ESANN'1995 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 19-20-21 April 1995, D-Facto public., ISBN 2-9600049-3-0, pp. 247-252

Improvement of EEG classification with a subject specific feature selection

M. Pregenzer, G. Pfurtscheller and C. Andrew

Ludwig Boltzmann Institute of Medical Informatics and Neuroinformatics and the Department of Medical Informatics, Institute of Biomedical Engineering, Graz University of Technology, Brockmanngasse 41, A-8010 Graz, Austria

Abstract. This paper describes the application of the Distinction Sensitive Learning Vector Quantizer (DSLVQ) to select optimal features for an EEG-based Brain Computer Interface (BCI). In one experiment DSLVQ is used to find optimal electrode positions to differentiate between planning of left and right finger movement. In a second experiment, optimal frequency bands are selected to discriminate between 3 different brain patterns. Both experiments show that the optimal selection is strongly dependent on the specific subject.

1. Introduction

For the construction of an Electroencephalogram (EEG)-based Brain Computer Interface (BCI) to help people with severe motor impairment (Wolpaw et al. 1991, Pfurtscheller et al. 1993b) it is necessary to classify spatiotemporal EEG patterns related to specific "thoughts" on-line. These EEG patterns are recorded on the intact scalp during pure mental activity and classified with a learning classification method which is able to adapt to each specific subject. The classification results of the classifier can then be used to generate control signals for different tasks. The number of available features is considerable: in practice the number of electrodes for EEG recording can be high (up to 30 and more) and different frequency bands can be used for classification. However, the number of examples which can be recorded from one subject is limited. Since the number of examples which are necessary to train a learning classifier properly generally increases with the complexity and dimensionality of the classification problem, it is important to pre-select the most distinct features and feature combinations. It has been demonstrated by Flotzinger et al. (1994a) that selection of appropriate features can improve the classification accuracy of a BCI considerably. Pfurtscheller et al. (1994) show that the Distinction Sensitive Learning Vector Quantizer (DSLVQ, Pregenzer et al. 1994a, 1995) is an appropriate tool for single trial based data analyses. In this study DSLVQ is used for feature selection in two different experiments: in the first experiment, the most distinct electrode positions to distinguish between two different types of movement-planning are selected from a large number of possible positions. It is shown that different subjects require different electrode positions in order to get good classification rates with a learning classifier. The second experiment uses DSLVQ to analyze the importance of 1-Hz bands of EEG power spectra for the prediction of three different types of movement. This experiment shows for the first time that for

[1] Supported by the "Fonds zur Förderung der wissenschaftlichen Forschung", projects P9043 and P10000, in Austria

ESANN'1995 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 19-20-21 April 1995, D-Facto public., ISBN 2-9600049-3-0, pp. 247-252

the prediction of the type of movement from EEG recordings, the optimal spectral bands are also strongly dependent on the subject.

2. DSLVQ

The probability density function of a vectorial input variable can be approximated through a finite number of codebook vectors with the classical Vector Quantization (VQ, cf., e.g., Gray 1984) method. Learning Vector Quantization (LVQ, Kohonen 1990) is a classification method based on VQ: an input vector is classified according to the shortest distance to codebook vectors from different classes. Optimal positions for the codebook vectors are approximated in a supervised iterative learning process. Flotzinger et al. (1992, 1994b) show that LVQ is a very powerful classifier for single trial EEG data. Compared with other methods such as Backpropagation Neural Networks (Rumelhart et al., 1986) one major advantage of the LVQ classifier are its good generalization ability (Peltoranta and Pfurtscheller 1994).

In common classification problems different feature values of a data vector are differently important. The DSLVQ algorithm is an improved LVQ classifier which uses a weighted distance function and adjusts the influence of different input features through a supervised learning algorithm. The influence of a single feature is modified according to its contribution to correct/wrong classifications of the system. Pregenzer et al. (1994b) show the efficiency of this implicit scaling method. To store and control the influence of each feature, DSLVQ uses a scaling or weights vector $w = (w_1,w_2,..,w_n)$. A value w_i of this weights vector reflects the importance of the corresponding feature i of the data vectors. In this study the DSLVQ weight values are analyzed to pre-select the most significant features for a BCI.

3. Experimental Results

For the following experiments a subject was sitting in front of a computer screen which instructed the type of the next movement. EEG was recorded before an acoustic stimulus asked the subject to actually perform the movement, i.e. during the planning phase of the movement. In the first experiment, the side of left and right index finger flexion is predicted from EEG data recorded during a one second period before movement onset. The information, which is exploited for classification, is the amplitude attenuation or event-related desynchronisation (ERD) which can be observed before hand movement. EEGs are recorded with 56 scalp electrodes placed at an inter-electrode distance of 2.5 cm over areas relevant for movement. The signals are digitally filtered, two frequency bands, 10-12Hz and 20-24Hz, are analysed separately. For a detailed description of the experiment setup and the preprocessing steps see Pfurtscheller et al. (1993a). DSLVQ is applied to 56-dimensional data vectors, where each feature value is the pre-processed EEG signal from a specific electrode position. Patterns with strong artifacts, which may be caused, for example, by eye-movements, are discarded. After artifact correction,

ESANN'1995 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 19-20-21 April 1995, D-Facto public., ISBN 2-9600049-3-0, pp. 247-252

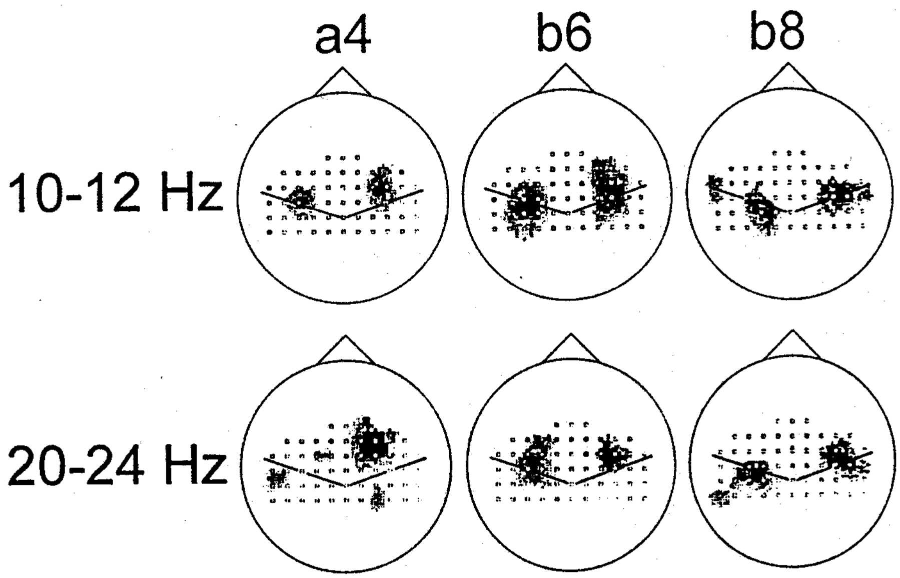

Fig. 1. Schematic view of the head with the nose above and electrode positions arranged in distances of 25 mm. The average weight values of the DSLVQ classifier are linearly interpolated for 2 different frequency bands in three subjects (a4,b6,b8). "Black" indicates areas important for discrimination between preparation of left and right hand movement. Scale in arbitrary units. (modified from Pregenzer et al. 1994b)

between 76 and 156 patterns were available for each of the three subjects. Fig. 1. shows average DSLVQ weight values from 10 runs of the DSLVQ classifier.

		56 electrodes	11 electrodes expert selection	11 electrodes DSLVQ selection
a4	10-12 Hz	74.2 (±2.8)	***77.3 (±2.1)***	76.1 (±1.5)
	20-24 Hz	59.3 (±2.6)	56.3 (±2.5)	***61.4 (2.7)***
b6	10-12 Hz	***87.2 (±1.8)***	80.8 (±1.6)	86.1 (±0.1)
	20-24 Hz	78.5 (±2.8)	82.3 (±2.7)	***88.6 (±1.2)***
b8	10-12 Hz	81.8 (±4.8)	69.5 (±3.8)	***82.4 (±3.8)***
	20-24 Hz	75.2 (±4.2)	68.5 (±2.2)	***79.6 (±1.7)***

Table 1. Classification results of LVQ for left and right hand movement with 56 electrodes and two different sub-sets of electrodes. The first sub-set has been selected from an expert and is identical for all 6 cases (3 subjects: a4, b6, b8; 2 frequency bands). The second sub-set has been selected with DSLVQ specifically for each individual subject. The figures show the average recognition accuracy and the standard deviation of 100 test runs with LVQ; the percentage of correct classifications on a testing set is shown; the rest were false classifications.

ESANN'1995 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 19-20-21 April 1995, D-Facto public., ISBN 2-9600049-3-0, pp. 247-252

Averaging over multiple runs of DSLVQ, where different initialization values and training sequences are used, results in more reliable weight values. Selection of the most important features can be done easily from DSLVQ weight values by thresholding. For all three subjects the found electrode positions form areas over the left and right primary sensorimotor hand regions. This is physiologically highly reasonable and verifies the DSLVQ selection. It is important to note that the selected areas are, however, not identical for all three subjects: the exact locations of the most important electrodes are dependent on the individual subject and on the frequency band. This is also reflected in the classification results: table 1 shows the results of an LVQ classifier for three subjects and two frequency bands. In 5 out of these 6 cases feature selection improves the classification rate. The selection of the expert fits perfectly in one case (a4, 10-12Hz), but drops important information in the other 5 cases. The LVQ classifier performs best on the sub-set of electrodes which have been selected by DSLVQ for each specific subject. In 4 out of the 6 cases the DSLVQ selection grants the best recognition results; in the other 2 cases, the results with the DSLVQ selection were only about 1% under the better of the two other results. This shows clearly the importance of a subject specific electrode selection.

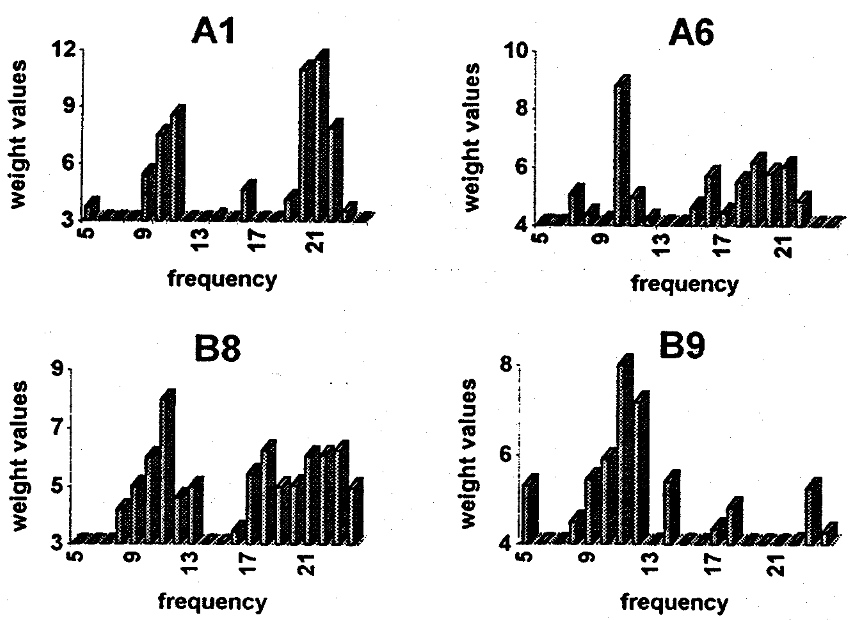

Fig. 2. Average DSLVQ weight values for 1-Hz spectral bands from 5 to 24 Hz; 4 different subjects (A1,A6,B8,B9). High relevance for the discrimination of the type of a planned movement is reflected in high weight values.

In the second experiment three different types of movement (left hand, right hand and foot) have to be discriminated. Three bipolar EEG signals were recorded from sensorimotor areas. After calculation of power spectra 1-Hz bands from 5 to 25 Hz are analyzed separately. DSLVQ is employed to select the most relevant spectral bands. Fig. 2. shows the DSLVQ weight values for 4 different subjects. The reported values are average weight values from 10 runs of DSLVQ and all 3 EEG signals. For all the subjects the significance of mu rhythms (10-12 Hz) can be seen clearly. However, similarly to the results of the first experiment, an optimal feature selection is strongly dependent on each individual subject. Notable differences between the 4 subjects are: a significance of lower mu rhythms (8-10 Hz) can only be observed for two subjects and the relevance of the central beta rhythms (20-23Hz) varies strongly: central beta rhythms are even more informative than mu rhythms for one subject (A1), but for another subject (B9) they are not relevant at all.

These differences between the four subjects could also be verified with a performance increase / decrease of an LVQ classifier when different bands were tested with different subjects. However, compared with the classification rates on band power values over wider frequency bands, the improvement obtained through a precise selection of small spectral bands was not remarkable. Unlike selection of a sub-set of electrodes, the selection of certain spectral bands does not simplify the experiment. Therefore, the additional computational effort seems not justified. However, to determine the optimal bandwidth of a band-pass filter, subject specific spectral analyses could be very important.

4. Conclusion

Single trial EEGs can be used to predict different types of movement. One problem for a learning classifier is the huge number of possible features such as different electrode positions and frequency bands. The results from two independent experiments show significant relevance differences among these features. However, the most important electrode positions and frequency bands are not identical for different subjects. Therefore, it is necessary to select the features individually for each subject. DSLVQ is shown to be an appropriate feature selection method for single-trial EEG data. The results on the sub-set of electrode positions selected by DSLVQ for each individual subject were considerably better than the results without pre-selection of important positions and on the sub-set selected by an expert. DSLVQ weights calculated for different spectral bands can be used to select the optimal bandwidth of a band-pass filter which is also dependent on a specific subject. Compared with other feature selection methods such as genetic algorithms, one major advantage of DSLVQ is its speed (Flotzinger et al., 1994a). Compared with averaging methods for EEG data analyses such as ERD mapping (Pfurtscheller and Berghold 1989), which has also been applied for subject specific feature selection, a major advantage of DSLVQ is that it is based on single trials (Pfurtscheller et al. 1994). Averaging weight values from different runs of DSLVQ, where different initial values and training sequences can be used, improves the reliability of the DSLVQ feature selection. This is one reason why feature selection is still important even if DSVLQ is also used as the subsequent classifier.

ESANN'1995 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 19-20-21 April 1995, D-Facto public., ISBN 2-9600049-3-0, pp. 247-252

References

Flotzinger, D., Kalcher J., and Pfurtscheller, G., (1992): EEG classification by Learning Vector Quantization. *Biomed. Technik*, 37: 303-309.

Flotzinger, D., Pregenzer, M., and Pfurtscheller, G., (1994a): Feature selection with Distinction Sensitive Learning Vector Quantisation and Genetic Algorithms, *Proc. IEEE Int. Conf. on Neural Networks, ICNN-94*: 3448-3451.

Flotzinger, D., Pfurtscheller, G., Neuper, Ch., Berger, J., and Mohl, W., (1994b): Classification of non-averaged EEG data by learning vector quantisation and the influence of signal preprocessing. *Medical & Biological Engineering & Computing*, 32: 571-576.

Gray, R. M., (1984): Vector Quantization. *IEEE ASSP Mag*, 1: 4-29.

Kohonen, T., (1990): The self-organizing map, *Proc. IEEE*, 78: 1464-1480.

Peltoranta , M., and Pfurtscheller, G., (1994): Neural network based classification of non-averaged event-related EEG responses, *Medical and Biological Engineering and Computing*, 32: 189-196.

Pfurtscheller, G., and Berghold, A., (1989): Patterns of cortical activation during planning of voluntary movement, *Electroenceph. clin. Neurophysiol.*, 72: 250-258.

Pfurtscheller, G., Neuper, Ch., and Kalcher, J., (1993a) 40-Hz oscillations during motor behavior in man, *Neuroscience Letters*, 162/1-2: 179-182.

Pfurtscheller, G., Flotzinger, D., and Kalcher, J., (1993b): Brain-computer interface - a new communication device for handicapped persons, *J. of Microcomp. Applications*, 16: 293-299.

Pfurtscheller, G., Pregenzer, M., and Neuper, C., (1994): Vizualization of sensorimotor areas involved in preparation for hand movement based on classification of mu and central beta rhythms in single EEG trials in man, *Neuroscience letters*, in press.

Pregenzer, M., Flotzinger, D., and Pfurtscheller, G., (1994a): Distinction Sensitive Learning Vector Quantization - A noise insensitive classification method, *Proc. IEEE Int. Conf. on Neural Networks, ICNN-94*: 2890-2894.

Pregenzer, M., Pfurtscheller, G., and Flotzinger, D., (1994b): Selection of electrode positions for an EEG-based Brain Computer Interface. *Biomed. Technik*, in press.

Pregenzer, M., Pfurtscheller, G., and Flotzinger, D., (1995): Automated Feature Selection with a Distinction Sensitive Learning Vector Quantizer, *Neurocomputing*, in press.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J., (1986): Learning internal representations by error backpropagation, in: Rumelhart, D. E. and McClelland, J. L., eds., *Parallel Distributed Processing* (MIT Press, Cambridge): 318-362.

Wolpaw, J. R., McFarland, D. J, Neat, G. W., and Forneris, C., (1991): An EEG-based brain-computer interface for cursor control, *Electroencephal. and clin. Neurophysiol.*, 78: 252-259.