COMPUTER TECHNOLOGY

Data acquisition system for precision eye-fixation studies

M. BUCHSBAUM, J. S. BRYAN,
A. PFEFFERBAUM*, G. M. CARMAN**,
and R. C. STILLMAN

*Department of Health, Education and Welfare
National Institute of Mental Health
Bethesda, Maryland 20014*

A system has been developed to sense and record eye point-of-gaze for on-line computer processing, using a Mackworth camera in conjunction with a closed-circuit television system. System optics and electronics are described, together with computer software for system calibration and automatic artifact removal.

The difference between looking you in the eye and looking at the tip of your nose (a difference in point-of-gaze as small as 0.75 deg) can be detected by human Os in face-to-face contact (Cline, 1967). Direct observation of the eye has been used in the study of fixation and movement (Libby, 1970), but it is less accurate when the S looks elsewhere than directly at the O and rapid sequences of fixations cannot be recorded. A variety of instruments have, therefore, been developed both to measure eye position (point-of-gaze or fixation) and eye movements (see review by Young, 1963). All of the techniques have some tradeoff between speed of measurements, accuracy, range, and convenience to S and E. Studies of individual differences in perceptual and cognitive behavior (e.g., Gardner & Long, 1962) and in the sequencing of eye fixations for pattern recognition (Noton & Stark, 1971) suggest the need for a system that enables the investigator to determine where a S is looking, how long he gazes, and where he looks next. Relatively accurate vertical and horizontal coordinates, as well as the length of time that they eye remains directed at a point, are needed in a form that conveniently allows statistical analysis. In addition, precise control of stimulus delivery and convenient calibration procedures are necessary.

The vertical and horizontal coordinates of point-of-gaze have generally been measured with motion picture recording techniques, using a corneal reflective marker or analogue recordings from the diffuse scleral reflection technique. The tremendous amount of data produced are very tedious to quantify and reduce. Mackworth (1970) measured fixations from Polaroid

photographs of the corneal reflected spot by hand. Noton and Stark (1971) obtained data by FM tape-recording the X and Y analogue from their diffuse scleral reflection technique; this was digitized by computer and converted back into analogue form for tracing by an x-y plotter on reproductions of their visual display material. Fixations were numbered by hand as they were plotted. Our experience with the diffuse scleral reflection technique disclosed rapidly shifting accuracy and nonlinearity on the vertical axis. Further, the artifacts produced by squinting and blinking were often difficult to discriminate. The system herein described was developed in order to retain the accuracy of the Mackworth camera while providing a high resolution output for on-line computer data collection and analysis.

The method chosen uses a closed circuit television camera to produce an analogue, rather than digital, signal output for both horizontal and vertical axes. Since many modern small computers are equipped with analogue-to-digital converters, the system becomes readily adaptable to a number of such machines without constructing special-purpose digital input interfaces. We were faced with a need to produce an apparatus to work both with the classic LINC and the Systems Engineering Laboratories 810B system.

APPARATUS

A Mackworth camera, supplied by the Polymetric Company and already in use for eye-movement studies, was adapted by placing a Vidicon television camera (Ampex CC-324) at the camera viewing port, enlarging the optical stop somewhat to provide additional light and replacing the existing light source with a Bell and Howell slide projector with a 1-mm circular aperture in the film plane. Figure 1 shows a schematic representation of the optical system.

Eye position is indicated on the television monitor by a single bright spot produced by a virtual image of the light source filament reflected from the corneal surface and subsequently imaged on the photosensitive vidicon surface. Since the radius of curvature for the cornea is much smaller than its distance from the center of rotation for the eye, the image moves with the eye in a proportional and nearly linear fashion over a visual field of approximately 20 deg. [Mackworth (1967) gives a limit of accuracy of 20 deg, and Young (1963), 12.5 deg. We have remained within 10 deg in our studies, but our device has the same visual maximum capacity as

*Presently with the Stanford University Hospital, Palo Alto, California.

†Now with the National Institute of Neurological Diseases and Stroke.

Fig. 1. Schematic representation of the Mackworth optical system. The scan path is unused when working with computer display and data acquisition.

the conventional Mackworth photographic recording system.] In the conventional Mackworth system, this bright spot is superimposed on an image of the viewing material, and magnification ratios and static translations are adjusted to achieve alignment between the superimposed bright spot and the location of image material currently in central view by the S. These alignments are less critical with the computer acquisition system, however, since an initial calibration run for each S automatically adjusts origin, scale factor, and nonlinear corrections as well. Light levels are adjusted so that superimposed image material, though still present, falls below the noise level in the television presentation. Thus, the Mackworth system is used primarily for the convenience of the camera stand, lens system, and head holding.

Figure 2 is a block diagram of the system electronics. The television monitor produces a single bright spot, which moves with the eye, within the limitations of Vidicon response and scanning rate (conventional 525-line standard). Detection of the response is accomplished in a small separate unit by means of a

simple absolute threshold circuit which is referenced to the video black level. Analogue outputs proportional to horizontal and vertical position are obtained by two sample-and-hold circuits which sample linear sawtooth waveforms on the horizontal and vertical time bases, respectively. Additional logic (a flip-flop preset once each vertical field) insures that only one brightness threshold crossing is recognized per vertical field. Thus, the effective viewing point is the first threshold crossing recorded as the scan proceeds from top to bottom of the raster; this point is at the upper left boundary of the engaged spot. In practice, the spot is approximately 20 lines wide. Since one vertical output value is obtained for each vertical field (one-half frame), the effective vertical resolution is not 525 lines, but 262×0.84 , or 220 lines (including effects of blanking). If one chooses, successive pairs of values can be averaged to yield the full, available 440-line resolution, with a corresponding sacrifice in rise time.

Since Vidicon tubes normally produce bright peak artifacts at the top of the raster and at the beginning horizontal edge, additional gating circuits are provided

Fig. 2. Block diagram of the system electronics used to obtain vertical and horizontal analog eye-position values from a closed-circuit television system.

to disable recognition during the first few scan lines and during the first few microseconds after each horizontal retrace and unblanking. Additional brightness artifacts normally encountered at raster corners and caused by overscan of the photosensitive surfaces are avoided by reducing raster size to eliminate all overscan.

When the apparatus is used with light levels comfortable for the S, the signal corresponding to the imaged spot is just above the noise level of the camera tube used; a compromise is involved between S comfort and speed of response, since Vidicon response is sluggish at low light levels. A level of 600 lux incident flux density was found to be satisfactory. Camera tubes must be chosen which are free of white blemishes, since any bright spot will produce false detections unless the threshold adjustment is advanced beyond its peak value with substantial sacrifice in sensitivity. In practice, since the entire tube surface is not used with these optics, only two tubes were rejected because of developing blemishes.

We upgraded the camera recently with a Tivicon photodiode tube. Since the tube is more than 10 times as sensitive as the Vidicon in the red range, we have been able to use much reduced light levels and still keep the imaged spot well above tube noise levels. Further, the faster tube response (lag at 100% modulation—first frame down to 20%) has reduced response time and the loss of the spot during rapid movements.

RESPONSE TIME

Response time for the system was evaluated by driving a simulated cornea (small 6-V lamp housing) with a pen motor. Settling time for the mechanical drive and lamp was 10 msec. Figure 3 shows the system response to a 10-Hz square wave. The lack of symmetry in the resulting output waveform results from Vidicon screen persistence. The point on the screen first encountered by the scan holds the threshold detection after the optical excitation stops and until the screen response decays below threshold. The same point, however, captures threshold almost immediately, since rise time is shorter than decay. The error thus introduced by screen persistence is of the same order as the vertical scan period, or about 17 msec. With actual eye movements, the eye may move continuously for 16 or 32 msec and at no point reach detection threshold until the new fixation point is reached.

COMPUTER SOFTWARE

Eye-movement experiments involve the presentation of visual stimuli and recording of eye position for precise time intervals, subsequent calibration of the data points, and statistical analysis of patterns of movement. A process control computer with a visual display conveniently handles all these functions. Previously described systems have not utilized the visual display capacity or dealt with alignment, linear, and nonlinear distortion fully (Goode & Russo, 1970).

Fig. 3. Oscillographic recording showing system transient response to a 10-Hz square wave applied to an artificial cornea moving in the horizontal direction. Horizontal time base is 20 msec/cm. Vertical displacement represents approximately six peak-to-peak eye motions.

DATA COLLECTION PROGRAMS

The computer used is a classic LINC with 2,048 12-bit words of memory of which one-half can be used for programs. There are two digital display CRTs—one at the console and one slave run remotely at the eye-movement recorder. All programming is done in assembly code, using the LAP-6 compiler (Wilkes, 1970). The CERES remote monitor system (Gips et al, 1971) is used to call up eye-movement calibration and visual display programs used for collection of eye-movement data and to store collected data in files.

CALIBRATION OF SYSTEM

A digital oscilloscope program allows adjustment of x-y gain, proper system operation, and the absolute threshold for spot detection. The x-y gain need not be adjusted frequently, since the voltage change output to the computer associated with a given x-y coordinate on the screen is determined for every S each time he is run or removes his head from the head holder.

The calibration program begins with a central fixation spot to allow S alignment on the optical stand. It then displays a sequence of 9 points from right to left, top to bottom, which form a grid covering the LINC slave CRT. Each point appears for 1 sec, but data are sampled for only 640 msec at 20-msec intervals, starting 200 msec after the spot has appeared in its new position. The S is instructed to follow the moving spot with his eyes, attempting not to blink. A 10th point then appears in the center of the screen, and data are sampled as before in order to determine if the point of gaze returns exactly to the initial coordinates. The calibration procedure lasts 10 sec. The data are automatically displayed on the digital CRT, and the calibration can be repeated if all 10 points do not appear clearly or if the 5th and 10th points do not match (both screen center points). The

Fig. 4. Left: Data obtained from a standard calibration run. Raw, uncalibrated values from LINC A/D converter are plotted on same scale as digital CRT screen coordinates. Each of the nine circles represents the true successive screen coordinates of the visual target. Right: Correction values from calibration run applied to a second record in which the visual target moved to each of the same nine points, but in a random order. Data was sampled continuously, and the 0.5-deg criterion applied to choose a new point of gaze. Large circles are 0.5 deg in diam.

absolute threshold for spot detection must usually be adjusted for each S. We have found that Ss can easily keep up with the 1/sec moving spot and that problems with blinking and head movement have been minimal after one or two initial runs.

HEAD POSITIONING

The corneal reflective technique for measurement of point-of-gaze is extremely sensitive to head movements. While a biteboard can be used to stabilize the S's head, this is not only unpleasant for the S but distracting as well. This distraction, from fatigue and gagging, may prevent the occurrence of normal visual search patterns or preclude the testing of patients who may be less cooperative or experienced than a highly motivated eye-movement researcher. We have utilized a padded headholder-chinrest that holds the S's head still for the brief periods necessary for collection of eye data. The presentation of a spot in the center of the CRT for the S to fixate allows convenient correction for head position. Head position movement can be potentially quite small. The minimal error is assessed by using a cooperative S (M.B.), wearing eyeglasses with a large ball bearing firmly fixed to the lens over the left eye as in an artificial cornea. A standard calibration run is then done with the S following the moving dot with the uncovered eye and the reflective spot on the artificial cornea. Estimated head movement is less than 0.2 deg in 10 sec and consists of random tremor rather than a slow drifting of head position. This is only slightly above total system noise of less than 0.1 deg with a stationary artificial pupil fixed to the Mackworth camera stand.

REPRODUCIBILITY OF x-y COORDINATES

The 5th and 10th points displayed on the calibration run were used to assess reproducibility, since fixations on these two points appearing in the center of the

display should have identical x and y coordinates. The median x and y coordinates for the two points were calculated from 25 consecutive calibration runs. The median individual differed by 0.15 deg on the x axis and by 0.10 deg on the y axis on the two measurements, 5 sec apart. Ninety-two percent were accurate to better than ± 0.5 deg. Since the high-resolution fovea covers a visual angle of about 1 deg, this is quite adequate for sequential information search studies of gaze position.

DATA COLLECTION

At the beginning of each experimental run, a central fixation spot appears. The S is instructed to look at the spot, and eye position is sampled and stored to allow alignment at a later time. The visual display material is then presented, and the sampled x and y coordinates are stored in the upper page of memory and written onto magnetic tape files by the CERES monitor at the end of an experimental run. The 2K memory size of the classic LINC and our 1K experimental monitor limits the system to 10.2 sec of continuous data collection; buffered magnetic tape drives would allow continuous data collection. After each run, the S has 3-5 sec to blink and prepared for the next experimental run.

DATA EDITING

A master editing system allows review and manipulation of all data, including artifact removal, scaling aligning, and saccadic movement identification programs. The output of the editing program is the display coordinates and duration of each eye fixation.

(1) Artifact removal: When the reflective spot does not appear on the screen, as during a blink, or when an eye movement is so rapid that the video spot is never sufficiently intensified to trigger the detector circuit, the x-y value sampled is the first raster location, the upper right corner. These dropouts are replaced by using the last good (i.e., within the visual display areas) x-y coordinate pair. Ninety-five percent of data collection runs contain fewer than 5% dropout points. Blinks or off-field fixation may be identified by observing long periods of dropout (more than 60-100 msec).

(2) Axis alignment and computation of scaling factors: The x-y coordinates of each of the nine points in the calibration are identified by finding the median value during the 800-msec period. The distances between values are calculated and then written on the data tape in the alphanumeric header block for the calibration run. The editor then requests file numbers for the actual experimental runs for which the calibration applies and copies the point locations into their headers. The user then requests a single experimental run and the editor presents, in 10X real time or the digital CRT, the movements associated with the initial fixation spot. Cross-hairs are set on the spot, or, if no spot is present, the user can skip that trial and advance to the next. (A visual check was inserted here to insure that the correct center fixation point is selected.) The program then does

all subsequent steps automatically.

Each trial takes only a few seconds to score. First, the raw data are corrected by aligning the data to the observed $x = 0, y = 0$ point. Next, data are corrected for scale and distortion. The nine identified medians are treated as nine discrete points with known x and y error distance from the true nine points (calibration stimulus points). (See Fig. 4.) All data collected on subsequent runs are corrected with appropriate interpolation procedures, using the nine error distances.

(3) Definitions of new points-of-gaze: Digital records of eye position reveal a wide variety of rates of movement and total distances moved. Based both on reproducibility estimates and on foveal resolution width, we selected a criterion of 0.5 deg to define a new point of gaze. This criterion might also be expected to help differentiate saccadic and other types of movements. Smooth movements that include pursuit, compensation, vergence, and torisional movements are generally limited to a maximum velocity of 25-30 deg/sec (Stark, 1968; Young, 1963), which corresponds to a movement of 0.5 deg in 20 msec, our sampling interval. Thus, smooth movements should rarely reach a rate of change of 0.5 deg between samples, whereas rapid saccadic movements usually exceed 50 deg/sec and are often as high as 200-600 deg/sec (Stark, 1968). While the reflective spot is in motion at these high rates, it is not always sufficiently intense to raise the video level above the detector threshold. Figure 3 illustrates that the spot can be tracked at rates of about 120 deg/sec. Thus, smooth movements at 250 deg/sec or less would be correctly followed on each TV frame and would not spuriously appear as saccadic jumps.

Fig. 5. Time for eye to settle at new position for S distances between current and past position. Consecutive x-y coordinate pair differences are plotted beginning 40 msec before a detected new position and ending 100 msec after detection. Note that eye has stopped moving by 60 msec after the new position was detected.

Fig. 6. Frequency of occurrence of points-of-gaze of durations less than 200 msec, 200 to 300 msec, and longer than 300 msec for five distances of movement from the preceding point. Data collected during free scanning of size-estimation task on 10 Ss (total, 1,080 movements). The shape of these distributions is, of course, highly dependent on the nature of the visual display. (Compare with frequency distributions in Mackworth & Bruner, 1970.)

Fig. 7. Average consecutive x-y coordinate pair differences for eye movements of 1.6 to 3.2 deg. Vertical scale is the distance in visual angle between successive time samples of the x-y position of the reflective spot. Eye fixations were sorted by duration and averaged time-locked to the first pair difference that met the 0.5-deg criterion. For example, all eye movements in which the time to the next movement (time between coordinate pairs that exceeded 0.5 deg) was 200 msec were selected and the mean x-y coordinate pair difference was computed for each sampling time from 40 msec before the first movement to 40 msec after the second.

Fig. 8. Typical experimental data as plotted by computer. The S viewed the two squares on the CRT and was asked to judge whether the lower was smaller or larger than the upper square. Two typical patterns of eye movement are shown with each successive fixation represented by a letter.

Using the 0.5-deg criterion, each pair of x and y coordinates is compared with the preceding pair; if the present pair differs by 0.5 deg or more (diagonal distance), a new point-of-gaze is recorded. Figure 5 shows the x-y point pair differences for five distances of movement. By the time of the third sample after new position detection, the eye is again stable, and this sample is used as the coordinates of the new point-of-gaze. The point-of-gaze detection program compares each pair of x and y coordinates with the preceding pair; if the present pair differs from the preceding pair by 0.5 deg, then the next two pairs are tested and both also must differ by the criterion distance for a new point-of-gaze to be considered found. If position differences are less than 0.5 deg, the program advances to compare the present and the next x-y pair. This prevents slow drifts from being considered as movements.

Figure 6 illustrates the frequency of occurrence for point-of-gaze duration and distances between new points. Most points are held more than 200 msec and are more than 0.8 deg from the previous location. Consecutive x-y position changes for movements in the 1.6- to 3.2-deg category are illustrated in Fig. 7. While eye position jumps less than 200 msec apart are uncommon, they occur in small numbers following all distances moved.

The duration of each point-of-gaze is measured from the first x-y sample that deviates more than 0.5 deg from its predecessor. The time to move from one position to the next is lumped with the preceding position. The x, y and time coordinates of each new point-of-gaze are stored in the alphanumeric header block for the S for later statistical analysis.

(4) Data display: The data editing system allows the E to display the collected eye movements superimposed on the CRT display that the S was looking at during data collection. This display can represent all sampled data points "played back" in any ratio to real time with a digital millisecond clock showing elapsed time. Alternately, only the identified fixations can be displayed or plotted, as illustrated in Fig. 8.

(5) System flexibility: This system allows the collection of point of fixation data on large numbers of visual tasks and in large numbers of Ss with precise control over stimulus configuration and timing of presentation. For example, we have tested eight Ss on 20 to 30 8-sec visual task trials each in an afternoon. This eliminates tedious handwork between data collection and obtaining x, y and time coordinates in digital form for statistical purposes. It also allows convenient modifications of the parameters for definition of artifacts, fixations, and movements to suit experimental needs. Further, the on-line capacities can easily be utilized to collect other simultaneous physiological parameters (e.g., heart rate or average evoked responses), as well as data on S responses or tasks like reaction time. Modification of the stimulus display under control either of a S pressing a key or of an eye-movement parameter is also possible.

REFERENCES

- Cline, M. C. The perception of where a person is looking. *American Journal of Psychiatry*, 1967, 80, 41-50.
- Gardner, R. W., & Long, R. I. Control, defense and centration effect: A study of scanning behavior. *British Journal of Psychology*, 1962, 53, 129-140.
- Gips, J., Pfefferbaum, A., & Buchsbaum, M. Use of a small process control computer in a psychophysical laboratory. *Psychophysiology*, 1971, 8, 538-542.

- (Received for publication March 8, 1972;
revision received May 20, 1973.)