Information Transfer Rate in a Five-Classes Brain–Computer Interface

Bernhard Obermaier, Christa Neuper, Christoph Guger, *Associate Member, IEEE*, and Gert Pfurtscheller, *Member, IEEE*

***Abstract*—The information transfer rate, given in bits per trial, is used as an evaluation measurement in a brain–computer interface (BCI). Three subjects performed four motor-imagery (left hand, right hand, foot, and tongue) and one mental-calculation task. Classification of the electroencephalogram (EEG) patterns is based on band power estimates and hidden Markov models (HMMs). We propose a method that combines the EEG patterns based on separability into subsets of two, three, four, and five mental tasks. The information transfer rates of the BCI systems comprised of these subsets are reported. The achieved information transfer rates vary from 0.42 to 0.81 bits per trial and reveal that the upper limit of different mental tasks for a BCI system is three. In each subject, different combinations of three tasks resulted in the best performance.**

***Index Terms*—Brain–computer interface (BCI), electroencephalogram (EEG) classification, hidden Markov model, information transfer rate.**

I. Introduction

An INTERFACE between the human brain and the computer was recently introduced as a spelling device [1] for a patient suffering from amyotrophic lateral sclerosis (ALS). Such a spelling device allows the patient to select letters from an alphabet exclusively via mental activity [2]. The duration for the selection of one letter is reported to be 2 min and is based on a binary decision. The selection itself is done by successive isolation of the desired letter, starting with groups containing all letters and finally ending with the choice between the desired letter and another one. A feasible way to increase the speed of the system is by moving from a binary decision to a more diverse decision, giving a choice between more options.

In the Graz brain–computer interface (BCI), decisions are based on the classification of electroencephalogram (EEG) patterns recorded during N predefined different mental tasks. "Mental task" refers to an imagined performance of, e.g., a motor act, without, in fact, performing it. Studies done at our institute already showed that a real-time and online BCI system can be realized for two or three different mental tasks (see [3]–[5] for $N = 2$ and [6], [7] for $N = 3$). Anderson *et al.* proposed an offline BCI system for $N = 5$ [8].

In this work, a BCI system able to classify $N = 5$ different mental tasks is presented. Given such a BCI system, spelling can be done based on an alphabet divided into five groups, each arbitrarily associated with a specific mental task. The groups are displayed onscreen simultaneously and the selection of a letter contained in a certain group is done by the subject performing the corresponding mental task. Referring to the classification result (options one to five), the corresponding letter group is further split into five subsets containing one letter each. These letters are then again simultaneously displayed on the screen. Finally, the desired letter is selected by the performance of the assigned mental task. It is important to note that the selection of the letter is not based on evoked responses due to a displayed stimulus. In fact, the subjects are able to modulate their spontaneous EEG activity regarding the needed option. This method of spelling a letter [9], [7] was demonstrated to work properly in the case of a BCI system classifying two mental tasks.

Furthermore, these results were interesting because the spelling was based on a BCI system trained with EEG patterns captured during a training session different from the spelling procedure. During training, the subject is asked to perform a mental task associated with a displayed symbol. The recorded EEG is then used to setup the classifier of the BCI system. During spelling, first, the subject has to determine the mental task in order to select the desired letter and, then, actually perform it. There is no symbol determining the mental task, as in the training. Even though the reasons for performing the mental tasks are different (training—due to a stimulus, spelling—due to selection of a letter), there is no difference between the EEG data obtained during training and spelling conditions [9], [7]. This fact allows investigation of offline data captured in an training procedure like the one presented and still leaves the possibility of drawing conclusions about such a BCI system used for spelling.

It is obvious that the speed of spelling depends on the reliability of classification. It should be noted that an increase of the number of mental tasks decreases the reliability of classification because every additional EEG pattern to be classified results in decreased reliability. In order to quantify this reliability, the information transfer rate B introduced by Shannon [10] is used. Wolpaw [11] introduced the information transfer rate of a BCI as bits per trial (bits per decision) and bits per minute with 12 selections per minute (trial length is 5 s). It is of

Manuscript received December 30, 2000; revised May 9, 2001 and June 11, 2001. This work was supported by the Austrian "Fonds zur Förderung der wissenschaftlichen Forschung," Project P11208MED, and the "Steiermärkische Landesregierung."

B. Obermaier and C. Neuper are with the Ludwig Boltzmann Institute of Medical Informatics, Graz University of Technology, Graz A-8010, Austria (e-mail: obermai@dpmi.tu-graz.ac.at; neuper@dpmi.tu-graz.ac.at).

C. Guger is with G.TDEC, Graz 8051, Austria (e-mail: office@gtec.at).

G. Pfurtscheller is with the Ludwig Boltzmann Institute of Medical Informatics, Graz University of Technology, Graz A-8010, Austria, and the Department of Medical Informatics, Graz University of Technology, Graz A-8010, Austria (e-mail: pfu@dpmi.tu-graz.ac.at).

Publisher Item Identifier S 1534-4320(01)08916-1.

interest whether three, four, or even five different mental tasks (four different motor imageries and one mental calculation) can increase B. Based on the information transfer, a method is proposed to determine the types of mental tasks and their combination in order to achieve the maximum information transfer rate. For more details about the classification of EEG using hidden Markov models (HMMs), see Obermaier [7].

II. Methods

A. Experiment and Data Acquisition

Subjects: Three male subjects (ages 17–25 years, with the codes S1, S2, and S3) familiar with the Graz BCI [13] took part in the study. All were experienced with the BCI experiment with motor imagery focused on the upper limbs. They were paid for their participation and were free from medication and central nervous system abnormalities.

Procedure: The subjects were sitting in a comfortable armchair looking at the center of a monitor approximately 2 m in front of them. Each trial started with the presentation of a fixation cross at the center of the monitor, followed by a short warning tone at second 2. At second 3, a symbol representing one of the five different types of mental tasks was displayed at the center of the monitor for 1.25 s. These symbols and their corresponding imaginations were an arrow pointing to the left [left-hand movement (L)], an arrow pointing to the right [right-hand movement (R)], an arrow pointing down [foot movement (F)], a simple calculation [repeated subtraction of a constant number from a randomly chosen number (A)], and the picture of a tongue [tongue movement (T)]. At second 4.25, the fixation cross was presented again, lasting until the end of the trial at second 8, and was followed by a blank screen until the beginning of the next trial. The intertrial period (ranging from 0.5 to 2.5 s) as well as the sequence of mental tasks was randomized to avoid adaptation. Every session was divided into four runs containing 50 mental tasks each, so that each task was performed 40 times per session. The subjects had a 15-min break between runs, and sessions were performed on different days. This was to guarantee that the EEG patterns related to the mental tasks were stable over time.

Recordings: The EEG was recorded from 29 gold electrodes, 17 of which were placed according to the international 10–20 system (see Fig. 1). The ground electrode was glued to the forehead. The EEG signals were filtered between 0.5 Hz and 30 Hz and sampled at frequency of 256 Hz. EEG trials containing electromyogram (EMG) or electrooculogram (EOG) artifacts were excluded from the data sets. The number of artifact-free trials are summarized in Table I.

B. Data Preprocessing

The logarithm of the band power for two α bands (7–10 Hz, 10–13 Hz) and three β bands (16–20 Hz, 20–24 Hz, 24–30 Hz) were used as descriptive parameters for the EEG signal calculated for every channel. Frequency components in these bands are predominately involved in motor imagery [14]–[16]. The calculation was performed sample by sample in the time domain using a fifth-order Butterworth filter in a window from seconds 4 to 8 of each trial (see [12]).

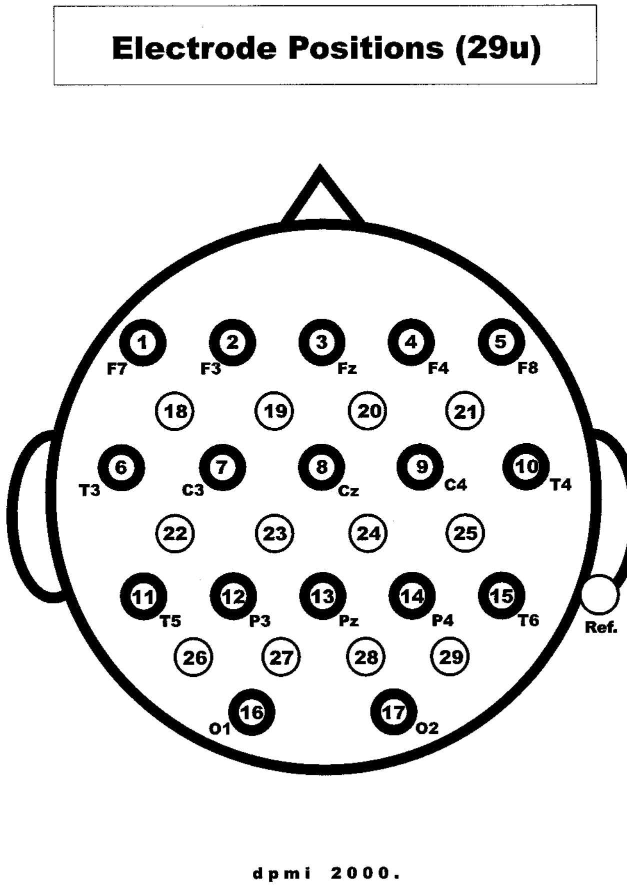

Fig. 1. Positions of the 29 electrodes used for the EEG recording. The electrode positions with bold circles belong to the international 10–20 system. The other positions are inserted in between, in order to increase spatial resolution.

Therefore, the feature vector describing all EEG signals from all electrodes had 145 components. In order to reduce the dimension of $\mathbf{X} = \{x_1, x_2, \ldots, x_{145}\}$, a sub-et of features $\overline{\mathbf{X}} = \{\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_n\}$ was selected based on class separability J [17].

$$J(\overline{X}) = \mathrm{tr}(\mathbf{S}_w^{-1}\mathbf{S}_b) \tag{1}$$

whereby $\mathbf{S}_w$ is calculated by

$$\mathbf{S}_w = \sum_{i=1}^{\text{classes}} (\overline{\mathbf{X}} - \mathbf{M_i})(\overline{\mathbf{X}} - \mathbf{M_i})^{\mathbf{T}}$$

a within-class scatter matrix showing the scatter of feature vectors $\overline{\mathbf{X}}$ around their respective class expected vectors $\mathbf{M_i}$.

$$\mathbf{S}_b = \sum_{i=1}^{\text{classes}} (\mathbf{M_i} - \mathbf{M_0})(\mathbf{M_i} - \mathbf{M_0})^{\mathbf{T}}$$

where $\mathbf{M_0} = E\{\overline{\mathbf{X}}\}$ is the expected vector of the mixture distribution and $\mathbf{M_i} = E\{\overline{\mathbf{X}}_\mathbf{i}\}$ the expected vector of the distribution of class i.

The determination of the subset $\overline{\mathbf{X}}$ was based on a step-by-step forward procedure where the number of parameters n was successively increased by one. The first feature,

TABLE I
AVAILABLE TRIALS FOR ALL FIVE CLASSES AND SUBJECTS AFTER REMOVAL OF TRIALS CONTAINING EMG OR EOG ARTIFACTS

Subject	Left hand (L)	Right hand (R)	Foot (F)	Arithmetic (A)	Tongue (T)
S1	115	113	103	107	107
S2	113	110	106	104	74
S3	94	94	98	93	70

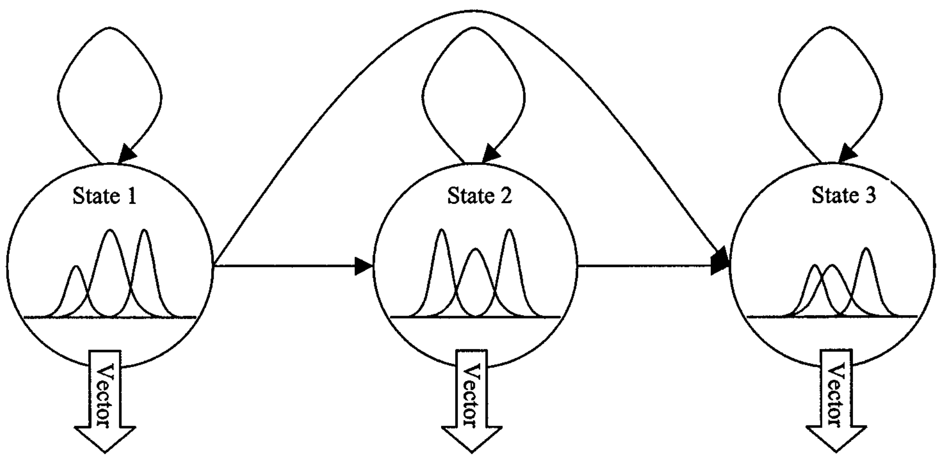

Fig. 2. A hidden Markov model (HMM) comprised of three states s and three Gaussian mixtures m per state, transitions are from left to right states only. The bold arrows indicate the emission of a feature vector at every state.

$\overline{x}_1$, was selected by $x_n|J(x_n) = \max(J(x_n)|_{n=1}^{n=145})$—the feature x_n with the highest separability. In the next step, $\overline{x}_2$ was selected in such a way that the separability of the combination $J(\overline{X} = \{\overline{x}_1, \overline{x}_2\})$ was again maximized for all possible combinations using the remaining features $X/\{\overline{x}_1\}$. This step-by-step procedure terminated when the separability between two successive steps changed less than $\varepsilon \leq 0.05$ or a maximum number of features was reached. The maximum number of 15 was determined in order to guarantee that the classificator (see Section II-C) could be sufficiently trained with the available EEG data.

C. Classification Methods

The basic principles of a hidden Markov model (HMM) will briefly be discussed throughout this section, whereas a detailed description can be found in [18], [19]. The HMM itself could be seen as a finite automata (see Fig. 2) containing s discrete states, emitting a feature vector that depends on the current state at every time point. Each feature vector is modeled using m Gaussian mixtures. The transition probabilities between states are described using a transition matrix.

During the training phase the expectation maximization (EM) algorithm introduced by Dempster [20] was used to estimate the transition matrix and the Gaussian mixtures. Based on randomly selected values for the transition matrix (upper triangle matrix) and an initial estimation of the mixtures, the EM algorithm was performed. The estimation formulae guarantee a monotonic increase of the likelihood $P(\overline{\mathbf{X}}|HMM)$ until a local or global maximum is reached, which finishes the training phase. The maximum number of states s was limited to five, which corresponds to physiological changes in the spatiotemporal patterns in a one second range [21]. The number of mixtures m was limited to eight, according to earlier studies made by the authors [12]. The Gaussian mixtures were approximated based on a k-means clustering of the feature vectors, whereas the number of mixtures corresponds to the number of estimated clusters. The mean and variance of all feature vectors belonging to a cluster were used to estimate one Gaussian mixture with a diagonal covariance matrix. The m Gaussian mixtures were weighted proportional to the number of vectors comprised in that cluster divided by the number of vectors comprised in all clusters of that state [18].

D. The BCI-HMM System

The structure of the BCI-HMM system used for classification was related to the number of different types of mental tasks to be bclassified. Every type of mental task was represented by a HMM trained on the respective data. Classification of an unknown trial was based on the likelihood $P(\overline{\mathbf{X}}|HMM)$ calculated separately for all HMMs given the unknown trial. The unknown trial was labeled with the mental task associated with the HMM of maximum likelihood.

E. Rating of Classifiers

The information transfer rate B (derived from [10]) in bits per trial was calculated by

$$B = \log_2 N + P \log_2 P + (1-P) \log_2 \frac{1-P}{N-1}. \qquad (2)$$

N is the number of different types of mental tasks and P the accuracy of classification. In Fig. 3, B is depicted for the four possible BCI systems ($N = 2$, 3, 4 and 5).

In order to determine whether an increase of the number of different mental tasks in a BCI system also results in an increase

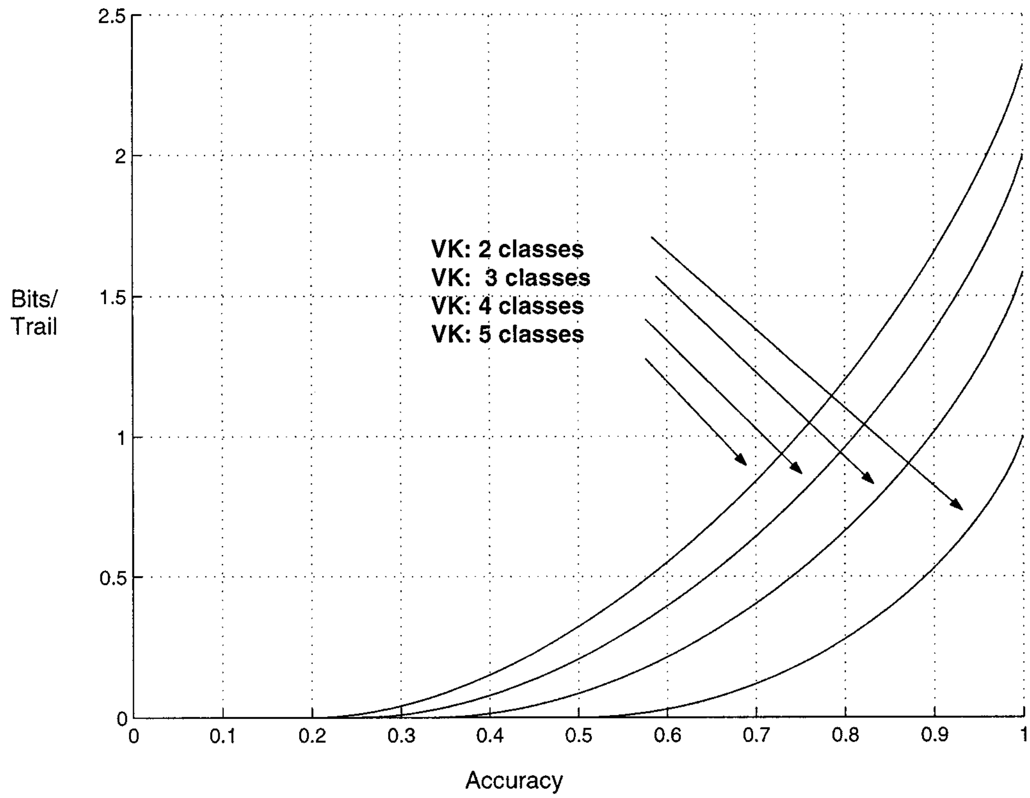

Fig. 3. Information transfer rate B for a BCI system comprised of a different number of mental tasks $N = 2, 3, 4, 5$. The information transfer rate is only shown for an accuracy $\geq 100/N$ (i.e., $\geq$chance).

of B, B was calculated for subsets of $N = 5$ ($N = 2, 3, 4$). The subsets were determined via the following procedure. Based on the separability [see (1)], the features $\overline{\mathbf{X}}$ were selected and used to train an HMM classifier (see Section II-C) comprised of five HMMs representing the five different mental tasks. To evaluate the classification accuracy, a 5×5 cross-validation test was performed. The available data were divided into five sets. Four sets were used to train the classifier and one set was classified. The classification accuracy was determined at the end of the trial (second 8) as the ratio of the number of misclassified trials to the total number of trials to be classified. This procedure was repeated five times so that every set was used as test set once. The overall procedure was repeated five times. Furthermore, a confusion matrix based on the classification accuracy was calculated in order to describe the classification results in more detail. It gives the relation as a percentage of how many times a given mental task was classified correctly and, if not, to which mental class it was assigned. The cross-validation test was repeated for different HMM structures (varying the number of states from one to five, and varying the number of mixtures from one to eight; see Section II-C). The classification accuracy was used to determine the optimal structure of the $N = 5$ classes HMM classifier. The confusion matrix was used to determine the class combinations for the $N = 2, 3, 4$ classes classifier. The two (3, 4) mental tasks having the highest values in the diagonal elements (the mental tasks most distinguishable) in the confusion matrix were combined to the $N = 2$ (3, 4) classes classifier. For the $N = 2, 3, 4$ classes classifier, the procedure of feature selection was performed again (see Section II-B).

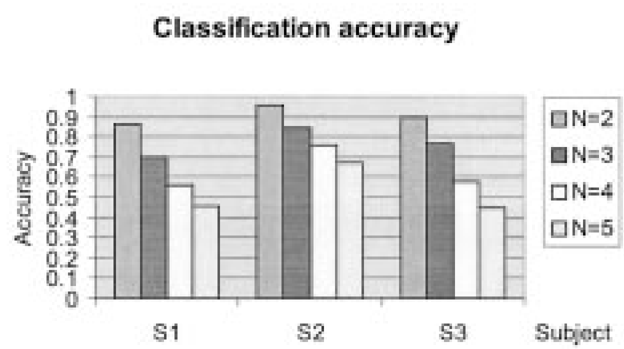

Fig. 4. Classification accuracy for the three subjects and the four types of classifiers ($N = 2, 3, 4, 5$).

III. Results

The classification accuracy for the three subjects and the four different classifiers ($N = 2, 3, 4, 5$) is presented in Fig. 4. All subjects demonstrated a steady decrease of the classification accuracy with an increasing N. For all subjects, the classification accuracy is at a maximum for $N = 2$ ($S1$: 0.8627, $S2$: 0.9609, $S3$: 0.8964) and at a minimum for $N = 5$ ($S1$: 0.4572, $S2$: 0.6718, $S3$: 0.4518).

The information transfer rate B for the three subjects and the four different classifiers ($N = 2, 3, 4, 5$) is presented in Fig. 5.

The information transfer rate B for two subjects (S1 and S3) shows a significant decrease for $N > 3$. For subject S2, B just slightly changes over all N. The maximum B for the subjects

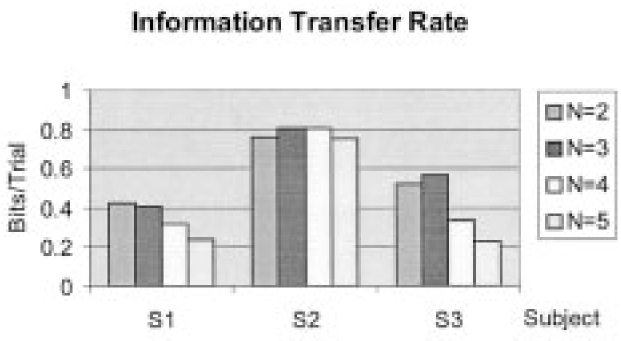

Fig. 5. Information transfer rate for the 3 subjects and the 4 types of classifiers ($N = 2, 3, 4, 5$).

TABLE II
CONFUSION MATRIX FOR SUBJECTS S1, S2, AND S3

Subject / Task	L	R	F	A	T
S1/L	45.4	17.2	12.2	16.2	8.9
S1/R	22.2	26.8	21.0	18.5	11.5
S1/F	16.7	8.5	58.0	5.7	11.1
S1/A	6.5	7.1	6.3	61.0	19.0
S1/T	9.6	7.7	12.7	26.9	42.9
S2/L	68.1	15.6	7.1	6.1	3.0
S2/R	18.4	73.9	2.4	2.9	2.4
S2/F	9.8	4.7	58.4	11.7	15.4
S2/A	8.9	5.2	12.1	57.7	16.1
S2/T	0.8	2.4	15.0	10.4	71.4
S3/L	22.2	26.9	20.1	18.6	12.1
S3/R	24.1	29.0	13.6	20.3	13.0
S3/F	16.4	10.4	55.4	9.3	8.5
S3/A	16.5	23.1	10.0	40.1	10.3
S3/T	6.9	6.3	5.5	1.3	80.0

were $S1$: 0, 42 ($N = 2$), $S2$: 0, 81 ($N = 4$), and $S3$: 0, 56 ($N = 3$).

The confusion matrices for the three subjects are given in Table II. Subject S1 is able to produce two, S2 produces all, and S3 produces two mental tasks with an accuracy above 50%. The maximum accuracies are 61% for S1/A, 73.9% for S2/R, and 80.0% for S3/T.

In Table III, the combinations of different mental tasks (see Section II-E) found for the $N = 2, 3, 4, 5$ classifier are presented. The combinations resulting in the classifier with highest classification accuracies are distinct for all subjects—no identical combinations exist ($N < 5$).

TABLE III
COMBINATIONS OF DIFFERENT CLASSES FOR $N = 2, 3, 4, 5$ (L: LEFT MOTOR IMAGERY; R: RIGHT MOTOR IMAGERY; F: FOOT MOTOR IMAGERY; A: ARITHMETIC TASK; T: TONGUE MOTOR IMAGERY)

Subject / N	S1	S2	S3
2	A-F	R-T	T-F
3	A-F-L	R-T-L	T-F-A
4	A-F-L-T	R-T-L-A	T-F-A-L
5	A-F-L-T-R	R-T-L-A-F	T-F-A-L-R

IV. DISCUSSION

The classification accuracy achieved for the three subjects demonstrated a common trend for all subjects (see Fig. 4). The highest classification accuracy was achieved classifying two classes ($N = 2$); an increasing number of classes resulted in a decrease of the classification accuracy. The decrease of accuracy between two adjacent classifiers ($N = n$ and $N = n+1$ ($n = 2, \ldots, 4$) for all classifiers and subjects is 12, $6\% \pm 3.5\%$. The decrease of classification accuracy is caused by the fact that every classification of a mental task is afflicted with a probability of misclassification due to overlapping features in the feature space. Increasing the number of mental tasks, therefore, increases the areas of overlap and decreases the classification accuracy.

The information transfer rate B (see Fig. 5) shows a high intersubject variability, and also a high intrasubject variability, for the subjects (S1 and S3). The maximum values for B vary from $0.42(S1, N = 2)$ to $0.81(S2, N = 4)$, although the classification accuracy does not demonstrate such high variability. This demonstrates the importance of accuracy in a BCI system, e.g., the accuracy difference of $(S1, N2)$ and $(S2, N2)$ is about 10%, whereas the difference for B is 34%. In other words, a slight improvement in classification accuracy can result in a significant improvement in the information transfer rate.

The intrasubject variability for subjects S1 and S3 is due to the decrease of B, for $N \geq 4$. This is caused by the additional types of mental tasks that cannot be distinguished from the others. For S1/R, S3/L, and S3/R (see Table II), the classification accuracies of the respective mental task differ just slightly from the remaining ones. Even though this is not the case for S1/T, the information transfer decreases after including S1/T. This demonstrates the problem of deciding whether a further mental task is added to a BCI or not solely based on the confusion matrix. For subjects S1 and S3, the number of mental tasks resulting in distinctive EEG patterns is limited to $N = 3$. A further increase of N decreases the information transfer rate, except for subject S2. This subject is able to produce up to five different mental task related EEG patterns. The classification accuracy decreases but the information transfer rate changes just slightly. The classification accuracies of the respective mental task differ a minimum of 40% to the closest classification rate of the remaining mental tasks (see Table II).

The combinations of different mental tasks (see Table III) reveal no common structures and differ from subject to subject. It is interesting to note that even though all subjects were familiar with the previous BCI paradigm, in which the task was to imagine a left- or a right-hand movement, none of the subjects showed a preference for these two tasks for the $N = 2$ classifier. This result is interesting in two ways: i) the already trained patterns of motor imagery do not result in maximum B and ii) no general conclusion can be made about which mental task results in the most distinctive EEG patterns. The optimal tasks have to be evaluated from subject to subject. This should be considered in future designs of a BCI system.

The proposed method allows us to find an optimal combination of mental tasks so that a BCI system achieves the maximum information transfer rate B. Based on the results and the experimental procedure, the use of three mental tasks can be recommended for a EEG-based BCI system. The main effort in future work should focus not only on the search of suitable mental tasks, but also on improving the classification methods [13], [5], because even a slight increase of the classification accuracy can have a great influence on the information transfer rate.

ACKNOWLEDGMENT

The authors wish to thank G. Müller and G. Schweizer for their suggestions.

REFERENCES

[1] N. Birbaumer, N. Ghanayim, T. Hinterberger, I. Iversen, B. Kotchoubey, A. Kübler, J. Perelmouter, E. Taub, and H. Flo, "A spelling device for the paralyzed," *Nature*, vol. 398, pp. 297–298, 2000.

[2] A. Kübler, B. Kotchoubey, H. P. Salzmann, N. Ghanayim, J. Perelmouter, V. Hörnberg, and N. Birbaumer, "Self-regulation of slow cortical potentials in completely paralyzed human patients," *Neurosci. Lett.*, vol. 252, pp. 171–174, 1998.

[3] G. Pfurtscheller, Ch. Neuper, A. Schloegl, and K. Lugger, "Separability of EEG signals recorded during right and left motor imagery using adaptive autoregressive parameters," *IEEE Trans. Rehab. Eng.*, pp. 316–325, Sept. 1998.

[4] Ch. Guger, H. Ramoser, and G. Pfurtscheller, "Real-time EEG analysis with subject-specific spatial patterns for a brain-computer interface (BCI)," *IEEE Trans. Rehab. Eng.*, vol. 8, pp. 447–456, Dec. 2000.

[5] B. Obermaier, C. Guger, and G. Pfurtscheller, "Online classification of single trial EEG data using hidden Markov models," in *Proc. RECPAD2000*, 2000, pp. 251–255.

[6] J. Kalcher, D. Flotzinger, Ch. 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. Comp.*, vol. 34, pp. 382–388, 1996.

[7] B. Obermaier, "Design and implementation of an EEG based virtual keyboard using hidden Markov models," Ph.D. dissertation, Univ. Technology, Graz, Austria, 2001.

[8] C. W. Anderson, E. A. Stolz, and S. Shamsunder, "Multivariate autoregressive models for classification of spontaneous electrencephalogramm during mental tasks," *IEEE Trans. Biomed. Eng.*, vol. 45, pp. 277–286, Mar. 1998.

[9] B. Obermaier, G. Müeller, and G. Pfurtscheller, "Virtual keyboard controlled by spontaneous EEG activity," presented at the ICANN 2001, Vienna, Austria, 2001.

[10] C. E. Shannon and W. Weaver, *Mathematical Theory of Communication*. Champaign, IL: Univ. Illinois Press, 1964.

[11] J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFaralnd, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, and T. M. Vaughan, "Brain–computer interface technology: A review of the first international meeting," *IEEE Trans. Rehab. Eng.*, vol. 8, pp. 164–173, June 2000.

[12] B. Obermaier, Ch. Guger, and G. Pfurtscheller, "Hidden Markov models used for the offline classification of EEG data," *Biomedizinische Technik*, vol. 44, no. 6, pp. 158–162, 1999.

[13] G. Pfurtscheller, C. Neuper, C. Guger, W. Harkam, H. Ramoser, A. Schlögl, B. Obermaier, and M. Pregenzer, "Current trends in graz brain-computer-interface (BCI) research," *IEEE Trans. Rehab. Eng.*, vol. 8, pp. 216–219, June 2000.

[14] G. Pfurtscheller and Ch. Neuper, "Motor imagery activates primary sensorimotor areas," *Neurosci. Lett.*, vol. 239, pp. 65–68, 1997.

[15] G. Pfurtscheller, Ch. Neuper, D. Flotzinger, and M. Pregenzer, "EEG-based discrimination between imagination of right and left hand movement," *Electroenc. Clin. Neurophys.*, vol. 103, no. 5, pp. 1–10, 1997.

[16] C. Neuper and G. Pfurtscheller, "Motor imagery and ERD," in *Handbook of Electroencephalography and Clinical Neurophysiology, Event-Related Desyncronization*, revised ed., G. Pfurtscheller and F. H. L. da Silva, Eds., 1999, vol. 6, pp. 303–325.

[17] K. Fukunaga, *Introduction to Statistical Pattern Recognition*. New York: Academic, 1990.

[18] L. Rabiner and B. H. Juang, *Fundamentals of Speech Recognition*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[19] J. R. Deller, J. G. Proakis, and J. H. L. Hansen, *Discrete-Time Processing of Speech Signals*. New York: Macmillan, 1993.

[20] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Royal Statist. Soc.*, vol. 39, no. 1, pp. 1–38, 1977.

[21] A. Schlögl, C. Neuper, and G. Pfurtscheller, "Subjects specific EEG patterns during motor imagery," in *Proc. 19th Int. Conf. IEEE/EMBS*, 1997.

Bernhard Obermaier was born in 1972. He received the Ph.D. degree in biomedical engineering from the Graz University of Technology, Graz, Austria, in 2001.

Currently, he is a Scientific Member at the Ludwig Boltzmann Institute of Medical Informatics and Neuroinformatics, University of Technology. His research interests include the real-time feature estimation and classification of EEG data in the environment of a brain–computer interface.

Christa Neuper received the Ph.D. degree in psychology from the University of Graz, Austria, in 1984.

Currently, she is a Research Assistant at the Ludwig Boltzmann Institute of Medical Informatics and Neuroinformatics, Graz University of Technology, Graz, Austria. Her research interests include the study of brain oscillations, their functional meaning, and neuropsychological applications.

Christoph Guger (S'97–A'00) was born in 1973. He received the M.Sc. degree in electrical engineering and the Ph.D. degree in biomedical engineering from Graz University of Technology, Graz, Austria, in 1997 and 1999, respectively.

Currently, he is an Assistant Professor at the Institute of Biomedical Engineering, Graz University of Technology. In 1999, he founded the medical engineering company G.TDEC. His research interests include real-time processing systems and brain–computer communication systems.

Gert Pfurtscheller (M'92) received the M.Sc. and Ph.D. degrees in electrical engineering from Graz University of Technology, Graz, Austria, in 1964 and 1966, respectively.

He is a Professor of Medical Informatics at the Institute of Biomedical Engineering, Graz University of Technology. Since 1982, he has been Head of the Deparment of Biomedical Informatics and since 1987, he has been 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 include biomedical signal processing with special emphasis on electrical brain activity analysis, brain–computer communication systems, and medical expert systems.