A ROBUST METHOD TO DETECT DRIVER DISTRACTION

Gert Weller & Bernhard Schlag
Traffic & Transportation Psychology. Technische Universität Dresden
01062 Dresden
gert.weller@tu-dresden.de

ABSTRACT: Distraction is regarded as the sole or most important contributing factor to accidents (McEvoy, Stevenson, & Woodward, 2007; Regan, Lee, & Young, 2009). Therefore, methods are needed which allow the reliable detection of distracted driving. For the results reported here, data were collected during on the road driving experiments with an equipped vehicle. Besides driving behaviour, gaze data was recorded during distracted and non-distracted periods of driving. The data were first analysed for differences between both conditions. Based on these results the variable "standard deviation of the gaze data" was selected for further analyses. By performing a binary logistic regression it could be shown that this variable could reliably predict states of distraction. With the technical progress in gaze detection these results are regarded as highly promising for future solutions of in-car detection of driver distraction.

1 INTRODUCTION

Distraction is regarded as the sole or most important contributing factor to accidents (McEvoy et al., 2007; Regan et al., 2009). Detecting periods of distraction could thus significantly reduce accidents, for example, by selectively increasing a car's preparedness to brake in emergency situations. Because gaze parameters change with distraction and are related to driving performance (Dukic, Hanson, & Falkmer, 2006) they are suitable candidates to detect distraction while driving. However, in order to be accepted by the drivers the algorithms used have to be reliable in addition (Muir & Moray, 1996; Wickens et al., 2009). In driving, this specifically requires that they are independent of road geometry. However, since the often cited study by Land & Lee (1994) it is known that road geometry does influence gaze behaviour. In contrast to preceding studies of gaze behaviour and distraction (e.g. Victor, 2005) this aspect was taken into account by including road geometry as an additional factor.

2 METHOD

2.1 The Test Drives

On-road-driving tests were conducted with a measurement vehicle equipped with the usual speed and position recording devices[1] . In addition, the vehicle was equipped with the contact-free eye-tracking system Smart Eye. The system specification we used was a two camera solution with an additional scenery

[1] Details for the measuring vehicle can be found at http://www.strassenentwurf.tu-dresden.de.

camera which recorded the road ahead of the driver. Gaze data was analysed with WatchOut (Schulz, 2007) and own algorithms (see chapter 2.2.).

A peripheral detection task (PDT) was presented to the drivers on a 19 inch touch-screen monitor which was installed near the dash board to the right of the driver. It was adjusted individually for each participant via a number of flexible joints in order to be easily reached. The signals presented to the participants were red rectangles (55 × 40 mm) which popped up at different locations across the whole screen. The location of the rectangles was quasi-random as they were taken from a file of previously randomly created positions. The reaction required from the participants was to touch these rectangles, which caused the rectangles to disappear. The time between onset of the rectangles and reaction was measured as reaction time (RT). The time between reaction to one rectangle and onset of the next rectangle was parameterised to 500 ms. This interval was chosen to allow an additional analysis of fixations between to stimuli. The PDT had to be performed by the participants at selected locations. Start and stop instructions for entire sections were given to the participants by the instructor, who was seated in the back seat of the car. To get familiar with the task, the participants performed the task at a standstill before driving and during the first 500 m of the test-drive. By analysing the development of the reaction time values throughout the course it was ensured that a baseline was reached before the data were used for experimental purposes.

Sixteen participants, who were recruited from the university staff, took part in the study. They were aged between 25 and 47 and were all active drivers (average annual mileage driven over the last three years was between 5,000 and 40,000 km).

The course driven consisted of rural roads and was selected based on accident criteria. Prior to the test drives it was driven with the equipped vehicle whereby pictures were taken every 10m. With the data collected during this drive the road geometry (horizontal and vertical alignment) was assessed and the data was integrated together with the pictures in a database administered by the custom made software RoadView (Dietze, 2007). The course was driven in the outbound and the inbound direction, amounting to a total of 80 km.

For the results reported here, two curves were selected with a radius of 129 and 198 m respectively. The fact that the course was driven in both directions allowed the selected curves to be driven once with and once without the PDT, albeit in different directions. In order to minimize the potential effect of the driving direction it was ensured that the characteristics of the approach zone to the curves were comparable in both directions, regarding geometry, speed limits and environmental characteristics.

2.2 The Gaze Date

Gaze behaviour was recorded with the contact-free two-camera solution of Smart Eye, which was already integrated into the measuring vehicle at the time of the experimental drives. Contact-free eye tracking has three major advantages:

- the driver is not constantly reminded that eye-movements are being tracked;
- longer drives are possible because an uncomfortable helmet with cameras does not have to be worn by the driver; and
- there are no liability issues because of additional danger to the eyes in the case of an accident.

The system used the cornea-reflex method, which is essentially based on the fact that light directed towards the eye is reflected from the cornea in a direction which depends on the position of the cornea. In order to compensate for different light conditions, the system was equipped with infrared light emitting diodes which were directed towards the eyes and were placed at the eye-tracking cameras. The system had to be calibrated for each participant in order for the system to identify the eyes. In a second step, the scenery camera was calibrated to the gaze data by asking participants to fixate pre-defined objects in the scenery. Both calibration steps were conducted with each participant prior to the experimental drive at a car park.

Besides the data collected in the log file, gaze direction was shown in the video of the scenery camera and could be monitored by the experimental leader in real-time.

The gaze data were analysed using the software WatchOut (Schulz, 2007), which was tailor-made at the Chair of Road Planning and Road Design at TU Dresden for the hardware in the measurement vehicle. WatchOut uses several variables and parameters of gaze behaviour. For this paper, the subsequent variables and parameters were used for the analyses and are described in detail below:

- fixations (number and duration);
- the scan path;
- the standard deviation of the gaze location.

All parameters were calculated and averaged by WatchOut for sub-sections of 25 metres within the experimental sections.

For the definition of fixations an algorithm developed by Jacob (1995) was used as basis for the algorithm used here:

- The start of a fixation is recorded when gaze direction lies within an area defined by 0.6 degrees for a duration of 100 ms (three data points with the 30 Hz used by Smart Eye).
- The averaged gaze data location is defined as the fixation location.
- The end of a fixation is defined as when gaze direction is outside 1.6 degrees of this fixation location for more than 100 ms.

With the assumed duration of one second needed to drive a sub-section of 25 metres and an average fixation duration in driving of approximately 400 ms (Velichkovsky, Rothert, Kopf, Dornhöfer, & Joos, 2002), approximately two to three fixations per sub-section could be expected for the data at hand. If fixations started at the end of one subsection but were terminated in the next sub-section, the fixation was assigned to the former only, with the duration of this fixation also including the time in the latter sub-section. Thus, data were not lost despite the division of the sections into sub-sections.

Based on the fixation locations, the scan path was calculated as the Euclidian distance between two consecutive fixations and was averaged for the subsections. Thus, the scan path took into account the temporal succession of fixations.

In order to be independent of the thresholds used to define fixations, additional parameters were used which were not based on fixations but on the raw data. One parameter was the standard deviation of the gaze data. In a first step, the centre of the gaze data was determined by averaging the X-values and by averaging the Y-values:

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n}; \overline{Y} = \frac{\sum_{i=1}^{n} y_i}{n} \tag{1}$$

where n equals the number of data points.
In a second step, the distance to this centre was calculated for each data point as:

$$d = \sqrt{(\overline{X} - X_i)^2 + (\overline{Y} - Y_i)^2} \tag{2}$$

this distance was averaged as:

$$\overline{d} = \frac{\sum_{i=1}^{n} d_i}{n} \tag{3}$$

and the standard deviation was calculated as:

$$SDXY = \sqrt{\frac{\sum_{i=1}^{n} (d_i - \overline{d})^2}{n-1}} \tag{4}$$

The interpretation of this parameter is similar to the interpretation of the standard deviation as described in Victor, Harbluk & Engström (2005) where the authors used glances as defined in (DIN EN) ISO 15007-1 (2003) instead of the raw data, and degrees instead of pixels as the unit of the standard deviation. While glances are appropriate to assess the effect of in-vehicle devices, as defined, for example, by the European Commission (2007), the use of raw gaze data is more appropriate for driving on rural roads. This is because driving is to a large degree governed by pursuit movements and peripheral visual information uptake which takes place independently of whether a fixation or glance is present or not.

For the variables of the gaze data, the parameters maximum and average were used for the statistical analyses. The minimum values were not used because aggregating the values for the minimum resulted in a value of zero for most participants in most sections which made a comparison between two or more curves based on the minimum inappropriate. The zero values were present in almost all curves somewhere in one of the sub-sections whenever drivers looked away from the cameras for a time period longer than it took to drive the sub-section.

2.3 Other Behavioural Data

As described above, driving data was collected with the measuring vehicle. The data used for the subsequent analyses were:

- speed (minimum, maximum, average and percentage change in speed)
- lateral acceleration (maximum)
- longitudinal deceleration (maximum).

Regarding the reaction times to the PDT, conditions driven with PDT could obviously not be compared to conditions driven without PDT.

2.4 Statistical Methods of Data Analysis

In a first step of the analysis several behavioural data were analysed for differences between the condition with and without PDT. The two locations described above were integrated as an additional factor. Two-factor ANOVAs for dependent samples were performed separately for each of the dependent parameters described in the preceding chapters.

In the second step, a binary logistic regression was used to assess the power of the selected variable to distinguish between the condition driven with and the condition driven without PDT. The robustness of this method (Backhaus, Erichson, Plinke, & Weiber, 2006), allows its application also in case not all preconditions are met. For this study, this concerned the dependent nature of the sample which is not corrected in binary logistic regression and the comparatively small number of subjects.

3 RESULTS

3.1 Selection of a variable suited to detect distraction

In a first step the behavioural variables described above were tested for differences between the condition with and without PDT whereby the two locations were integrated as an additional factor. Perhaps the most salient results were found for minimum speed which is shown in Table 1.

Table 1. Effects of the PDT on minimum speed in two curves: results of a two-factor repeated-measures ANOVA

	Descriptive Statistics *M (SD)*		Results for the ANOVAs			
Curve	Without PDT	With PDT	Effect	*F*(1, 9)	*p*	η^2
One	79.92 (5.63)	72.60 (8.06)	Curve	6.08	.04	.40
Two	81.86 (6.86)	78.48 (6.92)	PDT	16.76	.00	.65
			Curve × PDT	1.30	.28	.13

The results for the other speed parameters and additional driving behaviour parameters referred to in the preceding chapters were of similar nature regarding the direction of differences and the level of significance and are reported in Weller (2009). The results for the gaze variable standard deviation of gaze location are given in Table 2.

Table 2. Effects of the PDT on maximum and average standard deviation of the gaze locations in two curves: results of two-factor repeated-measures ANOVAs

	Descriptive Statistics *M (SD)*		Results for the ANOVAs			
Curve	Without PDT	With PDT	Effect	*F*(1, 9)	*p*	η^2
			Maximum SD			
One	156.4 (98.0)	494.4 (105.7)	Curve	0.54	.48	.06
Two	219.9 (284.4)	505.2 (140.2)	PDT	34.90	.00	.80
			Curve × PDT	0.36	.56	.04
			Average SD			
One	31.3 (18.9)	284.7 (69.7)	Curve	0.03	.86	.00
Two	44.9 (61.4)	276.9 (82.7)	PDT	124.9 2	.00	.93
			Curve × PDT	0.25	.63	.03

The results for the variables scan path, fixation duration and number of fixations were:

- Average Scan path: Curve F(1, 9) = 0.00, p = .98, η2 = .00; PDT F(1, 9) = 30.14, p = .00, η2 = .77; Curve × PDT F(1, 9) = 2.13, p = .18, η2 = .19.
- Average Fixation Duration: Curve F (1, 9) = 2.79, p = .13, η2 = .24; PDT F(1, 9) = 5.09, p = .05, η2 = .36; Curve × PDT F(1, 9) = 0.48, p = .50, η2 = .05.
- Average Number of Fixations: Curve F(1, 9) = 5.78, p = .04, η2 = .39; PDT F(1, 9) = 14.69, p = .00, η2 = .62; Curve × PDT F(1, 9) = 5.98, p = .04, η2 = .40.

The analyses so far were used to estimate which parameter of which variable was best suited to distinguish between conditions with and without distraction. This parameter should additionally be independent of road geometry in order to allow a reliable detection of distraction independent of conditions not related to distraction. As is prototypically shown in Table 1, this was not the case for driving behaviour which was largely influenced by road geometry. Detecting distraction with variables of driving behaviour would therefore not only require knowledge of the road geometry but additionally a highly reliable determination of the vehicle's location. This is not the case for the variables of gaze behaviour which were to a large extent influenced by the PDT but were less or not at all influenced by road geometry. The decision in favour of a single parameter within the variables of gaze behaviour was based on the $\eta 2$ values. This resulted in the selection of the "standard deviation of gaze location" for which the parameter "average across all sub-sections" was used for the subsequent analyses.

3.2 Reliability of distraction detection

After having selected the variable "standard deviation of gaze location" and its parameter "average across all sub-sections" the question arose how this data could best be used to predict whether a driver is distracted or not. Statistically speaking the method of choice is the binary logistic regression. This allows the prediction of one out of two conditions based on the actual values of the variable of interest.

For the data used here, these two conditions are prototypically represented by the two conditions with or without PDT (coded 0 for the condition without PDT and 1 for the condition with PDT). The average standard deviation of the gaze location represented the sole covariate. In addition, a constant was included. The data were divided into two data subsets that were used for the development and validation of the model. The values for curve one were selected for the calculation of the model, whereby the values for curve two served as input for the validation of the model developed with the data for curve one. Because there was only one independent variable, it was necessary to calculate the model itself with the inclusion method.

The different statistics of the logistic regression indicated very good model fit (Omnibus test statistics: $\chi 2 = 29.41$, df = 1, $p \leq .001$; Nagelkerke's pseudo R2: .83; 2LL value: 12.18). The Wald statistics of the logistic regression were highly significant (Table 3; units used: pixel).

Table 3. Wald statistics and odds ratios for the gaze parameter and the constant in the equation of the logistic regression which was used to determine whether a driver was distracted or not

Variable	B	SE	Wald (df = 1)	p	OR	CI 95%	
Gaze data	0.03	0.01	8.73	.00	1.03	1.01	1.05
Constant	-3.62	1.29	7.80	.01	0.03		

The seemingly low value of the odds ratio for SD-gaze must be related to the units used. For the data at hand, these units were pixels. Applied to the odds ratio, this means that with every additional pixel in SD-gaze, the likelihood that the data belonged to the condition with PDT increased by 1.03. If the difference between the condition without PDT and the condition with PDT were just one pixel, this odds ratio would be meaningless. However, with the approximate average difference between both conditions having been around 200 pixels (Table 2), the finding is indeed meaningful, even at this stage. The quality of a logistic regression can be largely influenced by single cases. Despite the encouraging results reported above, two cases were wrongly classified in the dataset used to develop the function. This resulted in a correct classification rate of 93.3%. In detail, one participant was wrongfully assigned to the PDT-condition and one participant was wrongfully assigned to the non-PDT-condition. A case-wise analysis was thus conducted with the standardised residuals (Z-Resid). These values are shown together with the participant-wise SD-gaze values in Figure 1. There, the two outliers in the data which represent the two misclassified cases can clearly be identified from the Z-Resid values: participant No. 13 who was wrongfully classified as driving in the non-PDT condition, and participant No. 16 who was wrongfully classified as driving in the PDT-condition. Because SD-gaze was used as the single predictor in the logistic regression, the outliers in the Z-Resid data are matched by outliers in the SD-gaze data, highlighted by the grey circles in Figure 1.

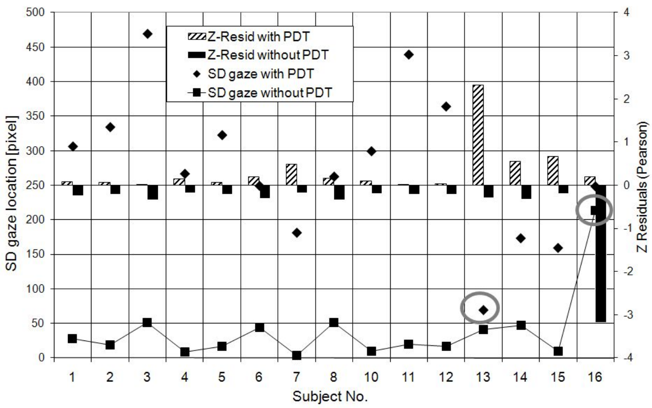

Figure 1. Z-Residuals for the binary logistic regression combined with the standard deviation (SD) of the gaze data to identify outliers.

No explanation was found for the two misclassified cases despite a secondary analysis of the videos and the reaction times. However, the video used for this purpose was taken from the scenery camera which did not include the driver's head. Therefore, despite analysis of the video, it cannot be ruled out that the participant wrongfully classified as being in the PDT condition was engaged in another secondary task. However, despite these two misclassified cases in the subset of data used to develop the function, the classification rate for the subset of data used to validate the function was still 100%.

4 CONCLUSIONS

The results are very encouraging regarding the development of an in-vehicle device to detect driver distraction. The robustness of the parameter is its most important feature when it comes to its suitability for its application in an accident-prevention device in driving. It does not require excessive calibration and can be derived from the raw data without additional processing steps such as needed for the calculation of fixations or saccades (Salvucci & Goldberg, 2000). This is especially important in time critical situations as present prior to (potential) accident occurrence. Furthermore, the method does not place additional demand on the driver. To circumvent the problem of unsatisfying explanations for outliers, future studies must place additional emphasis on recording all driver actions.

5 REFERENCES

Backhaus, K., Erichson, B., Plinke, W., & Weiber, R. (2006). Multivariate Analysemethoden - Eine anwendungsorientierte Einführung [Multivariate analysis methods - an application-oriented introduction]. Berlin: Springer.

Dietze, M. (2007). RoadView. Visualisation and analysis software for roads: TU Dresden; Chair of Road Planning and Road Design.

DIN EN ISO 15007-1. (2003). Messung des Blickverhaltens von Fahrern bei Fahrzeugen mit Fahrerinformations- und -assistenzsystemen. Teil 1: Begriffe und Parameter [Road vehicles - Measurement of driver visual behaviour with respect to transport information and control systems - Part 1: Definitions and parameters]. Berlin: Beuth Verlag.

Dukic, T., Hanson, L., & Falkmer, T. (2006). Effect of drivers' age and push button locations on visual time off road, steering wheel deviation and safety perception. Ergonomics, 49(1), 78–92.

European Commission. (2007). Commission recommendation of 22 December 2006 on safe and efficient in-vehicle information and communication systems: update of the European Statement of Principles on human machine interface (2007/78/EC). Retrieved April 5, 2009, from http://eur-lex.europa.eu/LexUriServ/site/en/oj/2007/l_032/l_03220070206en02000241.pdf

Jacob, R. J. K. (1995). Eye tracking in advanced interface design. In W. Barfield & T. A. Furness III (Eds.), Virtual environments and advanced interface design (pp. 258–288). Oxford: Oxford University Press.

Land, M. F., & Lee, D. N. (1994). Where we look when we steer. Nature, 369, 742–744.

McEvoy, S. P., Stevenson, M. R., & Woodward, M. (2007). The prevalence of, and factors associated with, serious crashes involving a distracting activity. Accident Analysis & Prevention, 39, 475–482.

Muir, B. M., & Moray, N. (1996). Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation. Ergonomics, 39, 429-460.

Regan, M. A., Lee, J. D., & Young, K. L. (Eds.). (2009). Driver distraction. Theory, effects, and mitigation. Boca Raton, FL: CRC Press.

Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. In Proceedings of the eye tracking research & application symposium 2000. New York: ACM Press (pp. 71–78) electronic version. Retrieved 24 July 2009, from http://www.cs.drexel.edu/~salvucci/publications/Salvucci-ETRA00.pdf

Schulz, R. (2007). WatchOut. Eye-movement analysis tool. Dresden: TU Dresden; Chair of Road Planning and Road Design.

Velichkovsky, B. M., Rothert, A., Kopf, M., Dornhöfer, S. M., & Joos, M. (2002).

Towards an express-diagnostics for level of processing and hazard perception. Transportation Research Part F, 5(2), 145–156.

Victor, T. (2005). Keeping Eye and Mind on the Road. Doctoral dissertation Uppsala University, Department of Psychology. Retrieved 2008.02.08, from http://uu.diva-portal.org/smash/get/diva2:167500/FULLTEXT01

Victor, T. W., Harbluk, J. L., & Engström, J. A. (2005). Sensitivity of eye-movement measures to in-vehicle task difficulty. Transportation Research Part F, 8(2), 167–190.

Weller, G. (2009). The Psychology of Driving on Rural Roads. Development and Testing of a Model. Unpublished Dissertation, Technische Universität Dresden, Dresden.

Wickens, C. D., Rice, S., Keller, D., Hutchins, S., Hughes, J., & Clayton, K. (2009). False Alerts in Air Traffic Control Conflict Alerting System: Is There a "Cry Wolf" Effect? Human Factors, 51, 446–462.