PRECISION AND ACCURACY OF VIDEO-BASED MEASUREMENTS OF DRIVER'S GAZE LOCATION

Frank Schieber, Michael L. Harms, Jan Berkhout and Don S. Spangler
Heimstra Human Factors Laboratories
Department of Psychology
University of South Dakota

Abstract

Off-line video analysis of driver eye gaze location is becoming an increasingly important tool for automotive ergonomics research. However, little information is available through which to estimate the reliability and validity of these techniques. A current investigation assessed the accuracy and precision of subject's judgments of driver eye gaze location from video taped records of controlled fixations across a vehicle's instrument panel. Results revealed an average accuracy for the assessment of gaze location of 8.45 cm with a corresponding precision of 12.2 cm. Accuracy, but not precision, was found to vary systematically as a function of where on the instrument panel the driver happened to be looking. These precision and accuracy values establish a benchmark which subsequent studies can use to estimate the appropriateness and limitations of the off-line video analysis technique.

Introduction

The emerging need to quantify and optimize the attentional demands (i.e., "workload") imposed by in-vehicle ITS technology has lead to an increased reliance upon the use of eye gaze position information (International Standards Organization, 1995; Tijerina, 1996). Although accurate eye gaze position information can be obtained using computerized eye tracking systems, this approach requires expensive instrumentation that is difficult to calibrate. In addition, the head-mounted visors and cameras needed to implement computerized eye tracking is uncomfortable and obtrusive. As a result, much of the recent work dealing with the workload demands of in-vehicle ITS systems has employed the less intrusive off-line analysis of video tape records of driver behavior in order to quantify gaze location, gaze fixation time and gaze frequency (e.g., Dingus, Antin, Hulse and Wierwille, 1989). This approach uses video records of the driver's head and face which are captured by small cameras that are usually located on top of the instrument panel or behind the rearview mirror. These records are then converted into digitized form which allows frame-by-frame access using a computerized video editing system. Experienced analysts can use these "time stamped" video records to estimate when, how often and how long a driver fixates a display and/or control during its use. Although labor intensive, the off-line video analysis technique for quantifying driver gaze location information is relatively inexpensive and much less intrusive than alternative techniques.

Due to a rich literature from the social psychology of face-to-face communication, much is known about the accuracy of a human "receiver's" ability to judge eye gaze position through the observation of the "sender's" face (e.g. Anstis, 1969; Gibson and Pick, 1962; von Cranach and Ellgring, 1973). However, a search of the literature revealed that there was insufficient information available to construct an estimate of the accuracy and precision of judgments of driver eye gaze location using the aforementioned video analysis technique. The purpose of the current investigation was to perform a preliminary assessment of the accuracy and precision of estimates of eye gaze position based upon off-line analysis of video taped records of driver behavior.

Method

Subjects. Participants in this study included 6 experienced members of the staff of the Human Factors Laboratory at the University of South Dakota (2 faculty; 4 senior graduate students). All participants had normal visual acuity and color vision.

Apparatus and Stimuli. A model driver was videotaped while shifting his gaze back and forth between the vanishing point of the roadway ahead and specific locations on (and near) the instrument panel of a late model Nissan automobile. The gaze locations within the vehicle occurred at 100 points along the intersections of an imaginary grid with 8 cm row and column spacing. A high resolution color camera was located just below the center mounted rearview mirror. The angle formed between the driver's straight ahead line-of-sight and the principal axis of the video camera was 65 degrees. The zoom level of the camera's lens was adjusted so that the model driver's head occupied approximately 70% of the total image area. The resulting video taped record was digitized and still frame images capturing gaze fixation were extracted for each of the target locations. Forty of these still frames depicting driver gaze fixations at known locations were selected as stimuli for the study. The 40 images were selected to cover the full range and extent of the instrument panel with oversampling in those areas most likely to contain emerging ITS displays and/or interfaces. These stimulus images were subsequently displayed to the research participants using an SVGA computer monitor. Each stimulus consisted of a 24-bit color image that was 364 pixels wide by 268 pixels high. The actual display size of each image was 16 cm by 12 cm.

A schematic model of the vehicle's interior was constructed for the purpose of providing a medium through which observer's could record their judgments of eye gaze location upon viewing a stimulus image (see Figure 1). This schematic drawing faithfully preserved the geometry of the original vehicle's instrument panel with a 10:1 reduction in scale (N.B., Figure 1 is not shown at its actual size).

Procedure. All of the research participants were familiar with the objectives of off-line video analysis of eye gaze position in driving-related research. At the beginning of each experimental session, the participants were given a series of calibration trials. Each calibration trial displayed a facsimile of the vehicle's instrument panel (similar to Figure 1) which contained a mark noting a specific location paired with the stimulus image showing the model driver gazing at that particular location within the vehicle. Seven such pairs of calibration stimuli were constructed to represent the four corners of the instrument panel along with three other points in the center of the panel. The research participants studied the calibration stimuli until they were ready to begin judging the actual experimental stimuli.

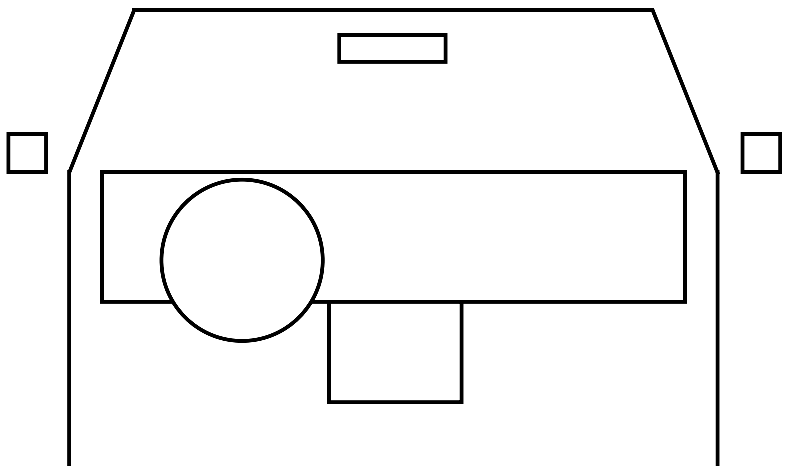

Figure 1.
Schematic representation of vehicle interior used to record gaze location judgments.

Next, the observer was presented with 40 self-paced eye gaze location estimation trials. Each trial consisted of the sequential presentation of two images: The first image always showed the model driver gazing through the windshield at the vanishing point of the road ahead. The second image consisted of one of the 40 test stimuli showing the model driver gazing at a predetermined location on (or near) the vehicle's instrument panel. Upon viewing this second image, the observer was required to make a judgment regarding the model driver's eye gaze location and to record this judgment by placing a mark on a sheet of paper containing a clean copy of the schematic drawing representing the vehicle's interior (see Figure 1). A new response sheet was provided for each eye gaze location estimate. Each session consisting of calibration trials and 40 experimental gaze location trials lasted approximately 30 min. The entire sequence was repeated on four consecutive days yielding a total of 160 eye gaze location judgments per observer (i.e., 4 replications at each of 40 gaze locations). [Note: Because of "technical difficulties" only the data collected from 38 of these locations were analyzed and reported below]

Results and Discussion

Each estimate of eye gaze location was converted into an (x,y) coordinate measure relative to the scaled geometry of the schematic response sheet. The average of the four estimates of gaze location (collected across successive days) for each of the positions tested was calculated for each experimental participant. This average gaze location estimate was then converted into "real world" coordinates. Based upon the difference between the actual predetermined location and the averaged estimate of eye gaze location, two indices of response accuracy were calculated: (1) the magnitude of the distance (in cm) between the actual site of fixation and the estimated location, and (2) the directional bias (in degrees) resulting from errors in estimation. These measures of response accuracy are summarized in Table 1.

In order to visualize how response accuracy might change as a function of position across the instrument panel, a spatial map of the accuracy data is depicted in Figure 2. Each average estimate of gaze location is shown as a separate vector. The origin of the vector is the actual gaze location; while the end of the vector (indicated via an arrow head) marks the site of the average estimated eye gaze location. The length of the vector graphically captures the magnitude of the error of estimate while its angular orientation depicts the directional bias. Each of the experimental locations is labeled with a number to facilitate discussion and cross-reference to Table 1.

Inspection of Figure 2 reveals that gaze localization accuracy appears to vary systematically as a function of where in the vehicle one happens to be looking. Our qualitative analysis of the accuracy map uncovered 8 distinct response clusters - which are labeled and summarized in Table 2. Cluster I (consisting of points 1-3 at the upper left region of the instrument panel) is characterized by good vertical accuracy but a distinct leftward error bias along the horizontal dimension with an average magnitude of 7.7 cm. Cluster II (the far left instrument panel) shows a pattern of estimation highlighted by excellent horizontal localization and a modest error along the vertical meridian averaging around 5.8 cm in magnitude. Cluster III (the steering wheel area) shows a distinct propensity for gaze fixations involving targets along the perimeter of the steering wheel to be misattributed to a site near the center of the steering wheel - resulting in a relatively large average error of 12.7 cm. Cluster IV (the upper central instrument panel) is characterized by a moderate leftward misplacement error averaging 7.9 cm. Cluster V (the central instrument panel) is characterized by excellent accuracy scores (mean = 0.9 cm). This high degree of accuracy appears to stem from the fact that these locations were in the same vertical plane as the video camera used to record model driver behavior. Accuracy in this region might be expected to change if a different camera angle (e.g. driver line-of-site) was employed. Cluster VI (the lower console area) gaze estimates yielded a consistent upward and leftward bias averaging 10.9 cm in magnitude. Cluster VII (upper right instrument panel) was characterized by a high degree of response accuracy - demonstrating an average error of only 4.6 cm. Cluster VIII (lower right instrument panel) performance is characterized by a distinctive and sizable overestimation toward the lower right averaging 10.4 cm in magnitude.

IP Position	Error Magnitude (cm)	Error Direction (degrees)	Precision (cm)
1	7.66	199	6.90
2	5.40	178	12.90
3	10.16	186	16.60
4	6.47	140	9.70
5	4.65	171	15.00
6	19.05	102	9.40
7	3.25	225	8.10
8	11.84	213	10.70
9	11.96	201	15.90
10	10.62	202	8.80
11	12.68	165	16.10
12	18.53	93	12.90
13	5.81	86	10.10
14	0.58	59	9.60
15	6.25	262	7.40
16	8.60	234	14.00
17	16.26	217	8.80
18	0.50	217	8.10
19	0.22	206	11.60
20	8.70	45	9.60
21	9.12	315	18.30
22	7.73	84	17.50
23	11.93	162	12.80
24	1.81	186	10.60
25	0.81	299	15.90
26	9.91	312	15.40
27	8.74	326	11.70
28	3.40	227	8.90
29	7.65	61	13.30
30	12.21	117	12.30
31	14.08	132	10.40
32	0.42	225	14.70
33	11.09	303	14.60
34	12.98	347	12.70
35	16.61	113	16.80
36	1.58	161	14.40
37	11.24	36	10.20
38	10.67	13	12.40
	mean = 8.45	mean =	12.2

Table 1.
Accuracy and precision of eye gaze location judgments for
38 locations across a vehicle instrument panel.

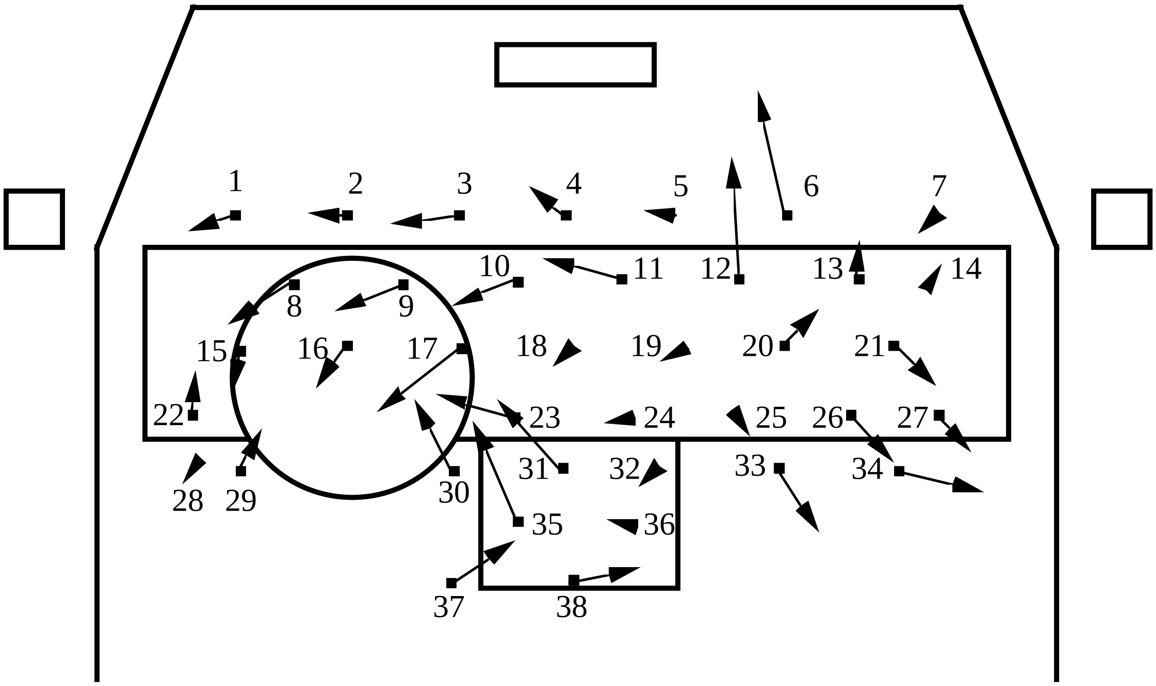

Figure 2.
Accuracy map of gaze direction estimates (see text).

Cluster	Positions	Accuracy	Precision
I. Upper Left	1,2,3	7.7	12.1
II. Left	15,22,28	5.8	11.3
III. Steering Wheel	8,9,10,16,17, 23,30,31,35	12.7	12.3
IV. Upper Central	4,5,11	7.9	13.6
V. Central	18,19,24,25,32,36	0.9	12.5
VI. Lower Console	37,38	10.9	11.3
VII. Upper Right	6,7,12,13,14,20	9.3	9.9
VIII. Lower Right	21,26,27,33,34	10.4	14.5

Table 2.
Cluster analysis of gaze estimation accuracy as a function of instrument panel location.

In addition to measures of accuracy, an index of response precision was also calculated. The degree of “spread” in the four estimates of gaze location collected from each participant at each site was quantified as follows: First, the center of gravity in 2-dimensional space between the four estimates was calculated. Then, the average distance (in cm) of the estimates from the center of gravity was calculated. Mean response precision values for each of the gaze locations (and clusters) are summarized in Tables 1 and 2. Unlike the accuracy measures which appear to vary systematically as a function of instrument panel location, the response precision data appears to be relatively consistent across gaze positions.

Summary and Conclusion

Estimates of driver eye gaze location accuracy yielded an average error of 8.45 cm (see Table 1). However, response accuracy varied as a function of region within the vehicle (see Figure 2 and Table 2). The direction of eye gaze location to targets within the central region of the vehicle's instrument panel was estimated with a high degree of accuracy. However, judgments involving the fixation of targets around the steering wheel tended to suffer from systematic biases. Average response precision was 12.2 cm. These data provide useful information for assessing the accuracy and precision of off-line video analysis techniques for the determination of eye gaze location in driving research environments. These findings are limited to situations where the video recording device is placed at or near the center mounted rearview mirror. Additional work is in progress to replicate the current findings and extend the results to driver's line-of-sight camera angles.

References

Anstis, S.M., Mayhew, J.W. and Morley, T. (1969). The perception of where a television portrait is looking. *American Journal of Psychology*, 82, 472-489.

Dingus, T.A., Antin, J.F., Hulse, M.C. and Wierwille, W.W. (1989). Attentional demand requirements of an automobile moving-map navigational display. *Transportation Research*, 23A(4), 301-315.

Gibson, J.J and Pick, A.D. (1963). Perception of another person's looking behavior. *American Journal of Psychology*, 76, 386-394.

International Standards Organization, Road Vehicle Ergonomics TICS-MMI Working Group ISO TC22 SC13 WG8 (1995). *Draft standard: Driver visual demand measurement method.*

Tijerina, L. (1996). *Final report -- Program executive summary -- Heavy vehicle driver workload assessment.* Washington, DC: National Highway Traffic Safety Administration [DOT-HS-808-467].

von Cronach, M. and Ellgring, J.H. (1973). Problems in the recognition of gaze direction. In M. von Cronach and I. Vine (Eds.), *Social communication and movement.* London: Academic Press.