Classification of Eyetracking Signals with Vector-Based Neural Networks

Martin Golz, David Sommer
University of Applied Sciences Schmalkalden
Department of Computer Science
D-98574 Schmalkalden, Germany
{golz, sommer}@informatik.fh-schmalkalden.de

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 picked out and the measured signals in the preceding eight seconds were analyzed. The spectral power densities of these segments were classified using Learning Vector Quantization, Self-Organizing Map and Growing Cell Structures. For the latter two networks the supervised and the unsupervised version were applied. Best results were obtained with a modified LVQ3 network.

KEYWORDS: Learning Vector Quantization, Self-Organizing Map, Growing Cell Structures, Eyetracking

INTRODUCTION

Many authors suggested the measurement of the pupil size and the eye movements to estimate a subject‘s alertness level [5, 9, 10, 11, 12]. The first three groups used electrooculography (EOG); the later two groups used infrared corneal reflection as measurement principle.
The present study employed an eyetracking systems based on the analysis of the combined corneal and foveal reflection [1]. Our intention was not to estimate the alertness level at a time scale of some minutes, but to explore characteristics of the eyetracking signals immediately before the onset of a microsleep.
Five individuals with an age between 19 and 28 years participated in a monotonically overnight driving simulator study. Every hour from 1 a.m. until 7 a.m. one driving session of 25 min length was carried out. The portrait of the driver and the right eye was video recorded. The eyetracker was working in the near infrared with an accuracy of 0.65 deg and measured the pupil diameter (D), and the horizontal (X) and the vertical (Y) component of the eye gaze point in the plane of the driving simulator screen with a sampling rate of 30 Hz.
Microsleep events (MSE) were visually scored off-line by an expert using the video recordings and simultaneously EEG recordings. 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 substituted with Beziér spline interpolation. Additionally outlier elimination was necessary, especially for the Y-signal immediately after an eye blink.

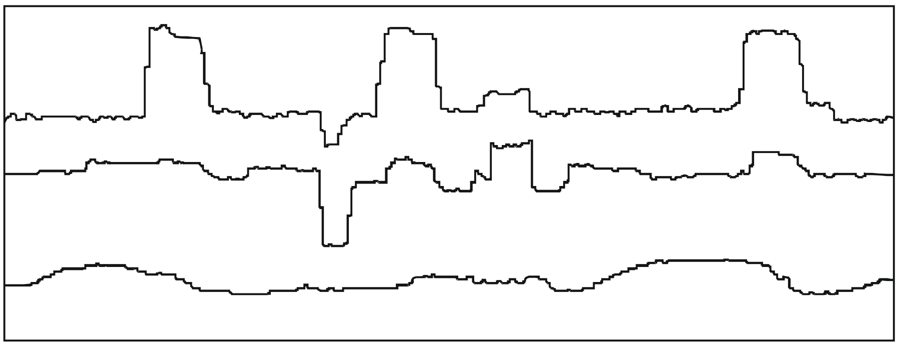

Figure 1: An 8 sec segment of the pupil diameter D (lower graph) and of the eytracker signals X and Y (upper graph).

603 segments of the length of 8 sec were taken out of all three signals (X,Y,D) immediately before a MSE and before a NMSE (Fig. 1). 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.
For each signal we got 80 spectral power density samples in the range from 0 to 9.9 Hz with a resolution of 0.125 Hz. All spectral samples were used as an input vector for the neural networks.
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) [7], the Self-Organizing Maps (SOM) [6] and the Growing Cell Structures (GCS) [3].
The LVQ networks are trained supervised; here the binary information MSE / NMSE was used as 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 neurons in interclass regions. LVQ3 additionally allows a slight adaptation of weight vectors in intraclass regions [8].
The SOM and the GCS networks are trained unsupervised. After training, both network types were calibrated with the binary MSE / NMSE information. SOM tries to minimize the error of vector quantization and to some extend to find a discrete approximation of the probability density function of the input vectors. GCS are incremental neural networks and are with some restrictions capable 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. Two-dimensional rectangular and one-dimensional topologies were applied to the SOM networks.

RESULTS

Each network was trained with several parameter settings and with several initializations of the weight vectors. Before each training the learning set was randomly partitioned in training set (80%) and in test set (20%). After training had finished, the reclassification rate was estimated by the ratio of correct classified 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, like 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. Fig. 2 shows an example.

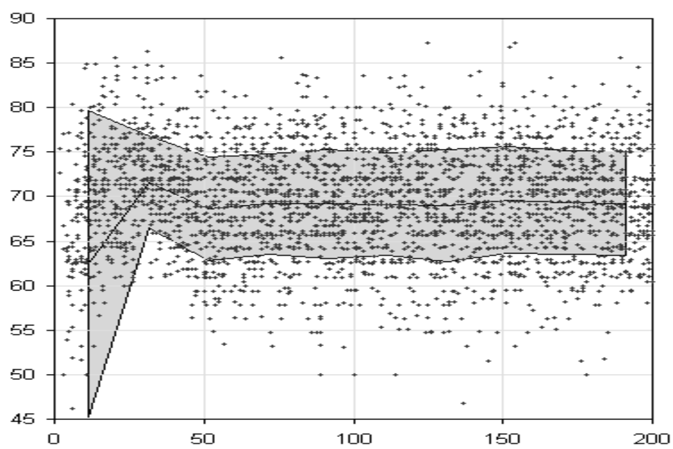

Figure 2: Test-set classification rate (in percent) versus number of neurons for an LVQ3 network. The input vectors contain spectral power densities of the pupil diameter D only. The lines indicate the mean ± standard deviation range.

The optimal number of neurons ranged between 8 and 20. With an increasing number of neurons the LVQ network shows a better adaptation to the training set, the reclassification rate is mostly above 90%; but it shows a decreasing ability to generalize indicated by decreasing classification rates.
The average maximum classification rate was obtained by searching the maximum of the mean + standard deviation (upper curve in Fig. 2) for all different settings of the LVQ networks (Tab. I). Initialization with median 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 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 [4].

Network	Initialization	Scaling	D	Y	X	DX	DXY
LVQ1	median	none	77	69	70	71	72
LVQ1	data driven	none	77	68	71	70	71
LVQ1	data driven	square root	76		71	72	
LVQ1	data driven	normalized	75		72	75	
LVQ2	median	none	77		75	74	
LVQ2	data driven	none	79		75	74	
LVQ3	median	none	80		75	74	
LVQ3	data driven	none	80		75	75	
LVQ3	data driven	square root	77		73	74	
LVQ3	data driven	Normalized	73		75	79	

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

We tried a number of different scaling, but we want to report only the results of no scaling, the square root of each input vector component and the normalization of each component with respect to the sum of all components (relative value). In the columns ‘D’, ‘Y’, ‘X’ the input vectors consisted only of the spectral power densities of the D, Y and X signal respectively. 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 were used. The best results were 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 the number of dimensions of the input space was obviously too much.

Typically SOMs calibrated with the MSE and NMSE information are shown in Fig. 3 and Fig. 4 using the U-matrix [13]. The U-matrix represents distances of topological neighbored weight vectors in the input space and is visualized as gray shades. Larger distances of neighbored weight vectors are visualized by darker gray shades. Weight vectors in the NMSE regions show larger distances than weight vectors in MSE regions. Under the assumption 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 probability density and is more compact. The Self-Organizing Map of Fig. 3 also shows that the NMSE input vectors are mapped preferably to the left part and to the lower part of the map, whereas the MSE input vectors are mapped to the right upper part of the map. Between both regions there is a region of overlapping classes.

The differentiated U-matrix (Fig. 4) 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 ability of generalization with increasing number of neurons and the onset of this effect at small numbers of neurons already.

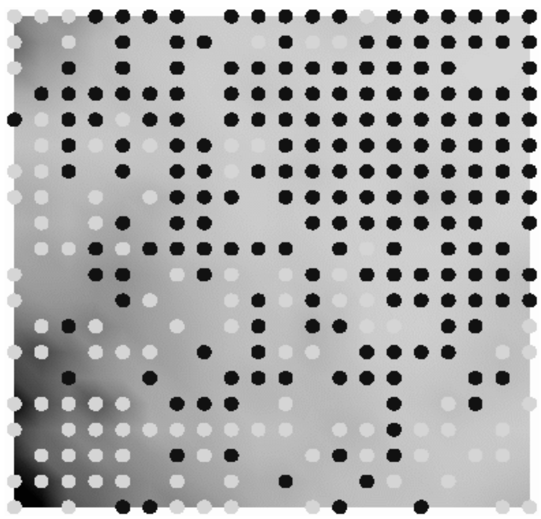

Figure 3: Typical calibrated SOM. Gray shades indicating the U-matrix. Microsleep events (dark nodes) and non-microsleep events (light nodes) are separable with some limitations. Vacancies indicating dead neurons.

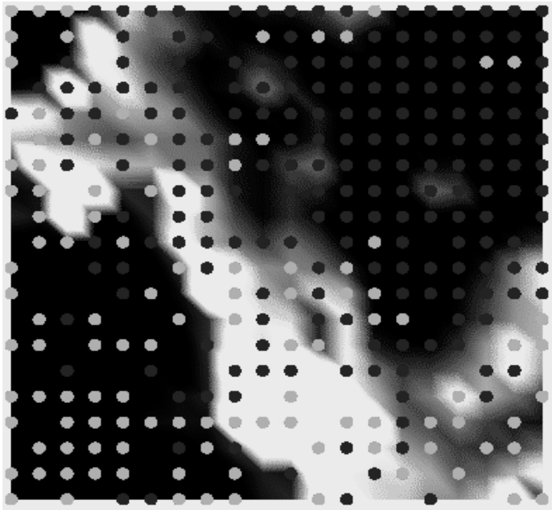

Figure 4: The same SOM with a differentiated U-matrix. Gray shades indicating the U-matrix. Microsleep events (dark nodes) and non-microsleep events (light nodes) are separable with some limitations. Vacancies indicating dead neurons.

The GCS networks were trained and tested with the same method as the SOM. Additionally, there is a fast learning by inserting and deleting neurons depending from a local criterion. Two criteria were proposed [2]: 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 [2].

Network	No. of neur.	dim.	Criter.	D	Y	X	DX	DXY
SOM	20 x 1	1		76		72	70	
SOM	20 x 10	2		74		68	68	
SOM	20 x 20	2		72		66	67	
sv SOM	20 x 1	1		70		67	65	
sv SOM	20 x 20	2		69		62	62	
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	
sv GCS	300	1	pdf	70		69	70	
sv GCS	300	1	vqe	68		66	65	

Table II:. Average maximum test-set classification rate (in percent) with SOM and GCS networks, different number of neurons applied on different feature sets. Supervised modifications are marked with 'sv'. (For details see text)

Both networks, the SOM and the GCS, came to lower average maximum classification rates compared to LVQ (Tab. II). This is not surprising because their training is unsupervised.
Following the suggestions in [8] a supervised training for SOM was applied. The input vectors are concatenated during training with a binary unit vector of encoded class number. Every class is assigned to one component of the unit vector. In the recall phase the input vectors are applied without the unit vector. In two- and three-dimensional examples one can see, that this modification leads to a higher density of prototype vectors in interclass regions. This idea was applied to the GCS network in the same manner.
Surprisingly, the supervised versions showed poorer classification rates, approximately 5% below the results of the unsupervised version. Compared to the unsupervised version the ability to generalize was poorer and the adaptation (training set classification rate) was improving.
With SOM and with GCS the best results were obtained processing D data only and mapping on one-dimensional topology. In this case it is not considerable to choose vqe or pdf as fast learning criterion function. If pdf and one-dimensional topology was chosen 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.

REFERENCES

[1] Cleveland, D; Cleveland, N (1992). Eyegaze Eyetracking System. Imagina – 11th Monte-Carlo International Forum on New Images; Monte-Carlo, USA.

[2] Fritzke, B (1992). Wachsende Zellstrukturen – ein slebstorganisierendes neuronales Netzwerkmodell. University of Erlangen, PhD thesis (german).

[3] Fritzke, B (1994). Growing Cell Structures - a self-organizing network for unsuper-vised and supervised learning. Neural Networks, 7, 1441-1460.

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

[5] 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.

[6] Kohonen, T (1982). Self-organized formation of topologically correct feature maps. Biol Cybern, 43, 59-69.

[7] Kohonen, T (1988). Learning Vector Quantization. Neural Networks, 1 (Suppl.1), 303

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

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

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

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

[12] 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

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