DRIVER DISTRACTION DETECTION WITH A CAMERA VISION SYSTEM

Matti Kutila[1], *Maria Jokela*[1], *Gustav Markkula*[2], *Maria Romera Rué*[3]

[1] VTT Technical Research Centre of Finland
[2] Volvo Technology Corporation
[3] Centro Tecnico SEAT, S.A.

ABSTRACT

Driver assistance systems and electronics (e.g. navigators, cell phones, etc.) steal increasing amounts of driver attention. Therefore, the vehicle industry is striving to build a driving environment where input–output devices are smartly scheduled, allowing sufficient time for the driver to focus attention on the surrounding traffic. To enable a smart human–machine interface (HMI), the driver's momentary state needs to be measured. This paper describes a facility for monitoring the distraction of a driver and presents some early evaluation results. The module is able to detect the driver's visual and cognitive workload by fusing stereo vision and lane tracking data, running both rule–based and support-vector machine (SVM) classification methods. The module has been tested with data from a truck and a passenger car. The results show over 80% success in detecting visual distraction and a 68–86 % success in detecting cognitive distraction, which are satisfactory results.

Index Terms— Stereo vision, classification, vehicle, distraction detection, camera

1. INTRODUCTION

Modern vehicles are full of driver-assistive electronics (multimedia displays, navigator, climate control, parking radar, etc.). In addition, third-party entertainment facilities—such as music players, PDA devices, mobile phones, etc.—are also siphoning off a growing part of the driver's attention, increasing the number of traffic incidences and even accidents. The study [11] indicates that wireless devices, passenger-related distraction (mostly conversation), and in-vehicle distraction sources are the most frequent reasons for incidences. Consequently, the automotive industry has paid more interest in controlling in-vehicle human–machine interface (HMI), including third-party products, in order to make driving more comfortable and more importantly to accentuate traffic safety.

The AIDE project[1] [4] is a response from the European Commission that takes into account available driver attention and the new time-sharing requirements between driver information and assistive systems. The key idea is not to warn the driver but to schedule input–output devices to allow the driver to concentrate more on the driving task when necessary. This is done, for example, by providing low-priority messages only when the driver is not required to give full attention to the surrounding traffic.

The key issue for estimating a driver's momentary state is to monitor a driver's behaviour in real-time. This paper focuses on monitoring a driver's visual and cognitive distraction. The Cockpit Activity Assessment (CAA) module has been built for and partially tested in detecting a driver's momentary state (see. Figure 1) [7], [8]. Visual distraction in this context is, roughly, a measure of how much the driver's attention is diverted from the road ahead, which obviously is the main target (i.e. most attention should be focused on the road). Cognitive distraction is related to reductions in the driver's awareness of the surrounding environment and is therefore only indirectly measurable. Examples of cognitive workload are daydreaming, thinking hard and conversations with passengers.

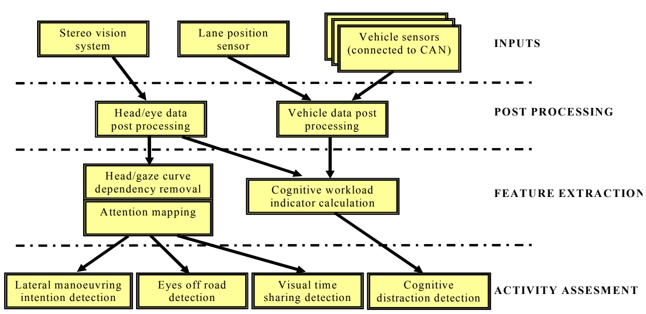


Figure 1. The CAA module architecture for monitoring driver's activity and intention

2. PRIOR KNOWLEDGE

The existing monitoring systems can be basically divided into two branches: drowsiness and distraction detection systems. However, the distinction between them is not clear since cognitive distraction may in some cases be linked to the driver's vigilance (e.g. daydreaming). They both influence the driver's physiological state by impairing alertness and thus, increasing the reaction time.

Early driver monitoring methods were being tested already 20 years ago in the aviation industry [1], [3]. They typically involve measuring the heart rate, eye blinks and EEG so as to estimate the stress level of a pilot. However, intrusive driver monitoring techniques are not suitable for an in-vehicle environment and therefore, camera-vision-based systems are preferred by the

[1] AIDE (Adaptive Integrated Driver-vehicle InterfacE) is the project initiated by the European Commission in the FP6. The project identity number is IST-1-507674-IP. There are 28 partners, including all the major automotive manufacturers in Europe, involving this activity.

automotive industry. A driver is not expected to wear special equipment when driving a car. Consequently, for example, detecting Percent Eye Closure (PERCLOS) and eye blinks are favoured methods for detecting fatigue in the vehicle environment [2], [6].

A number of studies have shown that a driver's behaviour changes due to workload, which can be observed by monitoring the driver directly or by following the vehicle's dynamics. The studies [5], [6], [12], [13] indicate that the driving-related parameters are changed due to cognitive workload. The result is that the distribution of gaze and head orientations over the time window is narrower (i.e. the variation of the parameters decreases) due to degraded situation awareness. The same influence has also been observed for lane-keeping performance and steering activity.

As a result of the prior knowledge, we state as the research hypothesis that: "*The driver's visual and cognitive distraction levels are effectively detected by using multiple data sources, including machine vision and a fusion of separate classification methods.*" Thus, the hypothesis is closely related to the selected classification techniques, which are the key issue of this paper.

3. VISUAL DISTRACTION DETECTION

The attention mapping algorithm is based on the driver's head and gaze directions (yaw and pitch angles). The view from the cockpit is divided into four clusters of interest: road ahead, windscreen, and left and right exterior mirror. The cluster sets – separately for the head and gaze signals - were manually defined and evaluated by examining the driver's behaviour and attention direction on pre-recorded videos of the test drives. The result cluster is determined by examining in which cluster the driver's gaze points [7].

The output from the attention mapping algorithms—telling whether or not the driver's attention is momentarily directed towards the road ahead—is used to estimate two output parameters related to visual distraction: *driver eyes off road*, and *driver visual time sharing*. The former is obtained by applying a noise-reducing filter to the attention mapping output, and it is a suitable parameter for use for example in combination with Advanced Driver Assistance Systems (ADAS), which can be adapted to where the driver's momentary attention is directed (the typical example being a forward collision warning given earlier when the driver is not looking at the road ahead). The latter parameter is, simply put, similar but further filtered to have even slower dynamics. The purpose of this is to detect when the driver is continuously dividing his attention between the road ahead and something else (e.g. a "secondary task"). As one part of the calculations, a simple model of driver visual awareness is used, modelling the ability of the driver to maintain a mental model of the surrounding traffic for a short time even when not directly looking at it.

As indicated in Figure 1, the attention mapping output is also used to generate an estimate of driver intent for lateral manoeuvring (e.g. lane changes). This is based on detection of repeated mirror checks.

4. COGNITIVE DISTRACTION DETECTION

The support vector machine (SVM) is a machine learning algorithm which was first introduced by the Russian scientist Vladimir N. Vapnik. The basic idea of SVM is to nonlinearly map the training data to a higher-dimensional feature space where it can be separated linearly. A kernel function $K(\mathbf{x}, \mathbf{x'})$ is used for mapping.

The separating hyperplane is generated by maximizing the margin between positive and negative classes, which leads to an optimization problem. For soft margin SVM, the dual form of the problem is given by,

$$\max W(\alpha) = \sum_{i=1}^{m} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{m} \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j) \qquad (1)$$

$$\textit{subject to } 0 \le \alpha_i \le \frac{C}{m} \textit{ for all } i = 1, \ldots, m$$

$$\textit{and } \sum_{i=1}^{m} \alpha_i y_i = 0 \cdot$$

Here, the parameter C determines the trade-off between minimizing the training error and maximizing the margin. The classification result of SVM is based on which side of the hyperplane the sample belongs. Thus, the decision function can be defined as the sign of the classifier,

$$f(\mathbf{x}) = \operatorname{sgn}\left(\sum_{i=1}^{m} y_i \alpha_i K(\mathbf{x}, \mathbf{x}_i) + b\right) \qquad (2)$$

where the kernel is assumed to be symmetric and of the form $K(\mathbf{x}, \mathbf{x}_i) = \Phi(\mathbf{x}) \cdot \Phi(\mathbf{x}_i)$. As can be seen from the decision function, only support vectors $\mathbf{x}_i$ are needed for classification. These samples actually construct the hyperplane, meaning the rest of the training samples become unnecessary during a classification task.

Using an SVM-type classification method for detecting cognitive distraction is not well tested in state-of-the-art studies. The SAVE-IT project [9] has proposed detecting the driver's distraction level with Hidden Markov Models (HMM) and mentioned Support Vector Machines (SVM) as an alternative solution. However, the project's experiments are executed using HMM and using only eye movements as an indicative factor. The advantage of HMM is that it takes into account the transitions from one state to another (e.g. sleep is a transition from a drowsy state and not directly from an alert state). However, SVM can adapt better to momentary changes. Ultimately, the cognitive distraction may occur rapidly, e.g. the mobile phone rings and the conversation steals the driver's attention. Moreover, the assumption is that one parameter alone does not reveal the distraction, but rather by fusing many parameters, the robustness of the detection can be improved. On this basis, we selected the SVM for our application.

5. TEST ARRANGEMENTS AND IMPLEMENTATION

The test data were gathered with a SEAT passenger car and a Volvo's truck. Tuning and testing the developed algorithms took place remotely in the office. The truck data was gathered by recruiting 12 professional drivers including one female to drive the Volvo FH 12 truck, which was equipped with the stereo camera system and a special data logging system [10]. The average age of the drivers was 40 (youngest 21, oldest 59). They had between 2 and 39 years of experience as professional truck drivers. The SEAT data included three ordinary drivers that had 5–10 years driving experience. The data gathering for the passenger car was not as exhaustive as with the truck since the experiments with the truck were partially used for the passenger car adaptation too and the purpose was more to fine tune the algorithms.

The stereo vision system used in the prototype vehicles is a commercially available product, faceLab of Seeing Machines, since the purpose of this work was to create algorithms for detecting driver distraction. Figure 2 shows the hardware needed by the eye-tracking system. The data post-processing unit (faceLab computer) calculates the gaze and head orientations from among the multiple other measures (e.g. saccades, eye blinking, head position, etc.). Classification and recognition of distraction is performed in separate computers, though this could be merged in future applications. The classifiers were built to the industrial-PC as Matlab/Simulink binaries, which run stand-alone and process input data in real-time (60 Hz). Furthermore, an MS Windows application was created to make the classifier adaptation easier and to execute a validation of the performance. The tool allows a drag-and-drop type definition of the attention clusters and is useful for visualizing the driver's momentary attention targets. The tool also creates a semi-automatic SVM parameter tuning capability and visualizes the borders of the SVM clusters by projecting them onto a 2D chart.

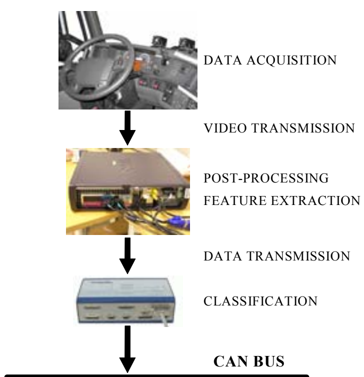


Figure 2. The equipment installed in the test vehicles (the Volvo truck and the SEAT passenger car)

The test roads include sequences of different types of environments like motorways, rural and city driving. The driving distractions in the test data sets were artificially induced: Cognitive distraction was induced by asking the drivers to perform mental arithmetic (repeated integer subtractions), while visual tasks were initiated by requesting the driver to read sequences of numbers from stickers attached to the radio, speedometer, mirrors, etc..

6. RESULTS

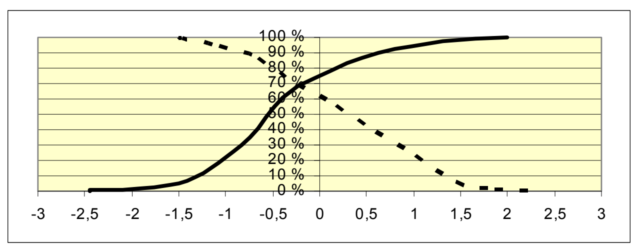
The performance of visual distraction detection is determined in terms of how well the algorithm can detect glances towards various clusters in a cockpit. The implemented attention clusters consist of left- and right mirrors and windscreen, which also included the road-ahead cluster. Table 1 shows the manually evaluated results in the truck application. During the tests, the model was re-adapted once to improve the windscreen cluster's borders. As can be seen, the road ahead cluster is well detected since that is directly in the camera view, thus rarely loosing eye tracking. Performance of the mirror detection is not as good as expected due to the large gaze and head movements in the truck. Unfortunately, passenger car tests are not available for reporting here. However, the most important thing is that the road-ahead detection can be performed with some 84% accuracy, which promises a good outcome for the visual distraction detection too.

Table 1. The results of the truck tests for capturing the different clusters in the truck's cockpit

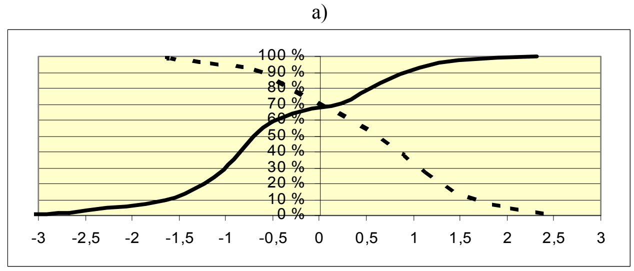
DRIVER	ROAD AHEAD	LEFT MIRROR	RIGHT MIRROR	WIND-SCREEN
D3	99 %	37 %	60 %	12 %
D4	91 %	27 %	66 %	10 %
D5	100 %	20 %	30 %	7 %
D6	98 %	67 %	69 %	4 %
D7	90 %	56 %	7 %	12 %
D8	99 %	7 %	64 %	10 %
COCKPIT MODEL RE-ADAPTED				
D6	85 %	51 %	24 %	37 %
D9	95 %	49 %	59 %	62 %
D10	50 %	32 %	38 %	56 %
D11	48 %	62 %	76 %	43 %
D12	76 %	58 %	24 %	35 %

The features used by the cognitive distraction detection module are: gaze angles, head rotations and lane position. The standard deviations of the above features are used as indicative measures of the driver's activity. Additionally, there are three quality parameters for estimating completeness of gaze and head angle data and the face tracking performance. Overall, six possible features are optionally selectable in the application.

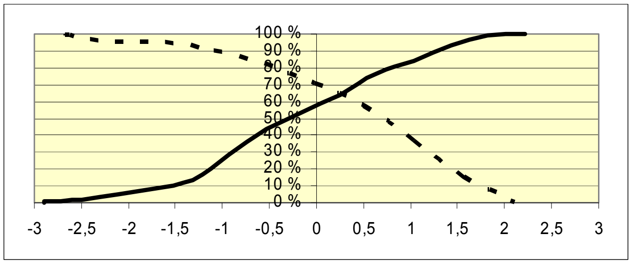
Figure 3 shows the results of the cognitive distraction detection evaluation. There are the three samples a, b and c presented with different input feature configurations. As the tests show, the optimal alternative is to use all the aforementioned features despite the lane-keeping measurement having a strong influence. Indeed, it stabilizes the output function, which improves the robustness of the algorithm by making it more predictable. Further, the test indicates that the lane-keeping measurement does not interfere with the performance of the cognitive distraction detection.



a)



b)



c)

Figure 3. Results of the truck tests for detecting cognitive distraction with a SVM-type classifier a) all six features are implemented, b) lane position measurement is excluded, c) the data quality factors are not included. Horizontal axis is the threshold between cognitive and non-cognitive output and the vertical axis is hit-rate. The hit-rate for non-cognitive detections is shown as a continuous line while cognitive detection is a dashed line.

The above graphs provide a realistic picture of the performance in the truck. The overall detection performance is some 68%. However, it is anticipated that further tests in a passenger car environment will give improved results. In the office, an 86% hit rate was achieved for the passenger car, which is a very high rate, especially if taking into account the lack of lane-position measuring equipment.

7. CONCLUSIONS AND FUTURE WORK

The attention mapping algorithm works well, providing an 84% detection of attention targets in a cockpit. Adding some filters to prevent suspiciously large head and gaze movements would probably improve the results. The performance achieved for the cognitive distraction detection is encouraging, especially in the passenger car case (86%). However, the outcome of the truck application (68%) is not as good as expected but is nonetheless promising. The hit rate has improved in recent tests after excluding cognitive distraction detection in a city environment (when the speed is below 60 km/h). In a city, the cognitive distraction is considerably more difficult to detect, and arguably not as commonly present since driving demand is higher due to manoeuvring.

The achieved results are sufficient in the case of the AIDE project since the objective is to schedule information flow of the in-vehicle HMI. For AIDE, 70% accuracy is sufficient and 85% would be good performance so that the driver does not realise the HMI scheduling. However, the issue would be very different if warning messages are provided since even 5% false alarms would frustrate human.

We stated in the hypothesis that image analysis accomplished with other driver- or driving-related measures can be used for detecting whether the driver is in a distracted state or not. The experiments have proven the hypothesis to be true. The distraction level can be estimated by using two different types of classification methods (rule based and SVM) and by utilising driver gaze orientations and lane-keeping measurements to estimate the visual and cognitive distraction levels.

Overall the classification algorithms seem to be working to a satisfactory level, though perhaps not yet sufficient for in-vehicle applications, but certainly close. However, work is still needed to reduce the total price of the system (now some 35 000 EUR). A vehicle is a very dynamic environment and special attention is needed to achieve robustness in the equipment to suit for example varying lighting conditions and driving habits. Nevertheless, the starting point is very good since we have already earned experience of the basic requirements (e.g. robustness, input features, etc.) for the monitoring facility.

8. REFERENCES

[1] W. B. Albery, D. W. Repperger, G. B. Reid, C. Goodyear, L. E. Ramirez and M. M. Roe, “Effect of noise on a dual task: subjective and objective workload correlates [pilot performance in fighter aircraft cockpit],” IEEE National Aerospace and Electronics Conference: NAECON 1987. Dayton, OH, USA. Vol. 4, pp. 1457-1463, 18-22 May 1987

[2] L. M. Bergasa, J. Nuevo, M. A. Sotelo, R. Barea and M. E. Lopez, “Real-Time System for Monitoring Driver Vigilance,” *IEEE Transaction on Intelligent Transportation Systems*, Vol. 7, Issue 1, pp. 63-77, 2006.

[3] J. A. East K. W. Bauer Jr. and J. W. Lanning, “Feature selection for predicting pilot mental workload: A feasibility study,” *International Journal of Smart Engineering System Design*, Vol. 4, Issue 3, pp. 183-193, Jul./Sep. 2002.

[4] J. Engström, J. Arfwidsson, A. Amditis, L. Andreone, K. Bengler, P. C. Cacciabue, W. Janssen, H. Kussman, and F. Nathan, “Towards the automotive HMI of the future: Mid-term results of the AIDE project,” 10th International Forum on Advanced Microsystems for Automotive Applications (AMAA). Berlin, Germany. Apr 25–27, 2006.

[5] J. Engström, E. Johansson and J. Östlund, “Effects of visual and cognitive load in real and simulated motorway driving,” *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 8, Issue 2, SPEC. ISS., pp. 97–120, 2005.

[6] R. Grace, V. E. Byrne, D. M. Bierman, J.-M. Legrand, D. Gricourt, B. K. Davis, J. J. Staszewski and B. Carnahan, “A drowsy driver detection system for heavy vehicles,” Proceedings of the 17th DASC. AIAA/IEEE/SAE. Digital Avionics Systems Conference. Bellevue, WA, USA. 31 Oct - 7 Nov 1998. Vol. 2, Issue 36, pp. 1-8, 1998.

[7] M. Jokela, “Implementation and Analysis of The Cockpit Activity Assessment Module,” Master of Science Thesis, Tampere, Finland: Tampere, University of Technology. 61 p., 2006.

[8] M. Kutila, “Methods for Machine Vision Based Driver Monitoring Applications,” Dissertation thesis, Tampere. Finland, VTT Publications series, 2006.

[9] J. Lee, M. Reyes, T. Smyser, Y. Liang and K. Thornburg, “SAfety VEhicles using adaptive Interface Technology (Task 5) Final Report: Phase 1 of the SAVE-IT project,” Available at: http://www.volpe.dot.gov/opsad/saveit, 2004.

[10] G. Markkula, M. Kutila, J. Engström, T. W. Victor and P. Larsson, “Online Detection of Driver Distraction – Preliminary Results from the AIDE Project,” Proceedings of the 2005 International Truck and Bus Safety and Security Symposium. Washington, Alexandria, Virginia, U.S.A., Pp. 86-96, 14-16 Nov 2005.

[11] V. L. Neale, T. A. Dingus, S. G. Klauer, J. Sedweeks and M. Goodman, “Overview of the 100-Car Naturalistic Study and Findings,” National Highway Traffic Safety Administration, Washington DC, U.S.A., 2005.

[12] T. Victor, “Keeping Eye and Mind on the Road,” Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences 9, Uppsala: Acta Universitatis Upsaliensis, 2005.

[13] J. Östlund, L. Nilsson, O. Carsten, N. Merat, H. Jamson, S. Jamson, S. Mouta, J. Carvalhais, J. Santos, V. Anttila, H. Sandberg, J. Luoma, D. de Waard, K. Brookhuis, E. Johansson, J. Engström, T. Victor, J. Harbluk, W. Janssen, R. Brouwer, “HASTE Deliverable 2: HMI and Safety-Related Driver Performance,” EU funded project GRD1/2000/25361 S12.31962, 2004.