NASA CONTRACTOR REPORT

NASA CR-11
C.1

NASA CR-1121

RESEARCH STUDY OF A FUNDUS TRACKER FOR EXPERIMENTS IN STABILIZED VISION

by D. H. Kelly and H. D. Crane

Prepared by
STANFORD RESEARCH INSTITUTE
Menlo Park, Calif.
for Ames Research Center

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • SEPTEMBER 1968

NASA CR-1121

RESEARCH STUDY OF A FUNDUS TRACKER FOR EXPERIMENTS IN STABILIZED VISION

By D. H. Kelly and H. D. Crane

Prepared under Contract No. NAS 2-3995 by
STANFORD RESEARCH INSTITUTE
Menlo Park, Calif.

for Ames Research Center

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

For sale by the Clearinghouse for Federal Scientific and Technical Information
Springfield, Virginia 22151 – CFSTI price $3.00

ABSTRACT

An image stabilization method is discussed that involves the tracking of retinal blood vessels and other structures of the back of the eye--i.e., a fundus tracker. The method involves projecting a scanning pattern onto the retina, and detecting the translational and rotational movements of the reflected pattern by means of a certain type of high-speed (real-time) correlation processing of the video signal. A particularly simple, circular scan method for this purpose is implemented by inverting a standard fundus camera in order to project the scan pattern onto the retina. The correlation processing is presently simulated on a digital computer at slow speed. This simulation program will be used for overall evaluation of the tracking system as well as to help in design of the required high-speed correlation equipment. This fundus tracker technique may provide greater precision of tracking than is provided by contact-lens techniques, as well as the potential for greater convenience of use.

CONTENTS

ILLUSTRATIONS

I INTRODUCTION

A. Retinal Signal Processing

Optical images formed in the plane of the retina are converted into spatio-temporal neural patterns in the optic nerve. However, these signals have already lost certain selected kinds of information, as a result of the signal-processing operations at the retinal level. Generally speaking, the retina acts as a spatial and temporal bandpass filter. The retinal circuits completely, or almost completely, block information contained in zero-frequency spatial and temporal signals. Thus, if the retinal image were uniform in intensity and of infinite extent, no information about absolute light intensity would be transmitted; also, if one could eliminate temporal changes in the light intensity on all parts of the retina, no signals about spatial distribution would be transmitted. At the other extreme, the eye also blocks high spatial and temporal frequencies, so that very fine details cannot be seen, and very fast periodic changes in intensity cannot be discriminated from steady illumination.

The statements above are qualitative, and at the present time they cannot be given the precision desirable. We cannot yet state precisely the temporal and spatial transfer characteristics of the human visual system, for reasons that will now be explained.

Even when an observer tries to fixate steadily, his eyes are in constant motion. This motion is referred to as involuntary eye movement, or physiological nystagmus. Each movement of the eye shifts the retina under the retinal image. Therefore, when an observer looks at a patterned field, each movement of the eye causes a temporal change in the intensity of light falling on many receptors. Furthermore, eye movements are irregular and unpredictable. Thus, if one wishes to study the response of the visual system to spatial or temporal changes in intensity, he must eliminate the motion of the image with respect to the retina. The only way to eliminate the effects of these movements,

without interfering with the normal motor system of the eye, is to move the retinal image along with the eye so that it remains fixed with respect to the retina. Such a condition is called a stabilized image.

In this sense, the blood vessels and nerve fibers lying in front of the retina, through which light must pass to reach the receptor cells, are "stabilized." Their unchanging nature is the main reason why these retinal features are normally invisible to the subject, though he can be made aware of them as transient phenomena, by various techniques--for example, by sudden changes in the wavelength or polarization of the stimulus. Some attempts have been made to use these methods to study retinal responses, but they suffer from the problem that these entoptic stimulus patterns cannot be controlled or changed.

B. Previous Image Stabilization Techniques

There are two principal ways in which researchers have attempted to produce stabilized images. The most common way is to attach a tightly fitting contact lens to the eye, and then either attach the test object to the lens or reflect an image from a mirror attached to the lens and through an optical system such that the retinal image of the object is fixed with respect to the retina. For many purposes, such an opto-mechanical system is adequate; but its degree of stabilization depends critically upon how closely the contact lens follows the angular movements of the visual axis of the eye. It has been shown that such a system cannot achieve stabilization sufficient for very precise work.

The second procedure for producing a stabilized image consists of a device for tracking the movements of the eye and moving an object or image proportionately. At the present time, the best way to track movements of the eye is to track movements of a contact lens attached to the eye; this method is therefore no more precise than the first technique described above.

C. Fundus Tracker

The eye-tracking method that is the subject of this report involves the electro-optical tracking of retinal blood vessels and other structures

at the back of the eye--i.e., a fundus tracker. The potential of the present technique is twofold: potentially greater precision of stabilization than provided by contact-lens techniques; and the potential for greater convenience of use, so that larger numbers of subjects can be tested. (Good stabilization with contact lenses requires a very accurately fitted individual lens for each subject; this requirement is both costly and difficult, and sometimes painful as well.)

The overall program has three major goals:

(1) To determine the optical properties of the retinal structures, primarily with respect to wavelength and spatial distribution, which govern the precision of "lock-on" for various automatic tracking devices; and to select the optimum wavelengths and the size and location of the retinal area to be tracked.

(2) To develop a working model of a system for automatically tracking extremely fine, angular motions of the visual axis (the operation of this system should preferably not be distracting to the subject).

(3) To demonstrate the usefulness of such a prototype tracking device by coupling its output to correspondingly deflectable stimulus patterns (generated, for example, with a CRT), to provide stabilized retinal stimulus images of greatly improved stability.

This report is primarily concerned with Objective (1). Some discussion of the principles, algorithms, and designs which have been evolved for use in proceeding toward Objective (2) is also included. Before treating technical details, however, we discuss below some of the kinds of experiments that could profitably be conducted with a successful instrument of this type. These experiments have generally been conducted in the past with contact-lens stabilization methods, and as often as not their results have been inconclusive because of the above-noted limitations of the contact-lens techniques.

D. Applications to Basic Experiments in Vision

"Spontaneous" Reappearance. Studies performed with contact-lens stabilization procedures have all indicated that the fixated object disappears, and then reappears from time to time. Two alternative hypotheses have been offered to account for this reappearance. It has been suggested that the reappearance is a manifestation of some unknown neural process which re-establishes vision. According to this interpretation, the form of the regenerated images may furnish insights into the nature of neural mechanisms. Alternative explanations are that the reappearances are caused by slippage of the contact lens (and hence are artifacts of the apparatus), or by lack of rigidity of the eyeball itself.

It has been demonstrated that slippages as small as a few seconds of arc are capable of causing a stabilized image to reappear. Since it is impossible to be sure, with a contact-lens system, that errors this small will not occur, the cause of image reappearance cannot be unequivocally determined without an improved stabilization device.

Closely related to this problem is the question of how long a stabilized image takes to disappear. With current devices, this time seems to change with the parameters of the stimulus pattern, and the results are highly variable. This would be expected if the process that causes an invisible image to reappear also operates while the image is visible. For example, it is reasonable to assume that slippage of the lens must prolong the time to disappear. Further, the influence of any such artifact would depend upon the parameters of the stimulus itself (e.g., a given slippage will be more effective in preventing a high-contrast image from disappearing than a low-contrast one). Measurements of the time of disappearance could be made with much more precision under conditions of optimum stabilization.

Spatial and Temporal Response Functions. The action of the visual system in transmitting temporal and spatial information may be evaluated by measuring the spatial and temporal frequency response curves for human vision. However, this procedure again cannot be carried out

adequately unless the uncontrollable changes produced by eye movements are eliminated.

From analyses of the spatial and temporal response functions, it is possible to infer many of the physiological and anatomical characteristics of the human visual system. For example, the spatial transfer function obtained with short flashes (which contain all temporal frequencies) indicates that there is strong lateral inhibition in the human visual system, and yields quantitative measures of the spatial parameters of that inhibition.

Spatio-Temporal Interaction. On the available evidence, the spatial and temporal frequency-response curves are not independent of each other. At low frequencies, for example, the temporal response is strongly dependent on the spatial pattern of the stimulus, while at high frequencies it is not. The most complete set of data that could be taken under these circumstances consists of the spatial response curve for every temporal frequency (or vice versa). Apparatus for mapping this complete threshold surface is available, but only with an unstabilized retinal image.

It seems significant that the velocity of the so-called "drift" component of involuntary eye movements is approximately equal to the ratio of the most effective temporal frequency to the most effective spatial frequency (in the various spatial and temporal response experiments that have been conducted so far). This suggests that the spatial response in the presence of image motion is at least partly controlled by the temporal response. It would be highly desirable to separate the effects of image motion from the other properties of the visual system by performing spatio-temporal experiments with a stabilized image.

Color-Vision Mechanisms. There is good in-vitro anatomical evidence that the human retina contains three populations of cone receptors, each population containing a different one of three visual pigments. However, no one has yet been successful in demonstrating the perceptual correlates of this fact. For example, the simplest theory would predict that the stimulation of any single cone would give rise to the

perception of one of three colors. However, it may be true that color perception occurs only when particular combinations of cones are stimulated simultaneously. The excitation of small sets of cones has been attempted by stimulating the eye with points of light in very short flashes, but the results are equivocal, possible because of the shortness of the flashes. The experiment has also been attempted with longer exposures in a contact-lens stabilized image, but again the results are equivocal, probably because of the imperfect stabilization achieved. The results of this type of experiment would be more clear-cut with more precise image stabilization.

Stereopsis. The motions of the two eyes are correlated, but the correspondence is not perfect. (The variance of the difference in direction of motion of the two eyes is about the same as the variance of the motion of each eye alone.) Therefore, stereoscopic vision may be strongly affected by eye movements. Attempts to study this question by producing a stabilized image in both eyes at the same time, using two contact lenses and associated optics, have also yielded equivocal results.

Eye-Movement Control Mechanisms. In order to study the fine control of eye movements themselves, it is frequently desirable to "open the loop"--that is, to interfere with the operation of the physiological eye-movement control system by preventing the movements from producing their usual visual consequences. The contact-lens stabilization technique may be precise enough for such studies, but only when the contact lens is very carefully fitted. Thus it is difficult to study a large number of observers. The device that we are developing may be usable with a large number of observers and, for this reason, may provide useful information about individual differences in eye-movement control. This should be of particular interest with respect to neurological diseases.

Flicker Sensitivity. At low frequencies and high brightnesses, more flicker sensitivity is obtained with a small circular target (on a dark surround) than with a wide-field target. To test the hypothesis that this behavior is due to eye movements across the sharp edges of

the small target, the same experiment has also been conducted with the target stabilized (by the contact-lens technique). The results were negative, but again the stabilization was probably imperfect (e.g., the target did not completely disappear during the several seconds in which the threshold settings were made).

It has also been shown that subliminal flicker in an annular surround can affect the low-frequency threshold in a small circular field. These experiments did not use a stabilized image, and the target field was only 40 minutes in diameter (which is not much greater than the largest involuntary eye movements). It would be of considerable significance to conduct this type of experiment with a precisely stabilized image.

E. Experiments of Interest in Applied Fields

Information Displays. The bandpass characteristics of the human visual system may be used to optimize certain forms of visual displays. For example, if a display were introduced to an observer through an image-stabilizing device, all aspects of the display that were not changing in time would disappear, and only changing signals would be visible. Moving objects would thus be particularly noticeable. If photographs taken on successive days were presented in movie form through an image-stabilizer, objects new on the scene would be immediately observed. Similarly, selected stationary objects could be made to change in brightness in a stabilized display, so that they alone would be visible.

Medical Applications. Very precise tracking without the necessity of fitting a contact lens may be useful in medicine. For example, optically induced lesions (e.g., with laser light) may be made with greater precision. The laser could be triggered from an eye-tracking device so that its rays will strike one precisely defined part of the retina.

Ophthalmoscopic examination of the retina currently provides a great deal of information about neurological, circulatory, and other

general conditions of the patient. However, the resolution of such an examination is limited by the movements of the patient's eyes. If the examination were carried out through an eye tracker, this limitation would be eliminated, and microscopic structure could be examined in much more detail.

The resolution of retinal photographs is similarly limited. The light must be delivered in a very short flash, in order that the photograph is not blurred by eye movements. As a consequence, the flash must be exceedingly intense. Photographs through an eye tracker could be taken with longer exposures, thus increasing the resolution and decreasing the required intensity of illumination.

Bandwidth Compression. The transmission and storage of visual information (as in the form of motion pictures or television signals) at present assumes that the spatial and temporal response coordinates of the ultimate receiving system--the human eye--are independent. In fact, this is not the case, as much of the work mentioned above confirms. Thus it may not be necessary to provide a completely independent (spatial) picture element for each new (temporal) sampling instant, or vice versa, as present transmission and storage systems now do. In order to design suitable encoding systems, however, it is necessary to know the exact nature of the interdependence between these coordinates in the visual system. In other words, we must obtain direct and complete measurements of spatio-temporal interaction, both with and without eye movements.

II BASIC SCAN TECHNIQUE

Direct tracking of the fundus pattern may be appealing, but the question is just how to go about achieving such tracking. An analysis by Cornsweet,* of the possibility of tracking the edge position of a single retinal blood vessel with a very small, intense spot of light, showed that to achieve the required signal-to-noise ratio for satisfactory tracking, the spot would have to be so intense that there would be serious risk of tissue damage. The same amount of energy spread out over the entire retina would be quite harmless, however, and therefore useful for some form of "correlation tracking." In the latter case the tracking information could be obtained from a large area of the retina--e.g., the blood vessel pattern, and perhaps other structures as well. A device for tracking the retina in this way might operate in the following fashion:

(1) A real image of the retina is formed externally, as in a fundus camera; motions of this image then correspond to motions of the eye.

(2) A previously recorded image of this retinal pattern is stored in some type of memory.

(3) An image-correlation device is used to compare the stored pattern with the present pattern.

In general, such a device cannot measure the displacement directly, but can only decide whether a significant displacement has occurred. It must therefore be incorporated into a servo loop, which moves the real image of the retinal pattern until it is in registration with the stored pattern. The x-, y-, and torsion-component motions that are

---

*T. N. Cornsweet, Recent Developments in Vision Research, M. A. Whitcomb, Ed., Armed Forces--NRC Committee on Vision, Publication 1272 (National Academy of Sciences, National Research Council, Washington, D.C., 1966).

required to keep the real-time pattern aligned with the stored pattern are the three measures that we seek. Simultaneously, these component motions could be applied to any desired target, providing automatic compensation of its image motion at the retina.

In explaining the basic scanning system that has evolved from our studies, we will not be concerned initially with energy level, noise, or other requirements for use with a real eye. First, we will simply assume that we have a two-dimensional object, such as a photograph, whose position we wish to track. As noted above, to detect movement we must compare the present position of the pattern, or at least some portion of it, with a previous measure of its position. This requires a storage function of some sort, with facility for subsequent comparison. The brute-force approach would be to record an entire copy of the two-dimensional scene--obtained, for example, by a TV scan--and compare the entire pattern frame-by-frame. In order to minimize bandwidth, storage capacity, and cost, however, let us consider whether a one-dimensional scan could not provide all three motion components--i.e., x, y, and α (torsion). We will now show how a single-line circular scan, as in Fig. 1(a), can in fact provide the necessary information.

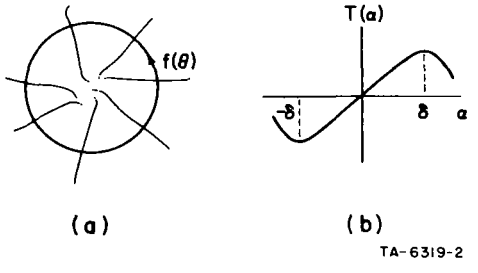


FIG. 1 (a) CIRCULAR SCAN RASTER ON SCHEMATIC BLOOD VESSEL PATTERN
(b) FORM OF OUTPUT CONTROL CHARACTERISTIC WITH TORSIONAL MOVEMENT

Let $f(\theta)$ be the real-time video signal obtained from such a circular scan, and assume for the moment that we are interested only in detecting torsional changes--i.e., rotation of the pattern about the center of the

circular scan. Let $f_o(\theta)$ be the video signal that we obtain when the pattern is in its nominal position, so that for a pure rotation α, the resulting signal is $f(\theta) = f_o(\theta - \alpha)$. The problem of course is to determine α.

If we simply form the cross-correlation:

$$C(\alpha) = \int_0^{2\pi} f_o(\theta) f(\theta)\, d\theta = \int_0^{2\pi} f_o(\theta)\, f_o(\theta - \alpha)\, d\theta \qquad (1)$$

then we cannot find the direction of rotation (i.e., the sign of α), because the cross-correlation function $C(\alpha)$ is an even function of α--i.e., $C(\alpha) = C(-\alpha)$. But if the output signal is to be used in a closed-loop servo system to stabilize the pattern position, the direction of movement is in fact even more important than the magnitude. To determine the polarity of the signal we need two reference signals to compare against, so that we can note whether the movement brought us closer to one or the other. Suppose, for example, that instead of storing just the function $f_o(\theta)$, we initially store two copies, $f_o(\theta - \delta)$ and $f_o(\theta + \delta)$, slightly rotated from each other. Now we form the function:

$$T(\alpha) = \int_0^{2\pi} f_o(\theta - \delta)\, f_o(\theta - \alpha)\, d\theta - \int_0^{2\pi} f_o(\theta + \delta)\, f_o(\theta - \alpha)\, d\theta \qquad (2)$$

where T is the function of torsional change, α. For $\alpha = 0$, we see that $T(\alpha)$ is also zero, since the basic cross-correlation is an even function of δ--i.e.,

$$\int_0^{2\pi} f_o(\theta - \delta)\, f_o(\theta)\, d\theta = \int_0^{2\pi} f_o(\theta + \delta)\, f_o(\theta)\, d\theta \quad . \qquad (3)$$

More generally, we see that

$$T(\alpha) = -T(-\alpha) \quad .$$

That is, T is in fact an odd function of α, which yields the desired polarity information, as shown in Fig. 1(b).

The magnitude of delay δ used in forming the two storage functions $f_o(\theta - \delta)$ and $f_o(\theta + \delta)$ adjusts the sensitivity of the system. In the present case, the magnitude of δ would probably be set close to the

resolution limit of the system--namely, about one minute of arc. (General discussion of system specifications is delayed until Sec. IV-A.) But regardless of the choice of S, for small δ, the torsional function will have the form sketched in Fig. 1(b).

Actually, the torsional control function $T(\alpha)$ as expressed in Eq. (2) would be very difficult to obtain directly, for physical reasons. In the present application, we are looking for extreme sensitivity (δ set near the resolution limit of the system), and the light levels at best will be very low, so that we expect a relatively poor signal-to-noise ratio in the video signal. Thus, the difficulty is that $T(\alpha)$ as expressed in Eq. (2) is a small difference between two relatively large numbers, so that the direct approach of storing two noisy copies, $f_o(\theta - \delta)$ and $f_o(\theta + \delta)$, obtained initially from two separate scans, and then forming two cross-correlation functions in two independent devices, and finally differencing, could hardly be expected to work.

Fortunately, the function in Eq. (2) is distributive, and can be rewritten in the form

$$T(\alpha) = \int_0^{2\pi} [f_o(\theta - \delta) - f_o(\theta + \delta)]\, f_o(\theta - \alpha)\, d\theta \qquad (5)$$

which shows that we need only perform the delay and differencing of a single initial scan for the storage function, and use only a single correlator to achieve the control function. If we define f_δ as the storage function--i.e.,

$$f_\delta = f_o(\theta - \delta) - f_o(\theta + \delta) \qquad (6)$$

then Eq. (5) can be rewritten as

$$T(\alpha) = \int_0^{2\pi} f_\delta\, f_o(\theta - \alpha)\, d\theta \qquad (7)$$

or more generally,

$$T(\alpha) = \int_0^{2\pi} f_\delta\, f(\theta)\, d\theta \quad . \qquad (8)$$

A system for instrumenting the function $T(\alpha)$ in this fashion is shown schematically in Fig. 2, where, for simplicity of notation, the

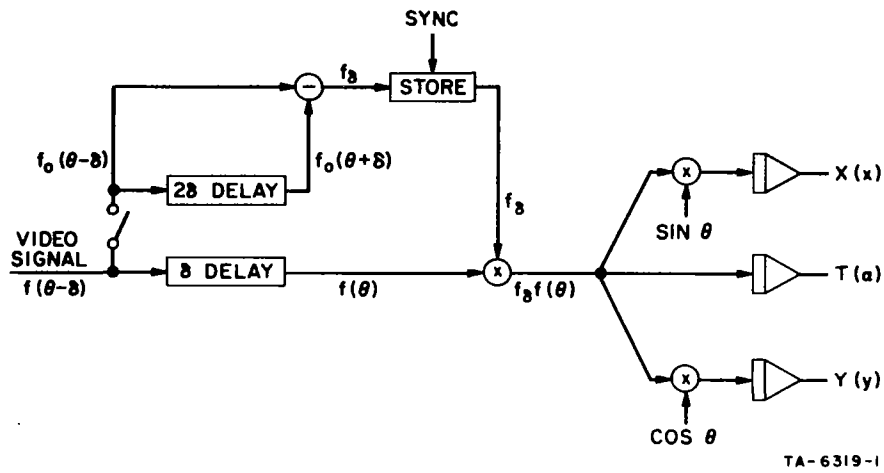


FIG. 2 CROSS-CORRELATING THE VIDEO SIGNAL FROM THE SCANNER AGAINST A PREVIOUSLY STORED MASTER, TO DETECT X, Y, AND T MOVEMENTS

input video signal is labeled $f(\theta - \delta)$. With the use of a 2δ delay line we can form the stored reference, f_δ. This is done only once, initially, as indicated by the switch. (Note that the reference function, f_δ, now has no dc component to be stored, which is an important advantage.)

While tracking, f_δ is continuously correlated with the real-time video signal, $f(\theta)$ (i.e., multiplied and integrated over the duration of one frame) to determine the torsion feedback signal. As we saw in Fig. 1(b), the extreme positive and negative values of this torsion signal occur when the fundus has rotated by just the amount $\pm\delta$, but a roughly proportional correction will be applied everywhere within this range.

Thus far we have discussed only the torsional component of motion. Actually, we can determine the x and y components of motion in the same manner as the torsional component, if before integrating we take the proper components of the video signal--multiplying by $\sin\theta$ or $\cos\theta$, as shown in Fig. 2--so that we have

$$X(x) = \int_0^{2\pi} f_\delta \, f(\theta) \, \sin\theta \, d\theta \tag{9}$$

and

$$Y(y) = \int_0^{2\pi} f_\delta \, f(\theta) \, \cos\theta \, d\theta \quad . \tag{10}$$

The control functions X(x) and Y(y) have the same basic form as that shown in Fig. 1(b), except that the extreme positive and negative values occur for fundus movements in the x or y direction of magnitude $\pm R\delta$, where R is the radius of the scanning circle. Again, a proportional correction will be made for smaller displacements, limited only by the signal-to-noise ratio.

To show that these quadrature components contain the maximum available position information (and only that), let us follow through a simplified case of X displacement, as shown in Fig. 3. Here the video signal is assumed to consist of two broad pulses occurring approximately at the top and bottom of the scanning circle. The first four lines of the figure illustrate the formation of the stored difference signal f_δ, from the reference frame, f_o. Note that f_δ contains a positive pulse at each leading edge and a negative pulse at each trailing edge of the original broad pulses of f_o.

The fifth line shows a real-time video signal $f(\theta)$, in which a horizontal displacement of $+R\delta$ has now occurred. This advances the first broad pulse so that it correlates with the first positive pulse of f_δ, and retards the second broad pulse so that it correlates with the second negative pulse of f_δ, as shown in Line 6 of the figure. In this case, the product $f_\delta \cdot f(\theta)$ would integrate to zero; as it should, since we have postulated no torsion. But if this product is first multiplied by

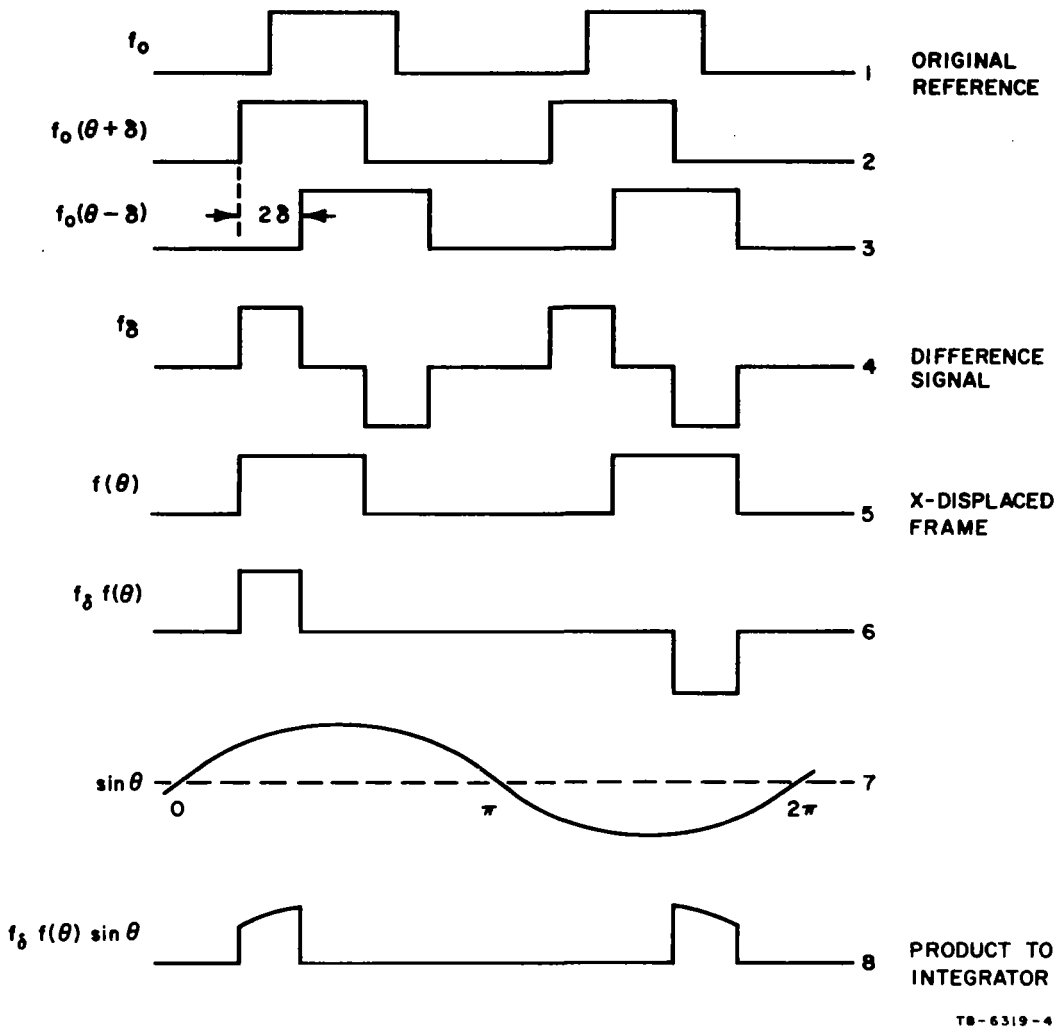


FIG. 3 TRACING THE CROSS-CORRELATION PROCESS FOR DETECTING HORIZONTAL MOVEMENT

$\sin\theta$, the sign of the negative pulse is reversed, and the integral of the last line is roughly proportional to the net positive x movement.

If the x displacement had been in the opposite direction, both pulses in the last line (and hence its average) would have been negative. An entirely analogous derivation of the Y displacement signal is obtained by integrating the product $f_\delta\, f(\theta)\cos\theta$. In the illustrated case of x movement only, this product would of course integrate to zero.

As in any correlator, major noise suppression results from integration over a complete frame; the use of the quadrature components makes this averaging as effective as possible without throwing away any significant position information. Thus, the video signal is most sensitive to x-displacement at $\theta = \pi/2$ and $3\pi/2$--i.e., the x-component of the scan velocity is greatest at these points--and these are just the points at which the integration is most heavily weighted by the sine wave in Fig. 3. If the product $f_\delta \cdot f(\theta)$ were multiplied by a square wave instead of the sine wave, this would not necessarily produce erroneous position information, but it would introduce considerably more noise in the vicinity of $\theta = 0$, $\theta = \pi$, etc.

With these principles governing the proposed scanning system, let us now consider how we may achieve such a scanning operation in the present application. In Sec. IV we will return to a more detailed set of specifications for the scanning process, and in particular to some initial results from a computer simulation of the process.

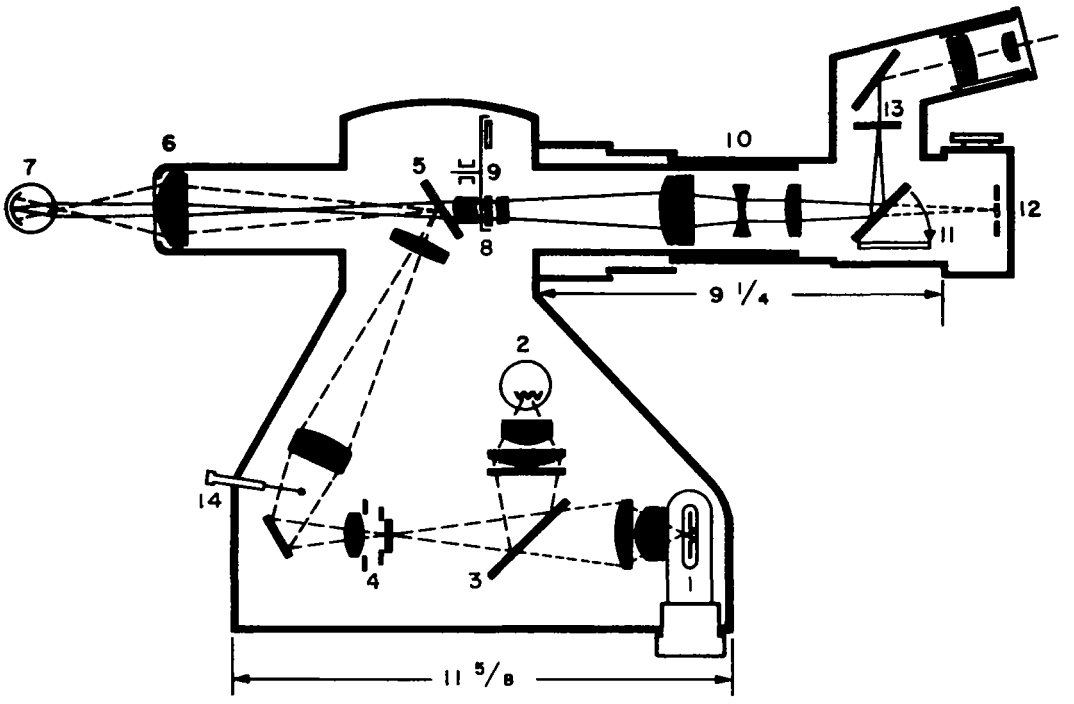
III THE FUNDUS CAMERA

In Sec. II we developed a set of control algorithms for eye-tracking, on the basis of a single-line circular scan of the fundus image. Let us now consider how we can physically realize such a scanning process, in real time, in the case of the human retina. A straightforward approach would be to use the optics of a fundus camera, but with a TV camera tube to scan the retinal image that is normally formed on the film. Though this technique might work, a significant improvement can be obtained by inverting the process--i.e., using a flying-spot, circular-scanning light source in the normal film plane (to be projected onto the retina), and putting a broadband photodetector in place of the normal light source. To appreciate the nature of the improvement, let us first review the basic operation of a high-quality fundus camera.

A. Basic Operation

The optical configuration of the new Zeiss fundus camera is shown in Fig. 4. In Fig. 5 we have separated the input and output paths to help explain the operation. The basic idea is to form an external image of the retinal pattern by means of input light that is reflected from the fundus and passes back out through the eye. The main problems in the design of such an instrument are (1) to provide enough light, (2) to maintain high resolution, and (3) to eliminate unwanted light scatter and reflections (primarily from the cornea) that might cause glare in the resulting photographs.

In this instrument, reflected light from the cornea is eliminated by using geometrically separate portions of the pupil for the input and output light. In normal use, the input light passes through an annular region of the pupil, where the inner diameter of the annulus is 2.3 mm. Reflected light from the retina, which forms the external image in the film plane, passes only through the central region of the annulus. This arrangement has the advantage that "high-resolution light"



1) ELECTRONIC FLASH BULB
2) DOUBLE FILAMENT BULB
3) GLASS PLATE
4) LUMINOUS FIELD DIAPHRAGM
5) MIRROR
6) ASPHERIC OBJECTIVE
7) PATIENT'S EYES
8) CORRECTION LENS
9) COMPENSATOR FOR ASTIGMATISM
10) 30° OBJECTIVE
11) TILTABLE MIRROR
12) PHOTOGRAPHIC FILM PLANE
13) EYEPIECE RETICULE
14) FIXATION TARGET

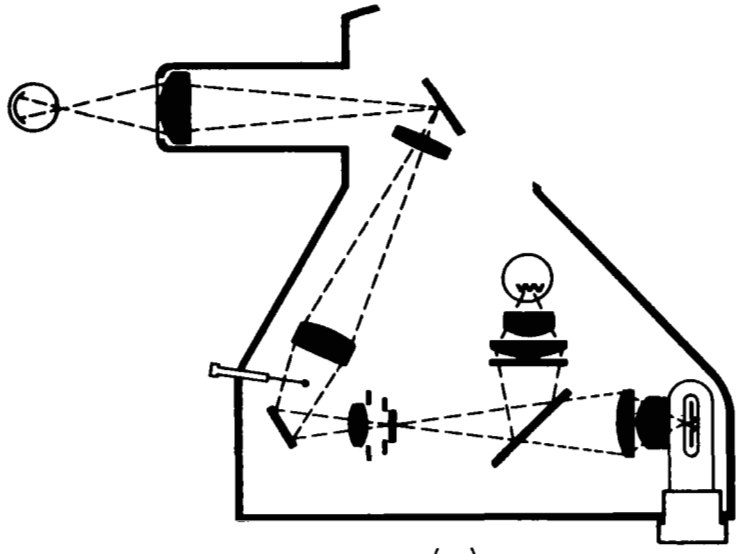
SOURCE: Zeiss Fundus Camera Operating Instructions (1965), issued by Carl Zeiss, Oberkocken, W. Germany.

TA-6319-3

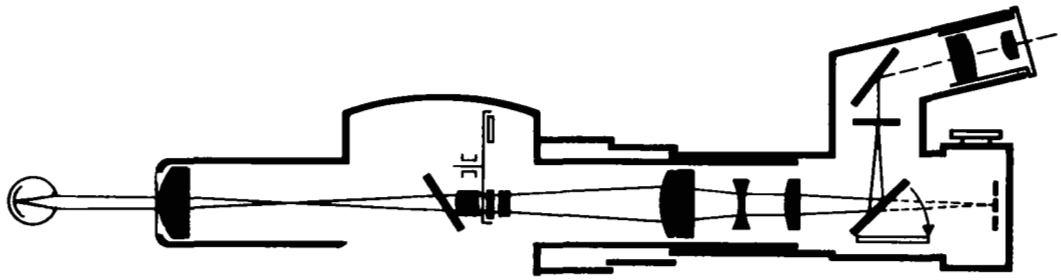
FIG. 4 OPTICS OF THE ZEISS FUNDUS CAMERA

for forming the external image passes only through the central region of the eye lens, which tends to minimize optical aberrations.

Let us now consider how this geometry is actually achieved. In Fig. 4, the input and output light paths separate at the mirror labeled (5), which has an elliptical hole that appears circular when projected toward the eye (7). In the input path, the flash tube (1) is imaged in the region (4) where the size of the annulus can be further limited by selection of one of several built-in aperture stops. This annulus in turn is imaged at the mirror (5) where the central zone is not reflected; the



(a)



(b)

SOURCE: Zeiss Fundus Camera Operating Instructions (1965), issued by Carl Zeiss, Oberkocken, W. Germany.

TA-6319-5

FIG. 5 SCHEMATIC OF THE INPUT AND OUTPUT OPTICAL PATHS OF THE FUNDUS CAMERA

resulting annular pattern is then focussed at the front of the eye, nominally in the subject's pupil plane. Light entering the eye through this annular zone illuminates the retina quite uniformly.

In the output system, Fig. 5(b), light can reach the camera plane (12) from three sources: unwanted light reflected from the objective lens (6); unwanted light reflected from the front surfaces of the eye; and the desired light originating from the retina. Lens (6) is designed to minimize any infocus reflections, so that very little reflected light can pass back through the small opening of the mirror (5). To understand how light reflected from the cornea is eliminated, consider the path of such reflected light. Neglecting aberrations, collimated input light is reflected from a convex mirror as though from a virtual source in the r/2 plane, where r is the radius of the mirror, in this case the radius of the cornea [see Fig. 6(a)]. Conversely, light converging to any point in the r/2 plane is reflected as a collimated beam. If the converging beam is displaced from the axis, as in Fig. 6(b), then the return beam will be at an angle to the optical axis. This angle is readily determined by noting that an input ray directed at the center of curvature is reflected on itself. Thus, the nominally collimated reflected beam will be parallel to the ray that passes through the point of convergence and the center of curvature [dashed ray in Fig. 6(b)]. The angle of reflection is simply

$$\beta = \tan^{-1} \frac{y}{r/2} \tag{11}$$

where y is the vertical displacement of the convergence point from the axis. To see the effect of these reflected beams, consider in Fig. 6(c) the system containing the eye together with the input lens (6) of the camera, shown in Fig. 4. Assume that the input annulus image is formed in the r/2 plane of the cornea, which almost exactly coincides with the real pupil plane of the eye. Assuming, further, that the eye is focussed for infinity, then light from the retina is collimated as it leaves the eye and is imaged in the focal plane of lens (6). Light from the corneal reflection is similarly collimated and therefore also forms an image

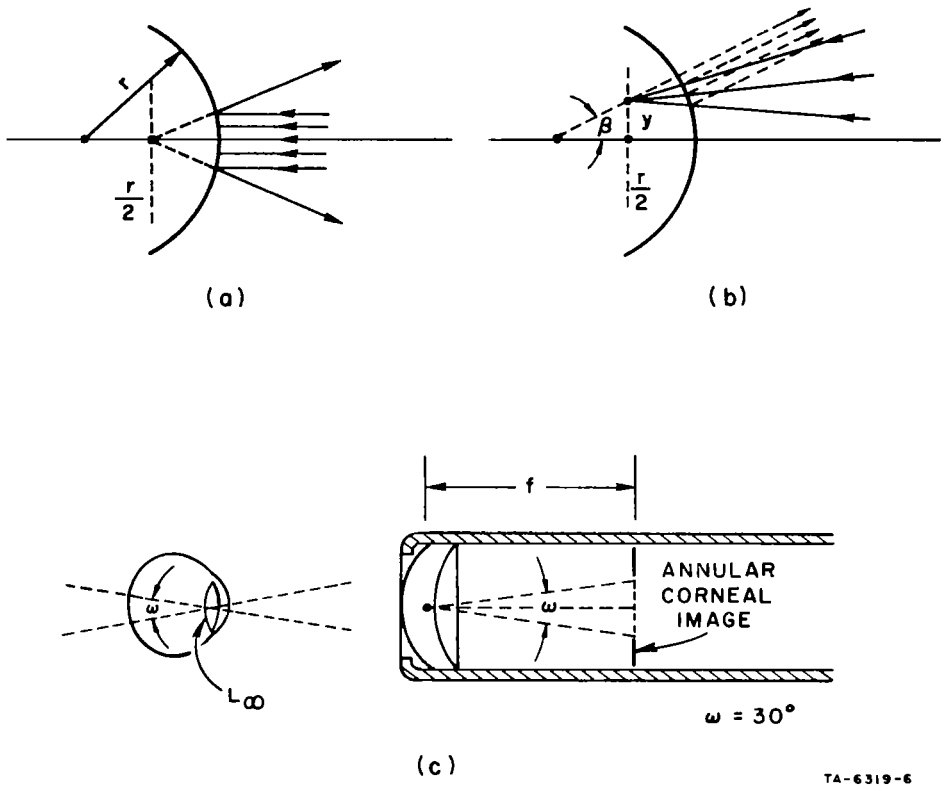


FIG. 6 HOW REFLECTED LIGHT FROM THE CORNEA IS BLOCKED FROM THE OUTPUT PATH

in the focal plane of lens (6). However, because of the annular shape of the input pattern, the reflected corneal light forms an annular image in the focal plane of lens (6), with an inner radius determined from Eq. (11). For an inner input diameter of 2.3 mm at the pupil plane, $y = 1.15$ mm and $r/2 = 3.9$ mm (the radius of curvature of the cornea is typically taken as 7.8 mm), or $\beta = 16.5$ degrees. In other words, for a 2.3-mm central hole in the input pattern, reflected light from the cornea is at an angle ≥ 16.5 degrees from the optical axis. This leaves a central clear area of at least $\omega = 30$ degrees for the retinal image, which is the field size of the Zeiss instrument.

Thus it should be clear that elimination of the corneal reflection is a rather delicate business, requiring reasonable alignment of the subject, both axially and laterally. Of course, alignment would be

simpler with a larger annulus hole, but this would result in decreased input light. A smaller annulus hole, in order to increase the amount of input light, would result in a smaller retinal field free of reflected corneal light, and poorer resolution. The design size of 2.3 mm seems to be a good compromise between a large field and reasonable tolerance in alignment.

Two other features of the fundus camera (Fig. 4) should be mentioned: the viewing path, via lens (13); and the input focussing light (2). For focussing the system in normal use, the focussing light (2) is energized and the retina is viewed directly through lens (13). [The focussing light enters the system through the beam-splitter (3).] After manual focus adjustment, the mirror (11) is depressed and the flash tube is energized for a very short, intense exposure.

B. Photographic Study of the Fundus

Using the fundus camera in its normal mode, we took a series of photographs using different spectral filters, in order to determine the effect of wavelength on contrast. (The subject's eyes were dilated with neosynephrine for these experiments.) In Fig. 7 is a series of black and white photographs taken on Eastman Type 5224 film, through a set of five narrow-band filters (bandwidths 30 to 40 mμ), centered at about 420, 480, 535, and 560 mμ. Four different intensity levels of the xenon flash, giving a total exposure range of 8 to 1, were used with each filter. From these experiments we concluded that:

(1) Greatest contrast is obtained in the wavelength range between 400 and 520 mμ. Longer wavelengths show a steady decrease in contrast toward the red end of the spectrum.

(2) The greatest amount of detail consists of blood vessels and other features in the vicinity of the optic disk. Much less detail can be seen in the vicinity of the fovea.

(3) The xenon flash provides enough energy for a finer-grained (and hence slower) film than Eastman Type 5224;

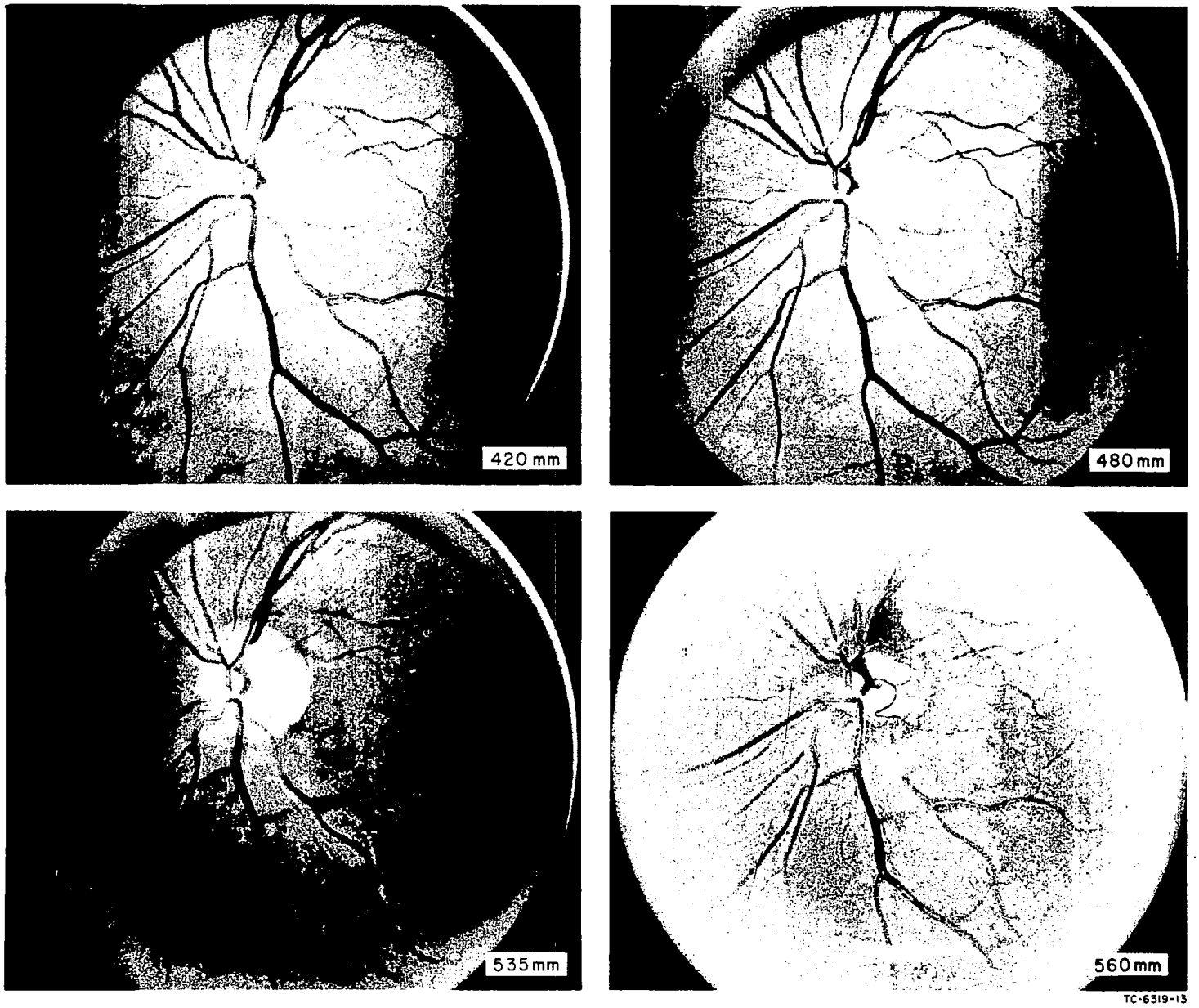


FIG. 7 FUNDUS PICTURES TAKEN WITH VARIOUS SPECTRAL BANDS

this is desirable to minimize granularity at the 2.5X magnification of the Zeiss camera. The second and third series were therefore exposed on Panatomic-X film.

In the second experiment, the optic disk was centered in the field of the Zeiss camera, and exposed through a series of broader blue filters, all at the highest intensity of the xenon flash. The short-wavelength cutoff was varied from 360 to 400 mμ and the long-wave cutoff from 460 to 560 mμ. From these experiments it was concluded that:

(1) A bandwidth of up to 100 mμ can be used in the blue region without degrading the contrast significantly.

(2) The short-wavelength cutoff should not be shorter than about 400 mμ, and the long-wavelength cutoff should not be longer than about 530 mμ.

(3) Extremely critical focus adjustment is required to take advantage of the finest details present in these fundus patterns.

In the third experiment, a dozen frames were exposed through each of the two best filter combinations, with a slightly different focus adjustment for each frame. The focus imcrements were small enough to be visually undetectable (in the view finder of the Zeiss camera). These negatives were then inspected microscoptically to select the sharpest frames for printing; the latter were enlarged 8 times (a total magnification of 20X from the original fundus pattern). A sample photograph of this type is shown in Fig. 8, at approximately 16X magnification.

Some of the details visible in the original picture appear to be quite close to the diffraction limit for a 2.3-mm pupil at 450 mμ--i.e., of the order of 1 minute of arc (or about 5 microns at the fundus). This photograph also confirms that the optic disk is a desirable location for the circular scanning locus. Fortunately, this relatively remote location will tend to minimize its interference with stabilized stimuli located at or near the fovea. Also note that a broad-wavelength

FIG. 8 A 16X MAGNIFICATION FUNDUS PHOTOGRAPH, SHOWING DETAILS CLOSE TO THE DIFFRACTION LIMIT FOR A 2.3 mm PUPIL AT A WAVELENGTH OF 450 MILLIMICRONS (About 1 minute of arc)

band in the blue region of the spectrum is readily attainable with existing phosphors and photocathodes.

C. Inverting the Optical Path

For the purpose of eye-tracking, we could in principle use the fundus camera in its normal configuration and simply scan the retinal image formed in the film plane. However, there are potentially great advantages to inverting the system. To appreciate these advantages, let us first consider a scanner using the fundus camera in its normal photographic mode. Consider, for example, the following arrangement:

(1) A steady light source of the proper spectral content in place of the flash tube in Fig. 4.

(2) An additional stop containing an annular opening placed in the input system, so that it is in focus on the retina (in this way we only illuminate a small annular region of the retina).

(3) A vidicon camera tube in the film plane. The image of the annular region of illuminated retina, which falls on the target of the vidicon, is scanned in a circular path by an electron beam centered within this annular zone.

There are two potential difficulties with this system. First, the annular zone must be considerably larger than the circular scanning spot, in order to permit reasonable tolerances on alignment and drift. This results in much more total energy into the eye than is actually used in the measurement. This would tend to close down the pupil of the eye more than necessary; but even more important, the unused light would decrease the contrast of the retinal image if there is any scattering, either within the retinal layer or optical media of the eye, or any reflections from the internal walls of the eyeball. Any light that increases the background level necessarily decreases the contrast.

Let us now see how these potential difficulties are eliminated by inverting the system. In the inverted system, we place a flying spot in

the plane of the film and a sensitive photocell in place of the flash tube. The normal high-resolution output system thus becomes a high-resolution input system that focusses an image of the flying spot onto the fundus, by way of light passing through the central zone of the pupil. Light reflected from the fundus passes out through the annular zone of the pupil and is collected by the photocell. Thus the temporal signal is a measure of the reflectivity of the fundus along the scanning locus. In this way we put into the eye no more light than necessary for the measurement, and internal scattering after the fundus reflection now merely affects the net gain of the system, not the picture contrast.

A photograph of the inverted scanning system is shown in Fig. 9. The cathode ray tube is a very-high-resolution tube from Westinghouse, with an experimental blue phosphor known as PX38. This tube has almost exactly the spectral characteristics we desire, as well as a very short persistence (nominally 0.6 μs). The tube is designated WX-5062PX38. The photomultiplier is an Amperex 56VP unit with an S-11 photocathode (in a PA56 mounting assembly), which is a very-high-gain tube.

Using the apparatus shown in Fig. 10, with the "model eye" shown in Fig. 11, we obtained the "video signal" shown in Fig. 12. Though a quite preliminary result, this photograph demonstrates that signals are at least obtainable with this geometric configuration. The model eye consists simply of a lens of suitable focal length (adjusted so that reflections from its surfaces are eliminated, in the manner of corneal reflections from the real eye), and a "retina" which is simply a photographic print of the fundus picture of Fig. 8. The signal recording of Fig. 12 is a relatively long scope-camera exposure, in an attempt to average the noise-in-signal. The result of this averaging is the line-widening seen in the trace. An estimate of signal-to-noise ratio is obtained by comparing the line width to the amplitude of the signal variations.

While Fig. 12 provides evidence that signals can actually be obtained with this configuration, there is still considerable work required to "tune up" the system--e.g., adjusting the field size of the

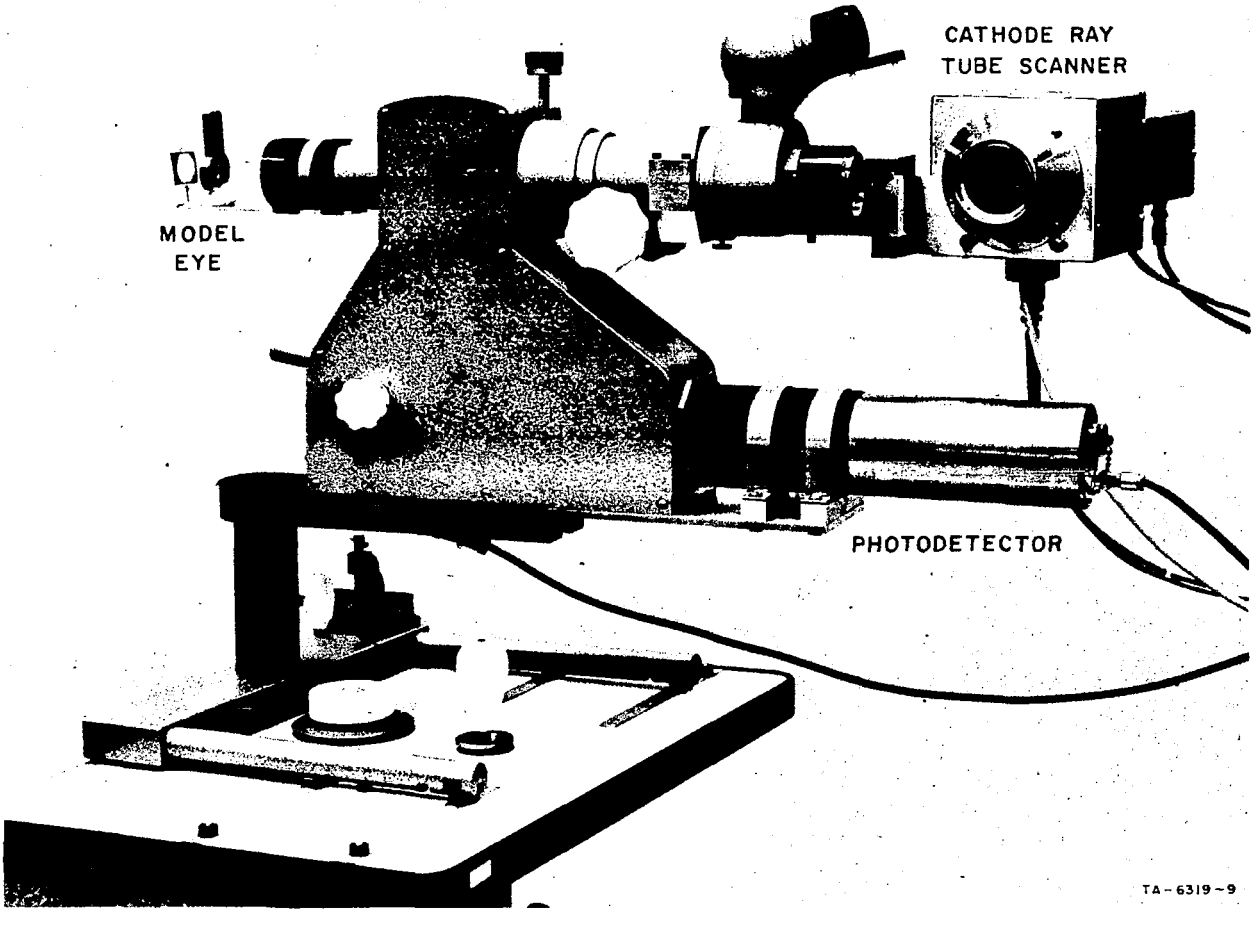


FIG. 9 MODIFIED FUNDUS CAMERA WITH A MODEL EYE IN PLACE OF A REAL EYE

photodetector system for optimum signal-to-noise ratio, adjusting the cathode ray tube for minimum spot size, and obtaining a higher-resolution "retina" (i.e., more like a real eye).

Although such tuning of the system seems rather straightforward, it is quite a delicate matter. For example, increasing the intensity of the cathode-ray-tube spot tends to improve signal-to-noise ratio; but on the other hand, the effective spot size is an increasing function of brightness, so that increasing the brightness may result in a decrease in resolution. Determining the optimum settings cannot simply be done by observing raw signals on an oscilloscope. Rather, it depends finally only on obtaining the best servo-control curves of the type shown in

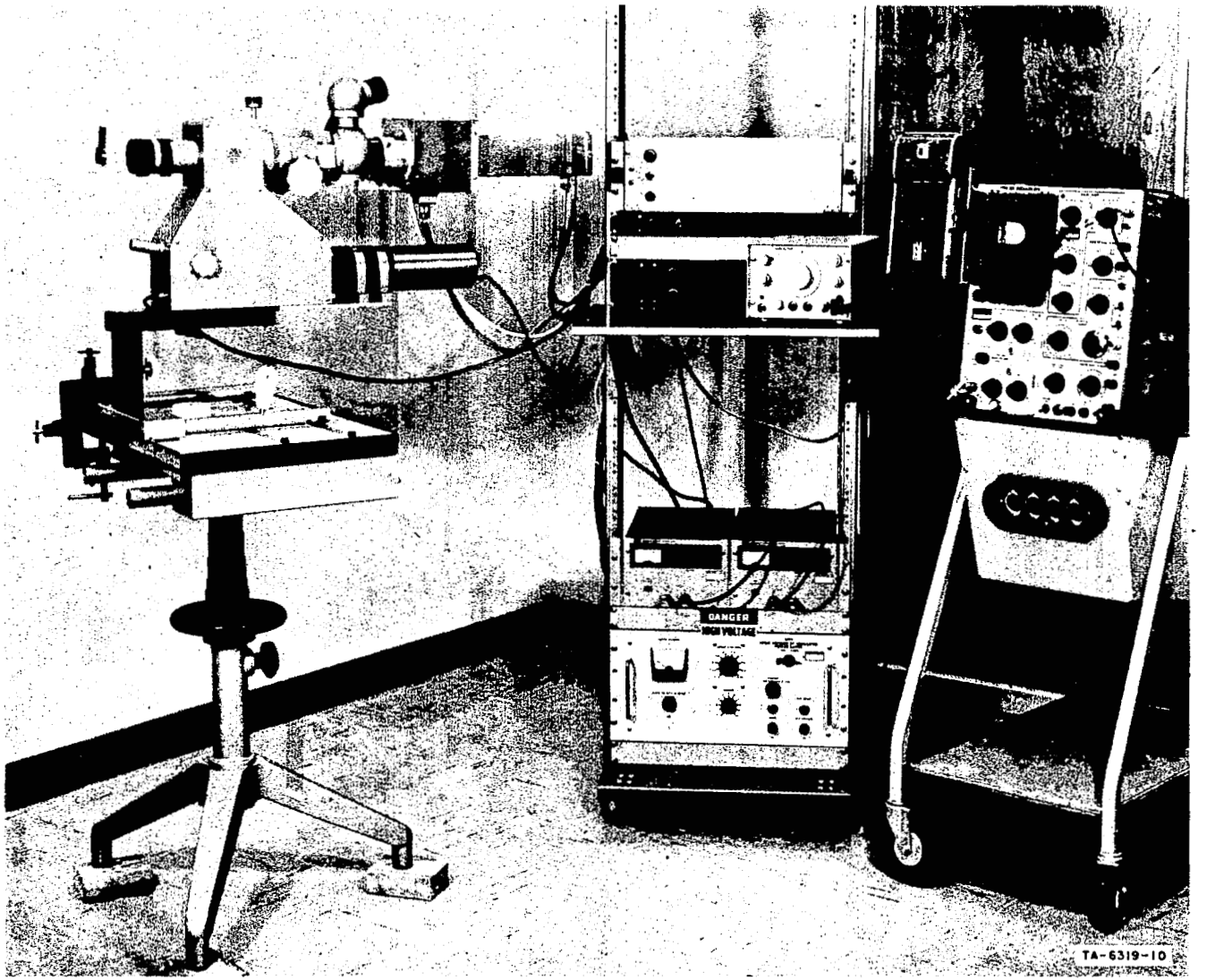


FIG. 10 OVERALL VIEW OF MODIFIED FUNDUS CAMERA AND ELECTRONIC CIRCUITS FOR DRIVING THE SCANNER CRT AND HIGH-GAIN PHOTOMULTIPLIER

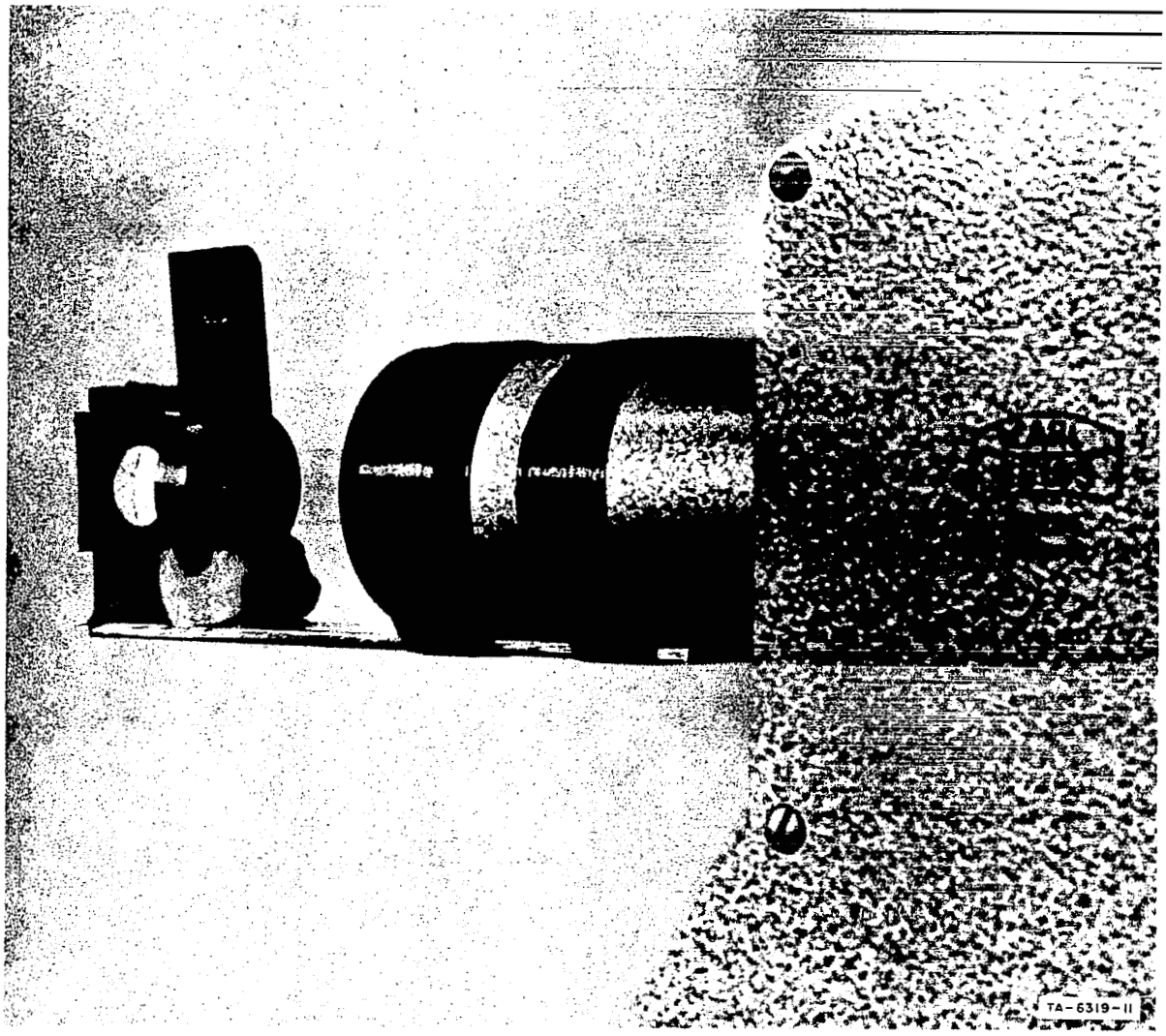
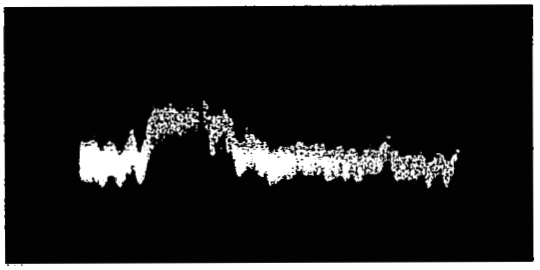


FIG. 11 CLOSE-UP VIEW OF THE MODEL EYE

Fig. 1(b), and this requires processing the signal according to the control algorithms of Eqs. (8), (9), and (10). It is for this reason that simulation of the final device is very important--i.e., a simulation system in which the signal is processed as it will be processed ultimately. The simulated processing may be performed very slowly (i.e., not in real time); this is not crucial. We have now developed such a computer simulation system, and the next major step in the program will be the transmission of signals from the fundus scanner



TA-6319-12

FIG. 12 A TIME EXPOSURE OF A VIDEO SIGNAL OBTAINED FROM THE MODEL EYE OF FIG. 11

to the simulator for the purposes noted above. The simulator program is discussed in Sec. IV.

IV SIMULATING THE SYSTEM

In Sec. II we discussed a technique for tracking the translational and torsional movements of an arbitrary two-dimensional image. In Sec. III we showed how such a technique could be applied to tracking the fundus pattern of a real eye. In particular, we indicated how a standard fundus camera could be adapted for this purpose, and showed signals obtained with such an apparatus in conjunction with a model eye. However, we have not as yet built the electronic hardware neccesary to process these signals in real time according to the control equations of Sec. II.

In view of the magnitude of the undertaking to build such high-speed correlation equipment, and also because we do not as yet have enough information to completely specify the requirements for such a system, it is highly desirable to simulate the system on a digital computer at much slower speeds, but with signals as realistic as possible. For example, as noted below, we would like to operate the real scan at 1 kc/s, so that we can make a correction of eye position every millisecond. However, if we are limited by signal-to-noise ratio, we might in fact achieve equivalent results by scanning at say, 200 cycles per second, in which case there is 5 times as much light per unit area per unit time, with a corresponding improvement in signal-to-noise ratio and reduced bandwidth requirement. Thus, though we might make only one-fifth as many corrections per unit time, each correction may be sufficiently improved so as not to warrant the much greater expense of the higher-speed correlation circuitry.

Experiments with the simulator system will help us determine realistic estimates for the performance that we can ultimately expect, and will significantly aid in the specific design approach for the signal-processing system. (In Sec. II we noted another important purpose for simulation--namely, as an aid in parameter optimization in the scanning system.)

In this section we will describe the basic simulation program. First, however, let us summarize some of the requirements for the real instrument, as we presently see them. This will keep the simulation program as realistic as possible.

A. Design Requirements for an Actual Eye-Tracker System

The design requirements for an actual eye-tracker system can be summarized as follows:

(1) As illustrated in Fig. 8, the resolution obtainable with the fundus camera is on the order of 1 minute of arc, which corresponds to an Airy disk on the retina about 5 microns in diameter. This implies that the flying spot should be about 1 mil in diameter (25 microns) when operating at the maximum 5X magnification from retina to film plane (which now contains the flying spot) available with the Zeiss optical system.

(2) The intensity (brightness) of the flying spot should be as high as possible, in order to obtain the maximum signal-to-noise ratio, which at best will be low at these scanning rates.

(3) The duration of a single scanned frame should be less than 5 milliseconds (preferably 1 millisecond), since this is the speed at which we need to make real-time position corrections.

(4) During a single frame time, the scanning spot should traverse as many independent cells (or Nyquist intervals) of the fundus pattern as possible. This is necessary in order to obtain the maximum amount of position information with the minimum concentration of energy per unit area at the fundus or at the phosphor.

(5) A circular scanning pattern, centered on the optic disk, is desirable (though not essential), since this would

cause the flying spot to cross many blood vessels and nerve fibers at approximately right angles.

(6) With a CRT scanner, the persistence of the phosphor should be as short as possible, since this will limit the number of resolvable picture elements per frame. For example, the 1-microsecond persistence of our experimental CRT phosphor will permit about 1000 picture elements per millisecond frame.

(7) Since a 1-mil CRT spot corresponds to about 1 minute of arc at the retina, a single circular line whose circumference equals 1000 of these resolvable elements will have a diameter of 5.3°, roughly the size of the optic disk.

Thus, we visualize the real system operating at a rate of at least 200 scans per second (preferably closer to 1000), where the scan has a diameter on the order of 5 to 10 degrees centered on the optic disk. The fundus camera optics have the required resolution, the experimental phosphor has the required spectral characteristics and short persistence, and the experimental phototube has very high gain in this spectral region.

Possibly all we need to do is to design and build the equipment for sampling, digitizing, and storing the signal from a single line scan, and the correlation equipment to operate at the necessary rates, plus the "servo" equipment for controlling the x- and y-position of the scanning circle on the cathode ray tube. But if we are to enter this construction phase with any reasonable confidence of success, we need the results from further simulation studies to aid in the final design. In fact, we still need to verify that there exists adequate signal-to-noise ratio to meet the tracking requirements. In the next subsection, we describe this basic simulation program.

B. Computer Simulation

The simulation effort has proceeded in two steps. To study the basic scanning technique, we first set up a simple optical system to

scan an enlarged (15X) transparency of the retina, much like the photograph shown in Fig. 8. This experiment used a very-slow-speed scan, on the order of one cycle every few seconds, with the "video signal" being digitized and then processed in the manner discussed in connection with Eqs. (8), (9), and (10). With this technique we were able to demonstrate a torsion-control characteristic of the type shown in Fig. 1(b). However, the signal-to-noise ratios of the "video signals" obtained in these experiments were much better than we could expect from the real eye at a high scan rate. Rather than simply adding arbitrary amounts of noise, we decided to delay further simulation studies until we could obtain actual signals from the fundus camera and could therefore get more realistic estimates of signal-to-noise ratios.

The first step in the next phase of the program will be to resume the simulation studies with signals obtained from the fundus scanner, as described in Sec. III. These signals will be recorded at high speed on an instrumentation recorder, and played back at suitably low speed into a computer, for the simulation program. Following is a brief description of this simulation program.

An analog-to-digital converter changes the video signal into digital form for entry into the computer. The computer also receives a sync pulse, phase locked to the flying-spot scan, to provide a fixed reference point, and measures the scan rate by timing the interval between successive sync pulses. Parameter values are entered manually from the computer console to control the details of the analysis. One can specify independently the sampling rate and the number of samples to be used in the computation. For example, the entry (1000, 900) indicates that the scan is to be divided into 1000 equal elements and that 900 of these are to be used in the computation.

The basic procedure is to make a "master" frame, against which each real-time frame is compared, in order to determine its x-translation, y-translation, and torsion. The master frame is the difference between two scans that are identical except for a specifiable delay 2δ. The control characteristics are then obtained by correlation. For example,

the torsion control characteristic is obtained by correlating a test frame directly with the stored master difference frame in the manner of Eq. (8); this correlation should be zero with the test scan delayed by δ (i.e., halfway between the two components of the master frame), and it should become increasingly positive for values of θ greater than this, and negative for lesser values. The x- and y-translation control characteristics are obtained similarly, except that the product of the master-difference frame and the real-time frame is multiplied by $\sin\theta$ or $\cos\theta$ before integrating.

The simulation program, written for the SDS 910 computer, is organized in several numbered subprograms, each of which can be called at any time from the console. These steps are shown in the computer printout for one case shown in Fig. 13.

Step 1: Clear halt to measure frame time. This program measures the actual scan rate in terms of computer memory cycles. In this case there are 407,185 cycles per scan cycle, or 3.257 seconds.

Step 2: Accept new parameters. NDIV specifies the number of uniformly spaced samples into which a single scan is to be broken, and NSMP specifies the number of samples that we wish to use in the correlation. In this case NDIV = 1000 indicates 1000 sample points, or 407 computer cycles per sample point. This number is used internally to set the sampling interval of the incoming signal. In other words, every 407 computer cycles, the input "video signal" is sampled, digitized, and stored. NSMP = 1000 implies that we will correlate over exactly one scan cycle.

Step 3: Accept master difference frame. Parameter KTHETA1 simply indicates where we wish to start sampling with respect to the sync pulse. Parameter DELTA (δ) specifies the separation

```
RETINAL TRACKING SIMULATOR


1.  CLEAR HALT TO MEASURE FRAME TIME
 407185 CYCLES
  3.257480 SECONDS


 2.  ACCEPT NEW PARAMETERS

NDIV, NSMP
1000,1000,

THETA/SAMPLE =     407 CYCLES
                21.58844 MINUTES
                 0.006279 RADIANS

MINUTES/CYCLE =  0.05304284


3.  ACCEPT MASTER DIFFERENCE FRAME

KTHETA1,  DELTA,  NO. REPEATS
10000,1221,1,


4.  ACCEPT TEST FRAME NO.   1

ΔTHETA,  NO. REPEATS,
0,1,


SPECIFY ΔTHETA-  MIN, MAX, INCR, IN CYCLES
-1221,1221,407,

  ΔTHETA        TOR              X               Y

  -1221     -1324821.      15877.930      235035.602
   -814     -1022966.      -8069.642      166004.362
   -407      -502787.     -31339.177       85819.049
      0        18610.     -54650.904       14454.514
    407       517765.      20174.919      -69212.719
    814      1032062.       -646.782     -155021.413
   1221      1430375.     -91618.999     -285254.496
```

FIG. 13 A SAMPLE OUTPUT FROM THE COMPUTER SIMULATION PROGRAM IN RESPONSE TO TORSIONAL MOVEMENT

distance for forming the master frame in the sense of Eq. (6). The master frame is formed by taking in one complete video frame, then a second frame displaced by 2δ (in this particular case, 2 times 1221 cycles, or 6 sampling intervals), and then subtracting one master frame from the other, point by point. Parameter NO. REPEATS indicates the number of times that this process is repeated with new frames, and the results averaged, for purposes of noise reduction. In this case, the process was done only a single time.

Step 4: Accept test frame. This procedure allows us to plot out the torsion-control characteristic by correlating new frames in different phase relationship with the stored master. Parameters MIN, MAX, INCR specify the phase range (in cycles) that is to be tested in steps specified by INCR. For each new phase relation a completely new test frame is brought in for correlation against the stored master. In this case, we test the torsion output over a phase range of (1221/407) = -3 sampling intervals to +(1221)/407 = +3 sampling intervals in one-interval steps. The resulting control characteristic is shown in Fig. 14. Note that the X and Y outputs are relatively insensitive to phase shift, or pattern rotation, which is of course as it should be.

For this simulation, the photographic transparency of the fundus was a 15X enlargement, and the scanning circle had a diameter of approximately 3 centimeters, which is a realistic scaling of 1 min, or 5μ, sampling intervals on the retina. The actual distance between samples on the photographic plate was therefore ≈100 microns. To test X and Y control characteristics we would therefore have required translation of this relatively large glass plate accurately and reproducibly with

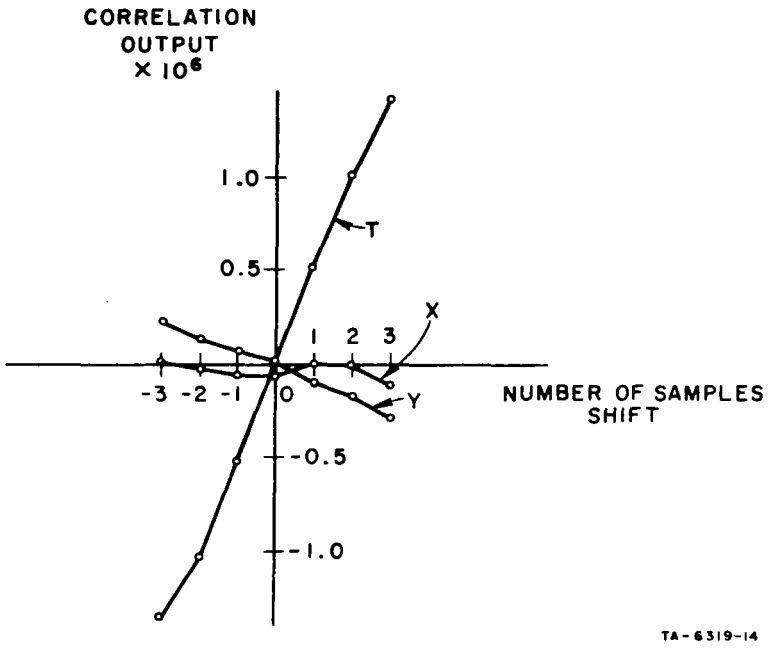


FIG. 14 A PLOT OF THE X, Y, AND T OUTPUTS FOR THE SIMULATION CASE OF FIG. 13

micron tolerances. It was deemed unwise to invest in the effort to arrange for X and Y control with the first-phase simulation. It was considered more reasonable to wait for real signals from the fundus tracker in which X and Y position can be shifted easily and accurately by means of fine control of the position of the scanning raster (deflection coil bias).

In summary, we have verified, with this simulation program, that the retinal photograph contains adequate information to yield satisfactory control characteristics within the desired accuracy range, measured in terms of torsion control only. What is not yet known is how these control characteristics will hold up under the relatively poor signal-to-noise ratio that we will obtain with high-speed scanning of a real eye. This is the next major step in the program.

We might note one simple, but very useful variation of the simulation program--namely, the ability to plot the autocorrelation function of a sampled frame, after averaging any specifiable number of times in order to average out noise. Such a plot, for just a single input frame,

is shown in Fig. 15. The abscissa is in terms of sample distances. The relatively steep slope of the curve--i.e., relatively large change in output for a shift of one sample distance, indicates that the signal contains considerable information in the desired range, at least for this situation of a greatly enlarged transparency of the fundus. This autocorrelation feature of the simulation program should be very useful for aligning our fundus camera system, as noted at the end of Sec. III. From autocorrelation plots such as this, for various settings of the control parameters (e.g., brightness of the scanning spot) we should be able to select optimum settings that maximize signal content and signal-to-noise ratio.

SPECIFY ΔTHETA- MIN, MAX, INCR, IN CYCLES
-4550,4550,455,

AUTOCORRELATE, NREP=
1,

ΔTHETA	TOR	X	Y
-4550	26097.	0.000	0.000
-4095	135018.	0.000	0.000
-3640	238853.	0.000	0.000
-3185	414997.	0.000	0.000
-2730	619168.	0.000	0.000
-2275	779143.	0.000	0.000
-1820	899685.	0.000	0.000
-1365	1034191.	0.000	0.000
-910	1472486.	0.000	0.000
-455	2220852.	0.000	0.000
0	2905763.	0.000	0.000
455	2219180.	0.000	0.000
910	1474201.	0.000	0.000
1365	1031452.	0.000	0.000
1820	895648.	0.000	0.000
2275	780741.	0.000	0.000
2730	626663.	0.000	0.000
3185	416860.	0.000	0.000
3640	238144.	0.000	0.000
4095	136804.	0.000	0.000
4550	30939.	0.000	0.000

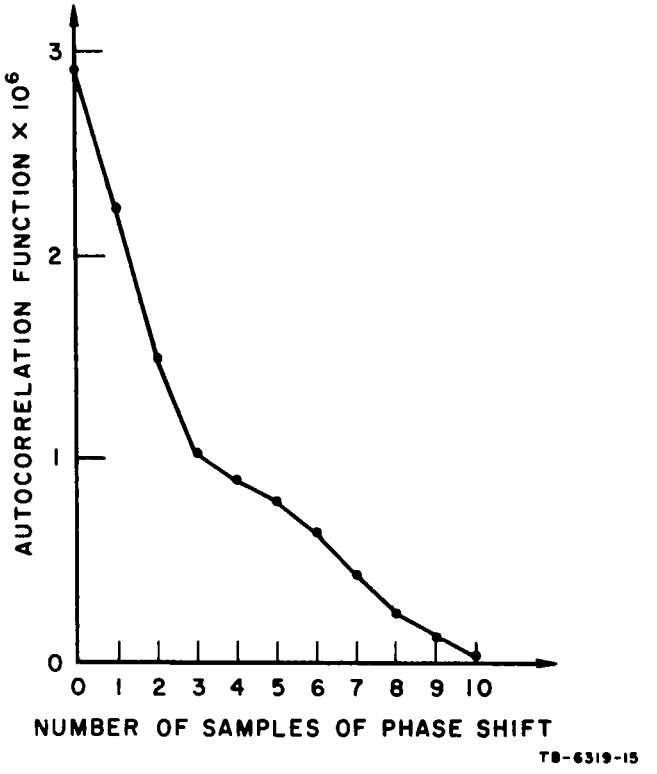


FIG. 15 A SAMPLE AUTOCORRELATION OUTPUT FROM THE COMPUTER SIMULATION PROGRAM

V STATUS AND PLANS

In this report we have discussed an eye-movement tracking technique that may yield a tracking accuracy on the order of 10 seconds of arc. The method we are developing is based on direct tracking of the retinal pattern (i.e., the fundus) itself. For test purposes, we have inverted the optical system of a diffraction-limited fundus camera, replacing the film by a flying-spot scanner, and the flash-tube by a photodetector. Thus we project a scanning pattern directly onto the fundus, and we plan to track the resulting image in space. With the tracking technique being developed, we should be able to detect torsional movements of the retina as well as vertical and horizontal translation movements.

Accomplishments during this first phase of the program can be summarized in four categories:

(1) A study of the spectral characteristics of the fundus, to determine the spectral range that yields the highest contrast images.

(2) Conceptual development of relatively simple scanning and correlation signal-processing techniques for determining translational and rotational movements of a generalized two-dimensional image.

(3) Realization of an optical configuration for implementing the scanning technique in the case of the two-dimensional fundus image from a real eye.

(4) Development of a computer program for simulating the correlation processing at slow speed for the purpose of overall evaluation of the tracking system as well as to help in design of the required high-speed correlation equipment.

In addition to various minor tasks, the remaining major tasks of the program are as follows. An extensive simulation study of the

system should be made, leading to realistic estimates of the performance that can eventually be expected. If these estimates indicate a potential performance that is unacceptable (but not limited by fundamental laws), then it would be necessary to look for alternative and improved techniques. On the other hand, if the projected system performance appears satisfactory, we can then develop the necessary hardware for processing the signals in real time, according to the control algorithms of Sec. II.

At the moment, in view of the signals already obtained from the fundus scanner with a model eye, and the control curves generated by the simulation system thus far, we feel optimistic about the image-stabilization possibilities of the present approach.

ACKNOWLEDGMENTS

The authors wish to express their appreciation of the efforts of a number of their colleagues--in particular, to Tom Cornsweet, who together with Don Kelly, initiated this particular study; Ken Kerwin, who helped in the design of the electronics; and Pete Johanson, who wrote the simulation program for the SDS 910 computer.