# Frequency Component Selection for an EEG-Based Brain to Computer Interface

Martin Pregenzer and Gert Pfurtscheller

***Abstract*— A new communication channel for severely handicapped people could be opened with a direct brain to computer interface (BCI). Such a system classifies electrical brain signals online. In a series of training sessions, where electroencephalograph (EEG) signals are recorded on the intact scalp, a classifier is trained to discriminate a limited number of different brain states. In a subsequent series of feedback sessions, where the subject is confronted with the classification results, the subject tries to reduce the number of misclassifications. In this study the relevance of different spectral components is analyzed: 1) on the training sessions to select optimal frequency bands for the feedback sessions and 2) on the feedback sessions to monitor changes.**

***Index Terms*— Brain to computer interface (BCI), classification, distinctive sensitive learning vector quantization (DSLVQ), feature selection.**

## I. Introduction

THE IDEA of direct brain to computer communication was first introduced by Vidal [1]. Meanwhile, different types of electroencephalogram (EEG)-based brain to computer interface (BCI) systems were suggested: Donchin, and Farwell and Donchin, used an external stimulation and analyzed evoked potentials [2], [3]. Hiraiwa *et al.* classified the "Bereitschaftspotential" before movement with artificial neural networks [4]. Bierbaumer *et al.* investigated negative shifts associated with sensory and cognitive processing [5] and Kotchoubey *et al.* utilized this for a direct brain to computer communication system [6]. Keirn and Aunon [7] investigated the discrimination of different mental tasks through the EEG and Sutter [8] introduced a system which is based on voluntary control of visually induced brain responses. The subjects of Wolpaw *et al.* and Wolpaw and McFarland learned to voluntarily generate distinct EEG patterns [9], [10]. Pfurtscheller *et al.* used learning classification methods to discriminate mu-rhythm differences during movement preparation [11]. Differences of central beta rhythms have been shown by Pfurtscheller *et al.* [12]. The major advantage of EEG-based BCI systems is that no motor output (including speech) is required. This could open a new communication channel for patients with severe motor deficiencies like, for example, patients in a late stage of amyotrophic lateral sclerosis (ALS), who have lost control over every motor output and, therefore, every communication with their surroundings.

Direct brain to computer communication requires EEG patterns which can be consciously generated from the subject and clearly distinguished from a computer system. Pfurtscheller *et al.* investigated movement-related EEG signals recorded over motor-, premotor-, and somatosensory areas [13]. During preparation of movement, the event-related desynchronization (ERD, [14]) of mu and central beta rhythms shows a different spatial localization dependent on whether, e.g., right or left hand movement is being planned [15]. Flotzinger *et al.* [16] and Pregenzer *et al.* [17] optimized classification methods to predict the side of movement. Kalcher *et al.* [18] and Pfurtscheller *et al.* [19] showed that similar classification results can be achieved also when the subject only imagines the movement. Pregenzer *et al.* [20] showed that the classification results depend strongly on an appropriate preselection of EEG parameters. In the present study the relevance of spectral components is analyzed: 1) on a first block of training sessions to optimize the online classification system and 2) on a second block of feedback sessions, where the subject is confronted with the online classification results, to investigate if the online feedback causes relevance changes.

## II. Data Recording

Ten student volunteers (seven females; age 20–28) participated in preliminary recordings. Five subjects showed prominent interhemispheric ERD differences and were analyzed with single-trial methods. Out of this group, three subjects (three females, age 20–27, right handed, free of central nervous system abnormality) achieved error rates below 35% in a single-trial classification. Only these three subjects were selected to complete the experiment with feedback sessions, since feedback sessions with higher error rates are often demotivating for the subject (and hence fruitless). Some of the other subjects could possibly improve their single trial accuracy in additional training sessions without feedback. In this study, this was, however, not investigated in more detail.

During the recording, the subjects looked at a computer screen from a distance of about one meter. Fig. 1 shows the timing of the experiment: each trial started with the presentation of a fixation cross in the middle of the screen. After two seconds a short warning tone ("beep") followed. After one more second, an arrow was superimposed onto the

Manuscript received October 15, 1997; revised October 22, 1998 and June 3, 1999. This work was supported in part by the Fonds zur Förderung der wissenschaftlichen Forschung Project P11208-MED, the Steiermärkischen Landesregierung, the Österreichischen Nationalbank, and the Allgemeinen Unfallversicherungsanstalt (AUVA) in Austria.

The authors are with the Department of Medical Informatics, Institute of Biomedical Engineering, Graz University of Technology and Ludwig-Boltzmann Institute for Medical Informatics and Neuroinformatics, Graz A-8010 Austria.

Publisher Item Identifier S 1063-6528(99)07102-5.

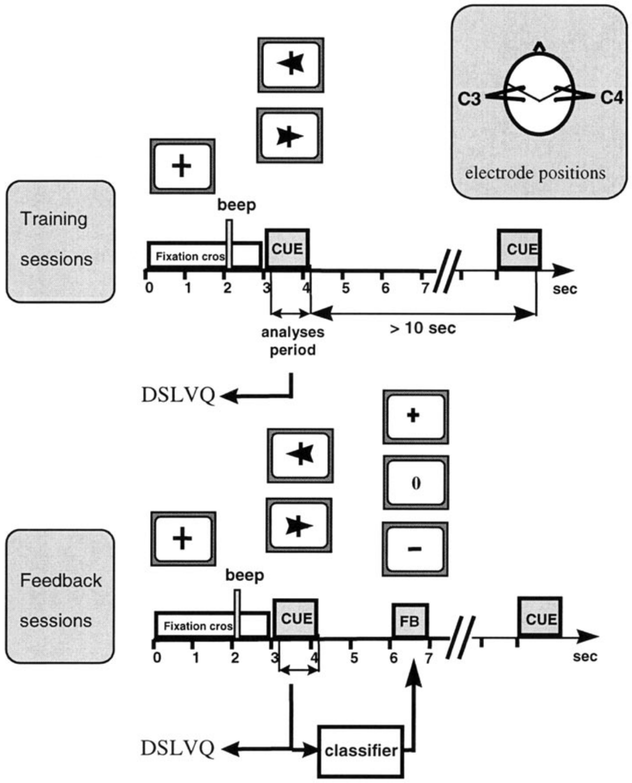


Fig. 1. Timing of training sessions and feedback sessions. The fixation cross is visible for 3 s from the beginning of a trial. At second 2 an acoustic warning stimulus is given; the visual cue stimulus in the form of an arrow, which is pointing either to the left or right side, is given at second 3. The interval between 3.25 and 4.25 is used for DSLVQ analysis and for the feedback, which is given at second 6 in feedback sessions. Inset: recording of bipolar channels: C3: C3a–C3p, and C4: C4a–C4p. C3a (C4a) and C3p (C4p) refer to an electrode placement 2.5 cm anterior and 2.5 cm posterior to C3 (C4).

fixation cross. The arrow in the center of the screen instructed the subjects in the following way: depending on the direction of the arrow (left or right), the subjects should imagine either a left or right hand movement. The arrow was always shown at the same position on the screen. Only the direction of the arrow varied with the instruction. Before the experiment, the subjects were given the opportunity to practice and actually perform the movements (dorsiflexions) of the left and right hand according to the stimulation on the screen. During the experiment, the subjects did not perform real movements, they only imagined the movements. This was supervised with EMG recordings of the extensor muscles of both hands.

Two bipolar EEG channels were recorded with four sintered Ag/AgCl scalp electrodes (Fig. 1, inset) which were placed 2.5 cm anterior and posterior to C3 and C4 (international 10–20 system, [21]). These optimal electrode positions were found in [22]. The EEG was amplified between 0.5–35 Hz (3 dB points, 16 dB/octave) by a Nihon Khoden amplifier and then sampled at 128 Hz. EOG was derived from two electrodes, one placed medially just above the right eye and the other laterally just below the right eye, in order to detect both vertical and horizontal eye movements. These signals were used to screen the EEG recordings for eye movement artifacts: trials with eye movement or muscle artifacts were removed.

The experiment consisted of two types of sessions: 1) initial training sessions, where the subject was repeatedly performing the given tasks as constantly as possible to get consistent data to build a subject-specific classifier and 2) test or feedback sessions, where the classification results were immediately reported to the subject. Between three to four training sessions (without feedback) were recorded for each subject. The sessions were recorded on different days and consisted of four experimental runs of 40 trials (20 of each side). A session lasted for about 1 h and the sequence of left and right trials and the length of the breaks between consecutive trials were randomized. The training data was then analyzed with distinction sensitive learning vector quantization (DSLVQ, [23]) to determine the band-pass filter settings for the online classification system. These filter settings were individually chosen for each subject. Another three to four sessions were then recorded with each subject and feedback was given. The feedback, which was presented at second 6 of each trial, was one out of five symbols ("+," "+," "o," "-," "–"). It indicated if the computer could classify the present EEG pattern correctly or not and also how clear the decision was. Details of the online classification system can be found in [24].

## III. Frequency Band Optimization

A 1-s time window starting 0.25 s after cue presentation was used for classification (see Fig. 1). The online system could process two frequency bands, which could be individually selected for each subject. To optimize this parameter the data from the classification window was transformed in the frequency domain by fast Fourier transformation (FFT). The window length of 1 s resulted in a 1-Hz frequency resolution. In previous studies it was shown that mu and central beta rhythms are relevant to classify movement preparation [20]. The analysis in this study could, therefore, be restricted to frequency components from 9 to 28 Hz (20 components).

To determine the relevance of the single components, DSLVQ [23], [25] was used. DSLVQ is a modified version of Kohonen's learning vector quantization (LVQ, [26], [27]) classifier: it incorporates an additional feature scaling to compensate for relevance differences among the input features (Fig. 2). DSLVQ uses the standard LVQ strategy to divide a classification problem into subproblems and to find an optimal linear approximation for each subproblem. DSLVQ then analyzes the importance of the input features for the subproblems. The average importance is stored in a relevance vector (**RT** vector) and used for scaling of the input space. The scaling of the input data leads to a better problem representation with the LVQ codebook and improved relevance analyzes. Compared to other parameter selection methods, which are often also much slower, a major advantage of the DSLVQ method is that the features are not analyzed individually and that features which are relevant only in combination with other features are also found. Theoretical analysis and a detailed description of the algorithm can be found in [23].

In this study, DSLVQ was applied to the 40 dimensional (20 frequency components, two electrodes) vector of spectral

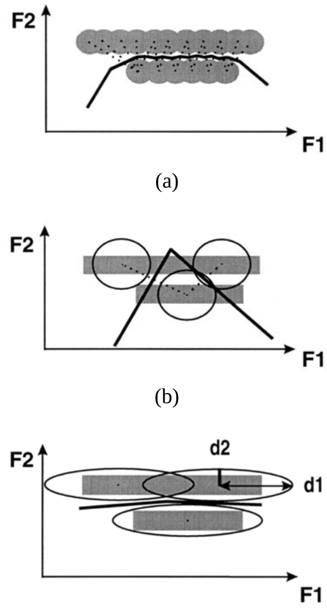


Fig. 2. Two classes have to be discriminated in a two-dimensional space. While feature 2 (F2) is highly informative for this discrimination task, feature 1 (F1) contributes only a little to a more reliable classification. (a) The problem could be solved with LVQ and a large number of codebook vectors. (b) With a limited number of codebook vectors, the equal influence of both features on the distance calculation causes classification errors. (c) Unlike LVQ, DSLVQ employs a weighted distance function: the influence of feature 1 is reduced (larger distances d1 can be accepted) and the problem can be solved with a small number of codebook vectors.

components. Before the training, the data was normalized to zero mean and unit variance to compensate for relevance differences due to different absolute power values. All trials (from different training sessions with the same subject) were mixed together and a 50% subset of the examples was arbitrarily selected for DSLVQ training; the remaining 50% were reserved for testing. The system was initialized with $k$-means clustering (three codebook vectors per class). The initialization was repeated when one cluster represented less than three examples (i.e., if a cluster represented only a few of outliers). The system was then fine-tuned with type-C DSLVQ training [23]. During the 10 000 training iterations, the learning rate $\alpha(t)$ decreased to zero from an initial value $\alpha(0)$ of 0.05. For the DSLVQ relevance values, a learning rate $\alpha'(t) = 0.1 * \alpha(t)$ was taken. After the training was finished the relevance values ($RT_n$ values) for the 40 input features were analyzed. Since DSVLQ training is very quick, the process could be repeated 100 times with a different splitting into training and testing data: the median value of these 100 runs was then analyzed. Additionally the variance was controlled. (A small variance is an evidence that the feature selection is clear without ambiguity.)

## IV. Results

### A. Selection of Frequency Components

The box plots in Figs. 3–5 visualize the distribution of the $RT_n$ values from 100 iterations of DSLVQ. From these

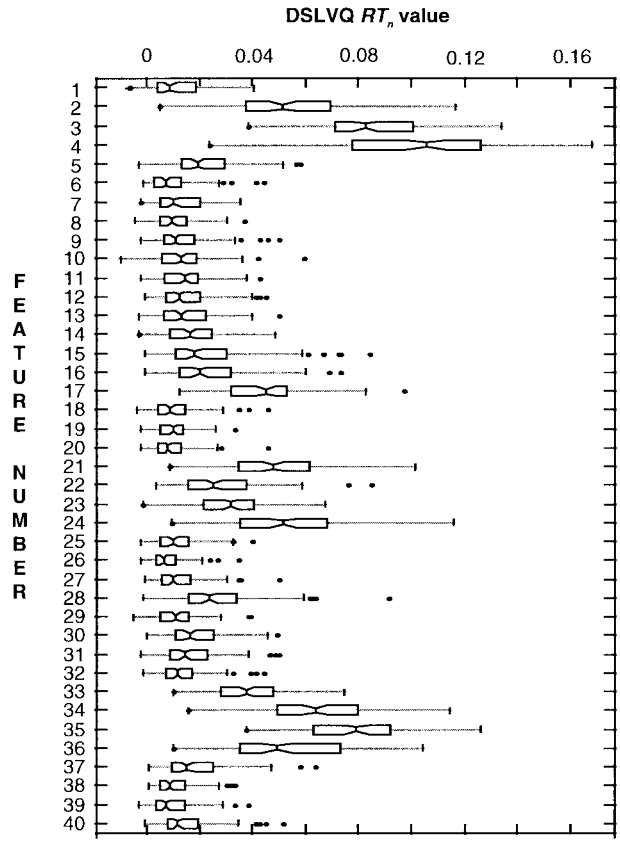


Fig. 3. Feature relevance for subject f3. Data from the training phase was analyzed with DSLVQ. Features 1–20 and features 21–40 are the 1 Hz spectral components between 9–28 Hz from C3 and C4, respectively. The box plot shows the distribution of the DSLVQ $RT_n$ values from 100 runs of the algorithm on different training sets. Each box has lines at the lower quartile, median and upper quartile values. The "whiskers," the lines extending from each end of the box, show the extent of the rest of the data. Outliers are indicated separately with black dots. The "notches" in the box are a graphic confidence interval for the median of the sample. It can be seen clearly that features 2–4 (representing the band 10–12 Hz at C3) and the features 33–36 (frequency band 21–24 Hz at C4) are most relevant. The notches for these features are not overlapping with the notches for little relevant features, which indicates a high significance level. Note that the feature relevance is not identical at C3 and C4: the relevance of a mu rhythm at C4 and central beta rhythms at C3 is much less pronounced.

plots, the most relevant frequency components were manually selected: a feature $n_0$ was selected if the average $RT_{n_0}$ value was large and the variance was small; in unclear cases it was analyzed if the frequency component lies within a relevant frequency band or if it might be an outlier. Note that the presence of relevant frequency bands in the data can be seen clearly in Figs. 3 and 4: since DSLVQ treats all the features equally, the similar $RT_n$ values of neighboring features cannot be induced by the method. The frequency components which were selected for the individual subjects are shown in Fig. 6. These components were used for the offline testing in this study.

### B. Classification Performance

To verify the DSLVQ feature selection, different classification methods were applied offline on all 40 features and on the selected subsets. Additionally to the DSLVQ

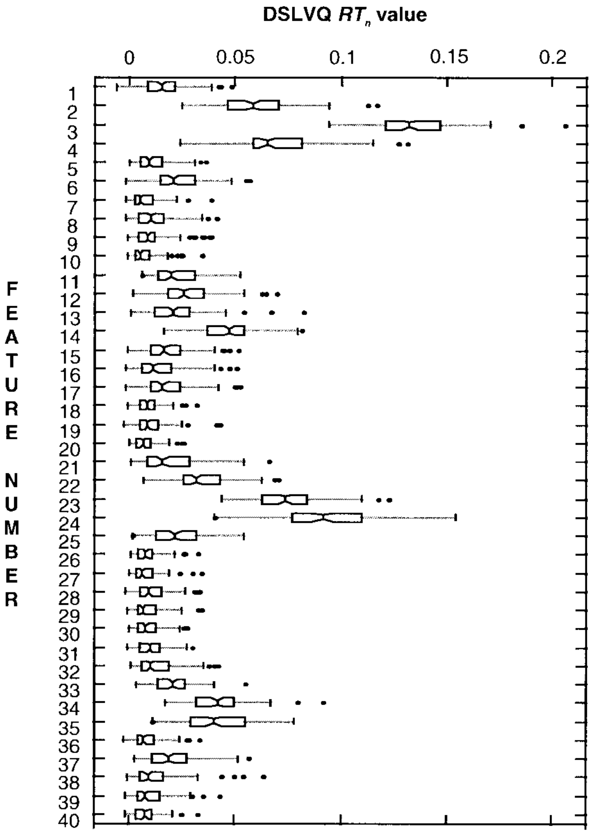


Fig. 4. Feature relevance for subject f5 during training sessions. For details of the box plot see Fig. 3. It can be seen clearly that a mu rhythm from 10 to 12 Hz is most relevant on both sides. (C3: features 2 to 4, C4 features 22 to 24). The box plot shows that this finding is highly significant. Additionally, a beta band around 22 Hz (features 14 and 34) is found relevant on both sides.

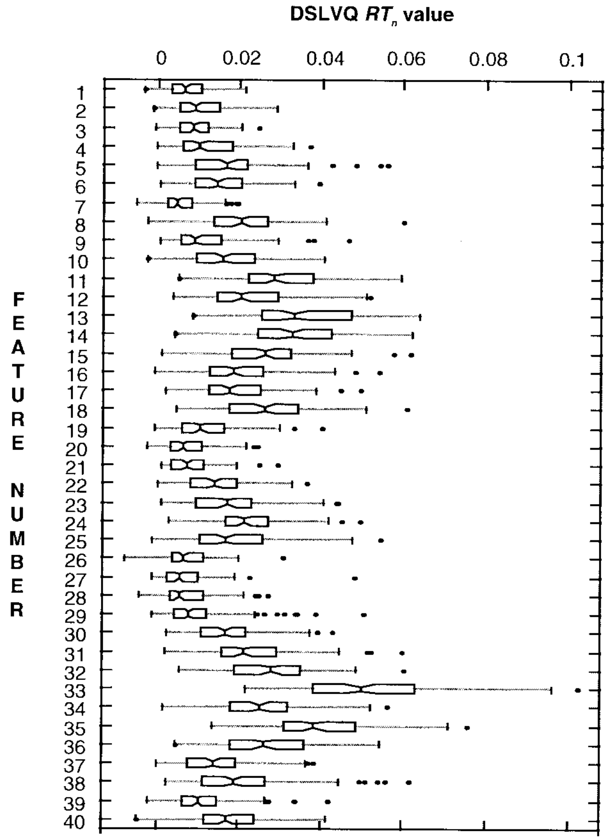


Fig. 5. Feature relevance for subject f7 during training sessions. For details of the box plot see Fig. 3. It can be seen that central beta rhythms around 22 Hz (features 13 and 33) are most relevant. The relevant band is broad (from 18 to 26 Hz, features 10 to 18 and 30 to 38) and similar at C3 and C4. The spread of the data is relatively large, the number of outliers is larger than in Figs. 3 and 4.

classifier, an LVQ1 classifier (6 codebook vectors, $\alpha(0) = 0.05$, $10\,000$ iterations) and k-NN were applied: LVQ, since it performed best on this type of EEG data in a comparative study of Peltoranta and Pfurtscheller [28], nearest neighbor (NN) classification, since the error probability of 1-NN classification is asymptotically bounded above by twice the Bayes error probability [29]. The results on all the features and the selected subsets are summarized in Tables I–III. Three observations are made as follows.

1) For all three subjects, LVQ1 and DSLVQ achieve error rates close to half of the 1-NN benchmark value, which is an indicator for very good results.
2) DSLVQ (with implicit feature weighting) outperforms LVQ (without feature weighting) when the features are not preselected. For subjects f3 and f5 a tailed $t$-test can exclude at a significance level of $p = 0.001$ that the DSLVQ results are equal or worse than the LVQ1 results. For subject f7 there is no difference between DSLVQ and LVQ.
3) For all the subjects it could be observed that the feature selection significantly improved the classification performance. Only in two out of 12 cases (f7 + LVQ1 and f7 + DSLVQ) the difference was not significant (at $p = 0.01$ level).

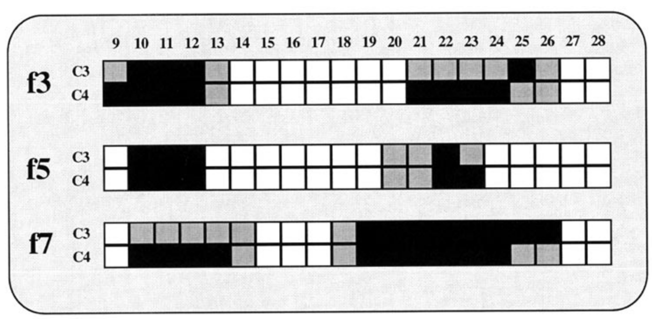


Fig. 6. Result of the manual feature selection from the DSLVQ relevance values (Figs. 3–5). Frequency components are shown on the $x$-axis. The most relevant components are marked black for the three subjects (f3, f5, and f7) and for the two electrode positions (C3 and C4). The grey extensions mark the features which had been used for the online feedback. Note that the online system required equal bands for both channels, i.e., that the perfect symmetry of the grey extensions is artificial. Similarities can be seen between the two electrodes and also between the three subjects. There exist, however, also clear differences between the two electrode positions (compare, e.g., subject f7, 10–13 Hz) and between the three subjects (compare, e.g., subjects f5, f7, beta band).

The online classification system did not allow the selection of single frequency components. A maximum of two frequency bands had to be specified and the bands had to be identical for the two electrodes. The bands which were used for the online classification are marked with gray color in Fig. 6.

TABLE I
CLASSIFICATION RESULTS FOR SUBJECT f3. DATA FROM TRAINING SESSIONS, RANDOM SPLITTING INTO TRAINING AND TESTING DATA (1 : 1). THE MEAN VALUE ($\pm$STANDARD DEVIATION) OF CORRECT CLASSIFICATIONS FROM 100 ITERATIONS IS REPORTED FOR ALL THE FEATURES (COLUMN 2) AND FOR A SUBSET OF SELECTED FEATURES (COLUMN 3). THE FEATURES THAT HAVE BEEN SELECTED ARE INDICATED IN FIG. 6. THE HYPOTHESIS THAT THE RESULTS WITH THE REDUCED FEATURE SET ARE BETTER THAN THE RESULTS WITH ALL THE FEATURES HAVE BEEN TESTED WITH A TAILED $t$-TEST: THIS HYPOTHESIS CAN BE ACCEPTED AT A HIGH SIGNIFICANCE LEVEL FOR ALL CLASSIFICATION METHODS (COLUMN 4)

| Algorithm | all features | selected features | significance level of improvement |
|---|---|---|---|
| 1-NN | 64.1 (±3.8) | 72.1 (±2.9) | $p < 0.001$ |
| 15-NN | 75.2 (±3.3) | 79.5 (±2.9) | $p < 0.001$ |
| LVQ1 | 77.0 (±2.8) | 80.9 (±3.1) | $p < 0.001$ |
| DSLVQ | 78.5 (±2.9) | 80.7 (±2.9) | $p < 0.001$ |

TABLE II
CLASSIFICATION RESULTS FOR SUBJECT f5 (PARALLEL TO TABLE I). LIKE FOR SUBJECT f3, THE FEATURE SUBSET SELECTION IMPROVED THE PERFORMANCE INDEPENDENT OF THE CLASSIFICATION METHOD

| Algorithm | all features | selected features | significance level of improvement |
|---|---|---|---|
| 1-NN | 67.1 (±2.6) | 76.0 (±2.5) | $p < 0.001$ |
| 15-NN | 75.3 (±2.7) | 80.5 (±2.3) | $p <0.001$ |
| LVQ1 | 77.6 (±2.2) | 80.9 (±2.0) | $p < 0.001$ |
| DSLVQ | 79.9 (±2.2) | 80.8 (±2.0) | $p = 0.0015$ |

TABLE III
CLASSIFICATION RESULTS FOR SUBJECT f7 (PARALLEL TO TABLES I AND II). OPPOSITE TO THE OTHER TWO SUBJECTS, FOR SUBJECT f7 THE DSLVQ SELECTION DOES NOT RESULT IN A CLEAR PERFORMANCE INCREASE. ONLY FOR NEAREST NEIGHBOR CLASSIFICATION A SIGNIFICANT IMPROVEMENT CAN BE OBSERVED. FOR LVQ1 AND DSLVQ DIFFERENCES BETWEEN ALL AND SELECTED FEATURES ARE NOT SIGNIFICANT. NOTE THAT ALSO THE RELEVANCE ANALYSIS (FIG. 5), WHICH WAS THE BASIS FOR THE FEATURE SELECTION, WAS NOT CLEAR FOR SUBJECT f7

| Algorithm | all features | selected features | significance level of improvement |
|---|---|---|---|
| 1-NN | 59.0 (±2.4) | 59.8 (±2.3) | $p = 0.01$ |
| 15-NN | 68.0 (±2.3) | 69.7 (±2.1) | $p < 0.001$ |
| LVQ1 | 70.5 (±2.0) | 70.7 (±2.2) | - |
| DSLVQ | 70.5 (±2.2) | 69.9 (±2.0) | - |

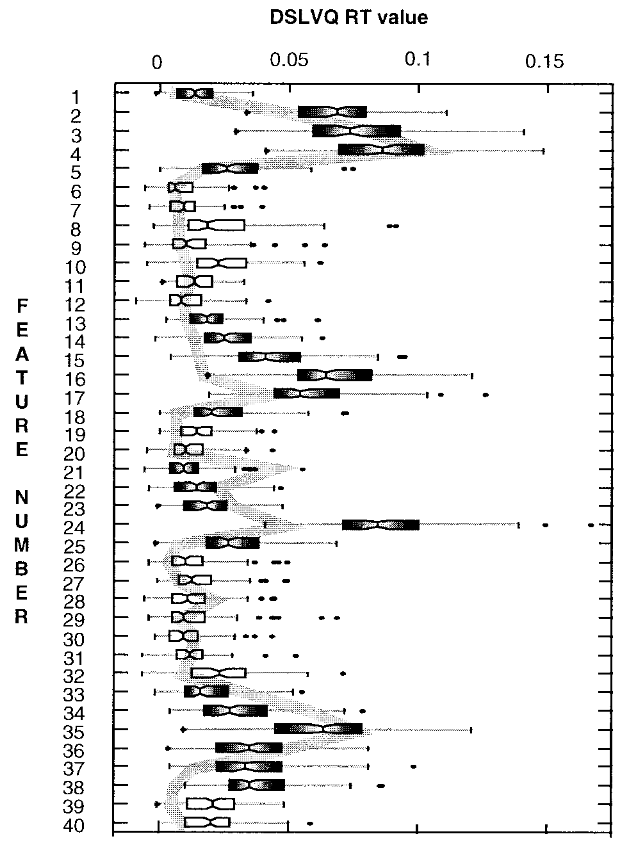


Fig. 7. Subject f3, DSLVQ $RT_n$ values from the feedback sessions (100 different runs). The results should be compared to Fig. 3 ($RT_n$ values from the training sessions). For easier comparison, the median values from the training sessions are indicated in this figure through the thick grey line. The filled boxes in this figure mark which features have been used for the online-feedback. It can be observed that all the features which show a large relevance increase (features 15, 16, 24, 38) were also used for the online-feedback.

### C. Feedback Sessions

The results of relevance analyzes on the feedback sessions are shown in Figs. 7–9. When these results are compared with the results on the training sessions (Figs. 3–5), it can be seen that for one subject (f5), the relevance of the features did not change: exactly the same frequency components as in the training sessions also stayed relevant in the feedback sessions. For the other two subjects (f3 and f7), the relevance varied between the training sessions and the following feedback sessions. The changes were particularly interesting with subject f3: while a single band was relevant at each electrode position (at C3 10–12 Hz, features 2–4, and at C4 21–24 Hz, features 33–36) during the training sessions, the relevance patterns became more symmetrical in the feedback sessions (attention should be paid to the relevance increase of features 13–16, which are symmetrical to features 33–36, and feature 24, which is symmetrical to feature 4).

## V. DISCUSSION

### A. Selection

The DSLVQ $RT_n$ values show clear relevance differences between the frequency components (Figs. 3–5 and 7–9). The box-plots indicate that the variation of the $RT_n$ values over multiple runs on different training data is small compared to the differences among the frequency components: this indicates that the results are highly significant. The clearest divergence between training and testing can be observed for subject f3: while in the training sessions, mu rhythms were dominant at C3 and beta rhythms at C4, both rhythms became equally relevant at both electrode positions in the feedback sessions. The feedback session differed from the training sessions in two ways: first, the subject saw a feedback symbol (between second 6 and 7, compare Fig. 2). Since the classification interval was much earlier (between 3.25–4.25) this can, however, not visually induce the difference. Second, the feedback sessions were recorded after the training sessions and a possible development or learning effect may be visible. This interpretation is very plausible: when the subject obtains feedback and tries to optimize his/her strategy this will necessarily change the EEG patterns and hence also the relevance of the different frequency components.

The box plots in Figs. 3–5 show a hemispheric symmetry of the feature relevance in two subjects (f5 and f7). For subject f3, the relevant frequency components are, however, not symmetrical for C3 and C4. This, itself, is a very interesting

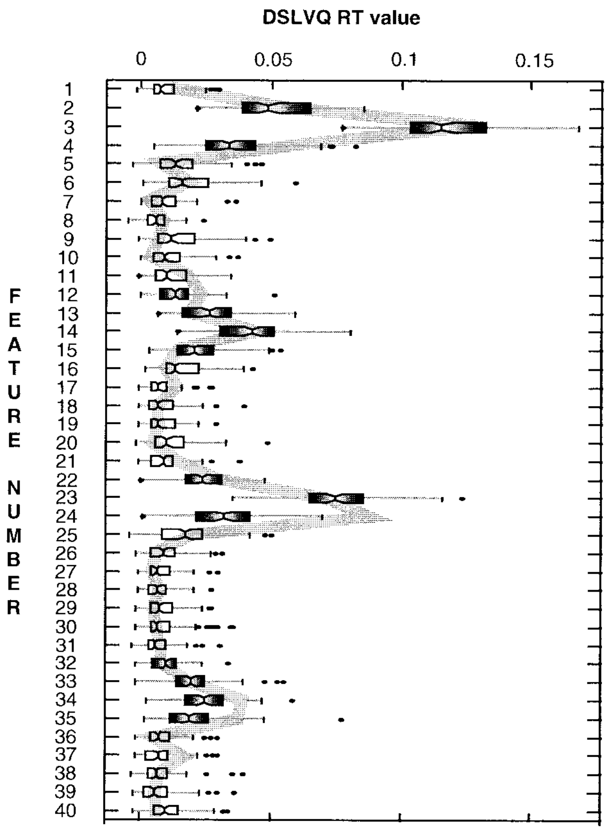


Fig. 8. Like Fig. 7, but for subject f5: DSLVQ $RT_n$ values from the feedback sessions. It can be seen that the results are very similar to the results from the training sessions.

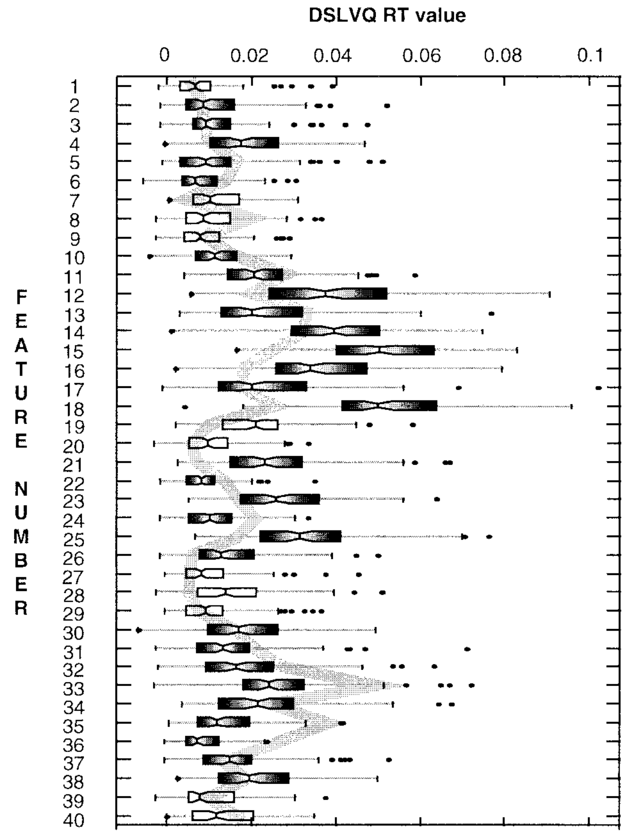


Fig. 9. Like Fig. 7, but for subject f7: DSLVQ $RT_n$ values from the feedback sessions. Similarly to the results from the training sessions (Fig. 5) the feature relevance is not very clear: no clear bands can be observed (bands would be physiologically plausible), and the variance of the DSLVQ $RT_n$ values is very high. A direct comparison of the median values shows large differences between training and feedback sessions. The features which have been used for online feedback are indicated with a filled box. Like for subject f3 (Fig. 7), it can be observed that the features with the clearest relevance increase have all been used for feedback.

observation, which cannot be explained with the misplacement of an electrode (the data was recorded in different sessions on different days). An explanation could be that the gyri and sulci in both hemispheres are not symmetrical and that the electrodes could be placed over different parts of the Rolandic region. Most interesting for the BCI is, however, that the relevance changed when feedback was given. It can be seen clearly in Fig. 7 that features which were of little relevance in the training sessions but were used for feedback (like features 13–16) became significantly more relevant in the feedback sessions. A similar observation can be made for subject f7: the relevance of many features varied considerably between training and testing sessions, but again, the features with the largest relevance increase (features 15, 16, 18, 21, 25 in Fig. 9) had all been used for feedback. The probability that this is observed by chance is only $p = 0.003$.

A plausible interpretation for the significant relevance changes is that the feedback increased the importance of specific frequency components, i.e., that the subject adapted to the classification system. The increase of the mu rhythm by biofeedback is already known and used to reduce the number of seizures in epileptic patients [30]. The present results suggest that biofeedback can increase the relevance of both mu and central beta rhythms. It will have to be investigated carefully which feedback should be given to the subject: the adaptation of the subject could possibly be assisted with a more specific feedback (i.e., from isolated bands).

### B. Classification

The results clearly indicate that the classification accuracy on selected frequency components is generally better than on all the frequency components: a removal of about half of the features could significantly improve the classification accuracy for all subjects. It can be concluded that: i) certain frequency bands bear all the relevant information and ii) only these bands should be used online since this also simplifies the system (less features, reduced response time of the system).

The improved classification results also confirm the DSLVQ relevance estimates: DSLVQ unquestionably found those features which provided the relevant information. In certain cases the improvements from the feature selection were greatest (subjects f3 and f5, 1-NN and 15-NN). The 1-nearest neighbor algorithm (which showed the largest improvements with 8 and 9% for subjects f3 and f5, respectively) is the most noise sensitive algorithm. Since the DSLVQ feature selection removes noise from the data, the 1-NN algorithm profits most. The DSLVQ feature selection was clearer for subjects f3 and f5 than for subject f7 (compare Figs. 3–5): the $RT_n$ values for subjects f3 and f5 showed one or two very distinct and plausible bands; the $RT_n$ values for subject f7 showed a broader band, which was also much less pronounced. Either the DSLVQ selection was better or the relevance differences in the data were more pronounced for subjects f3 and f5. It is plausible that the feature selection is less important when the band is very broad and multiple features reflect similar or equal information.

## VI. Conclusions

For the construction of a brain to computer interface based on motor imagery, the relevant spectral components must be carefully selected. This study pointed out that the selection should be individual for each subject and also for each electrode position. It was found here that the most relevant frequency bands are not necessarily symmetrical for the homologous electrode locations C3 and C4.

It was also observed in this study that biofeedback can increase the relevance of certain frequency components. This was found not only for mu rhythms but also for central beta components. The new results suggest that frequency selection might need dynamic adaptation.

The (possibly also disturbing) impact of feedback on the classification is presumably one of the key issues in future BCI research. It must be carefully investigated which parts must be learned on which side of a brain to computer interface.

## References

[1] J. Vidal, "Toward direct brain-computer communication," *Annu. Rev. Biophys. Bioeng.*, pp. 157–180, 1973.
[2] E. Donchin, *Event Related Potentials in Man.* New York: Academic, 1978, pp. 349–441.
[3] L. Farwell and E. Donchin, "Taking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials," *Electroenceph. Clin. Neurophysiol.*, vol. 70, pp. 510–523, 1981.
[4] A. Hiraiwa, K. Shimohara, and Y. Tokunaga, "EEG topography recognition by neural networks," *Eng. Med. Biol.*, pp. 39–42, 1990.
[5] N. Birbaumer, L. E. Roberts, W. Lutzenberger, B. Rockstroh, and T. Elbert, "Area-specific self-regulation of slow cortical potentials on the sagittal midline and its effects on behavior," *Electroenceph. Clin. Neurophysiol.*, vol. 84, pp. 351–361, 1992.
[6] B. Kotchoubey, H. Schleichert, W. Lutzenberger, and N. Birbaumer, "A new method for self-regulation of slow cortical potentials in a timed paradigm," *Appl. Psychophysiol. Biofeedback*, vol. 22, pp. 77–93, 1991.
[7] Z. A. Keirn and J. I. Aunon, "Man-machine communications through brain-wave processing," *IEEE Eng. Med. Biol.*, vol. 3, pp. 55–57, 1990.
[8] E. Sutter, "The brain response interface: Communication through visually-induced electrical brain responses," *J. Microcomput. Appl.*, vol. 15, pp. 31–45, 1992.
[9] J. R. Wolpaw, D. J. McFarland, G. W. Neat, and C. Forneris, "An EEG-based brain-computer interface for cursor control," *Electroenceph. Clin. Neurophysiol.*, vol. 78, pp. 252–259, 1991.
[10] J. R. Wolpaw and D. J. McFarland, "Multichannel EEG-based brain-computer communication," *Electroenceph. Clin. Neurophysiol.*, vol. 90, pp. 444–449, 1994.
[11] G. Pfurtscheller, D. Flotzinger, W. Mohl, and M. Peltoranta, "Prediction of the side of hand movements from single-trial multi-channel EEG data using neural networks," *Electroenceph. Clin. Neurophysiol.*, vol. 82, pp. 313–315, 1992.
[12] G. Pfurtscheller, M. Pregenzer, and C. Neuper, "Visualization of sensorimotor areas involved in preparation for hand movement based on classification of mu and central beta rhythms in single EEG trials in man," *Neurosci. Lett.*, vol. 181, pp. 43–46, 1994.
[13] G. Pfurtscheller, J. Kalcher, C. Neuper, D. Flotzinger, and M. Pregenzer, "On-line EEG classification during externally-paced hand movements using a neural network-based classifier," *Electroencephal. Clin. Neurophysiol.*, vol. 99, 416–425, 1996.
[14] G. Pfurtscheller and A. Aranibar, "Evaluation of event-related desynchronization (ERD) preceding and following voluntary self-paced movement," *Electroenceph. Clin. Neurophysiol.*, vol. 46, pp. 138–146, 1979.
[15] G. Pfurtscheller and A. Berghold, "Patterns of cortical activation during planning of voluntary movement," *Electroenceph. Clin. Neurophysiol.*, vol. 72, pp. 250–258, 1989.
[16] D. Flotzinger, G. Pfurtscheller, C. Neuper, J. Berger, and W. Mohl, "Classification of nonaveraged EEG data by learning vector quantization and the influence of signal preprocessing," *Med. Biol. Eng. Comput.*, vol. 32, pp. 571–576, 1994.
[17] M. Pregenzer, G. Pfurtscheller, and D. Flotzinger, "Automated feature selection with a distinction sensitive learning vector quantizer," *Neurocomputing*, vol. 11, pp. 19–29, 1996.
[18] J. Kalcher, D. Flotzinger, C. Neuper, S. Gölly, and G. Pfurtscheller, "Graz brain-computer interface II—Toward communication between humans and computers based on online classification of three different EEG patterns," *Med. Biol. Eng. Comput.*, vol. 34, pp. 382–388, 1996.
[19] G. Pfurtscheller, C. Neuper, A. Schlögl, and K. Lugger, "Separability of EEG signals recorded during right and left motor imagery using adaptive autoregressive parameters," *IEEE Trans. Rehab. Eng.*, vol. 6, pp. 316–325, 1998.
[20] M. Pregenzer, G. Pfurtscheller, and C. Andrew, "Improvement of EEG classification with a subject specific feature selection," in *Proc. European Symp. Artificial Neural Networks ESANN'95*, Brussels, Belgium, 1995, pp. 247–252.
[21] R. W. Homan, J. Herman, and P. Purdy, "Cerebral location of international 10-20 system electrode placement," *Electroenceph. Clin. Neurophysiol.*, vol. 66, pp. 376–382, 1987.
[22] M. Pregenzer, G. Pfurtscheller, and D. Flotzinger, "Selection of electrode positions for an EEG-based brain computer interface," *Biomed. Technik*, vol. 39, pp. 264–269, 1994.
[23] M. Pregenzer, *DSLVQ.* Aachen, Germany: Schaker Verlag, 1998.
[24] G. Pfurtscheller, C. Neuper, D. Flotzinger, and M. Pregenzer, "EEG discrimination between imagination of right and left hand movement," *Electroencephal. Clin. Neurophysiol.*, vol. 103, pp. 642–651, 1997.
[25] M. Pregenzer, D. Flotzinger, and G. Pfurtscheller, "Distinction sensitive learning vector quantization—A noise insensitive classification method," in *Proc. IEEE Int. Conf. Neural Networks ICNN'94*, Orlando, FL; Piscataway, NY: IEEE Service Center, 1994, pp. 2890–2894.
[26] T. Kohonen, "The self-organizing map," *Proc. IEEE*, vol. 78, pp. 1464–1480, 1990.
[27] T. Kohonen, *Self Organizing Maps.* Heidelberg, Germany: Springer, 1995.
[28] M. Peltoranta and G. Pfurtscheller, "Neural network based classification of nonaveraged event-related EEG responses," *Med. Biol. Eng. Comput.*, vol. 32, pp. 189–196, 1994.
[29] T. M. Cover and P. E. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inform. Theory*, vol. 3, pp. 21–27, 1967.
[30] W. Kuhlman, "EEG feedback training of epileptic patients: Clinical and electroencephalographic analysis," *Electroenceph. Clin. Neurophysiol.*, vol. 45, pp. 699–710, 1978.

**Martin Pregenzer** was born in Innsbruck, Austria, in 1969. He received the M.Sc. degree in Telematik (computer science) and the Ph.D. degree in neuroinformatics from the Graz University of Technology, Austria, in 1993 and 1997, respectively.

He is currently working as a Research Assistant at the Institute of Biomedical Engineering, Graz University of Technology, where he teaches a course on neural networks. His research interests focus on instance-based classification and feature selection.

**Gert Pfurtscheller** received the M.Sc. and Ph.D. degrees in electrical engineering from Graz University of Technology, Austria, in 1964 and 1966, respectively.

He is a Professor of medical informatics at the Institute of Biomedical Engineering at the Graz University of Technology. Since 1982, he has been Head of the Department of Medical Informatics and since 1987, Director of the Ludwig Boltzmann-Institute for Medical Informatics and Neuroinformatics. He has authored more than 300 scientific articles and four books. His scientific interests and current research work include biomedical signal processing with special emphasis on electrical brain activity analysis, brain computer communication systems, and medical expert systems.