Application of Vector-Based Neural Networks for the Recognition of Beginning Microsleep Episodes with an Eyetracking System

Martin Golz[1], David Sommer[1], Arne Seyfarth[1], Udo Trutschel[2], Martin Moore-Ede[2]

[1]University of Applied Sciences, Department of Computer Science, Postfach 182
Schmalkalden, D-98574, Germany
{golz, sommer}@informatik.fh-schmalkalden.de
[2]Circadian Technologies Inc., 24 Hartwell Avenue
Lexington, MA 02421, USA

Abstract

The eye gaze point and the pupil size of five subjects were recorded during an overnight driving simulation task. By scoring the recorded videos, clear microsleep events (MSE) and clear non-microsleep events were detected, and the measured signals in the preceding five seconds were analyzed. The spectral densities of these segments were classified using learning vector quantization, self-organizing feature map and growing cell structures.

Keywords: Eyetracking, microsleep episode, learning vector quantization, self-organizing feature map, growing cell structures

1 Introduction

The assessment of an individual's sleepiness and the recognition of beginning microsleep is important for activities requiring sustained attention, such as driving a vehicle or working at night as a machine operator, where falling asleep can be hazardous. A microsleep episode can be defined as a somewhat unexpected short sleep lasting between 2 and 30 seconds that occurs in the midst of ongoing wakeful activity [Thorpy & Yager, 1991].

The most widely used objective measure of sleepiness is the Multiple Sleep Latency Test [Cascardon & Dement, 1982], but, as in many other test procedures, there are the practical disadvantages of requiring sophisticated laboratory equipment and trained operators. Furthermore, this test does not supply a continuous measure of the subject's perform-ance and is not applicable without interrupting the night work, which leads to an unintentional break in monotony.

Many authors suggest the measurement of pupil size and eye movements to estimate the subject's alertness level. A lot of work has been done with the pupillary light reflex (PLR), which is the pupil diameter response of fully dark-adapted eyes to a light flash [Lowenstein et al, 1951] [Lichstein et al 1992] [McLaren et al, 1992] [Lichstein & Johnson, 1994] [Grünberger et al, 1994] [Keegan et al, 1995]. In a recent paper [Ranzijn & Lack, 1997] it has been stated that PLR cannot be used to measure sleepiness.

The measurement of eye movements has been investigated by [McPartland & Kupfer, 1978], [Ogilvie et al, 1988], [Hyoki et al, 1998], [Schmidt et al, 1979] and [Saito, 1992]. The first three used electro-oculography (EOG); the latter two used infrared corneal reflection as the measurement principle.

The present study employed an eyetracking system based on the analysis of combined corneal and foveal reflection [Cleveland, 1992]. Our intention was not to estimate the sleepiness level over time, but to explore the characteristics of the eyetracking signals immediately before the onset of a microsleep episode.

2 Experiments

Five individuals aged between 19 and 28 years were recruited by public announcements and paid between $80 and $120 to participate in an overnight driving simulator study. Their task was intentionally monotonous, simply to avoid major lane deviations. One driving session of 25 min length was carried out every hour from 1 a.m. until 7 a.m. The face of the driver and the region of the right eye were video-

recorded separately; additionally, EEG, EOG and ECG were recorded, but these data are not considered in this paper. The right-eye video was captured with the sophisticated eye-tracking system mentioned above. The eyetracker (Figure 1) operates with light in the near-infrared spectral regione with an accuracy of 0.65 deg and measures the pupil diameter (D) and the horizontal (X) and vertical (Y) component of the eye gaze point in the plane of the driving simulator screen. The sampling rate is 30 Hz.
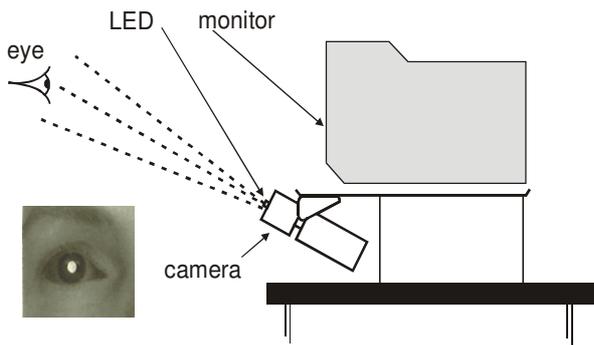

Figure 1: The eyetracking system using the pupil-center / corneal-reflection method; the infrared light source is placed on the center of the camera lens

Figure 2: Driving simulator

3 Analysis

Microsleep events (MSE) were visually scored by an expert using the video recordings and the EEG recordings simultaneously. Clear non-microsleep events (NMSE) were scored in the same manner.

The X- and Y-signals had a series of missing values during eye blinks. They were replaced by Beziér spline interpolation. Additionally, outlier elimination was necessary, especially for the Y-signal immediately after an eye blink.

603 segments of 8 sec length were taken out of all three signals (X, Y, D) immediately before a MSE and before a NMSE (Figure 3). Afterwards any linear trend was eliminated and a Welch window was applied to improve the results of the following discrete Fourier transform. The reduction of the total power density due to the windowing was corrected using the Parceval theorem.
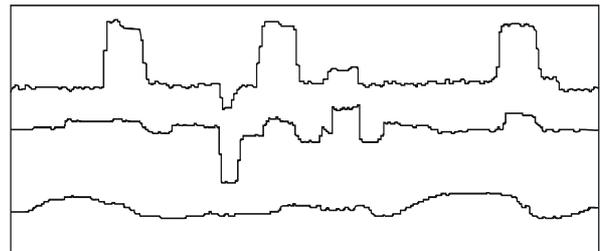

Figure 3: An 8 sec segment of the pupil diameter D (lower graph) and of the eyetracker signals X and Y (upper graph).

For each signal we got 80 spectral power densities in the range from 0 to 9.8 Hz with a resolution of 0.125 Hz. The input vector for the neural networks consisted of the spectral power densities. To build up a classifier for input vectors of the MSE- and of the NMSE-class, we applied three types of vector-based neural networks: the learning vector quantization network (LVQ) [Kohonen, 1988], the self-organizing maps (SOM) [Kohonen, 1982] and the growing cell structures (GCS) [Fritzke, 1994].

LVQ is a network with supervised training. The binary MSE / NMSE information was used as a teaching input. Kohonen suggested three modifications: LVQ1, LVQ2 (LVQ2.1) and LVQ3. The first modification uses an adapted step size, whereas LVQ2 leads to an adaptation of the neurons in interclass regions. LVQ3 additionally allows a slight adaptation of weight vectors in intraclass regions [Kohonen, 2000].

The SOM and the GCS networks were trained without supervision. After the training, both network types were calibrated with the binary MSE / NMSE information. SOM tries to minimize the error of vector quantization and to some extent to find a discrete approximation of the probability density function of the input vectors. GCS are incremental neural networks and are with some restrictions, able to approximate the probability density function of the input vectors. The topological structure is a k-simplex. We chose k=1 and k=2 to be able to visualize. For the SOM networks we used one-dimensional topology and two-dimensional tetragonal topology.

4 Results

Each network was trained with several parameter settings and with several initializations of the weight vectors. Before each training pass, the learning set was randomly partitioned in training set (80%) and in test set (20%). After training was finished, the re-classification rate was estimated by the ratio of correct classifications to all applied input vectors of the training set. The classification rate was estimated in the same way with input vectors taken from the test set.

We calculated the classification and reclassification rates in $1.7 \ 10^6$ different network simulations with different parameter settings, e.g. number of neurons, learning rate factor, and parameters of the neighborhood function and different variables selections for the input vectors and different learning set partitions. Figure 4 shows an example.

Network	init	scaling	D	Y	X	DX	DXY
LVQ1	MED	---	77	69	70	71	72
LVQ1	DAT	---	77	68	71	70	71
LVQ1	DAT	SQR	76		71	72	
LVQ1	DAT	NRM	75		72	75	
LVQ2	MED	---	77		75	74	
LVQ2	DAT	---	79		75	74	
LVQ3	MED	---	80		75	74	
LVQ3	DAT	---	80		75	75	
LVQ3	DAT	SQR	77		73	74	
LVQ3	DAT	NRM	73		75	79	

Table 1: Average maximum test-set classification rate (in percent) with different LVQ networks, different initializations and different scaling applied to different feature sets. (For details see text)

The optimal number of neurons ranged between 8 and 20. With an increasing number of neurons, the LVQ network showed a better adaptation to the training set. The reclassification rate was mostly above 90%, but it shows a decreasing ability to generalize, as indicated by decreasing classification rates.

The average maximum classification rate was obtained by searching the maximum of the mean + standard deviation (upper curve in Figure 4) for all different settings of the LVQ networks (Table 1). Initialization with median (MED) means that we assigned to each component of the weight vectors the median value of this component over all input vectors. During data driven initialization (DAT) each weight vector was assigned to a randomly selected input vector. Furthermore, in the first 30% of all training iterations the network was trained disregarding the

class membership of an input vector to diminish the variance of the classification rate described elsewhere [Golz et al, 1998].
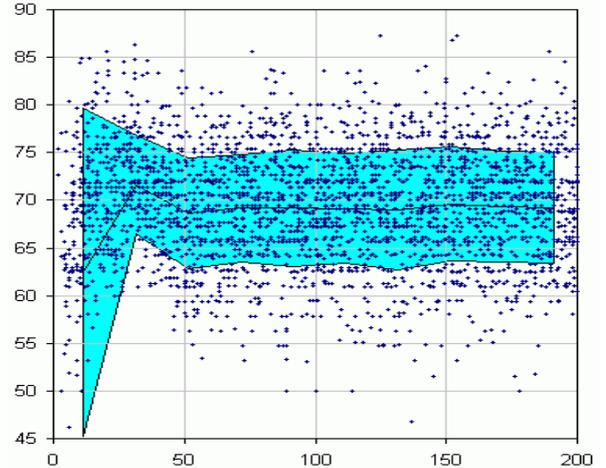

Figure 4: Test-set classification rate (in percent) versus number of neurons for an LVQ3 network. The input vectors contain spectral power densities of the pupil diameter D only. The lines indicate the mean ± standard deviation range.

We tried a number of different scaling techniques, but we want to report only the results of no scaling(--), the square root (SQR) of each input vector component and the normalization (NRM) of each component with respect to the sum of all components (relative value). In columns "D", "Y" and "X" (Table 1), the input vectors consist only of the spectral power densities of the D, Y and X signals. In column "DX" all spectral values of the D and X signal, and in "DXY" all spectral values of the D, X and Y signal are used. The best results are obtained with the set of input vectors obtained from the D signal only. Apparently, if we add further components to the input vectors as in the columns "XD" and "XYD", the results are not improvable. On the one hand we presented supplementary and independent information to the neural networks, but on the other hand there were apparently too many dimensions of the input space.

A typically SOM calibrated with the MSE and NMSE information is shown in Figure 5. Large distances of neighbored weight vectors are visualized as gray shades using the U-matrix [Ultsch & Siemon, 1989]. The NMSE input vectors are mapped to the left lower part of the map, whereas the MSE input vectors are mapped to the right upper region. The weight vectors in the NMSE region have larger distances visualized by darker shades. Assuming that the SOM has found a correct approximation of the probability

density function of the input vectors, this indicates that the MSE class has a higher density and is more compact.
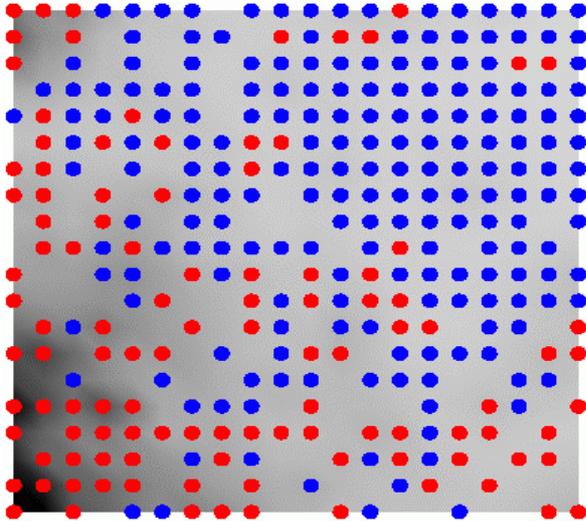

Figure 5a: Typical calibrated SOM. Gray shades indicate the U-matrix. Microsleep events (dark nodes) and nonmicrosleep events (light nodes) are separable with some limitations. Vacancies indicate dead neurons.
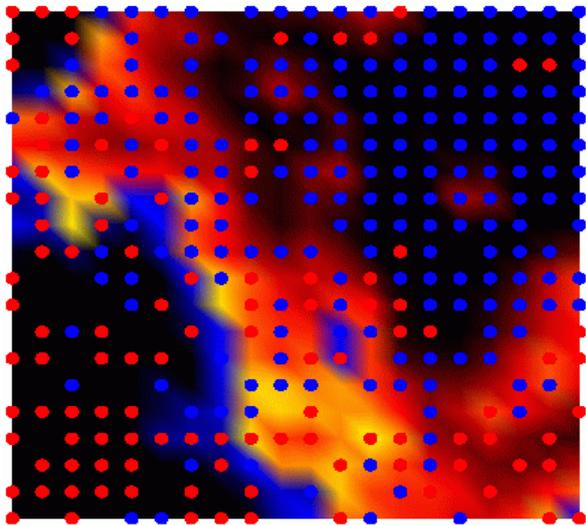

Figure 5b: The same SOM with a differentiated U-matrix. Microsleep events (dark nodes) and non-microsleep events (light nodes) are separable with some limitations. Vacancies indicate dead neurons.

The differentiated U-matrix (Figure 5b) roughly shows the region of overlapping classes with light shades. The two classes are distributed in only two more or less compact and overlapping regions in the input space. This could explain the decreasing generalization ability with increasing number of neurons and the onset of this effect even with relatively few neurons.

The GCS networks were trained and tested with the same method as the SOM. Additionally, fast learning is characterized by inserting and deleting neurons depending on a local criterion. Two criteria have been proposed [Fritzke, 1994]: the mean vector quantization error (VQE) and the local probability density function (PDF). For the calculation of the PDF, the volume of the n-dimensional Voronoi cell was approximated with the volume of the n-dimensional hypercube, generated with the mean local weight vector distance [Fritzke, 1994].

Both networks, the SOM and the GCS, manifested lower average maximum classification rates (Table 2). This is not surprising because training of these networks was unsupervised.

network	no. of neur.	dim.	criter.	D	X	DX
SOM	20 x 1	1		76	72	70
SOM	20 x 10	2		74	68	68
SOM	20 x 20	2		72	66	67
GCS	300	1	PDF	74	74	70
GCS	300	1	VQE	75	69	69
GCS	300	2	PDF	74	69	69
GCS	300	2	VQE	74	68	68

Table 2: Average maximum test-set classification rate (in percent) with SOM and GCS networks, different number of neurons applied to different feature sets. (For details see text)

With SOM and with GCS, the best results were obtained by processing D data only and mapping on one-dimensional topology. In this case the choice of VQE or PDF as fast-learning criterion function is not significant. With PDF and one-dimensional topology, the results were about equal for signal D and for X.

The visualization of the topology yielded no results. Between one and three separate topological nets grew during training. No net contained a large majority of input vectors of the MSE class.

5 Conclusions

The best results were obtained with learning vector quantization, especially LVQ3. The unsupervised SOM and GCS methods were understandably poorer. The spectral power of the pupil diameter signal contained information, which enabled us to get classification rates up to 80%. In addition, the

inclusion of spectral components of the gaze point signal does not improve the classification.

By visualizing the differentiated U-matrix of the SOM, we conclude that the MSE- and the NMSE-input vectors constitute two clusters with different probability densities.

References

McPartland, RJ; Kupfer, D (1978). Computerized measures of electro-oculographic activation during sleep. *Int J Biomed Comput*, 9, 409-419.

Ogilvie, RD; McDonagh, DM; Stone, SN; Wilkinson, RT (1988). Eye movements and the detection of sleep onset. *Psychophysiology*, 25, 81-91.

Cascardon, MA; Dement, WC (1982). The multiple sleep latency test: What does it measure? *Sleep*, 5, 67-72.

Cleveland, D; Cleveland, N (1992). Eyegaze Eyetracking System. *Imagina - 11th Monte - Carlo International Forum on New Images*; Monte - Carlo, Monaco.

Fritzke, B (1994). Growing Cell Structures - A self-organizing network for unsupervised and supervised learning. *Neural Networks*, 7, 1441-1460.

Golz, M; Sommer, D; Lembcke, T; Kurella, B (1998). Classification of the pre-stimulus-EEG of k-complexes using competitive neural networks. Proc. of the 6th European Congress on Intelligent Techniques and Soft Computing. *EUFIT'98*. Verlag Mainz, Aachen, Germany, Vol. 3, 1767-71

Grünberger,J; Linzmayer, L; Grünberger, M; Saletu, B (1994). Eine neue Methode zur Messung der zentralen Aktivierung: Fourieranalyse der Pupillen-oszillationen bei depressiven Patienten. *Wien. Klin. Wochenschr.*, 106/6, 164-170 (in German).

Hyoki, K; Shigeta, M.;Tsuno, M; Kawamuro, Y; Kinoshita, T (1998). Quantitative electro-oculography and electroencephalography as indices of alertness. *Electroenc. Clin. Neurophysiol.*, 106, 3, 213-219.

Keegan, AP; Oroujeh, AM; Meritt, SL; O'Neill, WD; Mercer, PW (1995). Two methods of measuring sleepiness by quantifying noise in pupillometry. *Sleep* Research, 24.

Kohonen, T (1988). Learning Vector Quantization. *Neural Networks*, 1 (Suppl.1), 303

Kohonen, T (2000). *Self-Organizing Maps. 3rd ed.* Springer Series in Information Sciences, Vol. 30, Springer, Berlin, Heidelberg, New York.

Löwenstein, O; Loewenfeld, IE (1951).Types of autonomic innervation and fatigue: pupillographic studies. *Archives of Neurology and Psychiatry*, 66, 580-599.

Lichstein, KL; Johnson, RS; sen Gupta, S; O'Laughlin, DL; Dykstra, TA (1992). Are insomniacs sleepy during the day?: A pupillometric assessment. *Behaviour Research and Theory*, 30, 283-292.

Lichstein, KL; Johnson, RS (1994). Pupillometric discrimination of insomniacs. *Behav. Res. Ther.*,32,1,123-129.

McLaren, JW; Erie, JC; Brubaker, RF (1992). Computerized analysis of pupillograms in studies of alertness. *Investigative Ophthalmology & Visual Science*, 33, 3, 671-676.

Newman, J; Broughton, R (1991). Pupillometric assessment of excessive daytime sleepiness in narcolepsy-cataplexy. *Sleep*, 14, 2, 121-129.

Ranzijn, R; Lack, L (1997). The pupillary light reflex cannot be used to measure sleepiness. *Psycho-physiology*, 34, 1, 17-22.

Saito, S(1992). Does fatigue exist in a quantitative measurement of eye movements? *Ergonomics*, 35, 5 / 6, 607-615.

Schmidt, D; Abel, FA; Dell'Osso, LF; Daroff, RB (1979). Saccadic velocity characteristics: Intrinsic variability and fatigue. Aviation, *Space and Environmental Medicine*, 50, 1979, 393-395

Thorpy, MJ; Yager, J (1991). *The Encyclopedia of Sleep and Sleep Disorders*. NewYork: Facts on File.

Ultsch, A; Siemon, HP (1989). Exploratory data analysis: Using Kohonen networks on transputers. Univ. of Dortmund, *Technical Report 329*, Dortmund, Dec. 1989.