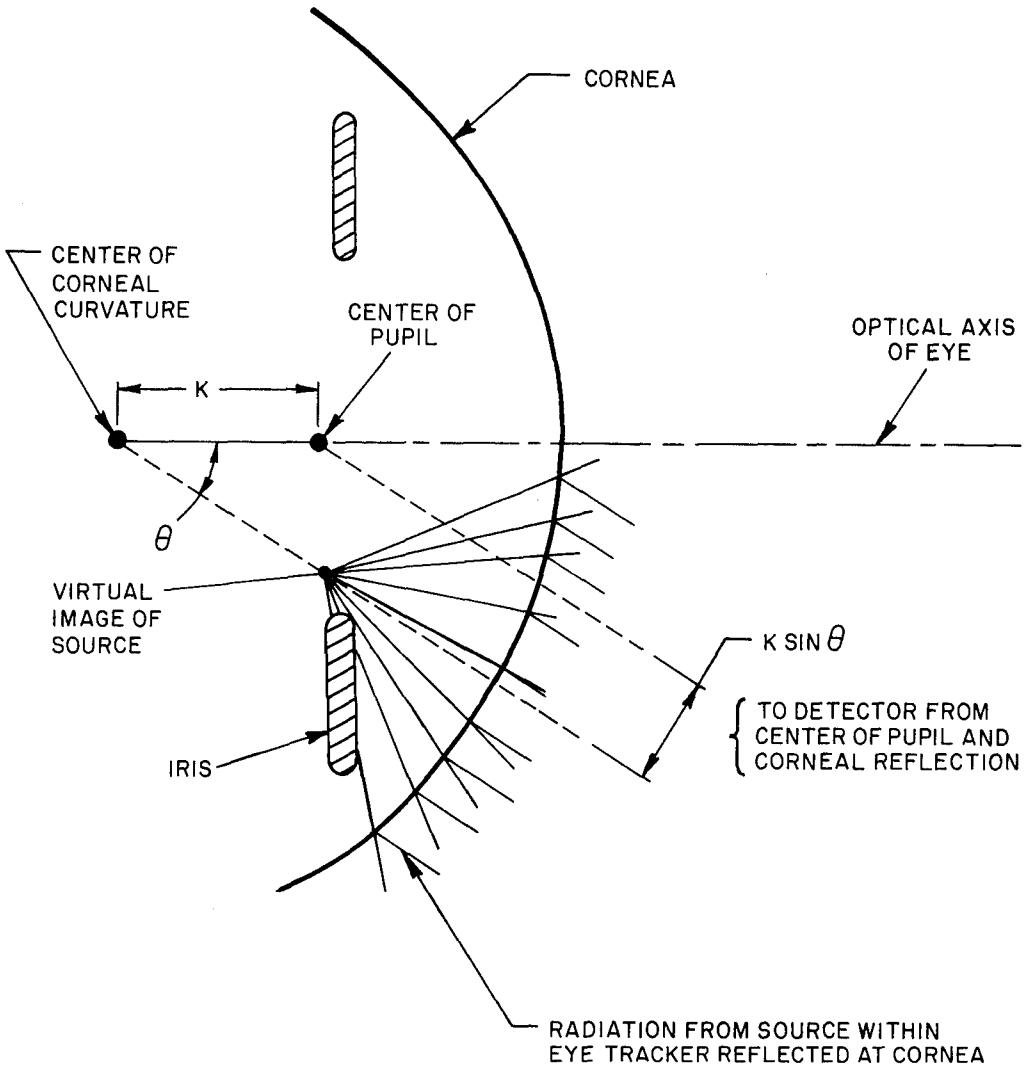


Figure 4.4 Position of the pupil center and corneal reflection with respect to a detector and coaxial illuminator, after Merchant, Morrissette, and Porterfield (ref. 8).

When the eye translates with respect to the detector, the relative position of the pupil center and CR does not change. In terms of equation (10), Δd does not change.

The pupil-to-CR technique can be used with either a dark or bright pupil. Under most conditions, the CR will appear significantly brighter than a bright pupil as well as the iris, sclera, and eyelids. The range is generally limited by CR detection in the horizontal axis and by eyelid occlusion of the pupil or CR in the vertical axis. In the case of a coaxial illuminator and detector, the range is usually about +25 degrees horizontally and about +30 degrees to -10 degrees vertically with respect to the detector. These values can vary by at least 5 degrees, depending on the specific implementation, and can also vary by several degrees between subjects. The lower vertical value is especially variable because it depends on upper eyelid position.

The potential accuracy of the pupil-to-CR technique cannot be stated precisely because there are no precise data available quantifying the stability of the pupil center with respect to the eye optical axis. Currently available systems that employ this technique all claim accuracies on the order of 1 degree visual angle.

Because of the relative ease with which both the pupil and CR can be detected, the technique lends itself to very unobtrusive implementations. Measurement bandwidth is limited by the need to detect and process enough optical information to accurately determine the pupil center.

The pupil-to-CR technique alone can be used to directly measure point-of-gaze on a scene directly, even in the presence of head motion, if the following conditions are met: 1. the illuminator and detector are far from the subject compared to the amount of head motion; 2. the scene being viewed is far from the subject compared to the amount of head motion. If these conditions are not met, then it is necessary to independently measure the length and direction of the eye-to-detector optics vector. This is shown for one plane in figure 4.5. The angle θ in figure 4.5 is the quantity determined by the pupil-to-CR technique as shown in figure 4.4. Note that, in order to define point-of-gaze explicitly on the scene (x), it is also necessary to know d_1, d_2, and ϕ. Two of these quantities, d_1 and ϕ, will vary with head motion.

If the detector and illuminator are fixed to the head, then the pupil-to-CR technique provides a direct measure of eye line-of-gaze with respect to the head fixture.

4.5.2 Dual Purkinje Image

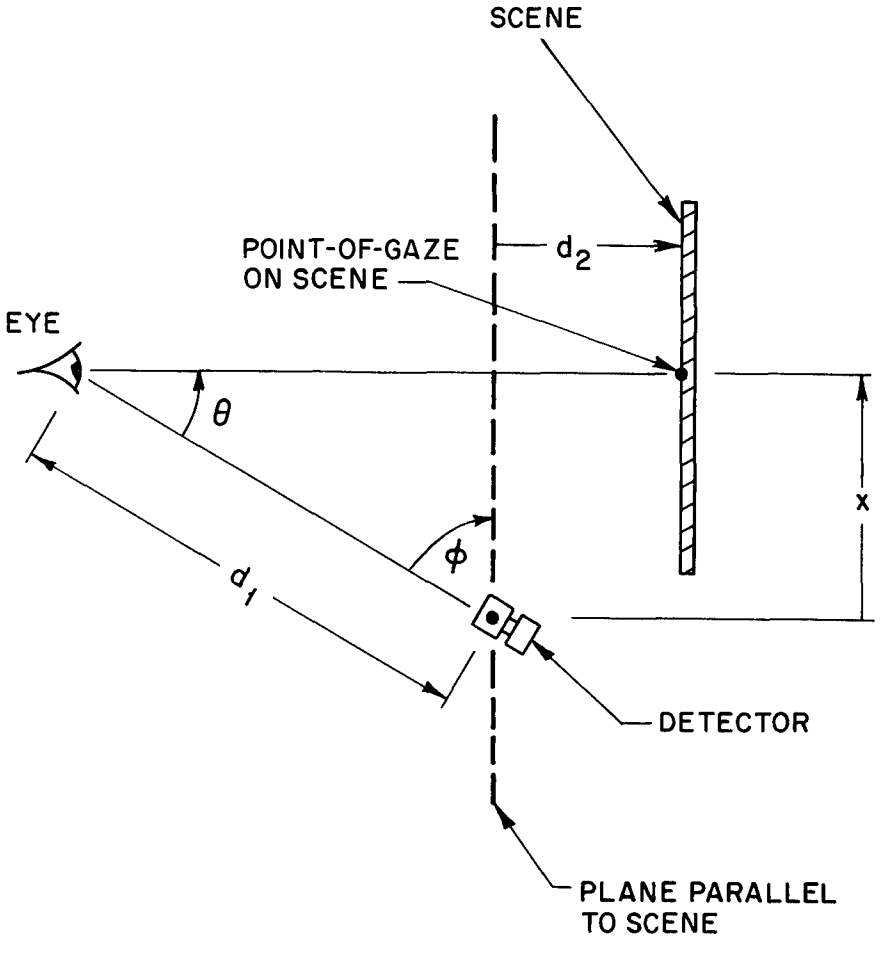
The CR (1st Purkinje image) and 4th Purkinje image move differentially with eye rotation and move together with eye translation. These two features can be used in much the same way as the pupil-to-CR method to measure line-of-gaze without errors due to translation. From sections 4.3 and 4.4

$$d_{cr} = (5.6 \text{ mm})\sin\theta \quad (11)$$

$$d_{p4} = (11.5 \text{ mm})\sin\theta \quad (12)$$

$$\Delta d = d_{p4} - d_{cr} = (6 \text{ mm})\sin\theta \quad (13)$$

where d_p is pupil displacement parallel to the sensitive plane of the detector, d_{p4} is 4th Purkinje image displacement parallel to the detector plane, θ is the angle between the detector optical axis and the eye optical axis, and Δd is the relative displacement of the pupil center and 4th Purkinje image.



$$x = \frac{d_1 \text{SIN}\,\theta - d_2 \text{COS}\,(\phi+\theta)}{\text{SIN}\,(\phi+\theta)}$$

Figure 4.5 Relationship in one plane between point-of-gaze on a flat scene and relative eye, detector, and scene positions.

The dual Purkinje image method allows significantly greater accuracy than the pupil-CR method because both Purkinje images have a more fixed position with respect to the eye optical axis than does the pupil, and because their positions can be detected very precisely. Accuracy on the order of 1 arc minute and frequency response up to about 500 Hz have been achieved with the dual Purkinje image technique. The 4th Purkinje image, however, is very dim and, therefore, difficult to detect. For this reason, a complex and relatively large optical apparatus is usually required. Measurement range is restricted to about ±10 degrees, primarily because of fourth Purkinje image occlusion by the iris. The range can be expanded to about ±15 degrees if the subject's pupils are artificially dilated.

4.6 Eyelid

The lower eyelid tends to move proportionally to static vertical eye position. The function is reasonably linear over about ±15 degrees of vertical eye motion. If a detector tracks the boundary of the lower lid with respect to the head, and if some simple linearization is also performed, vertical eye position with respect to the head can usually be measured with at least a 2-degree accuracy. With some subjects, accuracies of about 1 degree are possible. There is often some drift over time and there is probably some lag or overshoot, especially during saccadic motions, although the dynamics are not well documented.

Lower eyelid boundary position can be measured with simple photo cell detectors and lower eyelid tracking is often used in conjunction with horizontal limbus tracking measurements.

4.7 Laser Doppler Velocimetry

The velocity of moving particles can be measured by detecting the frequency shift of laser light scattered by the particles. Rather sophisticated instruments have been developed using this technique for measuring shaft rotation velocities and other applications (ref. 9). Some prototype work has also been done to investigate possible measurement of corneal movement with this technique. Note that the cornea is not totally transparent and does scatter a small percent of incident light.

Laser doppler velocimeters have the potential to provide very accurate, high bandwidth eyeball velocity measurements, but cannot directly provide position information. Although position can be computed by integration, there is no way to correct for gradual position error accumulation or to reacquire position information after eye blinks. If combined with a relatively low bandwidth eye position measurement technique, laser doppler velocimetry might be an effective means for gathering high frequency information.

5.0 OPTICAL SENSORS

Virtually all of the optical sensors used in current eye tracking systems are solid state devices that rely on the photoelectric effect.

5.1 Photo Conductive Cells

The simplest optical sensors that are of significant use in eyetracking are single photocells which undergo a conductivity change proportional to incident radiation. Photo cells are available with three different semiconductor structures: photo transistors, photodiodes, and photoresistors. In all cases, if power is properly applied, a voltage proportional to incident radiation can be derived.

This type of sensor has been available for many years and has long been used, for example, in limbus and eyelid tracking devices. Because these are single output analog devices, high bandwidth is easily achieved, but the information content is quite limited. Phototransistors and photodiodes have more gain and faster response times than the older photoresistors and are more frequently used.

5.2 Quadrant and Bicell Detectors

Bicell or quadrant photo cells consist of two or four discrete photo cells separated by a small gap. If a uniform spot of light falls on all detectors, each photo cell will respond proportionately to the area of the spot that falls on it. In the case of a quadrant detector, the relative x and y position can be calculated as

$$x = K \frac{(B + D) - (A + C)}{A + B + C + D} \qquad (14)$$

$$y = K \frac{(A + B) - (C + D)}{A + B + C + D}$$

where A, B, C, and D are the respective quadrant signal outputs, as shown in figure 5.1. Total intensity information can be derived by simply summing the four elements.

Depending on the shape of the spot, the x and y positions, as calculated by equation (14), may not change linearly with true position, but will always be monotonic and will be zero when the spot is centered. If the spot is not uniform, the equation (14) x and y calculations will still be monotonic and will be zero when the light energy centroid is centered on the detector.

The detector will provide position information only when the spot of light partially covers all four detectors. In other words, the range is twice the spot diameter. The spot of light must also be smaller than the detector's sensitive area in order to measure position. Within this

range, quadrant detectors can provide resolution on the order of 0.1 micrometers (ref. 10).

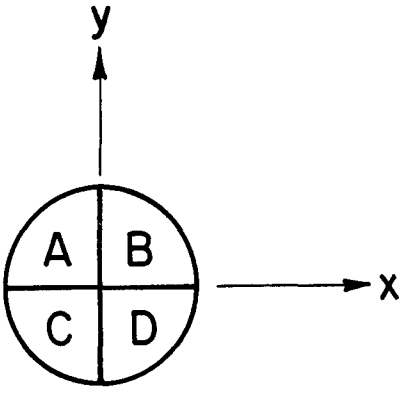


Figure 5.1 x and y axes with respect to quadrant photodetector cells A, B, C, and D.

Because they have excellent resolution, but a nonlinear input/output relation and a small effective range, quadrant detectors are most often used as nulling devices to center small spots of light. They can provide position as well as intensity information, but do not provide any detailed image information. Quadrant detectors have been very successfully used to implement the dual Purkinje image eye tracking method (see section 6.5).

5.3 Lateral Effect Photo Diodes

Lateral effect photo diodes (often called position-sensitive detectors or PSD's) are single element devices that produce continuous position data. Incident light on the sensitive area produces a charge which must travel through a resistive layer of material to electrodes. Ideally, the resistive layer is uniform, so the resulting current at any electrode is proportional to the distance of that electrode from the incident light. For a two-dimensional device there are typically 4 electrodes at the periphery of the sensitive surface and along two orthogonal axes.

Position information can be calculated by:

$$x = K \frac{A - C}{A + C} \qquad (15)$$

$$y = K \frac{B - D}{B + D}$$

where A, B, C, and D are signals at the four contacts, as shown by figure 5.2. These devices provide the average or centroid position of

light incident on the detector sensitive area. The position measurement is thus independent of light intensity and profile.

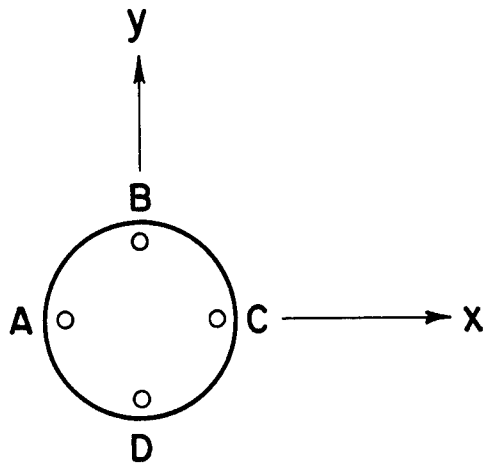


Figure 5.2 x and y axes with respect to four contacts on a lateral effect photodiode.

Because the conductive region cannot be made completely uniform, linearity is not perfect and usually varies from about 1 percent in the central region to several percent in the periphery. Position measurement resolution of about 5 micrometers can be achieved. The sensitive area for two-axis devices is typically about 1 cm^2, although both larger and smaller units are available.

Although not quite as sensitive as quadrant photo cells, position sensitive detectors are more versatile. They can be used for open loop position detection as well as nulling tasks. Like the quadrant photo cells and single photo cells, they do not provide any detailed information about the image.

Single photo cells, quadrant detectors and PSD's all have light sensitivities that are functions of wavelength. This function varies with the construction of the particular devices. A large portion of these devices use silicon as the sensitive element and have peak sensitivities in the 700 to 900 nm near-infrared wavelength region. The total spectral band is typically about 350 to 1100 nm.

In order to minimize the effects of external noise and amplifier drift, photocells and PSDs are often used in a pulsed mode. Typically, the light source is pulsed and the detector signal is processed by a synchronous amplifier, demodulator, and lowpass or bandpass filter circuit.

5.4 Linear and Two-Dimensional Array Detectors

More detailed image information can be acquired with solid state linear and two-dimensional sensor array devices. Arrays can be made by clustering photocells in tightly packed linear or two-dimensional

arrangements. As long as a relatively small number of elements are used, the individual elements operate as individual photocells and the electronics can remain relatively simple.

If an array is to contain thousands of elements, it becomes necessary to scan the array. Charge coupled devices (CCDs), charge injected devices (CID's) and other similar semiconductor arrays are designed to integrate for discrete time intervals and serially output accumulated charge packets from each array element. Relatively complex electronics are required, but the result is full gray scale information over either a single row or a two-dimensional array containing a very large number of pixels.

Linear arrays can be used to acquire information about a two-dimensional image, either by sweeping the image over the array or by placing a cylindrical lens over the array. The former technique results in full image information, but requires at least one moving part. A cylindrical lens, on the other hand, collects onto each pixel all the light from a strip of the image orthogonal to the array. In other words, the image dimension orthogonal to the array axis is collapsed onto the array. If a second array and cylindrical lens are placed orthogonally to the first, the resulting pixel data can be convolved in various ways to extract detailed spatial information. The cylindrical lenses effectively act as low pass spatial filters, however, and some high spatial frequency information is irretrievably lost.

A wide variety of linear arrays is available ranging from 256 or fewer pixels to well over 3000 pixels. The pixel elements are typically rectangular with dimensions in the 5 to 15 micron range.

Two-dimensional arrays can be used as cameras to acquire full two-dimensional gray scale image information, and are usually packaged as video cameras with all the necessary accompanying electronics. By using filtering and masking or multiple array techniques, full color can also be achieved. Color cameras will not be discussed here, however, because they are not usually relevant to eye tracking.

Because solid-state cameras are smaller, lighter, less expensive and do not suffer from as much lag as older vacuum tube cameras, there has been an enormous commercial demand stimulating their rapid development. Available resolution and sensitivity has been rapidly increasing, while size and cost have been decreasing. Since the greatest volume demand is for fairly traditional TV camera applications, most, but not all, commercially available solid state cameras are designed to produce an analog output signal that conforms to traditional video standards. Units intended for use in the U.S. generally have a 60 Hz field update rate, and produce a signal that meets the standard 525 line, RS170 format, regardless of the intrinsic resolution of the solid state array being used. (Units intended for sale in Europe generally conform to the 625 lines, 50 Hz, CCIR format.) Often, because of the logic imbedded in the sensor array chip, it is not possible to modify this format significantly, even by designing custom electronics to drive the sensor chip.

The sensor arrays usually have sensitive areas that conform to standard 2/3 inch or 1/2 inch video formats. Most currently available camera arrays have on the order of 500^2 pixels, although arrays of over 1000^2 can also be found. A large and rapidly increasing volume of "frame grabber" interfaces are available which digitize video signals with varying spatial and grey scale resolution, and store fields or frames in memory buffers (1 frame = 2 fields) for access through various standard computer busses.

Two-dimensional array chips that can be flexibly controlled for varying integration times, update rates, and output of selected pixel subsets are also available. The choice is more limited, however, and these more flexibly controllable chips are not yet as highly developed as those intended only for traditional video output. Since each sensor element effectively stores the radiant energy received, sensitivity (i.e., the minimum detectable flux) is a function of the integration time allowed. Sensitivity, therefore, decreases with increased update rates. Because it takes a finite time to transfer each packet of information out of the array chip, maximum update rate is inversely proportional to the number of pixels (i.e., resolution).

Photo detector arrays do exist in at least one other form. A device sometimes called an optical random access memory (RAM) is essentially a computer random access memory chip with the silicon memory sites exposed and positioned in a precise, known geometry. In fact, a crude device can sometimes be made by removing the protective covering from a standard RAM chip. If enough radiation is incident on a given element, enough charge will be created to turn it "on" so that it will read as a one. The resulting information is binary, since each element is read as a one or a zero, but elements are randomly accessible and can be addressed as memory by a computer. The threshold irradiance level for each element is proportional to the integration time allowed. Update rate is limited by the number of pixels used, the necessary integration time, and the minimum memory access time for the combination of chip, processor and bus. The authors are aware of only one commercially available chip designed for this purpose.

Array detectors provide the most complete image information and, therefore, the best potential for robust recognization of desired features in complex images. Array devices are the type most often used, for example, by eye tracking instruments that use the pupil-to-CR technique. By the same token, making use of such detailed image information imposes a very large processing burden.

Most solid-state array sensors are silicon-based and have peak sensitivities in the near infrared 700 - 900 nm wavelength region. This turns out to be very convenient for eye tracking applications. Human vision is minimally sensitive in this spectral region and, therefore, near-infrared eye illumination sources can be used without being overly obtrusive to subjects.

5.5 Optical Sensor Availability

A wide variety of phototransistors and photodiodes are available from most manufacturers of optoelectronic components, including General Electric, Hewlett Packard, United Detector Technology, Texas Instruments, TRW, Siemens, Motorola, EG&G Vactec, and others. Lateral effect photodiodes are available from United Detector Technology and Hamamatsu. Linear arrays are available from General Electric, Texas Instruments, Fairchild, EG&G Reticon and others.

A variety of solid-state cameras are available from RCA, Fairchild, GE, Sony, Hitachi, Cohu, NAC as well as a host of other manufacturers. A very flexibly controllable device capable of frame rates between 300 and 400 Hz with 128 x 128 pixel resolution is available from EG&G Reticon. General Electric offers a 512 x 512 pixel device with variable frame rates up to 300 Hz. The number of rows that can be scanned decreases as a function of increasing frame rate, so that at 240 Hz, for example, the effective resolution is 512 x 64 pixels. A device with higher resolution and a lower maximum frame rate is also available from the same source. Spin physics, NAC, and Video Logic have sensor arrays capable of 2000 Hz frame rates, but these are only available packaged in cameras designed for high speed video recording systems. An optical RAM chip is available from Micron Technology Inc.

6.0 CURRENT OPTICAL TECHNIQUE IMPLEMENTATIONS

The examples cited in the following subsections are representative, but by no means exhaustive.

6.1 Limbus Tracking Implementations

Limbus tracking techniques have been used since the early 1950's (ref. 11). Commercially available limbus tracking systems are currently offered by Applied Science Laboratories (ASL, Waltham, Massachusetts) and John Hains Optoelectronics (Havant, England).

The ASL system, based on a technique developed by Richter and Pfaltz (ref. 12) and Stark, Vossius and Young (ref. 13), uses a near-infrared light emitting diode (LED) to illuminate the eye and two phototransistors to detect the reflectivity from the iris-sclera boundaries (See figures 6.1 and 6.2). The reflected illumination from the region of each boundary changes differentially with horizontal eye motion and the photo-transistor signals are, therefore, subtracted to measure horizontal position. To measure vertical position, a similar set of sensors and an LED is aimed at the lower eyelid boundary on the other eye. The sensor signals are summed to measure the reflectivity change as the eyelid moves proportionately to vertical eye position. The sensor assemblies are mounted on a fixture that allows position adjustments in all three axes.

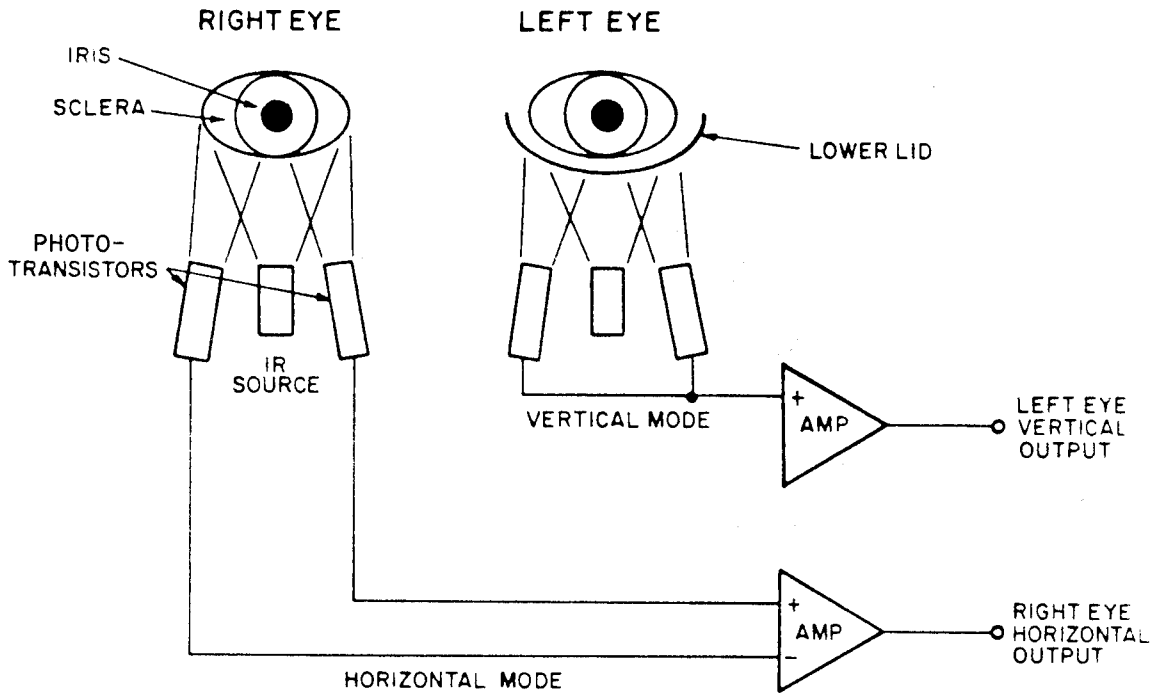


Figure 6.1 Schematic showing Applied Science Laboratories' photoelectric limbus and eyelid tracking system.

The LED's are pulsed at 2 KHz and the phototransistor signals are processed by an amplifier, a demodulator, and a set of filters to reduce noise and artifacts from other sources, including ambient illumination. In addition to mechanical sensor position adjustment, electronic offset, gain, linearity and vertical/horizontal crosstalk controls are used to linearize and calibrate the output. Accuracy of eye position measurement with respect to the head is about 1 degree visual angle along the horizontal axis and 2 degrees along the vertical axis; range is about $\pm$15 degrees in both axes; resolution is several minutes of arc; effective bandwidth is about 40 Hz; and sample rate is 1 KHz. Outputs are analog signals, parallel digital signals, and a point-of-gaze cursor superimposed on a video scene camera image.

The John Hains unit uses a similar sensor assembly, except that two infrared LED's are used with each set of sensors, such that one photocell/LED pair covers each "half eye." The horizontal measurement is made by subtracting sensor signals and the vertical measurement by adding them. The lower eyelid, however, apparently is not used for the vertical measure, but rather total reflectivity across the eye is measured. The LED's are pulsed and the sensor signals demodulated and filtered. The sensors are mounted on spectacle frames. The outputs are analog; the range is $\pm$15 degrees horizontal and +8 to -12 degrees vertical, and resolution is listed as better than 1 degree. Frequency response is not specified.

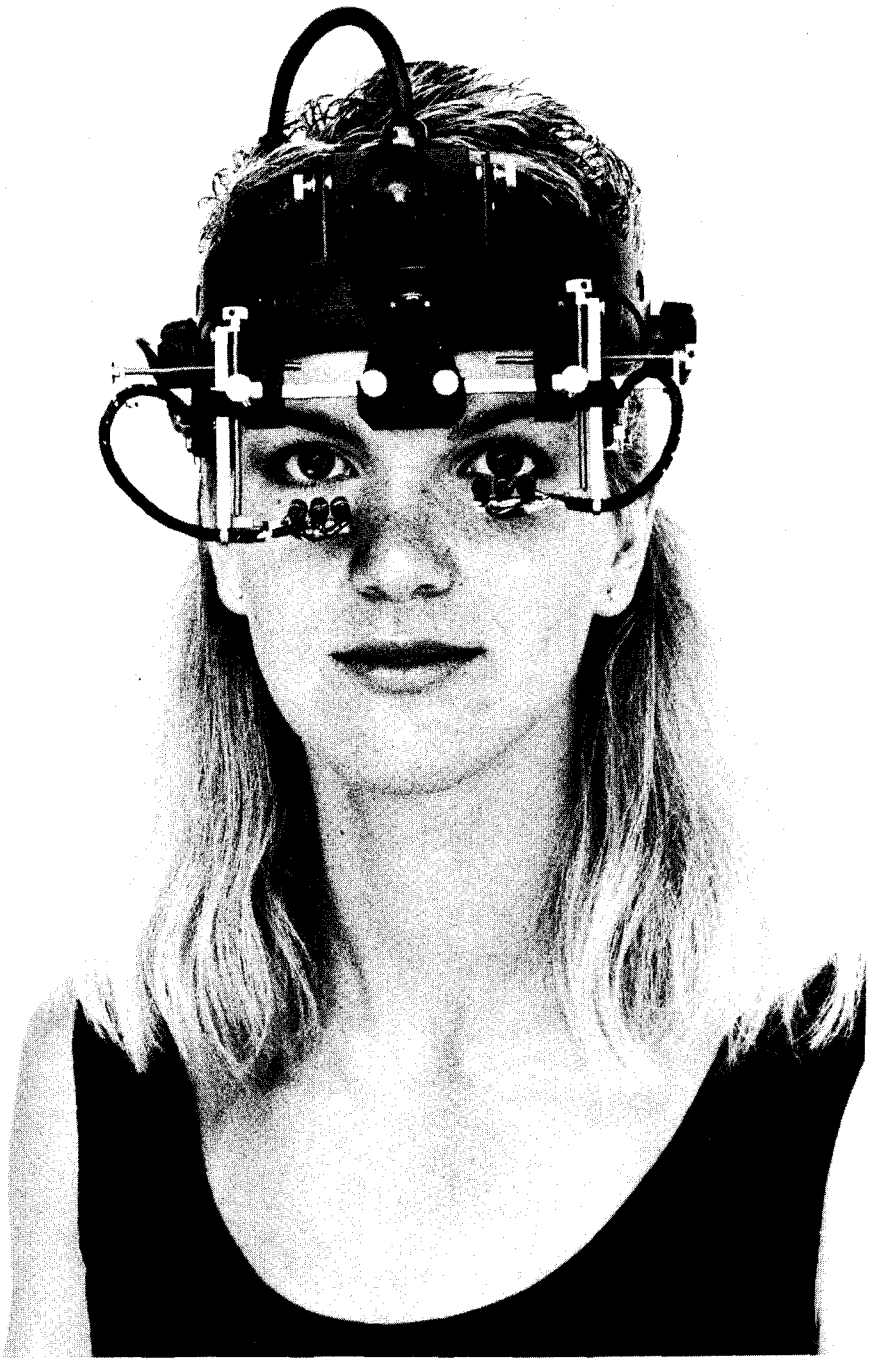


Figure 6.2 Applied Science Laboratories' limbus and eyelid tracker with head and mounted scene camera.

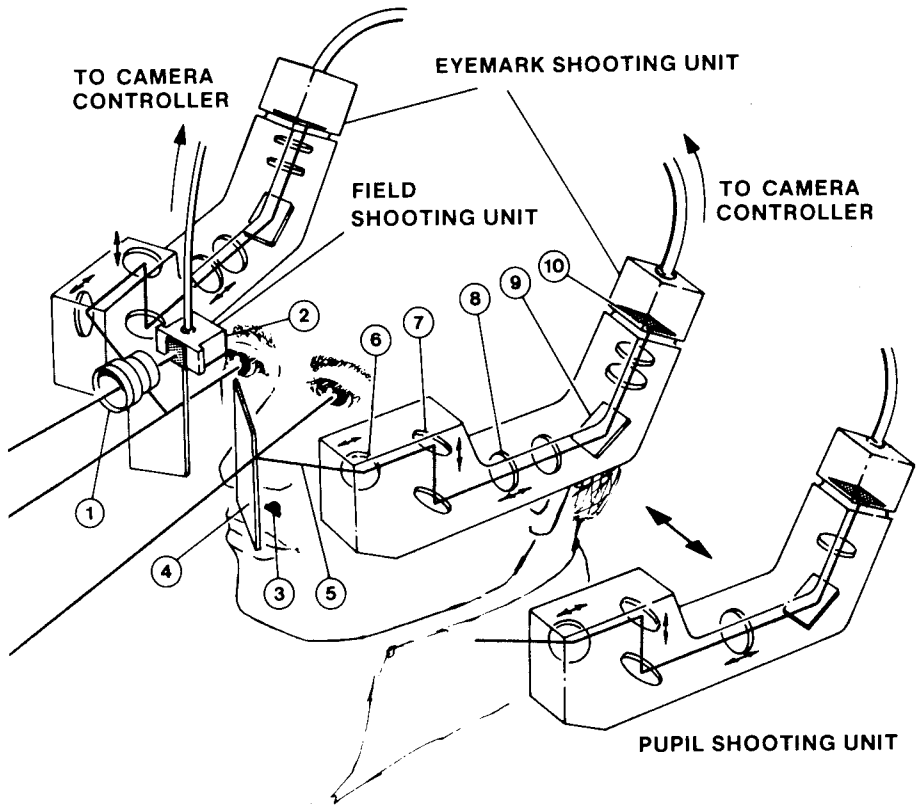
A limbus tracker design by Engleken, et. al. (ref. 14) uses essentially the same emitter detector arrangement as the ASL system for horizontal measurements. The modulation frequency used is 3 KHz. Performance tests with a model eye gave maximum deviations from a straight line curve fit of from 0.13 to 0.36 degrees over a ±25 degree range. Engleken, et al., estimate that accuracy with a human subject would be about 1 degree. The bandwidth of this instrument is 150 Hz. Measurement of vertical eye movements is not discussed.

None of the systems described in this section includes a computer or microprocessor. Of course further processing, including improved linearization, can be done with an external computer. All of these systems are also subject to errors arising from movement of the optics with respect to the head, as described in section 4.1.

Since the measurements are with respect to the head, line-of-gaze with respect to another reference frame (e.g. the room) can be computed only if the head is fixed in that frame, head position in that frame is independently measured, or the image from a head-fixed scene camera is used as the reference.

6.2 Corneal Reflex Tracking Implementations

The commercial corneal reflex measurement system shown on figures 6.3 and 6.4 is made by NAC (Japan). A near infrared LED light source, and a solid state camera sensor head are contained in a fixture that



① Field lens	⑥ X axis adjusting mirror
② Field camera (MOS)	⑦ Y axis adjusting mirror
③ IR Eye spot lamp	⑧ Focus lens
④ IR reflector	⑨ Reflection mirror
⑤ Eye mark optical axis	⑩ Eye mark camera (MOS)

Figure 6.3 Optical configuration or NAC corneal reflex tracking system, from reference 16.

Figure 6.4 NAC corneal reflex tracking system, from reference 16.

mounts on a subject's head. The corneal reflex image is relayed to the camera by a beam splitter in front of the subject's eye and by several mirror and lens elements. These optics are duplicated for the other eye. The corneal reflex positions on the detector arrays are electronically combined with the image from a scene camera also mounted to the head fixture, and are displayed as cursors on the scene image. Linearization and calibration are accomplished by mechanical adjustments in the optics. Optics modules can be added to also provide a video image of either eye. This eye image is not used for a measurement, but is displayed in a corner window on the scene image.

The resulting scene image can be video-taped, and the digital x,y position of the corneal reflections can also be recorded on the video tape during the vertical blanking periods. An optional output module

can be used to provide real-time or off-line analog and digital data. The data update rate is 30 Hz. System accuracy and range are not specified by the NAC literature. Based on the system configuration, the largest sources of error are probably slippage of the optics on the head (see section 4.3) and parallax due to the distance between the scene camera and the eye.

Frecher, Eizenman, and Hallet (ref. 15), at the University of Toronto, have designed an apparatus that makes an extremely precise measurement of CR position. The device images corneal reflex through cylindrical lenses onto two orthogonal linear arrays each composed of 20 phototransistors. The illumination optics and CR relay optics are designed so that the CR will have a bell-shaped intensity distribution along either axis which always covers at least three of the sensors on each array. Since the CR image covers more than one sensor in a statistically predictable way, it is possible to estimate CR center position to much better than one pixel resolution.

The sensor signals are sampled and electronically processed to achieve a reported 2% or better linearity over a 30-degree visual angle range with less than 30 arc seconds of noise, a velocity resolution of two deg/sec, and a one KHz update rate. These performance values assume virtually no motion of the head with respect to the optics and require head stabilization with a bite bar or similar restraint technique. The existing devices are laboratory units and are not commercially available.

6.3 Pupil Tracking Implementations

A system is offered by Dr Didier Bois in West Germany that tracks the "center of mass" (centroid) of an entire dark pupil eye image. It is included in this section because it seems likely that the position of the pupil is the major linear contribution to the measurement. The eye is illuminated by infrared LED's and imaged onto a sensor array. Neither the optics or sensor are specified in detail on the available product literature. A 4 KHz sample rate and five arc minute resolution are reported. Accuracy is not specified. The basic system is designed for head fixed operation, but is also available with an optical head motion detector that works by detecting the position of four infrared LED's fastened to a head-mounted fixture. In this case, the eye position measurement optics are also contained in the head fixture. The system uses a digital processor to compute direction of gaze with respect to a stationary scene in the presence of up to $\pm$10 cm head translation and $\pm$20 degrees of head rotation.

A prototype pupil tracker has been built by RTS laboratories (Gainsville, Florida) using an LED light source and an optical RAM detector, both mounted to spectacle frames. The binary image information is read by a computer which executes an algorithm to identify the pupil and find its center. Complete performance characteristics of this system are not yet known.

6.4 Pupil-to-CR Technique Implementations

The pupil-to-CR technique was first developed at Honeywell by Merchant, et al. (ref. 8). Commercially available systems employing the pupil-to-CR technique currently include systems produced by Applied Science Laboratories (Waltham, Massachusetts), ISCAN (Cambridge, Massachusetts), Micromeasurements (Berkeley, California), and Demel (West Germany). All of these systems illuminate the eye with near infrared light and use video cameras as detectors. They also all require computers or microprocessors to process the camera data.

All of these systems offer a configuration in which the optics are "room fixed" and view the subject's eye from a distance. The ASL and Demel systems offer tracking mirrors that maintain the necessary small eye camera field of view over the eye in the presence of significant subject head motions (on the order of one cubic foot). For the room-fixed optics configuration, all systems offer a means of superimposing a point-of-gaze cursor on a video image of the scene. Figure 6.5 illustrates a schematic of the pupil-to-CR method employed at ASL with floor-mounted optics.

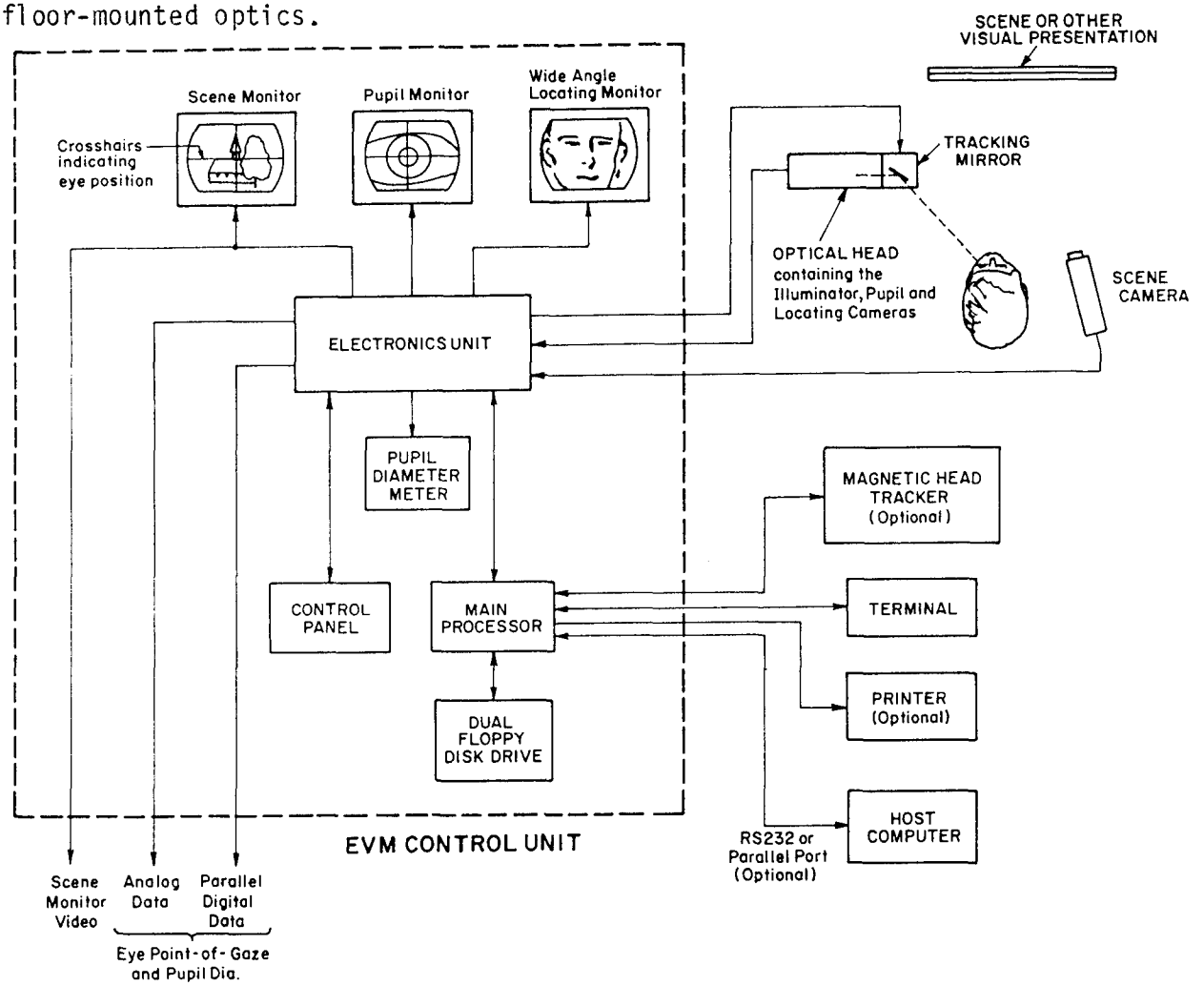


Figure 6.5 Schematic showing Applied Science Laboratories' pupil-to-corneal reflex method eye tracker with floor-mounted optics.

A helmet- or headband-mounted configuration with miniaturized optics and a head-mounted scene camera is also available as a catalog item from ASL (See figure 6.6). ISCAN and Micromeasurements offer head-mounted optics as a custom configuration. Because head mounted systems measure line-of-gaze with respect to the head, line-of-gaze with respect to some other frame can be computed only by stabilizing the head, independently measuring position and orientation of the head, or by using a head-mounted scene camera image as a reference. Slippage of the optics on the head is usually not a significant problem with the pupil-to-CR technique, unless the eye moves out of the camera field of view (see section 4.5).

The details of system functional organization, recognition, and linearization techniques vary somewhat among the different systems. The ASL system uses a hardware video preprocessor to supply edge information to a computer processor. The computer processor identifies the pupil and CR, finds the distance between their centroids, computes pupil diameter, linearizes the data, and computes point-of-gaze in the scene space. The same processor also handles calibration, data recording and a range of operator interaction functions. Pupil and CR recognition in the presence of artifacts relies primarily on size and shape criteria. Linearization and mapping are done with a polynomial curve fit technique.

The ISCAN system employs a specialized board or board set for pupil recognition and centroid determination and for CR recognition and centroid determination. The output from this package is pupil diameter, and pupil and CR position with respect to the eye camera field of view. A separate module is offered which subtracts pupil and CR position and performs linearization and mapping by interpolating between calibration target points. This module also displays a point-of-gaze cursor. Pupil and CR recognition techniques are considered proprietary.

The Micromeasurements system uses a windowing technique to help reject artifacts and to restrict the amount of video data that must be processed. Data from each field of video is examined over an area only slightly larger than the pupil and centered over the last-computed pupil position, thus excluding any artifact not close to the pupil. If the pupil is not found in the window, the window is moved in a search pattern until the pupil is found.

Details of the Demel system functional organization and recognition techniques are not available. Linearization and mapping are done by interpolation.

All systems so far mentioned, except for the ISCAN system, generally use bright pupil optics. The standard ISCAN system uses a dark pupil image. It is likely that, with suitable optics and some logic modifications, any of the systems could be switched from bright pupil to dark pupil recognition or vice versa.

Under good conditions, all of these systems probably exhibit accuracy of about 1 degree visual angle over about a 40-degree

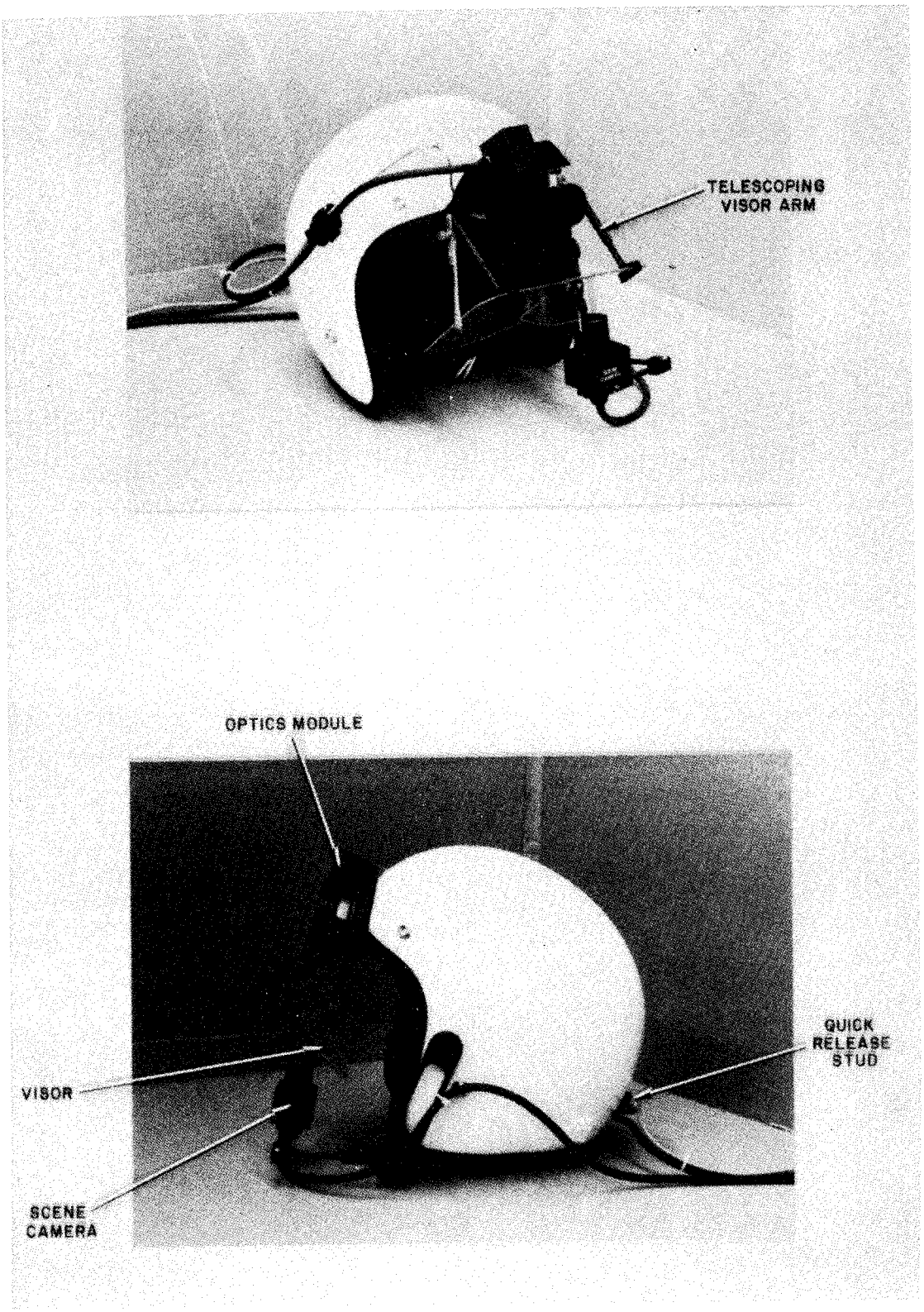


Figure 6.6 Applied Science Laboratories' helmet mounted optics for pupil-to-corneal reflex method eye tracker.

horizontal and 30-degree vertical range. The horizontal range is usually symmetrical with respect to the optical axis (see figure 4.3). The vertical range is usually biased by upper eyelid occlusion to about 25 degrees up (optical axis 25 degrees above the detector axis) and 5 to 10 degrees down. All of the systems have a standard 60 Hz update rate. ISCAN offers an optional high speed camera version that will work at sample rates up to 240 Hz, but with reduced resolution.

All of these systems operate by detecting features within a relatively complex image. Depending on the environment and the individual subject, the image may have varying contrast between desired features and background and may contain artifacts produced by partial eyelid occlusions, reflections from other light sources, etc. All of the systems are designed to handle, to some degree, artifacts and poor contrast. The relative robustness of the different systems can be determined only by careful side-by-side tests.

A system was recently in development by SRD (Israel) using a bright pupil-to-CR technique and using linear arrays as detectors. The system was designed to operate at up to 1 KHz sample rate. All other details have been considered proprietary and information concerning the current status of this system is not available.

A system is also in development by Dr. Moshe Eizenman and colleagues at the University of Toronto. The system recognizes two features and is designed to operate with at least a 240 Hz sample rate. Details concerning this system are also considered proprietary and are not available.

6.5 Dual Purkinje Image Implementation

A dual Purkinje image system, first developed by Cornsweet and Crane (ref. 7 and 17), is commercially available from SRI International (Menlo Park, California). The eye is illuminated by an infrared source beamed through a series of lenses, mirrors, and stops. The first and fourth Purkinje images are optically separated and imaged onto separate quadrant photocell detectors (See figure 6.7 and 6.8). The return Purkinje image paths, as well as the incident illumination beam, are directed by servo-controlled mirrors. Error signals from the detectors are used to drive the servo motors so as to keep the Purkinje images centered on the detectors. In other words, a closed loop nulling task is performed. The separation between the two Purkinje images is a function of the servo mechanism positions and this is output in analog form. An optical auto-focus technique is used to correct for fore-aft head motion.

The illumination source is pulsed at 4 KHz and, presumably, there is synchronous amplification and demodulation of the detector signals. Except for this 4 KHz chopping, it is a completely analog system.

The latest version of this system has a measurement noise level of about 20 arc seconds rms; a frequency response up to 500 Hz for eye movement up to several degrees, and an output signal delay of about 0.25 msec.

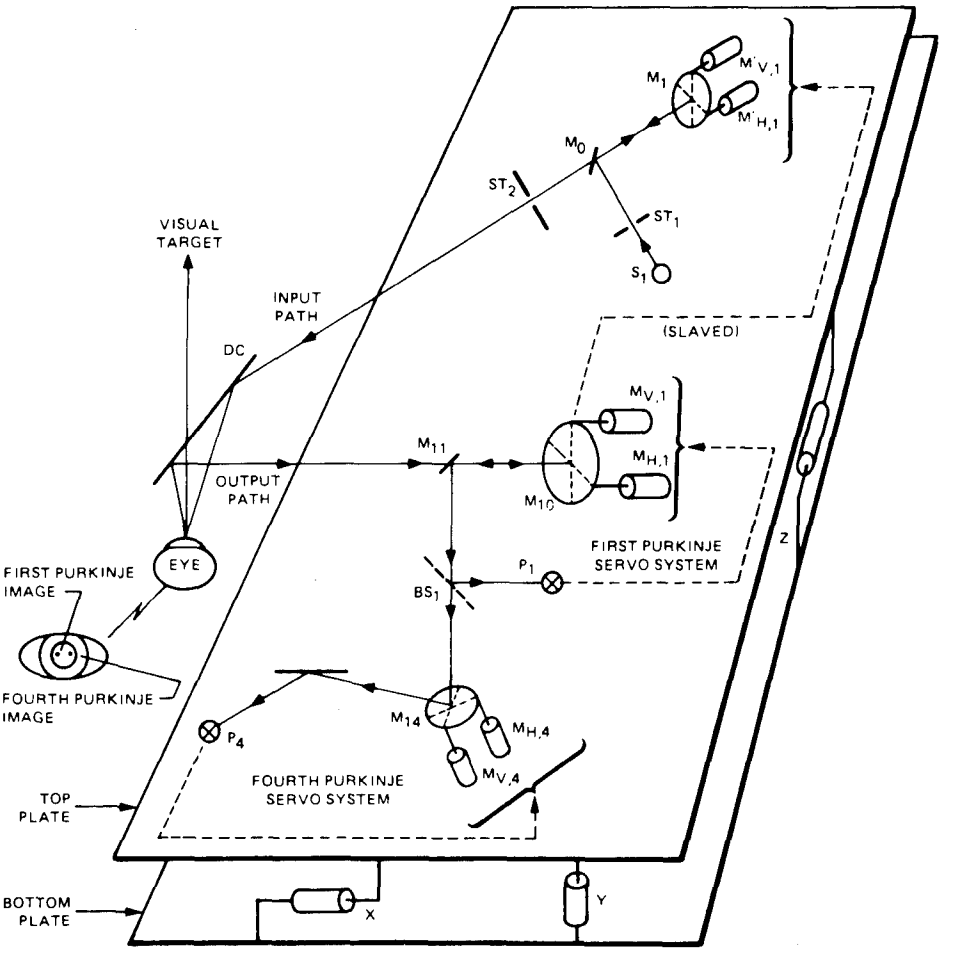


Figure 6.7 Simplified schematic of the SRI International dual Purkinje image eye tracker, from Crane and Steele (ref. 17).

The measurement noise level implies a potential accuracy of better than 1 arc minute. The measurement range is normally about ±10 degrees visual angle in both axes and ±15 degrees visual angle if drops are used to dilate the subject's pupil. Head motions of up to about ±2 mm can be tolerated. When an autostaging mechanism is activated, slow head motions can be tolerated up to ±25 mm along the fore-aft and horizontal axes, and ±12.5 mm along the vertical axis. Note that the system is far more precise and has a much higher bandwidth than available pupil-to-CR systems, but is also more obtrusive and does not permit as much subject freedom of motion.

7.0 CONCLUSIONS

The most accurate instruments available for measuring eye line-of-gaze are the scleral coil and the dual Purkinje image eye tracker, both of which can measure to within about one arc minute of visual angle. The scleral coil, however, is too invasive for nonlaboratory uses. The

dual Purkinje image technique has a very limited range (±10 degrees) and current implementations require a relatively large table-mounted optical apparatus.

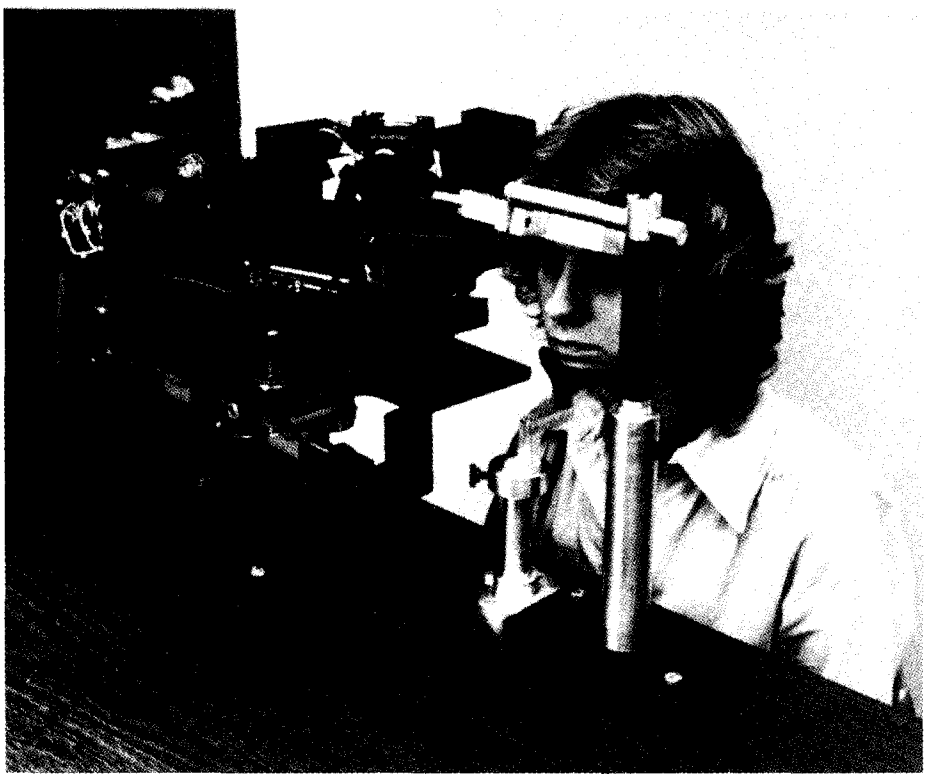


Figure 6.8 SRI International dual Purkinje image eye tracker, from Crane and Steele (ref. 17).

Electro-oculography is a useful technique for measuring eyeball dynamics but, because of drift problems, is not appropriate for measuring absolute eye position.

It is possible to make very precise measurements with some single feature tracking techniques, for example the University of Toronto corneal reflex tracker, but only if the head is very rigidly stabilized or head position is very precisely measured (to within better than 0.1 mm) with respect to the optics.

The pupil-to-CR technique is not as precise as the scleral coil or dual Purkinje image techniques, but, like these techniques, it will work in the presence of some head motion with respect to the optics. Current implementations have accuracies of about one degree visual angle. Successful implementations of this technique require processing of relatively complete eye image information. High frequency response (measurement update rate > 60 Hz) is, therefore, more difficult to achieve than with some of the other techniques, but continuing improvements in solid-state camera technology and in digital computer technology are reducing this difficulty.

8.0 REFERENCES

1. Mowrer, O.H., Ruch, R.C., and Miller, N.E. 1936. The corneoretinal potential difference as the basis of the galvanometric method of recording eye movements. American Journal of Physiology. 114:423.

2. Shackel, B. 1967. Eye movement recording by electro-oculography. A Manual of Psychophysiological Methods. P.H. Venables and I. Martion, Eds. North-Holland Publishing Co., Amsterdam.

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

4. Robinson, D.A. 1963. A method of measuring eye movement using a scleral search coil in a magnetic field. IEEE Transactions on Biomedical Electronics. BME-10:137-145.

5. Collewijn, H., van der Mark, F., and Jansen, T.C. 1975. Precise recording of human eye movements. Vision Research. 15:447-450.

6. Roth, E.M. and Finkelstein. 1968. Light environment. Compendium of Human Response to the Aerospace Environment, Volume I, Section 2. E.M. Roth, Ed. NASA CR-1205(I).

7. Cornsweet, T.N. and Crane, H.D. 1973. Accurate two-dimensional eye tracker using first and fourth Purkinje images. Journal of the Optical Society of America. 63:921.

8. Merchant, J., Morrissette, R., and Porterfield, J.I. 1974. Remote measurement of eye direction allowing subject motion over one cubic foot of space. IEEE Transactions on Biomedical Engineering. BME-21(4):309-317.

9. Hercher, M., Wyntjes, G., and DeWeerd, H. 1987. Non-contact laser extensometer. Proceedings of the International Society for Optical Engineering. 746:185-191.

10. Light, W. 1982. Non-contact optical position sensing using silicon photodetectors. Technical Report, United Detector Technology, Hawthorne, CA.

11. Torok, N., Guilleman, V., and Barnothy, J.M. 1951. Photoelectric nystagmography. Annals of Octology, Rhinology and Laryngology. 60:917.

12. Richter, H.R. and Pfaltz, C.R. 1956. Apropos de l'electro-oculographie (repport preliminair sur une nouvelle methode). Confinia Neurol. 16:279.

13. Stark, L., Vossius, G., and Young, L.R. 1962. Predictive control of eye tracking movements. IRE Transactions on Human Factors in Electronics. HRE-3:52.

14. Engleken, E.J., Stevens, K.W., Wolfe, J.W., and Yates, J.T. 1984. A limbus-sensing eye movement recorder. USAFSAM-TR-84-29.

15. Frecher, R.C., Eizenman, M., and Hallet, P.E. 1984. High-precision real-time measurement of eye position using the first Purkinje image. Theoretical and Applied Aspects of Eye Movement Research. A.G. Gale and F. Johnson, Eds. Elsevier Science Publishers B.V., North-Holland.

16. Defining Vision, NAC Eyemark Recorder. 1987. Product brochure. Distributed by Instrumentation Marketing Corp, Burbank, CA.

17. Crane, H.D. and Steele, C.M. 1985. Generation V dual-Purkinje-image eye tracker. Applied Optics. 24:527-537.

☆U.S. Government Printing Office: 1989 — 648-002/00207