A Multiclass BCI using MEG

M. Bensch[1], J. Mellinger[2], M. Bogdan[3], W. Rosenstiel[1]

[1]Wilhelm-Schickard Institute, Department of Computer Engineering,
Eberhard-Karls-Universität, Tübingen, Germany
[2]Institute of Medical Psychology and Behavioral Neurobiology, MEG Center,
Eberhard-Karls-Universität, Tübingen, Germany
[3]Department of Computer Engineering, Universität Leipzig, Leipzig, Germany
bensch@informatik.uni-tuebingen.de

Abstract

Classification of MEG brain signals has been used in the past to drive brain-computer interfaces with healthy subjects. We present a multi-class study with offline results. We investigate which combinations of classes are suitable for classification and show that this result is subject-specific, although a general trend to particular combinations of patterns exists. For the current set of offline results, information transfer rate increases in eight of ten subjects going from 2-class to 3-class communication.

1 Introduction

Brain-Computer Interface (BCI) research is largely concerned with increasing the communication speed, or bit rate. Attempts to increase the bit rate of BCIs based on spontaneous brain activity by increasing the number of classification tasks naturally suffer from a higher misclassification rate. In an experiment with 10 subjects and 7 classification tasks, we find the best trade-off concerning the number of tasks for each subject.

High signal quality and sensor resolution inspired us to use MEG. Other BCIs based on MEG using binary classification have already been presented [1], [2]. The work presented here is a multi-class approach.

An extensive multi-class investigation is performed in [3]. A 3-class asynchronous BCI using three motor tasks for spelling was shown to produce an average of 1.99 letters/minute [4], but is difficult to compare with synchronous BCIs. A study using two motor and two non-motor imagery tasks concludes that non-motor task pairs can be discriminated easier than others [5].

The following sections explain our experimental setup and artifact analysis. Thereafter, a summary of the results obtained so far is given. The aim of this work is to find combinations of imagination tasks particularly well-suited to multi-class BCIs.

2 Methods

2.1 Experimental paradigm and data acquisition

Data was recorded with the BCI2000 software [6] at a sampling rate of 586 Hz from 10 subjects aged between 24 and 34 on two different days for each subject. Each of the two sessions included 3 runs of data acquisition, with 7 imagination tasks being presented in a block-randomised order. Cues were given in textual form on a screen positioned in front of the seated subject. Subjects were instructed to focus on a cross in the centre of the screen during the imagination phase to prevent movement artifacts. The trial structure is as follows: display of a fixation cross (“get ready”, 0.7 s)

Category	Class	Task	Description
Motor	1	Foot	Rotate both feet
Motor	2	Left hand	Open and close hand in repetitive motions
Motor	3	Right hand	Open and close hand in repetitive motions
Motor	4	Tongue	Lick ice-cream with tongue in repetitive motions
Non-motor	5	Subtraction	Start at 99 and repeatedly subtract 7 from answer. Do not visualise numbers, do the calculation each time. Start at a new number once the sequence has been memorised too well.
Non-motor	6	Navigation	Walk around a well-known location (e.g. the house you grew up in). Recognise objects in the rooms.
Non-motor	7	Visual scene	Imagine a green scene, e.g. a lawn or green landscape.

Table 1: Categorisation of the imagination tasks.

Subject	BCI experience	EMG measured	Handedness	Hardest task
A	0	Session 1	RH	5
B	0	Session 2	RH	-
C	0	-	RH	-
D	4	Session 1	RH	7
E	0	Session 2	RH	6
F	0	Session 1	RH	7
G	0	Session 2	RH	7
H	0	-	RH	-
I	>5	Session 1	RH	-
J	0	Session 2	LH	5,6

Table 2: Subject details. The second column indicates the number of times the subject had participated in previous BCI experiments. The "Hardest task" column lists the task that the subject reported as the most difficult to perform.

- text cue overlayed on fixation cross ("prepare for task", 2 s) – fixation cross ("mental imagery", 4 s) – blank screen ("inter-trial interval", 2 s). Due to the exploratory nature of this study, no feedback was given. A 275-channel whole-head MEG-system (VSM MedTech Ltd.) was used.

We chose 7 imagination tasks as a trade-off between high number of classes (to test which tasks work best for the participants) and being able to collect enough data for each task to train a classifier. With 7 tasks we were able to collect 102 trials per class and keep each recording session around two hours. Each of the three runs per session lasted 17 minutes, enough to make most subjects feel tired after every run. A description of the imagination tasks is given in Table 1. As suggested in [5], functional motor tasks (some related to a subject's specific skills) were endorsed to make the task more interesting.

Some subject-specific details are given in Table 2. The only female subject was F. Note the high number of first-time BCI users. EMG was measured for 8 subjects in only one of the two recording sessions due to the increased preparation time. Head position was measured before each run. To recreate the previous session's or run's head position, the subject was given the opportunity to reposition his/her head according to an online head position display (CTF Systems).

2.2 Artifact rejection

Artifacts could be a major contributor to the fact that healthy subjects mostly achieve higher bit rates than patients when using BCIs. We precluded the outer MEG channels close to eyes and neck muscles and focussed on the inner 168 channels. Additionally, we investigated the effect that trials containing electromyographic (EMG) hand movement artifacts had on classification

Subject	Total per run			1	2	3	4	5	6	7	Total
A	20	11	21	8	4	8	7	8	8	9	52
B	7	14	3	3	4	2	4	2	4	5	24
D	7	9	5	0	16	2	0	2	1	0	21
E	1	6	8	0	2	2	3	3	3	2	15
F	1	1	6	1	0	0	1	0	3	3	8
G	4	7	5	1	0	1	3	5	6	0	16
I	7	16	9	7	4	5	2	7	2	5	32
J	13	12	7	7	2	4	1	5	5	8	32
Mean	60	76	64	27	32	24	21	32	32	32	200

Table 3: Number of artifacts, listed separately for runs 1–3 of the EMG session (columns 2–4) and for each task. Note the tasks 2 and 3 are the hand motor imagery tasks.

performance.

In the sessions shown in Table 2, one pair of EMG electrodes was placed on each forearm on either the lateral or medial antibrachial muscle. The EMG time series was high-pass filtered at 0.5 Hz.

To find trials containing EMG artifacts, we used a threshold-based algorithm on the EMG time series. Because the noise level of the EMG signal varied with time (due to stress on electrode cables and other effects), we used a sliding window to determine the noise level for each trial separately and standardised it accordingly. The window size was 250 samples (0.43 s) and the step size was 50 samples. Each window's score was computed as the average of the 10 highest peaks. The window with the lowest value was assumed to be free of artifacts and therefore used as noise level for the scaling. If the maximum peak in the standardised signal of a particular trial was further than n standard deviations from the mean, the trial was labelled as an artifact. After a visual inspection of some artifacts, n was set to 3 for most subjects.

The contaminated trials found by this method are listed in Table 3. Refer to Table 2 to see in which sessions the EMG was measured.

2.3 Feature extraction and classification

Multiple cross-validation runs were done with the following features extracted from the data: Autoregressive (AR) coefficients (model order 2) exclusively, spectral bandpower features exclusively, phase locking features exclusively, AR coefficients combined with bandpower features, AR coefficients combined with phase locking features, bandpower features combined with phase locking features.

The feature selection algorithm worked independently of the type of feature in those validations where multiple feature types were used. Using all the 168 inner channels, there were 336 AR coefficients in total. The mean number of features selected was 7.7 ($\pm$5.4) in the 2-class case and 44 ($\pm$27) in the 3-class case.

The best results were obtained by using solely the AR features and will be presented in the next section.

We performed a nested cross-validation (CV) with a feature selection in the inner loop (recursive feature elimination, 10-fold CV) and 5 outer folds, similar to the method described in [1]. The outer fold consisted of a random split of the data into 80% train and 20% test set. To obtain 3-class results, we used the MATLAB spider toolbox [7] implementation of the one-vs-all SVM with ridge regularisation.

3 Results

The information transfer rate (ITR) measured in bits/minute (we use the Wolpaw bit rate [8]) for all the binary combinations of the 7 classification tasks are shown in Figure 1.

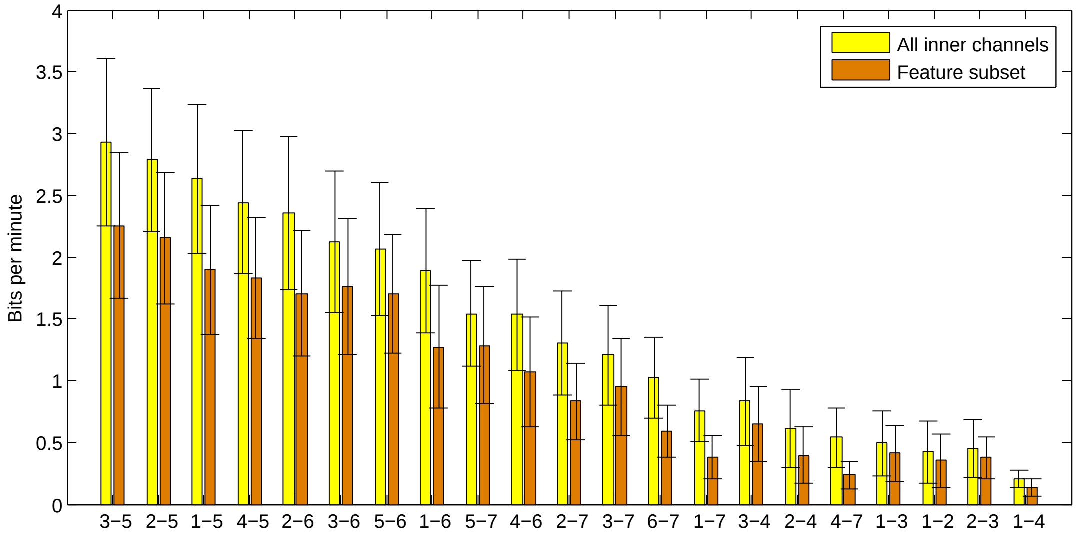

Figure 1: ITR in bits/minute (mean over 10 subjects) for each 2-class combination shown for all features and a feature subset. The error bars depict standard error. Bars are labelled with the class combination.

Subject-specific results for the 3-class and 2-class error estimates can be seen in Table 4. The column description "Inner channels" refers to the fact that we did not include the outermost MEG sensors. We do not expect the outer MEG sensors to contribute class-discriminative information, except possibly for artifacts, which we wanted to preclude anyway. The feature selection is needed for the online case, where processing speed is important.

	3-class Inner channels		3-class Subset of features			2-class Inner channels		2-class Subset of features		
Subject	Error	Comb.	Error	Comb.	ITR	Error	Comb.	Error	Comb.	ITR
A	0.34	1-5-6	0.40	3-5-6	1.50	0.19	3-6	0.23	3-6	1.53
B	0.16	2-5-6	0.23	1-2-6	4.00	0.08	2-6	0.11	3-5	3.37
C	0.29	1-5-7	0.29	1-5-7	2.98	0.15	2-5	0.21	2-5	1.77
D	0.32	2-3-6	0.39	3-5-6	1.61	0.22	3-5	0.24	3-5	1.41
E	0.25	2-5-7	0.23	2-5-6	3.94	0.10	1-5	0.14	2-5	2.90
F	0.23	4-5-6	0.23	1-5-6	3.94	0.10	3-5	0.15	3-5	2.65
G	0.16	1-5-6	0.21	4-5-6	4.39	0.06	2-5	0.10	2-5	3.68
H	0.16	1-5-6	0.20	2-5-6	4.70	0.05	3-5	0.09	5-7	3.90
I	0.48	2-3-5	0.50	3-5-6	0.59	0.26	3-5	0.31	2-5	0.75
J	0.09	3-5-6	0.12	3-5-6	6.47	0.03	3-5	0.06	3-6	4.64
Mean	0.25		0.28		3.41	0.12		0.16		2.66

Table 4: Results for the best combination of three classes per subject, in terms of classification error (and ITR when applying feature selection). The error estimate is by 5-fold cross-validation. The 2-class results are given as a comparison.

Removal of artifact-trials increased the offline classification error to 0.11 (subject J) and 0.35 (subject D) using all inner channels, and to 0.125 (subject J) and 0.41 (subject D) with feature selection.

A comparison with other multi-class BCI results is shown in Table 5. The results shown for our study are after feature selection (which we require for online operation of the BCI). The error using all inner channels is slighty lower.

Study	Classes	ITR: bits per trial worst	best	mean	ITR: bits per min worst	best	mean	N
Dornhege et al. [3]	3	0.5	1.15	0.78	6.7	15.3	10.32	5
Dornhege et al., feature comb.	3	0.6	1.19	0.92	8.0	15.9	12.2	5
This study	3	0.09	0.95	0.5	0.59	6.47	3.41	10

Table 5: Comparison of 3-class results. The three columns for the bit rates are worst/best/mean result of the subject population, whose size is given in the final column. Each subject's best combination is regarded.

4 Discussion

The analysis of artifacts showed that, with the exception of subject D, the EMG hand movement artifacts were evenly distributed across all tasks. This implies that hand movements did not unfairly bias the classifier during the offline analysis. However, we cannot rule out the possiblity that eye or head movement artifacts could have influenced the result. Furthermore, there is no guarantee that the subject did not move his/her tongue or feet during the recording.

The error estimates using feature selection are higher than without feature selection in the 2-class as well as the 3-class setting (Table 4). Feature selection is needed for the online BCI to reduce the computational load, especially if combinations of various feature types are to be used. The ITR estimates with feature selection are higher in the 3-class setting for all except the worst two subjects. An online spelling session was performed with one subject up to now.

The task combinations shown in Table 4 are the best on a per-subject basis. It is interesting to see that for most subjects, a motor task combined with the two non-motor tasks "subtraction" and "navigation" was the best combination (even though subjects reported difficulties with the non-motor mental tasks, see Table 2). This might have to do with the fact that cortical activation for these tasks is spatially far apart - we plan to investigate this in more detail. The best 2-class pairs were "subtraction" together with a motor task in eight of ten cases. This is in contrast to [5], where the most discriminable pair consisted of two non-motor tasks. The difference could be explained by the fact that (1) we did not include the "auditory" task, or (2) we allowed functional motor tasks (subjects A, B, F, H, J used functional motor tasks for the hand imagery). Our results suggest that some patterns of class combinations exist that generally lead to a higher discriminability (motor/non-motor task combination).

We acknowledge the fact that measuring MEG in people from the current target group for BCIs, namely paralyzed patients, is expensive and time-consuming. However, the focus of this study was to explore the multi-class paradigm. Furthermore, future BCIs may be based on a contactless measuring technique similar to MEG.

5 Conclusion

The bit rates found for the 3-class setting presented in this study are an improvement over the 2-class results, yet must still be compared with other studies. Additionally, a comparison with existing 3-class BCIs based on motor imagery alone has to be undertaken. This could show whether there is an advantage to using non-motor tasks combined with motor tasks, as suggested by the results presented here.

We found that it is possible to select a subset from a larger group of mental tasks individually to obtain a higher bit rate for each BCI user. In the offline analysis, eight subjects benefit by using a 3-task combination instead of binary decision tasks. The mean improvement in the ITR for these eight subjects is 32% when using 3 classes instead of 2. This motivates us to continue the online measurements where the subjects spell a word by selecting letters from a ternary decision tree.

6 Acknowledgements

The authors thank J. Hill for providing analysis software and M. Spüler for his assistance in preparing the results and programming the online interface.

References

[1] T. N. Lal, M. Schröder, N. J. Hill, H. Preissl, T. Hinterberger, J. Mellinger, M. Bogdan, W. Rosenstiel, T. Hofmann, N. Birbaumer, and B. Schölkopf. A brain computer interface with online feedback based on magnetoencephalography. In *ICML*, pages 465–472, 2005.

[2] J. Mellinger, G. Schalk, C. Braun, H. Preissl, W. Rosenstiel, N. Birbaumer, and A. Kübler. An MEG-based Brain-Computer Interface (BCI). *NeuroImage*, 36:581–593, 2007.

[3] G. Dornhege, B. Blankertz, G. Curio, and K. Müller. Boosting bit rates in non-invasive EEG single-trial classifications by feature combination and multi-class paradigms. *IEEE Transactions on Biomedical Engineering*, 51(6):993–1002, 2004.

[4] R. Scherer, G. R. Müller, C. Neuper, B. Graimann, and G. Pfurtscheller. An asynchronously controlled EEG-based virtual keyboard: Improvement of the spelling rate. *IEEE Transactions on Biomedical Engineering*, 51(6), 2004.

[5] E. Curran, P. Sykacek, M. Stokes, S. Roberts, W. Penny, I. Johnsrude, and A. Owen. Cognitive tasks for driving a brain computer interfacing system: a pilot study. *IEEE Transactions on Rehabilitation Engineering*, 12(1), 2003.

[6] G. Schalk, D. McFarland, T. Hinterberger, N. Birbaumer, and J. Wolpaw. BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Transactions on Biomedical Engineering*, 51:1034–1043, 2004.

[7] J. Weston, A. Elisseeff, G BakIr, and F. Sinz. The spider machine learning toolbox. http://www.kyb.mpg.de/bs/people/spider/, 2005.

[8] J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, 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 Transactions on Rehabilitation Engineering*, 8:164–173, 2000.