Is Posturography a Candidate for a Vigilance Test?

T. Schnupp[1], A. Schenka[1], D. Edwards[2], J. Krajewski[3] and M. Golz[1]

[1] Faculty of Computer Science, University of Applied Sciences Schmalkalden, Schmalkalden, Germany
[2] Machine Research, Caterpillar Inc., Peoria, Illinois, USA
[3] Work and Organizational Psychology, University of Wuppertal, Germany

Abstract— **Studies exhibit that between 20 and 40% of traffic accidents in Germany are related to driver's hypovigilance. Hypovigilance, as stated by some authors, effects driver's performance in a similar way as alcohol consumption does. But unlike blood alcohol level testing up to now there is no mobile, non-invasive vigilance test with low test durations. Posturography - a method to assess the balance control system quantitatively - may provide the possibility for such a vigilance test. In this paper we will investigate the discriminatory abilities of posturography using data recorded in conjunction with overnight driving simulation experiments commissioned by Caterpillar Machine Research. A total of 19 young adults volunteered to participate in balance assessment. Experiments included a driving simulation with repetitive assessment of subjective self-rating while driving using Karolinska Sleepiness Scale. As objective vigilance score the standard deviation of lateral position was chosen. Subjective and objective vigilance scores and the Time-Since-Sleep are utilized as labels for discriminant analysis. Two kinds of features were extracted from posturographic recordings. Among others, parameters of diffusion plot and sway density analysis were utilized as features in time domain. As spectral domain features power spectral densities were estimated and averaged in empirically optimized equidistant frequency bands. The usefulness of posturography as vigilance test was evaluated by the mean test set error of computational intelligence algorithms including artificial neural networks and Support-Vector Machines (SVM). These algorithms can be regularized between local and global decision finding. SVM using Gaussian kernel function achieve error rates of 8.8% (leave-one-out cross-validation). Considering some concerns regarding reliability and validity we conclude that it is possible to discriminate patterns of different vigilance levels using posturography.**

Keywords— **Posturography, Vigilance, Fatigue, Support-Vector Machines, Computational Intelligence**

I. Introduction

Human factors are a major cause for traffic accidents. It is expected, that 93.5% of all incidents are related to human factors. Beside others, hypovigilance - the lack of being capable to maintain attention during monotonous situations and reacting properly to weak stimuli - is one of these human factors. It can be deduced from accident reports that hypovigilance plays a role in at least 25% of all accidents. Means to counter hypovigilance in road traffic can be grouped into three types: educational, legal and technical means. Educational means are the most promising approach in order to raise people's awareness concerning the risks of driving impaired. The effort put into these means – besides some commonly event-related medial attention – is limited. From a legal point of view driving within an impaired state is prohibited. But law enforcement needs a quick, mobile and reliable vigilance testing method similar to the (blood) alcohol level testing. Vigilance tests are one possible technical mean supporting legal ones. Up to now no vigilance test suitable for comprehensive field usage is available. Among other reasons usual trial length is too long (e.g. psychomotoric vigilance test, PVT). In addition to enforcing legal means, the development of objective vigilance measures supports educational efforts. Objective measures help justifying subjective awareness. On a wider scope, vigilance testing can support individual alertness management providing circadian pace markers. In recent year's publications have shown that posturography issensitive to sleep deprivation and hypovigilance [1][2]. Posturography utilizes a force platform for non-invasive body sway measurements, providing one possibility to develop a vigilance test that meets the criteria mentioned. Data recorded in a pilot study aside overnight driving simulation experiments was used for evaluating the discriminatory abilities of posturography. Experiments were commissioned by Caterpillar Machine Research in order to benchmark different Fatigue Management Technologies [3]. Beside the Time-Since-Sleep (TSS), objective and subjective hypovigilance scores were used as class labels. The subjective hypovigilance scores are obtained using Karolinska Sleepiness Scale (KSS) reflecting the level of perceived sleepiness. Whereas both, the standard deviation of lateral lane position (SDLat) and posturography, are measures concerning physiological sleepiness. Benchmarking results are obtained using several classifiers with each discriminatory task, including Support-Vector Machines (SVM) and Learning Vector Quantization (LVQ).

II. Experiments

Study Design: Posturography recordings were obtained alongside overnight driving simulation experiments. A total of 19 (13 males, 6 females) young, healthy volunteers, aged 18 to 26 years, mean age 22.0±2.3 years, without known diseases or impairments of their visual, proprioceptive or vestibular system participated in this study. Adherence to an instructed sleep/wake regime prior to experiments was observed by actigraphy. Each subject conducted eight one hour lasting test-sessions. The temporal distribution can be obtained from Fig. 1. Sessions included 40 minutes of night-driving simulation, followed by two vigilance tests. The standard deviation of lateral lane position was utilized as an objective hypovigilance measure [4]. During driving simulation driver's KSS self-rating was assessed in three minute intervals. Driving performance was recorded by different variables including SDLat.

Posturography: In order to differentiate between test-sessions one to eight of the FMT study and sessions of the posturographic experiments latter are labeled A to C. Two posturographic trials were performed each prior to first (session A), between fourth and fifth (session B) and after last FMT test-session (session C, see Fig. 1). In the first trial without any impairment of visual feedback ("eyes open", EO) subjects were instructed to focus a marker on eye level in a distance of approximately two meters. The second trial was performed after a short break where altered instructions were given. With lights turned off subjects had to close the eyes, disabling visual feedback ("eyes closed", EC). In both trials 120 second stabilograms have been recorded. Subjects were instructed to stand upright, hands folded in a way that the tips of thumbs are placed on the navel. This way, a relatively fixed pose for all measurements can be achieved without utilizing Romberg pose. The exhausting effect of Romberg pose was expected to superimpose the effects of hypovigilance, e.g. by causing aversions. Possible implications of this decision will be discussed within the conclusion section. A self-developed 3-point force platform with 1506 Hz sampling rate was utilized. Quality control measurements following an standard procedure were performed and passed.

III. Methods

Preprocessing: Acquired sensory signals are transformed into stabilograms – a bivariate trajectory representing body sway in anterior-posterior (AP) and medio-lateral (ML) direction. In order to raise the number of patterns available for generalization, stabilograms have been segmented. Based on experience two different segmentation lengths (SL) were benchmarked. Therefore each posturographic measurement lead to six segments with 20 seconds length or three segments with 40 seconds respectively.

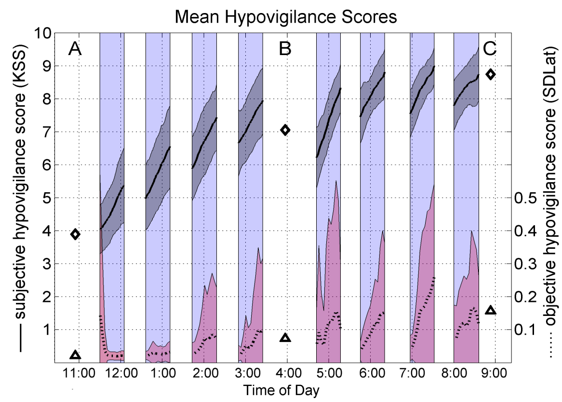

Fig. 1: KSS as well as SDLat values averaged over all subjects show clear Time-On-Task and Time-Since-Sleep effects. Diamonds and pyramids indicate the time of posturographic recording sessions. Marker's ordinate values represent the associated subjective respectively objective hypovigilance scores. Light colored areas indicate ±1 SD.

Feature Extraction: Two different sets of features have been derived from recorded stabilogram components. The first feature set consists of 23 time-domain features, including global, structural and physical parameters of balance. This set is a well-performing selection evaluated during earlier experiments. Time-Domain features include analysis of diffusion plot [5] and sway density curve [6]. Forsman et. al [2] identified the critical time interval (obtained by analysis of the diffusion plot) to be among the most significant single features for the regression of subject's TSS. The second set contains spectral domain features. In earlier examinations we have shown that time domain features are outperformed by those from spectral domain [7]. Power spectral densities (PSD) have proven to be reliable and robust measures of balance. PSD values are estimated and averaged in equidistant frequency bands. This reduces the dimensionality of the classification problem which is beneficial for some classifiers (e.g. LVQ). The amount of information lost by band averaging can be regarded as low, because adjacent frequency bins correlate significantly. Parameters of band averaging were optimized empirically utilizing OLVQ1 with 25 repetitions of random subsampling validation (80:20 partitioning). An optimal bandwidth was evaluated at 0.1 Hz. A lower cutoff frequency of 0 Hz and an upper cutoff at 11 Hz performed best (e.g. Fig. 3).

Data- & Feature-Fusion: In both trials (EO, EC) stabilograms where recorded. Each stabilogram consists of two components: AP and ML body sway. Best results have been observed by fusing all four sources into a single feature-vector (Fig. 3). Within other fields of biosignal analysis the fusion of different feature extraction methods show a significant increase in classification performance. Due to the pilot character of this study comprehensive optimization in order to maximize classification performance was not a primary objective of this work. Hence the fusion of time-domain features together with those from spectral domain and the usage of further methods (e.g. wavelet-decomposition or state-space features) was omitted.

Label Generation: Patterns obtained from posturographic recordings were separated into classes by four means. As a first approach three classes according to EO/EC trials, grouped by subject's TSS were defined. Using this separation it is possible to evaluate changes in EO vs. EC discrimination performance over night. A hypovigilance related influence on posturographic data may affect the features supporting discrimination between EO and EC trials. One possible expectation is, that hypovigilance alters the effectiveness of visual feedback interpretation, lowering the influence of visual feedback on postural equilibrium. Therefore the differences between EC and EO recordings is reduced, increasing the difficulty of EO/EC discrimination. This can be regarded as an indication for circadian rhythm based effects on human balance and would therefore endorse posturography as a mean of vigilance testing. In the following steps patterns were classified according to associated vigilance scores. The first approach utilizes TSS as labels. In the second approach SDLat was calculated for each driving session. Session A was linked to subject's SDLat during the first driving simulation task. Session B was associated with the average SDLat value from the fourth and fifth driving session. The last session C was linked to the SDLat measured during the eighth driving session. As expected to be demanded by later users of a vigilance test the associated labels were binarized. Binarization was obtained by applying a threshold of 17.4%. Samples linked to an objective score equal to 17.4% were omitted (less than 4%), leaving two equally sized classes. As a third vigilance score the subjective vigilance rating was used for pattern labeling. Subjective vigilance scores are obtained utilizing KSS prior to posturographic recordings. The binarization threshold was fixed to a KSS value equal 7. For both measures lower values are regarded as "mild hypovigilance", whereas higher values are regarded as examples of "strong hypovigilance". Omitting samples with an KSS score of exactly 7 (less than 4%) lead to two equally sized classes.

Discriminant Analysis: Distribution of values for single features shows a non significant shifting of mean values over night having a high variation and overlapping between different TSS labels (e.g. Fig. 2). Multivariate analysis utilizing soft computing may be able to generalize a function discriminating different levels of hypovigilance. Different methods were applied. As a representative for global decision finding Fisher's Linear Discriminant Analysis (LDA) was utilized. A purely local decision was obtained by the k-Nearest Neighbor algorithm (kNN) with k set to 1. Regularization between a global and local decision was achieved by varying k, using LVQ (especially OLVQ1) with a empirically optimized number of prototype neurons and by Support-Vector Machines (SVM) with empirically optimized hyper parameters. Testset error rates are obtained utilizing 25 fold random subsampling with an 80:20 separation (LDA, kNN, OLVQ1) and fast leave one out validation estimation (SVM, [8]) respectively.

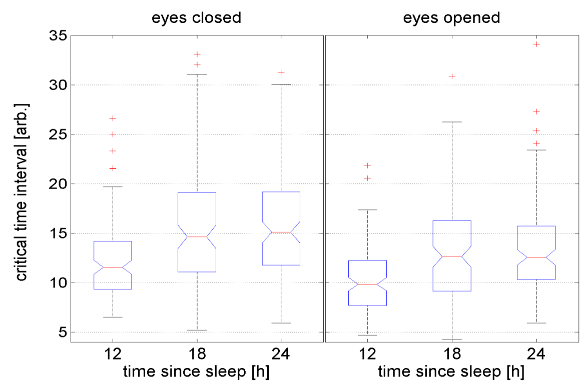

Fig. 2: Box plot of the critical time interval of the change from open-loop to closed-loop control versus Time-Since-Sleep. The increase of mean values between the first and second session is significant.

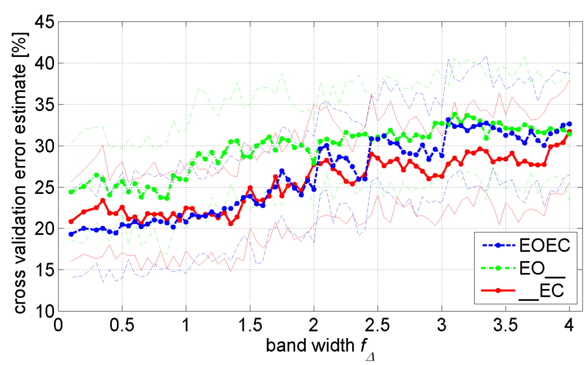

Fig. 3: Mean and standard deviation of test errors versus width of spectral bands. Three different PSD feature sets were used (see text). This is an example of the empirical optimization of free parameters which indicates low values of the band widths are optimal.

IV. Results

Discriminant Analysis - EO vs EC: Results show that the EO vs. EC discriminatory performance reduces between session A and B (e.g. 30.7±6.3 to 35.8±6.6%, PSD features, SL 20s). From other variables (e.g. SDLat, KSS) can be derived that subjects hypovigilance rises significantly between these sessions. Despite hypovigilance measure values rise significantly between Session B and C, no further drop of classification performance can be observed (e.g. 35.8±6.6 to 36.1±5.7%, PSD features, SL 20s). One possible explanation is that the last session is subjected to different effects. The most prominent ones are circadian rhythm and motivation due to the end of driving experiments. These effects may superimpose those of hypovigilance.

Discriminant Analysis - Hypovigilance Scores: A subset of the final results is shown in Table 1. The most prominent result of discriminatory analysis according labels defined by hypovigilance scores is that LDA fails in all trials indicating that hypovigilance discrimination cannot be regarded a linear separable problem. Finding a general function to discriminate between samples from session A and B was realized with acceptable error rates (8.8±0.0%, SVM). The difference between subject's hypovigilance level in session A and B is strong, with A being a quite alert state (Time-Since-Sleep: 12 hours, additional activation due to unknown laboratory situation) and B being a state near to the circadian trough. Session C is biased by alerting effects (e.g. end of experiments). All-in-all the obtained error rates are unsatisfying and have to be improved significantly in order to utilize posturography as a vigilance test.

Table 1: Mean and SD of Test Set Error Rates

Results obtained by discriminant analysis regarding different hypovigilance scores (class labels).

Set	Classifier	KSS	SDLat	A vs B
PSD	LDA	50.3 ± 6.7	48.3 ± 5.4	48.8 ± 7.2
SL 20s	OLVQ1 (n)	26.5 ± 3.9 (120)	22.9 ± 4.2 (165)	16.3 ± 5.8 (155)
	SVM (C; γ)	21.9 ± 0.0 (9; -2.25)	18.8 ± 0.0 (0.875; -2.75)	10.1 ± 0.0 (1.0; -2.5)
	kNN (k=opt)	14,0 ± 33.9 (199)	20.0 ± 9.3 (1)	17.2 ± 10.0 (1)
	kNN (k=1)	27.4 ± 8.2	20.0 ± 9.3	17.2 ± 10.0
PSD	LDA	49.3 ± 9.6	49.1 ± 8.8	50.5 ± 10.1
SL 40s	OLVQ1 (n)	18.2 ± 7.6 (185)	16.4 ± 5.9 (155)	11.1 ± 6.6 (170)
	SVM (C; γ)	24.1 ± 0.0 (8; -2.5)	20.8 ± 0.0 (2.0; -2.75)	8.8 ± 0.0 (1.25; -3.125)
	kNN (k=opt)	14.8 ± 10.8 (1)	12.7 ± 11.5 (2)	13.9 ± 15.0 (1)
	kNN (k=1)	14.8 ± 10.8	17.9 ± 14.0	13.9 ± 15.0

V. Conclusions & Outlook

One of the major disadvantages in this study is the missing baseline measurement for each subject. Despite a baseline depended approach will not be applicable to road side testing, it is essential to estimate the vigilance testing capabilities of posturography on an individual scope. Furthermore it is necessary to obtain more samples during the course of experiments in order to estimate the circadian influence on EO/EC discrimination performance. The decision not choosing Romberg Pose could have lead to less significant stabilograms by veering away from the model of an inverse pendulum. Especially model-based features may suffer from that decision. As already mentioned, aversions can emerge from keeping this exhausting pose for too long. The reduction of trial lengths down to 18 seconds has been proposed by Forsman et al. [9]. The application of shorter trial lengths may support acceptance of such kind of vigilance tests, but the demand of a distinct pose may be impractical for field use and remains as a critical matter in dispute.

In order to establish a new vigilance test it is important to validate this test against other measures, like e.g. PVT. Such well accepted tests should be included in future analysis to consolidate results. But it must be emphasized that up to now there exists no gold measure. Many methods of assessing hypovigilance suffer from several drawbacks. One main problem is the large inter-subject variability [4]. Future investigations based on much more data of different experimental sessions, different subjects and utilizing cross validation strategies should reveal intra- and inter-subject variability of our measure. With the discriminant analysis between mild and strong hypovigilance resulting in acceptable error rates, we conclude that posturography can be utilized for vigilance testing.

References

1. Liu Y, Higuchi S and Motohashi Y (2001) Changes in postural sway during a period of sustained wakefulness in male adults. Occup. Med. Vol. 51 No. 8, pp. 490-495
2. Forsman P, Hæggström E, Wallin A, Toppila E, Pyykkö I (2007) Daytime changes in postural stability and repeatability of posturographic measurements. J. Occup. Environ. Med.;49(6):591-6.
3. Edwards D, Sirois B, Dawson T et al (2007) Evaluation of fatigue management technologies. Proc 4th Int Sympos Driving Assessment, Stevenson WA, USA, pp 146-152
4. Ingre M, Åkerstedt T, Peters B, Anund A and Kecklund G (2005) Subjective sleepiness, simulated driving performance and blink duration: examining individual differences. J. Sleep Res. (2006) 15, pp. 47-53
5. Collins JJ, De Luca CJ, Burrows A and Lipsitz LA (1995) Age-related changes in open-loop and closed-loop postural control mechanisms. Exp Brain Res. 1995;104(3):480-92
6. Jacono M, Casadio M, Moraso PG, Sanguineti V (2004) The Sway-Density Curve and the Underlying Postural Stabilization Process. Motor Control, 8(3):292-311
7. Schnupp T, Holzbrecher-Morys M, Golz M (2007) Inter- and Intra-Subject Variability in Posturography. Proc Ann Conf Biomed Techn, BMT, Aachen, Germany, 2007, CD-ROM
8. Joachims T (2002) Learning to Classify Text Using Support Vector Machines. The Springer International Series in Engineering and Computer Science, Vol. 668, ISBN: 978-0-7923-7679-8
9. Forsman P, Hæggström E and Wallin A (2007) Reducing trial length in force platform posturographic sleep deprivation measurements. Meas. Sci. Technol. 18 2893-2897, doi: 10.1088/0957-0233/18/9/019

Author:	Thomas Schnupp
Institute:	University of Applied Sciences Schmalkalden
City:	Schmalkalden
Country:	Germany
Email:	schnupp@fh-sm.de