

## **STEERING WHEEL BEHAVIOR BASED ESTIMATION OF FATIGUE**

Jarek Krajewski<sup>1</sup>, David Sommer<sup>2</sup>, Udo Trutschel<sup>3</sup>, Dave Edwards<sup>4</sup> & Martin Golz<sup>2</sup>

<sup>1</sup>Experimental Business Psychology

University of Wuppertal, Germany

<sup>2</sup>Faculty of Computer Science, University of Applied Sciences

Schmalkalden, Germany

<sup>3</sup>Circadian Technologies Inc.

Stoneham, MA, USA

<sup>4</sup>Product Safety & Compliance, Caterpillar Inc.

Peoria, IL, USA

Email: krajewsk@uni-wuppertal.de

**Summary:** This paper examined a steering behavior based fatigue monitoring system. The advantages of using steering behavior for detecting fatigue are that these systems measure continuously, cheaply, non-intrusively, and robustly even under extremely demanding environmental conditions. The expected fatigue induced changes in steering behavior are a pattern of slow drifting and fast corrective counter steering. Using advanced signal processing procedures for feature extraction, we computed 3 feature set in the time, frequency and state space domain (a total number of 1251 features) to capture fatigue impaired steering patterns. Each feature set was separately fed into 5 machine learning methods (e.g. Support Vector Machine, K-Nearest Neighbor). The outputs of each single classifier were combined to an ensemble classification value. Finally we combined the ensemble values of 3 feature subsets to a of meta-ensemble classification value. To validate the steering behavior analysis, driving samples are taken from a driving simulator during a sleep deprivation study (N=12). We yielded a recognition rate of 86.1% in classifying slight from strong fatigue.

## **INTRODUCTION**

Determining driver fatigue through the analysis of biosignals poses great challenges. Not all biosignals demonstrate fatigue based changes to the same degree. Some biosignals contain more information about fatigue than others. Among them, EEG (reflecting cortical activity) and EOG (reflecting eye and eyelid movement dynamics) have the potential to be used as a laboratory reference standard and have been utilized e.g. in evaluating fatigue monitoring technologies (e.g. Golz, Sommer, Holzbrecher & Schnupp, 2007; Sommer, Golz & Krajewski, 2008). Even so, the requirements for these electrode based instruments are not conducive for being used out in the field. The major drawbacks are a lack of robustness against various environmental conditions (e.g. high humidity or temperature) and a lack of comfort and longevity due to electrode sensor application. In contrast to these electrode-based instruments, using steering wheel movement as an indicator for fatigue is more robust under these same operating conditions. Collecting this data is favourable for fatigue detection since it is non-obtrusive and uses cheap, durable, and maintenance free sensors that are already integrated into the steering wheel system.

The present paper will describe a steering behavior based approach to estimate a driver's level of fatigue. It is organized as follows: Section 2 discusses how steering behavior based measures are

correlates to fatigue. In Section 3 the procedure of computing features from time, frequency and state space is explained. Section 4 explains the design of the sleep deprivation study used for building a fatigue driver database. Having provided the results of the fatigue detection in Section 5, the paper closes with a conclusion and a discussion of the future work in Section 6.

## **FATIGUE INDUCED EFFECTS ON STEERING BEHAVIOR**

In general, steering behavior is influenced by characteristics of the driving task (e.g. speed, curvature, and lane width), driver traits (e.g. driving experience), and driver states (e.g. laxness, distraction or fatigue). Drivers are constantly judging the situation ahead and applying small, smooth, steering adjustments to correct for small road bumps and crosswinds by turning the steering wheel in small increments. This steering behavior can be impaired when drivers are becoming fatigued. A wide variety of steering wheel metrics have been suggested to measure steering behavior, from standard deviation of steering wheel angle (e.g. Liu, Schreiner & Dingus, 1999), steering wheel velocity (Peters, Kloeppe & Alicandri, 1999; Skipper, Wierwille & Hardee, 1984), steering wheel action rate (McDonald & Hoffman, 1980; Verwey, 2000), to more advanced metrics such as high frequency component of steering wheel angle (Östlund et al., 2004) and steering entropy (e.g. Godrich et al., 2004; Nakayama, Futami, Nakamura & Boer, 1999; Siegmund, King & Mumford, 1996). Fatigue-induced effects on steering behavior could be summarized follows: less small, smooth steering adjustments (micro-corrections), more zigzag and slow oscillation, greater steering entropy (measure of steering randomness), larger erratic steering movements (indicating e.g. overcorrecting for unanticipated road changes), lateral drift outside the driver's comfort zone, more large and fast steering corrections (Paul et al., 2006; van der Hulst, Meijman, & Rothengatter, 2001).

Until recently, the analysis of steering wheel movement was based on finding single features correlating with fatigue. The detection task was reduced to a simple thresholding of these features, but large inter- and intra-individual differences in fatigued driving patterns are an issue. Due to this overwhelming complexity, a sufficient and accurate classification of individual steering samples was not achieved. Thus, we propose a multivariate, data fusion, and machine learning based approach including the computation of a large, highly redundant feature set modelling fine grained temporal steering behavior structures (cf. Sayed & Eskandarian, 2001). This approach showed its superiority in many pattern recognition tasks (e.g. Ruvolo, & Movellan, 2008).

## **FEATURE EXTRACTION FOR STEERING BEHAVIOR**

Features of steering wheel movement can be divided into time, frequency and state space domains. This assignment follows the first processing step of computing frame level descriptors, independent of feature characteristics of the second, contour describing, functional based processing step.

### **(a) time domain features**

Within the time domain the following features can be extracted: regression descriptors (e.g. regression slope, intercept, maximum of regression error), class distribution measures (e.g. number of values within steering angle bin 0.0-0.1), peak amplitudes and distances (e.g. mean

distance of peaks; maximum of peak amplitude), entropy, zero crossing distances and slope (e.g. maximum of distance between consecutive zero crossings; mean velocity of steering angle in zero crossings)

(b) frequency domain features

To capture fine temporal changes of spectral descriptors we performed a framing and windowing of the signal (frame size = 512, frame shift = 256, hanning window), and computed the power spectral density per frame. The resulting frame-level descriptors (FLDs) were aggregated to FLD-contours. The next processing step captures temporal information of the FLD contours by computing functionals. Frequently used functionals are percentiles (quartiles, quartile ranges, and other percentiles), extremes (min/max value, min/max position, range), distributional functions (number of segments/ intervals/reversal points), spectral functionals (DCT coefficients), regression functions (intercept, error, regression coefficients), higher statistical moments (standard deviation, skewness, kurtosis, and zerocrossing-rate), means (arithmetic mean and centroid). This procedure of combining FLDs and functionals results in 803 frequency domain features as e.g.: relative and absolute power spectral density (PSD) of raw and first derivate contours in 30 spectral bands (e.g. minimum of relative PSD of first derivate of 0.6-0.7Hz spectral band contour), band energy ratios (PSD 1-5hz/ PSD 0-1hz), spectral flux (e.g. max of spectral flux = Euclidean distance of PSD between consecutive frames), and long term average spectrum descriptors (e.g. skewness of PSD distribution).

(c) state space

To extract the nonlinear properties of the steering angle signal, we computed a three-dimensional state space (steering wheel position, steering wheel velocity, steering wheel acceleration), and a reconstructed phase space ( $d = 3,4$  ;  $\tau = 1$ ; steering angle  $t_0$ , steering angle  $t_1$ , steering angle  $t_2$ ). The geometrical properties of the resulting attractor figures were described by trajectory based descriptor contours (angle between consecutive trajectory parts, distance to centroid of attractor, length of trajectory leg). The temporal information of the contours was captured by computing functionals.

## METHOD

*Subjects, Instruments, and Experimental Procedure.* Twelve participants took part in this study voluntarily. Initial screening excluded those having severe sleep disorders or sleep difficulties. The participants were instructed to maintain their normal sleep pattern and behavior. We conducted a within-subject sleep deprivation design (01.00 - 08.00 a.m). During the driving sessions a well established, standardized self-report fatigue measure, the Karolinska Fatigue Scale (KSS), was used orally by the subjects and 2 experimental assistants every 4 min. In the version used in the present study, scores range from 'extremely alert' to 'extremely sleepy, can't stay awake'. Given the verbal descriptions, scores of 8 and higher appear to be most relevant from a practical perspective as they describe a state in which the subject feels unable to keep stay awake.

During the night, the subjects were confined to the laboratory, conducting a driving simulator task and were supervised throughout the whole period. The driving simulation software system has been developed by the Department of Computer Science, Univ. of Applied Sciences

Schmalkalden, Germany. The fixed-based driving simulator consisted of a normal passenger city car (GM Opel Corsa) with original controls. The driving task involved a 40-minute night drive on a monotonous two-lane motorway course with simulated effects of headlights, i.e. involving a restricted range of sight. One round on the course with a speed of about 100 km/h took about 15 minutes. The course was mainly straight with sustained long curves. There were no obstacles or oncoming traffic.

*Driving data and Feature Extraction.* We conducted a validation experiment to examine whether automatically trained models can be used to recognise the fatigue of subjects. Our approach can be summarized in four steps: 1. Collect individual steering data and the associated fatigue ratings for each participant; 2. Extract relevant features from the steering data; 3. Build statistical models of the fatigue ratings based on the features; 4. Test the learned models on unseen steering data. The collection of data took place in a nearly dark laboratory room (sampling rate: 20 Hz). Furthermore the subjects were given sufficient prior practice so that they were not uncomfortable with this procedure. After elimination of values in the threshold region ( $7.5 \pm 0.5$  KSS) the binarization of the fatigue values leads to two classes: non-strong fatigue ( $KSS \leq 7$ ; Non-SF) and strong fatigue ( $KSS \geq 8$ ; SF). The total number of 714 steering samples was split into 390 samples of Non-SF and 324 samples of SF (KSS:= mean of one self-report and two observer report KSS-ratings;  $M= 7.06$ ;  $SD= 1.99$ ). To extract relevant features from the steering behavior data the data was split into segments 240 seconds in duration each for a total of 4800 data points per segment. As explained above, we estimated 88 time domain features, 803 frequency domain features, and 360 state space domain features. In sum, we computed a total amount of 1251 features per driving segment including also the standard steering behavior features as variance of steering angle, high frequency component, steering reversal rate, steering wheel action rate, peak steering deflection, number of steering holds. The potential advantage of the brute-force feature extraction approach proposed here (large collection of simple, light weighted features) is that it could cover a broad range of possible steering phenomena related to fatigue.

*Pattern Recognition.* To enhance robustness of the classification we feed these features separately into 5 sophisticated, but standard machine learning methods (support vector machine linear kernel (SVM\_l) and radial kernel (SVM\_r), 5- nearest neighbour (5-NN), decision tree (DT), and logistic regression (LR). By averaging the classification outputs, an ensemble classifier was created per each feature subset. In a last step the ensemble-classifier outputs from the feature subsets (time, frequency and state space) were combined (averaged and dichotomized) to a final fatigue prediction value.

## V. RESULTS

The raw steering angle signal, spectrogram, and state space plot in Figure 1 provide a first insight into possible fatigue sensitive steering features. In the fatigued driver the raw steering wheel position and velocity signal show the following correlations to fatigue: (a) time domain features:  $r_{\text{mean peak amplitude} - \text{KSS}} = .59$ ,  $r_{90\text{th percentile} - \text{KSS}} = .57$ ,  $r_{\text{mean absolute deviation of steering velocity} - \text{KSS}} = .52$  (see Fig. 1 top row), (b) frequency domain features:  $r_{\text{median spectral flux} - \text{KSS}} = .59$ ,  $r_{\text{median spectral flux of steering velocity} - \text{KSS}} = .59$ ,  $r_{\text{interquartile range of PSD in frequency band 1-3 Hz} - \text{KSS}} = .55$  (see Fig. 1 mid row), and (c) state space domain:  $r_{95\text{th percentile of distance to centroid of attractor (time lag} = 1 = 1, \text{ embedded dimension} = 4) - \text{KSS}} = .55$ ,

$\Gamma_{95}$ th percentile of distance to centroid of attractor (time lag = 1 = 1, embedded dimension = 3) – KSS = .54,  $\Gamma_{mean}$  distance to centroid of attractor (time lag = 1 = 1, embedded dimension = 3) – KSS = .50 (see Fig. 1 bottom row).

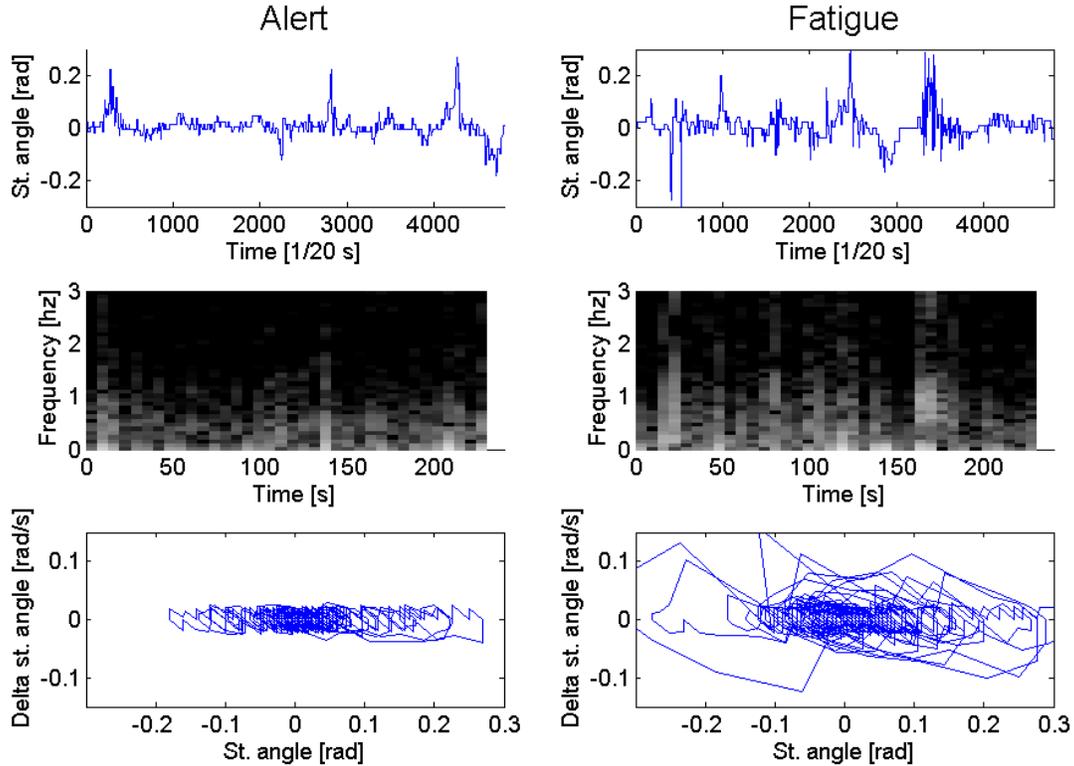


Figure 1. Steering angle signal, spectrogram and state space trajectory of non-strong fatigue (Non-SF) and strong fatigue (SF) driver. High power spectral densities (PSD) are coded white, low PSD are coloured grey within the spectrogram.

Applying the above described ensemble classification method, this feature set achieved 86.1% recognition rate (RR; ratio correctly classified samples divided by all samples) and 85.4% class-wise averaged classification rate (CL; average of sensitivity and specificity) in the classification of sleepy driver (sensitivity =77.4%; specificity 93.3%).

Table 1. Class-wise averaged classification rates (CL), recognition rates (RR), sensitivity (SE), and specificity (SP) (in %) on the test set using 3 different feature subsets (# = number of features) and an ensemble-classifier (SVM\_l, SVM\_r, LR, DT, 5-NN).

Feature Domains	#	CL	RR	SE	SP
Time	88	86.7	87.3	79.9	93.3
Frequency	803	77.6	78.2	71.3	83.8
State space	360	82.7	83.2	76.9	88.5
Combine all 3 subsets	1251	85.4	86.1	77.4	93.3

## DISCUSSION

Assessing driver fatigue using the steering wheel behavior analysis presented here yielded a recognition rate of 86.1%, which is among the highest values ever reached for a totally non-obtrusive system. The main findings referring to construction principles of this system may be summarized as following. First, the time frequency and state space domain features extracted from the driving simulator contains a substantial amount of information about the driver's fatigue state. Our analyses documented that time domain feature subsets yielded the best classification results, which might be in part due to imbalanced feature set sizes. Using different wrapper based feature selection and weighting methods would probably provide a more detailed view on useful feature subsets. Secondly, we found that the early fusion of 5 standard classifiers increases robustness and the discriminative classification power. Furthermore, the final meta-ensemble classifier combining fatigue prediction of 3 feature subsets (time, frequency and state space domain) on unseen data but known driver enhanced robustness of the classification.

Future research should address the topic of improving the fatigue monitoring by enriching the feature sets with information of lane deviation, time to lane crossing, and pedal movements. Furthermore, implementing separate fatigue prediction models for different driving tasks, geometrical road characteristics and environmental conditions (as e.g. speed, crosswind, street-granularity, curvature, or lane width) might improve the general fatigue prediction accuracy. Another promising improvement could be realized by using driver state data as e.g. personal driving performance baselines for the personalization of the prediction models. Moreover considering various driver states could be useful, because it is likely that especially combinations of activation and arousal states might be responsible for intra-individual variance in fatigue steering patterns. When combining a low energetic activation state such as fatigue with high self-induced (e.g. fighting against sleepiness) or external stressor-induced arousal (e.g. inner tension within high mental workload situations), the typical fatigue driving pattern might be masked. To handle this fact, explicit (for each internal state combination separately) or implicit (based on classifier immanent clustering) modelling is necessary. In both cases it is essential to feed the classifiers with sufficient learning data including these combinations of fatigue and arousal states.

## REFERENCES

- Golz, M., Sommer, D., Holzbrecher, M., & Schnupp, T. (2007). Detection and prediction of driver's microsleep events. *Proceedings Int Conf Road Safety on Four Continents*, 14.
- Goodrich, M. A., Boer, E. R., Crandall, J. W., Ricks, R. W., & Quigley, M. L. (2004). Behavioral entropy in human-robot interaction. *Proceedings of PERMIS*, Gaithersburg, Maryland.
- Hulst, van der, M., Meijman, T., & Rothengatter, T. (2001). Maintaining task set under fatigue: a study of time-on-task effects in simulated driving. *Transportation Research Part F*, 4 (2), pp. 103-118.
- Liu, Y.-C., Schreiner, C. S., & Dingus, T. S. (1990). Development of human factors guidelines for advanced traveler information systems (ATIS) and commercial vehicle operation

(CVO). *Human Factors Evaluation of the Effectiveness of Multimodality Displays in ATIS*. NHTSA FHWA-RD-96-150.

- McDonald, W. A., & Hoffman, E. R. (1980). Review of relationships between steering wheel reversal rate and driving task demand. *Human Factors* 22(6), 733-739.
- Nakayama, O., Futami, T., Nakamura, T., & Boer, E. R. (1999). Development of a steering entropy method for evaluating driver workload. *Proceeding SAE International Congress and Exposition*, Detroit, Michigan, USA.
- Östlund, J., Nilsson, L., Carsten, O., Merat, N., Jamson, H., Jamson, S., Mouta, S., Carvalhais, J., Santos, J., Anttila, V., Sandberg, H., Luoma, J., de Waard, D., Brookhuis, K., Johansson, E., Engström, J., Victor, T., Harbluk, J., Janssen, W., & Brouwer, R. (2004). *Deliverable 2 - HMI and safety-related driver performance. Human Machine Interface And the Safety of Traffic in Europe*. Project GRD1/2000/25361 S12.319626.
- Paul, A., Boyle, L., Boer, E. R., Tippin, J., & Rizzo, M. (2006). Steering entropy changes as a function of microsleeps. *Proceedings of International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, 3, 441-447.
- Peters, R. D., Kloeppe, E., & Alicandri, E. (1999). Effects of partial and total sleep deprivation on driving performance. *US Department of Transportation, Federal Highway Administration (Eds)*, Publication No. FHWA-RD-94-046.
- Ruvolo, P., & Movellan, J. (2008). Automatic Cry Detection in Early Childhood Education Settings. *IEEE International Conference on Development and Learning*, 7, 204-208.
- Sayed, R., & Eskandarian, A. (2001). Unobtrusive drowsiness detection by neural network learning of driver steering. *Proceedings of the Institution of Mechanical Engineers. Part D, Journal of Automobile Engineering*, 215, 969-975.
- Siegmund, G. P., King, D. J., & Mumford, D. K. (1996). Correlation of steering behavior with heavy-truck driver fatigue. *Society for Automotive Engineers Technical Report # 961683*.
- Skipper, J. H., Wierwille, W. W., & Hardee, L. (1984). An investigation of low-level stimulus-induced measures of driver drowsiness. Vehicle Simulation Laboratory, Human Factors Group (Eds.). *IEOR Department Report #8402*. Virginia Polytechnic Institute and State University, Blacksburg, Virginia.
- Sommer, D., Golz, M., & Krajewski, J. (2008). Consecutive detection of driver's microsleep events. In J. Vander Sloten, P. Verdonck, M. Nyssen, J. Haueisen (Eds.), *IFMBE Proceedings*, 22, (pp. 243-247). Springer: Berlin.
- Verwey, W. B. (2000). On-line driver workload estimation. Effects of road situation and age on secondary task measures. *Ergonomics*, 43, 187-209.
- Yoshihiro T., Y., & Furukawa, Y. (2005). Estimate of driver's fatigue through steering motion. *Proc. IEEE International Conf. on Systems, Man, and Cybernetics*, 2, 1765-1770.