Alertness Assessment using Data Fusion and Discrimination Ability of LVQ-Networks

Udo Trutschel[2], David Sommer[1], Acacia Aguirre2[2] Todd Dawson[2], Bill Sirois[2]

[1] University of Applied Sciences Schmalkalden, 98574 Schmalkalden, Germany
[2] Circadian Technologies, Inc., 2 Main Street, Suite 310 Stoneham, MA 02180
eMail: utrutschel@circadian.com

Abstract. To track the alertness changes of 14 subjects during a night driving simulation study traditional alertness measures such Visual Analog Sleepiness Scale, Alpha Attenuation Test (AAT), and number of Microsleep events per driving session were used. The aim of the paper is to assess these traditional alertness measures regarding their mutual correlations, revise one of them (AAT) and introduce new more general methods to capture changes in human alertness without too many constraints attached. The applied methods are utilizing data fusion methods and data discrimination capabilities via Learning Vector Quantification networks. The advantage of using more general data analysis methods which allows one to assess the validity of proposed alertness measures and opens possibilities to get a more comprehensive knowledge of obtained results.

1 Introduction

Recent technical developments have produced a 24-hour, advanced society that continues to grow on a global scale. Consequently, the basic human circadian rhythm ("working during the day and sleeping at night") is under constant siege. Because of the long working hours that eat up people's sleeping time, a general deterioration of people's daytime alertness and an increase in driver drowsiness is seen. Especially, accidents caused by drowsy drivers have a high fatality rate and high costs. To prevent these accidents a reliable tool to accurately measure human alertness levels is needed.
The first attempts to quantify human alertness were subjective reports that consisted of documenting the individual's self-assessment. The main measures include the Stanford Sleepiness Scale (SSS), the Visual-Analog Scale (VAS), and the Epworth Sleepiness Scale (ESS). More objective measures of human alertness can be derived from electroencephalogram (EEG) and electrooculogram (EOG) data. For example, the Multiple Sleep Latency Test (MSLT) measures the time to fall asleep while lying in a quiet, dark bedroom on repeated opportunities at 2 hours intervals throughout the day using EEG for sleep onset determination. The Maintenance of Wakefulness Test (MWT) requires that subjects sit in chairs in a darkened room and remain awake for 40 minutes. After applying different mathematical and statistical techniques, EEG-frequency bands (delta,

theta, alpha, beta, etc.) were used to define a variety of parameters, such as slow-wave activity, alpha slow-wave index, the alpha quotient, and others to estimate alertness. More and more, such alertness parameters have been introduced, in general using the Power Spectral Density (PSD) and multiple combinations of EEG-bands. A good review of the subjective and objective alertness measure can be found in [5] and [6]. Because of the shortcomings of these methods, a new test was developed by Michimori et al. in [1], the Alpha Attenuation Test (AAT) using the occipital (O1) - auricular (A2) EEG-derivation. The AAT is defined as ratio of Eyes-Closed (EC) to Eyes Open (EO) PSD of the alpha band (8 Hz - 12 Hz).
However, there are several drawbacks to all these proposed alertness measures. First, they are based on the countless definitions involving the PSD of a variety of EEG bands. Second, the definition of the EEG-frequency bands introduces artificial boundaries for the data analysis. Third, the separate analysis of the EEG-channels and frequency bands often leads to inconsistent results. Therefore, more general methods with less predefined assumptions are needed for comprehensive human alertness estimation. In order to identify insufficient perceptual capabilities (e.g. prolonged eye closure) and no ability to process external information (e.g. microsleep) reduced alertness should be defined as a combination of brain (EEG) and eye function (EOG). There are modern concepts of data fusion to combine multiple EEG and EOG signals in a way that ensures optimal information gain. For example, a Feature Fusion (FF) approach in combination with Learning Vector Quantization (LVQ) networks was already successfully applied by Sommer et al. in [3] for improving the detection of Micro-Sleep Events (MSE). The FF approach and the ability of LVQ networks to classify and discriminate data with low error rates [4] was utilized for the definition of generalized alertness measures. The high sensitivity of LVQ networks to small and unknown changes in the data is exactly what is required to detect variations in human alertness which are hidden in EEG and EOG.

2 Study Design and Data recorded

Fourteen young adults participated in nigh time driving study at the University of Schmalkalden. They arrived at the driving simulator facility in the evening, after a day of normal activity and at least 16 hours of continuous wakefulness, which was checked by wrist actigraph. After being wired up for EEG recordings, they started driving on a driving simulator at 1:00 A.M They had to complete seven driving sessions lasting 40 min, each followed by a 10 min period during which they estimated their subjective alertness using a VAS and performed a 5-minute AAT with five alternating 30 seconds Eyes-Open (EO) and Eyes-Closed (EC) episodes. Before the next driving session a 10-min break was scheduled. Experiments ended at 8:00 A.M. During the entire study seven EEG channels from different scalp positions (A1, A2, C3, C4, Cz, O1, O2) and two EOG-signals (vertical, horizontal) were recorded. The driving tasks were monotonous by design to induce drowsiness and Micro-Sleep Events (MSE). MSE are typically

characterized by driving errors, prolonged eye lid closures or nodding-off. Two experts performed an initial manual MSE scoring. Three video cameras were utilized to record (1) drivers face, (2) right eye region and (3) driving scenario. The number of MSE varied amongst subjects and increased during the night for all subjects indicating a clear deterioration in alertness.
The preprocessing of the all EEG and EOG data involved linear trend removal and applying the Hanning window to the data segments. Power Spectral Density (PSD) estimation was performed by the discrete Fourier transform. The so calculated PSD coefficients were averaged within 1.0 Hz wide bands. Further improvements in classification were achieved by applying a monotonic continuous transformation $log(x)$ to the PSD.

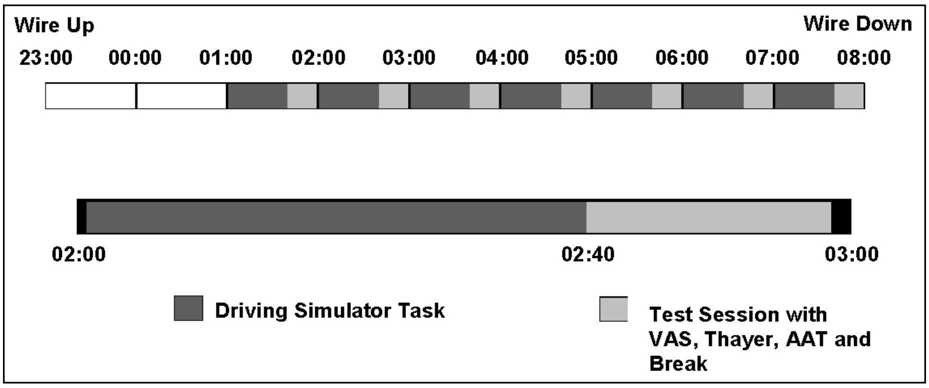

Fig. 1. Design of the driving simulator study.

3 VAS, MSE and AAT - Correlation Results

First, we present results obtained from VAS, MSE and AAT per driving session. Because of the different nature of the parameters, only the relative changes are meaningful as alertness measures. VAS score and number of MSE increase with reduced alertness, whereas the value of the so-called Alpha Attenuation Coefficient (AAC) decreases. The measures used are not only different in their correlation to alertness; they are different in the time period they cover. VAS reflects a punctual subjective alertness estimation. The AAT provides an objective alertness measure for a 5 minute period and the number of scored MSE per driving session gives an alertness score for a 40 minute period. A reliable EEG based alertness measure on a time scale of five minutes would be extremely useful for many applications. Therefore, we will focus on the Alpha Attenuation Coefficient (AAC), which was introduced by Michimori in [1] and is defined as ratio between the PSD of the alpha band (8 Hz - 12 Hz) for EC and EO episodes, respectively.

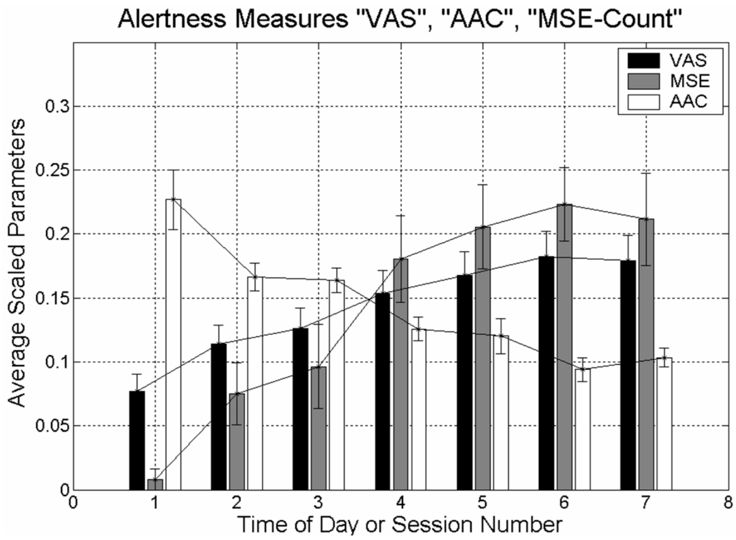

Fig. 2. Averaged, scaled results with Standard Error for VAS, MSE and AAC. Correlations: 'VAS-MSE'= 0.99; 'VAS-AAC'= -0.99; 'MSE-AAC'= -0.98.

Despite the different nature of VAS, MSE and AAC as alertness measure, the average scaled results show a remarkably similar alertness trend over the course of the night hours. Average results for all 14 subjects are shown in Fig. 2. This alertness trend is well-known from other studies and was thus expected [2]. The high VAS-MSE correlation will serve further as benchmark to evaluate the other alertness measures.

However, the good overall VAS-AAC and MSE-AAC correlation results hide the large individual variability and thus can not accurately estimate or predict individual changes in alertness. Therefore, individualized correlation results are shown in Fig. 3.

Considering the individual correlation results it appears that for subject '5' the AAC fails to reflect the alertness course established by VAS and MSE. This could because this individual did not produce alpha waves and/or by the restrictive definition of the AAC. Nevertheless, it would be extremely beneficial to have an objective, general valid test to measure the alertness for a broad population and not only for alpha wave producing individuals.

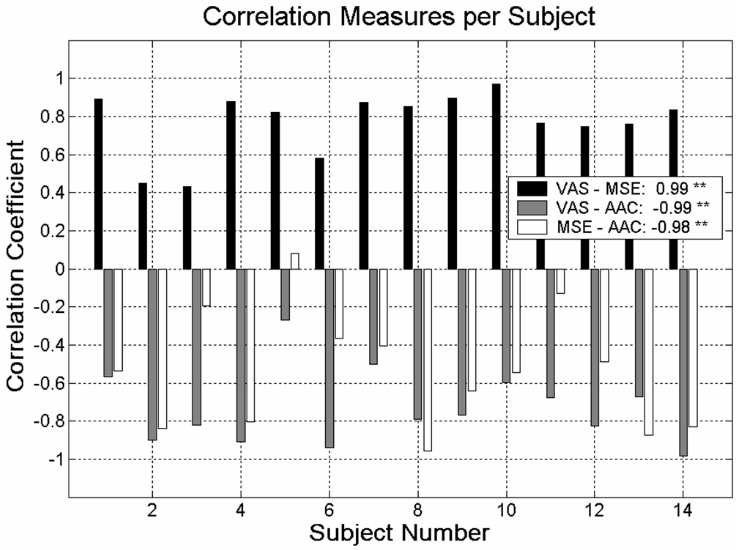

Fig. 3. Individualized correlation results between VAS, MSE and AAC.
Legend shows the overall correlations of averaged results for 14 subjects, $** = p < 0.01$.

4 "Alpha Attenuation Test" (AAT) - Revised

The AAT was developed based on two assumptions. First, the PSD of alpha waves changes in correspondence to the alertness level of the individual and second, the rate of change occurs in opposite direction for the AAT sections 'EC' and 'EO'.
To test the fundamental assumptions behind the AAT, the following generalized hypothesis should be investigated. Alertness would be high when the EEG-PSD differs substantially between EC and EO episodes, resulting in good LVQ discrimination and a low LVQ test error. Alertness would be reduced when the EEG-PSD during EC and EO are similar, resulting in low LVQ discrimination and a high LVQ test error. Therefore, the test error of the LVQ network should be directly correlated to the relative change in alertness. The EEG channel 'O1-A2' is used for the LVQ analysis allowing a direct comparison with the AAC. The overall LVQ correlation results (Table 1) are very close to the correlation for the AAC. Unfortunately, on the subject level there is disagreement. In addition to subject '5' now subject '7' shows no correlation any more to VAS and MSE (Fig. 4), if only the PSD of the alpha band is used (LVQ-A). The situation improves slightly if the band concept is abandoned and the PSD of the integer frequencies are used as features (LVQ). Still, there are troublesome indicators that the selected method is not able to correctly capture the alertness changes during the night. Adding additional EEG channels to the data set and using the FF results in further deterioration of the correlation to the VAS and MSE measures. For example, involving all EEG and EOG channels in the LVQ analysis reduces

the correlations for VAS-LVQ and for MSE-LVQ to 0.65 and 0.62, respectively. This should not happen to a generalized method. We concluded that using the differences and/or ratios of EEG-PSD between EO and EC episodes is not the most efficient way for detecting alertness changes.

Table 1. Overall Correlations of Alertness Measures. $** = p < 0.01$.

Correlations	VAS	MSE	AAC	LVQ (Alpha)	LVQ
VAS	1	0.99**	-0.99**	0.97**	0.96**
MSE	0.99**	1	-0.97**	0.95**	0.96**
AAC	-0.99**	-0.97**	1	-0.95**	0.96**
LVQ (Alpha)	0.97**	0.95**	-0.95**	1	--
LVQ	0.96**	0.96**	-0.96**	--	1

5 "Alpha Attenuation Test" (AAT) - Modified

From the results showed in the previous section it became clear that any ratios between EC and EO are not useful to extract information about alertness from the EEG data. As a modified approach, we propose a further simplification of the relative alertness estimation. We assume that EEG and EOG data during the first AAT probably reflect the highest alertness. Thus, the EEG and EOG of all other AAT sessions during the night will be compared by means of LVQ networks to the first AAT-session which is used as reference.
If the EEG and EOG data sets during a given AAT session are extremely different from the EEG and EOG data sets during the first AAT, then the LVQ network discriminates well, resulting in a low test error. A substantial change in alertness between different AAT sessions has then occurred. On the other hand, if the EEG and EOG data sets during a given AAT are similar to the EEG and EOG data sets during the first AAT, then the LVQ network can not discriminate well, resulting in a high test error. No significant change in alertness between different AAT sessions has then occurred. With this simple pairwise comparison of EEG and EOG for different AAT-times, we were able to obtain changes in alertness over the night which is correlating to a high degree with VAS and MSE (Fig. 5) on a subject by subject basis. The fact that a remarkable improvement in the correlation results was achieved by fusing the data from all EEG and EOG signals is encouraging as well.
Whereas the correlation results using only eye data (VAS-LVQ-EOG1, MSE-

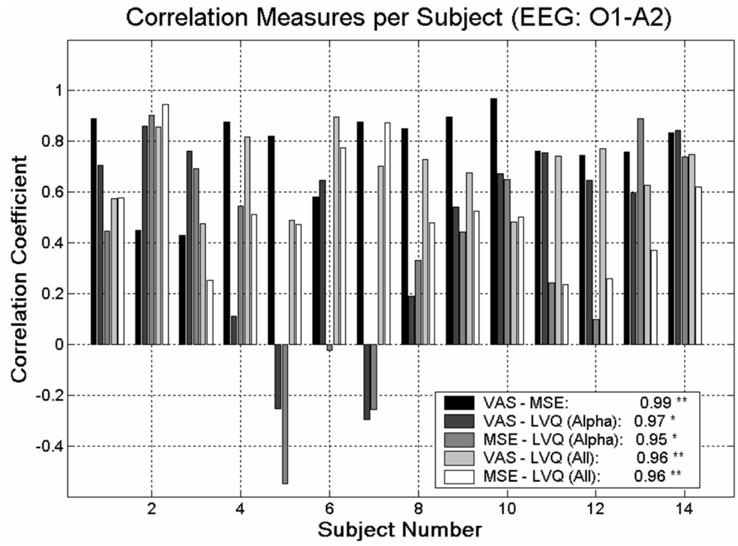

Fig. 4. Individualized correlation results between VAS, MSE and LVQ (EO v EC). Legend shows the overall correlations of averaged results for 14 subjects, $** = p < 0.01$.

LVQ-EOG1) for most subjects are good (except for subject '9', '13'), a substantial improvement can be achieved using all EEG and EOG data (VAS-LVQ-All, MSE-LVQ-All). The new proposed alertness measure based on LVQ pairwise discrimination AAT-session 1 (S1) versus AAT-session 2 (S2). to AAT-session 7 (S7), correlates well with VAS and MSE for all subjects.
Nevertheless, this method of tracking alertness changes based on data discrimination using LVQ, presented here for the first time, is not perfect. Our assumption that the point of the highest alertness is at the beginning of the night may be not valid for every subject. The potential of the approach would be further improved if a second marker for the lowest alertness could be added.

6 Conclusions

The newly introduced methods of estimating alertness are utilizing the advantages of Feature Fusion (FF) and the capabilities of LVQ neural networks to classify data with low error rates and good discrimination sensitivity. High sensitivity of LVQ network to small and unknown changes in the PSD of the EEG and EOG data was applied to trace subtle alertness changes. These proposed methods do not make any assumptions that certain EEG signals and predefined frequency bands has to be used in order to get a reliable alertness measure.
Our first approach relies only on the general assumption that the characteristics of the EEG-PSD between EO and EC episodes changes fundamentally when the alertness of an individual is changing over the course of time. Unfortunately, there are strong indications that the approach suffers from the same weaknesses

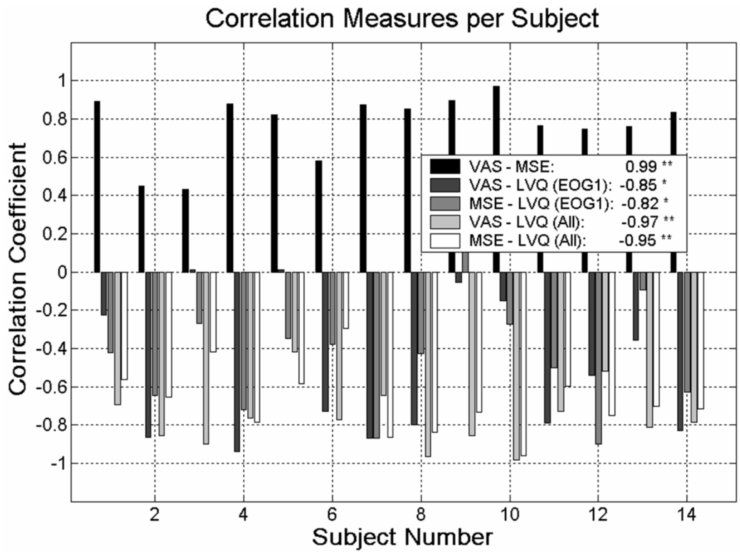

Fig. 5. Individualized correlation results between VAS, MSE and LVQ (S1 v S2 ... S7). Legend shows the overall correlations of averaged results for 14 subjects, $* = p < 0.05$.

than the AAT. The utilization of ratio between EO and EC episodes for EEG-PSD bands diminishes the sensitivity to alertness related changes in the signals. Our second approach relies on pairwise comparison of the PSD from well defined EEG and EOG data at different moments in time. This approach shows great potential and should be evaluated further using a second reference point of low alertness.

References

1. Michimori A, Stone P, Aguirre A, Stampi C., Analysis of the alpha attenuation test. Sleep Research, 23, (1994) 454
2. Trutschel, U., Guttkuhn, R., Ramsthaler, C., Golz, M., Moore-Ede, M.: Automatic Detection of Microsleep Events Using a Neuro-Fuzzy Hybrid System. Proc 6th European Congress of Intelligent Techniques Soft Computing (EUFIT98), Vol 3, Aachen (ISBN 3-89653-5005-5) (1998) 1762–66
3. Sommer, D., Chen, M., Golz, M., Trutschel, U., Mandic, D.: Fusion of State Space and Frequency-Domain Features for Improved Microsleep Detection. in W. Duch et al. (Eds.): International Conf Artificial Neural Networks (ICANN 2005); LNCS 3697, Springer, Berlin (2005) 753–759
4. Kohonen, T.: Self-Organizing Maps. Springer, Berlin, 3rd ed., (2001)
5. Sforza, E., Grandin, S., Jouny, C., Rochat, T., Ibanez, V.: Is waking electroencephalographic activity a predictor of daytimesleepiness in sleep-related breathing disorders? Eur Respir J (2002) 645–652
6. Leproult, R., Coleccia, E.F., Berardi, A.M., Stickgold, R., Kosslyn, S.M., Van Cauter, E.: Individual differences in subjective and objective alertness during sleep deprivation are stable and unrelated. Am J Physiol 284, (2002) R280–R290