Towards self-paced (asynchronous) Brain-Computer Communication: Navigation through virtual worlds

Reinhold Scherer, Felix Lee, Alois Schlögl, *Member, IEEE,* Robert Leeb, *Student Member, IEEE,* Horst Bischof, *Member, IEEE* and Gert Pfurtscheller *Member, IEEE*

***Abstract*— The self-paced (or asynchronous) control paradigm enables users to operate Brain-Computer Interfaces (BCI) in a more natural way: no longer the machine is in control of timing and speed of communication, but the user. This is important to enhance the usability, flexibility and response time of a BCI.**

In this work, we show how subjects, after performing a cue-based feedback training (smiley paradigm), learned to navigate self-paced through the "freeSpace" Virtual Environment (VE). Similar to computer games, subjects had the task of picking up items within a limited time period using the following navigation commands: rotate left, rotate right and move forward (3-classes). Since the self-paced control paradigm allows subjects to make voluntary decisions on time, type and duration of mental activity, no cues or routing directives were presented. The BCI was based on three bipolar electroencephalogram (EEG) channels and operated by motor imagery. Eye movements (electrooculogram, EOG) and electromyographic (EMG) artifacts were reduced and detected on-line. Results of three able-bodied subjects are reported and problems emerging from asynchronous control are discussed.

***Index Terms*— Brain-Computer Interface (BCI), Electroencephalogram (EEG), Motor imagery, Classification, Virtual Reality (VR), asynchronous operation mode**

I. INTRODUCTION

FOR severely paralyzed people, or patients in a "locked-in" state, direct brain-computer interaction represents one possibility to reestablish communication. A Brain-Computer Interface (BCI) recognizes voluntary changes in ongoing electrophysiological signals and translates different brain states into appropriate commands for communication and control. For a review see [1], [2], [3]. In this work the BCI terminology and definitions proposed in [4] were adopted.

Important aspects for BCI control paradigms are (i) the temporal availability of the system for the user, either periodically, requiring a cueing or synchronization mechanism, or continuously and (ii) the ability to support "non-control" (NC) [4]. NC, in contrast to intentional-control (IC), describes the ability of the system not only to discriminate between several voluntarily modulated brain patterns (IC), but also to detect that the user does not require IC. Nowadays the majority of BCI systems are operated using a synchronized (or cue-based) paradigm. Timing and speed of communication are preset by the paradigm. The most natural way of human-machine interaction is the self-paced mode. The BCI system is continuously analyzing the ongoing brain activity and can handle NC. Recently an increasing number of papers started to address the issues of constantly-engaged and self-paced BCIs [5], [6], [7], [8], [9], [10]. Another very important issue for BCI control, especially in early training sessions [11], is the automatic removal or detection of artifacts. Signals with origin other than in the central nervous system (e.g. muscle activity) must not have any influence on the BCI output.

R. Scherer, R. Leeb and G. Pfurtscheller are with the Laboratory of Brain-Computer Interfaces, Institute for Knowledge Discovery, Graz University of Technology, Krenngasse 37, 8010 Graz, Austria, E-mail: reinhold.scherer@tugraz.at.

F. Lee and H. Bischof are with the Institute for Computer Graphics and Vision, Graz University of Technology, Inffeldgasse 16, 8010 Graz, Austria.

A. Schlögl is with the Institute for Human-Computer Interfaces, Graz University of Technology, Krenngasse 37, 8010 Graz, Austria.

The Graz-BCI is based on the analysis and classification of sensorimotor electroencephalogram (EEG) patterns generated during imagination of specific movements (motor imagery, MI; e.g. hand or foot) [12]. In this work we present the new two-classifier based Graz-BCI designed for self-paced applications. The first of the two classifiers, CFR1, is set-up to discriminate between different MI tasks; the second classifier, CFR2, is trained to detect any MI-related brain activity in the ongoing EEG. By combining the results of both we create a system able to discriminate between several MI-modulated mental brain states (IC) and NC. As a first self-paced application participants had to navigate through the "freeSpace" Virtual Environment (VE) by using the following control commands: Rotate left, rotate right and move forward. Participants were placed into the "freeSpace" virtual park, presented in 2-D on a conventional computer screen, with the task of picking up items within a limited time. Results of three able-bodied subjects are reported.

II. METHODS

A. From externally-paced (or cue-based) to self-paced operation protocol

Subjects and BCI were trained (mutually adapted) in a three-step procedure. First, twenty-two monopolar EEG channels were recorded while subjects were performing a 4-class MI training without feedback (synchronized paradigm). Since one philosophy (endeavor) of the Graz-group is to reduce the number of channels, 3 bipolar channels were derived with best discrimination power between three of the four MI tasks (feature selection). CFR1 was trained with the obtained features and 3-class feedback training was performed (synchronized paradigm). Second, as soon as the 3-class classification accuracy was higher than 75% for each subject, classifier CFR2 was trained to classify between the enhanced MI brain patterns and non-control (NC). NC consisted of recordings were subjects were instructed to sit relaxed and perform "non-MI" mental activities. To evaluate the performance of CFR2

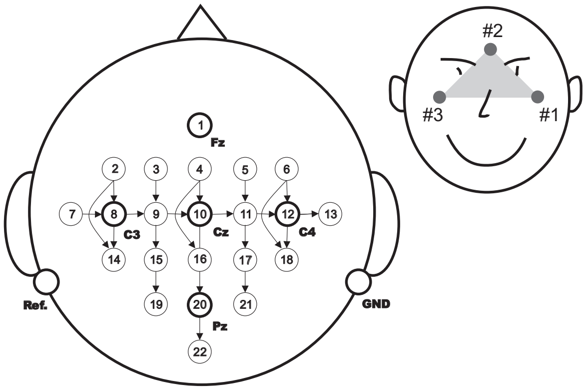

Fig. 1. EEG and EOG electrode placement. For both, reference was placed on the left and ground on the right mastoid. The arrows between EEG electrodes show the analyzed bipolar derivations ($\oplus \rightarrow \ominus$).

feedback training with longer inter-trial periods, representing longer periods of NC, were made. The third and last step was to train subjects to navigate self-paced through the ”freeSpace“ VE and to fine tune CFR2.

B. Subjects and data acquisition

Three healthy subjects (2 male, 1 female, right handed, age 24.1±1.9 years) participated in this study. Prior to the study presented in this work, subjects participated in prior experiments [13] and learned to operate the 2-class basket paradigm [14]. After 3 feedback training sessions the achieved classification accuracies were 71.4% for subject v4, 82.8% for v9 and 86.4% for x6.

Two different electrode arrangements were used. For the training without feedback twenty-two monopolar EEG channels (Ag-AgCl electrodes, extended 10-20 system, reference left mastoid, ground right mastoid) were recorded; Feedback experiments were performed using three bipolar EEG channels (ground position Fz) only. Additionally three electrooculogram (EOG) channels were acquired (see Fig.1). The signals were amplified, analog filtered between 0.5 and 100 Hz and recorded with a sample rate of 250 Hz.

C. Signal processing

1) The Graz-BCI: The Graz-BCI (rtsBCI [15]) works on a sample-by-sample basis. This means that feature extraction and classification is performed with the same rate of the data acquisition (250 Hz). This implies that the analysis performed and the results reported in this work are sample-by-sample based. The feedback is presented at a rate of 25 Hz (screen update).

Band power (BP) features were estimated from the ongoing EEG by digitally bandpass filtering (5^{th} order Butterworth IIR filter), squaring and averaging (moving average) the samples over the 1-s just passed. Consequently for each subject the only parameter to determine was the bandpass frequency band for a selected number of BP features. For classification Fishers linear discriminant analysis (LDA) was applied to the logarithm of the BP estimates.

An automated correction method was used to reduce the influence of EOG artifacts [16], [17]. It can be assumed that the recorded EEG Y is a superposition of the real EEG signal S and three spatial EOG components N (horizontal, vertical and radial) weighted by some coefficient b (accordingly $S = Y - N \cdot b$). The EOG was recorded with three monopolar electrodes, from which two bipolar EOG channels (covering the horizontal and the vertical EOG activity) were derived (see Fig. 1). The weighting coefficients were computed from a 1 minute recording where subjects were asked to perform eye movements. This eye movement recording was used to calculate the autocorrelation matrix C_{NN} of the bipolar EOG channels and the cross-correlation C_{NY} between recorded EEG Y and EOG N. The correction coefficients were obtained by $b = C_{NN}^{-1} C_{NY}$ (see [16] for more details).

We applied the method of inverse filtering [18] by estimating the autoregressive parameters of a 2-minute EEG segment without artifacts to detect muscle activity [17]. If an EMG artifact is superimposed to the EEG, the root-mean-square (RMS) of the inversely filtered process increases. Each time the increase exceeded the detection threshold of $5 \cdot RMS$ from artifact-free EEG, a warning message (marker) was presented for 1s on the screen. Subjects were instructed to relax (loosen the musculature) in order to make the warning disappear and continue with their task.

At the beginning of each feedback session a recording of approximately 4 minutes was performed (initial recording). Two minutes of EEG/EOG with eyes open, 1-minute with eyes closed and 1-minute with eye artifacts were recorded. During the first three minutes subjects were instructed to sit relaxed and to perform only a minimum of movements (eyes, swallowing, ...). For the last minute subjects were instructed to perform repetitively eye blinks, clockwise and counter-clockwise rolling of the eyes, repetitively horizontal and repetitively vertical eye movements for 15s respectively. The eyes should circumscribe the whole filed of view without moving the head. No limitations on the mental activity (thinking) was imposed. Written instructions were presented to the subjects on a computer screen for the duration of each individual task. At the beginning and end of each task a low and high warning tone was presented, respectively. The first two minutes were used to compute the inverse-filter coefficients and segment with the eye artifacts to set-up the EOG correction-coefficients.

2) Distinction Sensitive Learning Vector Quantization: For the selection of the most informative BP features, bipolar channel derivations and MI tasks, the Distinction Sensitive Learning Vector Quantization (DSLVQ), an extended version of Kohonens Learning Vector Quantization algorithm (LVQ), was used [19]. LVQ uses a reduced number of codebook vectors (labeled reference vectors) to approximate the optimal Bayesian decision borders between different classes. Each sample is classified to the label of its closest codebook vector according to a distance function (e.g. Euclidean distance); the influence of the features on the distance function is equal. DSLVQ introduces a weighted distance function which rates the influence of the features for classification: most informative features are upgraded, features that contribute to misclassification are discarded. The LVQ codebook splits the classification problem into sub-problems. By finding an optimal linear approximation for the sub-problems, the relevance of the features, which determines the correct classification, is analyzed [20]. The major advantage of DSLVQ is that it does not require expertise nor any a priori knowledge or assumption about the distribution of the data. Furthermore, not only relevant features, but also feature combinations, are identified.

In order to obtain a reliable feature relevance RT_n [19], the DSLVQ method was repeated 100 times (3 codebook per class, type C training, 10.000 iterations, learning rate decreased from $\alpha = 0.05$ to $\alpha = 0.0$ and $\alpha'(t) = 0.1 \cdot \alpha(t)$). For each run of DSLVQ classification randomly selected 50% of the BP features were used for training and the remaining 50% were kept to test the classifier.

D. Classifier CFR1: Design and Customization

The three class problem (number of classes $N_C = 3$) was solved by applying a majority voting to 3 pairwise trained LDA discriminant functions (see [8] for more details). In contrast to [8] only the sign of each LDA distance (classlabel information) was considered (< 0 class one, ≥ 0 class two) and the LDA distance information was discarded. The class with the highest frequency of occurrence (f_{occ}) within the past subject-specific N samples (maximum N=75 or 250 ms) was selected and a normalized distance (d_n) was defined in the following way:

$$d_n = (f_{occ} - \tfrac{N}{N_C}) / \tfrac{(N_C - 1) \cdot N}{N_C} \ \forall \ f_{occ} > \tfrac{N}{N_C} \text{ , or}$$
$$d_n = 0 \ \forall \ f_{occ} \leq \tfrac{N}{N_C}.$$

This normalization prevents quick changes of the classification result and enables a smooth transition (”zero“ crossing) between class-specific feedback. The disadvantage is the introduction of an additional feedback delay.

1) Training without feedback: The twenty-two monopolar EEG channel setup was used (Fig.1). Subjects were instructed to perform continuous kinesthetic MI [21] according to the instructions presented on the screen. The kind of movement was chosen by the subject according his/her preferences (e.g. playing a piano or swimming) and fixed before the recording started. Two sessions were recorded for each subject on different days. Each session consisted of 6 runs with 48 trials each (12 trials per class per run). At t=0s of each trial a short warning tone was presented and a fixation cross appeared in the middle of the screen. From t=2.00s to t=3.25s an arrow (cue) was shown, indicating the mental task to be performed (start MI). An arrow pointing to the left, to the right, downward or upward indicated the imagination of a left hand, right hand, foot or tongue movement, respectively. The order of appearance of the classes (cues) was randomized and at t=6s the screen was blanked (stop MI). To avoid adaptation to the timing a randomly selected inter-trial interval lasting between 1s and 2s was introduced after second 6.

2) Feature selection: For each subject individual 4-class DSLVQ classifiers were trained with BP features extracted with the same time lag (from t=0.0s to t=4.0s in steps of Δt=0.5s) to the cue presentation at t=2.0s. For each t and each trial, 15 non-overlapping BP features between 6 - 36 Hz with a bandwidth of 2 Hz were computed for each of the examined 23 bipolar channel derivations (Fig.1). The selected frequency range covers α and β bands, both relevant for the classification of MI [22]. By selecting 2 Hz frequency bands the resulting temporal delay (blur) and spectral resolution is acceptable. At the same time the total number of features was limited in order to avoid overfitting-effects.

The most relevant BP features were selected by evaluating the RT_n values (high value means important feature) from the DSLVQ analysis at t with the highest classification accuracy. BP features were selected manually according to the following criteria: (i) large mean RT_{n0} value and small variance (100 DSLVQ iterations) (ii) maximum number of bipolar channels = 3 (iii) maximum number of BP features = 6 (iv) symmetrically arrangement over sensorimotor areas (hemisphere). Two adjacent BP features were combined to one 4-Hz BP feature (e.g. 10-12 Hz and 12-14 Hz to 10-14 Hz).

With the identified features for each of the four 3-class MI combinations an independent CFR1 was trained (10x10 cross-validation) and a sample-by-sample on-line simulation was computed. The three MI tasks with the best classification accuracies within the feedback period (t=2.0s to t=6.0s) were selected and used for on-line experiments.

3) Feedback training: On-line feedback experiments were realized using the 3 subject-depended bipolar EEG derivations. Electrode position Fz served as ground.

Each session started (after the initial recording) with unguided practice (free training) lasting about 5 minutes. During this period subjects could test the classifier (continuous feedback) and the smoothing parameter N needed for d_n was adjusted according to the subjects preference (slow or fast reaction time). The feedback presented to the subjects was a smiley (see Fig.2.A).

Five (subject v4 and x6) and seven (v9) feedback sessions were recorded with at least 4 runs with 30 trials (10 per class) each. At the beginning of each trial (t=0.0s) the gray-colored smiley was positioned in the center of the screen. At t=2.0s a warning tone was presented followed by the cue at t=3.0s. From t=3.0s to t=7.5s subjects had the task to move the smiley according to the cue to the left/right/down(up) by performing left hand, right hand or foot (tongue) MI respectively (Fig.2.A). During the feedback period (t=3.5s to t=7.5s) the smiley changed to color green when moved to the correct direction, otherwise the color was red. The position of the smiley was set according to the classification result: The classlabel indicated the direction and d_n the distance to the origin. Furthermore d_n was mapped to the curvature of the mouth causing the smiley to be happy (corner or the mouth upwards) or sad (corner or the mouth downwards) according to correct classification or misclassification (see Fig.2.A). At t=7.5 s the screen was blanked and a random period between 1.5 s and 2.5 s was presented.

After each session feature selection was performed and the classifier was updated. Each time the accuracy of the updated classifier was higher than the on-line result, subjects tested the new classifier during the unguided practice period of the next session. When subjects achieved better BCI control

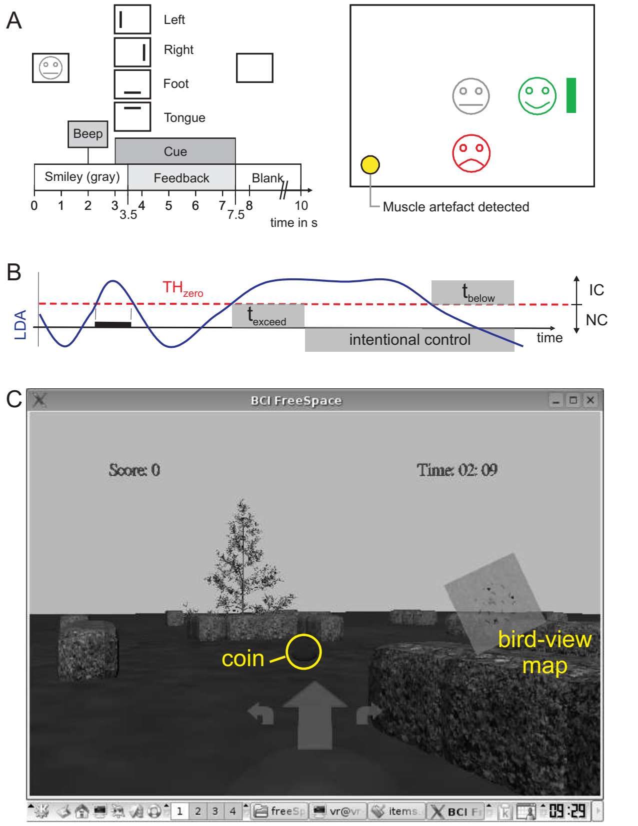

Fig. 2. A: Feedback timing and scheme of the Smiley paradigm. B: Asynchronous switch between intentional control (IC) and non-control (NC). Only when the classifier output exceeded (t_{exceed}) or did fall below (t_{below}) the predefined threshold (TH_{zero}) a switch between NC and IC or IC and NC was valid. C: First person view of the freeSpace virtual environment. A tree, some hedges and a coin (to collect) is visible on the screen. The big arrow represents the BCI classification result and indicates the navigation command (here forward movement). During the non-control state the three arrows had the same (small) size. On the upper left side the scoreboard and on the right side the elapsed time was presented. For an easier navigation on the right side a rotating bird-view map of the freeSpace was shown.

the updated classifier was used for feedback experiments, otherwise the old one was maintained.

E. Classifier CFR2: Design and Customization

One single LDA function was trained to discriminate between IC (3 MI tasks merged) and NC.

1) Feature selection: In order to obtain a more detailed spectral representation thirty-one overlapping (1 Hz) BP features between 6 - 36 Hz with a bandwidth of 2 Hz were extracted from each channel and analyzed by DSLVQ. The six most relevant features were selected to set-up the LDA.

Two BP feature vectors were extracted from the feedback interval (MI) around the best on-line classification accuracy (e.g. best classification at t=5.5s; BP extracted at t_1=5.5s and t_2=6.5s) from each trial of the last feedback training session (4 runs with 30 trial each). The resulting $4\, runs \cdot 30\, trials \cdot 2\, BP = 240$ BP samples were defined as IC. NC consisted of 120 BP samples extracted equidistantly from the EEG block of the baseline recording lasting 2 min with eyes open. Furthermore 120 feature samples were extracted from each trial at t=3.0s (before cue presentation). The latter time was selected with the intention to detect only MI specific patterns during feedback (after cue presentation) and not unspecific activations (e.g expectation) induced by the appearance of the fixation cross.

To make the CFR2 more robust and reliable, considering the non-stationarity and inherent variability of brain signals, one threshold (TH_{zero}) and two transition periods, one for the state switch NC to IC (t_{exceed}) and one for IC to NC (t_{below}) were introduced. Each time the distance between the BP features and the optimal LDA hyperplane (LDA distance) was higher than TH_{zero} for t_{exceed}, the state IC was detected. To get back to NC the LDA distance had to fall below TH_{zero} for t_{below} (see Fig.2.B). The system reaction time was modified by t_{exceed} and t_{below}; changing TH_{zero} meant moving the decision hyperplane towards IC or NC. The initial TH_{zero} used for feedback experiment was determined by Receiver Operator Characteristic (ROC) analysis. The value which maximizes the number of TP detections within the feedback period (t=3.5s to t=7.5s) and at the same time minimizes the number of FP detections everywhere else was selected (sample-by-sample).

By combining both, CFR1 and CFR2 we created a system able to discriminate between 4 classes: (i) left hand, (ii) right hand, (iii) foot (tongue) MI and (iv) non-control. The output of CFR1 was triggered by CFR2. Each time IC was detected the classification result of CFR1 was feed through. Otherwise the output was ”zero”.

2) Feedback training with longer inter-trial intervals: Two sessions with 5 feedback training runs (10 trial per class) were performed. The first 2 runs consisted of feedback training with CFR1 only (section II-D.3) to monitor the subjects actual performance. Thereafter subjects underwent an unguided practice lasting about 10 minutes to determine TH_{zero}, t_{exceed} and t_{below}. Depending on the statements of the subjects TH_{zero} was gradually increased or decreased. The criteria was that subjects were able to control the smiley, but at the same time the number of FP detections was a minimum. The transition times were set to 500ms and also gradually adapted. A maximum period of 1s was chosen for t_{exceed} and t_{below}. Subjects had to identify these values by themselves empirically. The values found were fixated and not changed during the remaining experiments of each session.

Training subjects to gain self-paced control, the feedback smiley paradigm from section II-D.3 was modified. The feedback smiley was presented and reactive during the whole run. Each run consisted of 30 trials (10 per class). From t=0.0s to t=8.0s the cue was presented and subjects had the task to stir the smiley to the indicated direction. After this period a random inter-trial period between 7.0s and 17.0s was presented (NC). At the beginning the gray smiley was positioned in the middle of the screen. During the transition times t_{exceed} or t_{below} the color of the smiley changed gradually from gray to green or green to gray, respectively. Additionally the smiley was moving according to CFR1 with the distance information weighted by the normalized transition time (from 0 to 1). In this way subjects were informed of a forthcoming state switch.

After the first session an additional DSLVQ analysis was performed for CFR2. Subjects tried the new CFR2 during the unguided practice period of session two and if the performance increased, the new classifier was used.

F. Evaluating self-paced control of CFR1 and CFR2

1) The "freeSpace" Virtual Environment: The Virtual Environment (VE) was created using the 3-D modelling software package Maya (Alias Wavefront, Toronto, Canada). Furthermore it was animated (collision detection) and visualized by the Qt application framework (Trolltech, Oslo, Norway). Since the VE was running on a separate personal computer, the communication with the BCI was realized using the User Data Protocoll (UDP) and updated 25 times per second. The virtual park, size 30×30 units, consisted of a flat meadow, several hedges and a tree placed in the middle for orientation. Three items (coins) were positioned on fixed predefined locations inside the park. Three navigation commands were implemented: turn left, turn right (angular velocity $45°/s$) and move forward (speed 1 unit/s). With this control, each part of the park could be reached. To help subjects not get lost and facilitate locating the coins, a bird view map of the VE, showing the actual position, was presented (see Fig. 2.C). Interaction with each of the existing virtual objects was possible. A sphere, representing the user in the VE, was used for collision detection. Each time the surfaces of two objects were intersecting an event was generated: Coins were collected and hedges or the tree had to be bypassed.

2) Experimental paradigm: Two sessions were recorded on two different days. Each session started with about 20 minutes of unguided practice (free training). Subjects could get familiar with the freeSpace VE and the navigation mechanism. In regard to the non-stationarity and inherent variability of EEG, TH_{zero} was adapted and fixated.

The VE was presented to the subjects in the first-person-view on a conventional computer screen (2-D). Subjects had the task of picking up the 3 coins within 3 minutes. From a randomly selected starting point (different positions for each run but same positions for all subjects), subjects could explore the park in the following way: left/right hand MI resulted in a rotation to the left/right whereas foot or tongue MI resulted in a forward motion. Whenever NC was detected, consequently no action was performed.

Six self-paced feedback training runs of 3 min each were performed. The first three runs served as training, run four to six were used to evaluate the performance.

For each subject the covered distance and resulting path was depending on the individual routing strategy (e.g. pickup order) and the ability to operate the BCI.

Since for this paradigm it was not possible to compute true positive of false positive detections, at the end of each sessions subjects were asked to self-report on the BCI performance.

III. Results

A. Muscular artifacts

By means of a 4 minute recording at the beginning of each session the proposed EOG reduction and EMG detection algorithms were initialized. An independent investigation revealed that 80% of EOG artefacts are removed by using the proposed method [16]. For each subject the percentage of samples classified as EMG artifact, and consequently the time of presentation of the warning message, was less than 0.9%. Additionally power spectral densities were computed for each channel and checked for muscle activity.

TABLE I
Motor imagery (MI) tasks (Left hand, Right hand, Foot or Tongue), bipolar channel setup (Bipolar), frequency components (in Hz) used for the discrimination between MI (CFR1) and intentional control vs. non-control state (CFR2) are summarized for each subject (ID).

ID	MI	Bipolar	CFR#1 (MI)	CFR#2 (IC vs. NC)
v4	L,R,T	+02-14	11-13, 25-27	12-14, 15-17, 20-22, 25-27
		+04-16	-	9-11, 21-23
		+06-18	11-13, 23-25	-
v9	L,R,F	+02-08	11-13, 12-14	12-14, 19-21, 27-29
		+04-16	11-13, 12-14	9-11, 11-13
		+06-12	11-13	21-23
x6	L,R,F	+02-08	10-12	8-10, 16-18
		+04-10	9-11	8-10
		+06-12	10-12, 19-23	15-17, 24-26

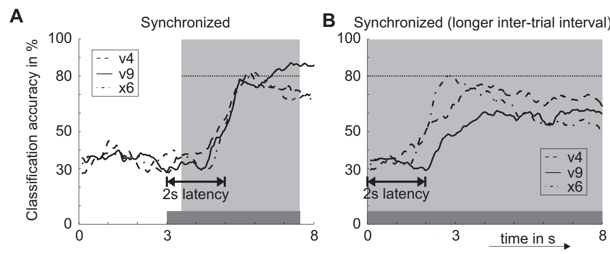

Fig. 3. Feedback experiments (synchronized protocol). The dark gray lines indicates the interval of cue-presentation, the light gray areas the feedback period. A: Three-class on-line classification performance. Four runs from 2 days were combined. Feedback was given from second 3.5 to 7.5. B: Mean classification performance of the on-line feedback training with longer periods without motor imagery. Subjects had the task from second 0.0 to 8.0 to move the feedback to the given direction.

B. Classifier CFR1

The most discriminative bipolar channels, BP features and MI tasks found by DSLVQ are summarized in Tab.I. As one could expect, α and upper β band of electrodes placed around electrode position C3, Cz and C4 (international 10-20 system) proved to substantially contribute to classification.

The on-line performance of the feedback training is shown in Fig.3.A. The curves show the classification accuracy of the 4 runs recorded at the beginning of the 2 feedback training sessions with longer inter-trial intervals (see section II-E.2). Classification accuracies around 80% were reached for all three subjects.

C. Classifier CFR2

Column CFR2 in Tab.I lists the BP features found by DSLVQ which most discriminate between IC (left hand, right hand and foot or tongue together) and NC. Off-line classification accuracies (10x10 cross-validation) of 77% for subject v4, 84% for v9 and 78% for x6 were computed.

The classification performance for CFR1 and CFR2 during the feedback training with longer inter-trial intervals days are presented independently. The mean classification accuracy of CFR1 during the active period is shown in Fig.3.B. Compared with the results in Fig.3. A similar characteristics can be observed. The mean latency from cue-presentation to a classification performance better than random was about 2s. For subject v4 and x6 the mean classification accuracies were 75% and 80%, respectively; for v9 the mean classification accuracy was 60%.

As performance measures true positive (TP) and false positive (FP) rates were computed (sample-by-sample) for CFR2. The evaluation criteria for TP and FP were very strictly: According to the feedback paradigm, samples from t=0.0s to t=8.0s, the period of target presentation, were defined as IC and consequently labeled as class 1. The remaining samples were NC labeled as class 2. Since cognitive processes (e.g. processing the visual cue, motor preparation, decision making) as well as digital signal processing requires time, and according to the time latency of about 2.0s, an additional evaluation with TP defined from t=2.0s to t=10.0s was computed. The TP/FP rates were computed by dividing the number of correctly classified samples within the TP/FP intervals by the total number of samples belonging to the TP/FP class. Tab.II summarizes these results. Mean FP rates (over subjects and sessions 2 hours NC and 1 hour IC) of 19.1% or 16.9% could be achieved. The mean TP rates for the 8 second action period were 25.1% or 28.4%. Column T_{IC} in Tab.II shows the number of transitions from NC to IC during the cue presentation. Subjects succeeded in 18.6 of 30 trials during the presentation of the cue (t=0.0 to t=8.0) to switch from NC to IC. The last column of Tab.II shows the time in seconds of NC for each run. The mean inter-trial interval was 12.5 s.

TABLE II

TRAINING WITH LONGER INTER-TRIAL INTERVALS. FOR EACH SUBJECT (ID), SESSION (S) AND RUN (R), THE DURATION OF THE RUN IN SECONDS (DUR.), THE NUMBER OF TRANSITIONS FROM NC TO IC (T_{IC}, MAXIMUM 30/RUN), THE TP/FP DETECTION RATES FOR THE 2 INVESTIGATED TIME PERIODS AND THE DURATION OF THE NON-CONTROL PERIOD IN SECONDS (t_{NC}) ARE SHOWN. ADDITIONALLY THE THE SESSION MEANS μ_1 AND μ_2 ARE REPORTED FOR EACH SUBJECT.

				0.0-8.0		2.0-10.0		
ID	S-R	Dur.	T_{IC}	TP	FP	TP	FP	t_{NC}
v4	1-1	602	13	18.4	16.9	20.0	15.8	362
	1-2	609	10	7.6	7.2	9.8	5.8	369
	1-3	615	16	17.8	12.2	21.1	10.1	375
	μ_1	*609*	*13*	*14.6*	*12.2*	*17.0*	*10.6*	*369*
	2-1	612	27	32.7	22.5	36.9	19.5	372
	2-2	612	29	33.5	23.5	37.1	21.2	372
	2-3	610	28	39.7	23.2	33.7	27.0	370
	μ_2	*611*	*28*	*35.3*	*23.1*	*35.9*	*22.6*	*371*
v9	1-1	608	20	24.1	16.6	32.9	10.8	368
	1-2	603	21	32.3	31.4	38.4	27.3	363
	1-3	636	21	23.4	8.8	31.6	3.8	396
	μ_1	*616*	*21.7*	*26.6*	*18.9*	*34.3*	*20.0*	*376*
	2-1	611	13	16.8	14.8	22.8	11.0	371
	2-2	620	20	38.7	24.5	43.1	21.7	380
	2-3	617	18	36.9	32.5	42.1	29.1	377
	μ_2	*616*	*17*	*30.8*	*23.9*	*36.0*	*20.6*	*376*
x6	1-1	602	16	22.1	11.6	21.4	12.0	362
	1-2	606	16	14.5	10.5	16.7	9.0	366
	1-3	608	15	16.1	15.7	19.2	13.6	368
	μ_1	*605*	*15.7*	*17.6*	*12.6*	*19.1*	*11.5*	*365*
	2-1	628	14	38.1	32.5	39.8	31.4	388
	2-2	632	18	18.9	16.2	19.6	15.8	392
	2-3	614	19	20.2	22.8	24.6	20.0	374
	μ_2	*625*	*17*	*25.7*	*23.8*	*28.0*	*22.4*	*385*
mean		613.6	18.6	25.1	19.1	28.4	16.9	373.6
std		9.9	5.3	9.8	8.0	10.0	8.2	9.9
median		611.5	18.0	22.8	16.8	28.1	15.8	371.5

D. "freeSpace" paradigm

The summary of the freeSpace experiment performance is given in Tab.III (best results are highlighted). The covered distance, number of collected items and pick-up times are shown for each of the 3 runs during the 2 sessions. Subject v9 and x6 were able to collect the three items within the 3 minutes time limit. Subject v4 was able to collect only 2 out of the 3 coins. While v4 and v9 could improve their performance (distance and collected items), the results of session two for x6 were poor compared to the first.

The routes (paths) of the best run for each subject are presented in Fig.4.A. Best results were achieved from each subject independently when starting from the same initial position. The paths show that each subject choose a different way to collect the coins. Fig.4.B shows the corresponding BCI classifier output (navigation commands) sent to the VE. The distribution of the BCI classification output is summarized in Tab.IV. Since it is not possible to report the percentage of erroneous navigation control signals (lucky errors) detected by the BCI which contribute to the collection of the coins, a "random walk" navigation was simulated. When starting from the position which is marked with "X" in Fig.4.A and randomly selecting the MI-states we obtained a zig-zag shaped route. The resulting course, however, is unidirectional. Accordingly, it was impossible to collect all three coins within the selected time limit without IC. The same results were obtained by increasing the frequency of occurrence of foot MI. For comparison also the shortest possible route was computed. With 100% classification accuracy about 110s were necessary to collect the three items.

The selected navigation strategy required that subjects were able to control at least two mental states: Either left or right for rotation and foot/tongue to move forward. The BCI classification results in Fig.4.B and the distribution in Tab.IV, however, show that all four classes occurred. Interviews with the subjects confirmed that all four mental states were deliberately used to navigate through the freeSpace. It was necessary that no navigation command was sent to the VE during non-MI related mental activity, like e.g. orientation or routing, or whenever subjects needed a break. For subject v4 and v9 the percentage of navigation commands increased from session 1 to session 2. Although subject x6 was satisfied with the achieved BCI control during the unguided practice period of session 2, a clear bias towards NC is visible during the evaluation.

IV. DISCUSSION

Asynchronous control, dealing with artifacts, reliable classification or a fast setup are some of the key-issues which contribute to make BCIs become a real alternative communication channel.

TABLE III

PERFORMANCE OF THE FREESPACE EXPERIMENT. FOR EACH SUBJECT (ID), RUN (R) AND SESSION, THE COVERED DISTANCE (DIST.) AND NUMBER OF COLLECTED ITEMS WITH CORRESPONDING PICKUP-TIMES [#ITEMS (TIME)] ARE SHOWN.

		Session 1		Session 2	
ID	R	Dist.	#items (time)	Dist.	#items (time)
v4	1	317	0	1172	**2 (0:46, 1:59)**
	2	544	1 (1:37)	523	0
	3	234	0	963	2 (0:53, 2:59)
v9	1	1223	2 (0:58, 2:21)	1781	**3 (0:16, 0:46, 2:33)**
	2	1573	1 (2:12)	1788	2 (1:10, 2:20)
	3	1438	1 (0:53)	1478	2 (0:39, 2:00)
x6	1	1520	2 (1:49, 2:34)	544	1 (2:19)
	2	1729	2 (0:41, 1:23)	807	1 (0:24)
	3	1635	**3 (0:20, 1:17, 2:26)**	468	1 (0:55)

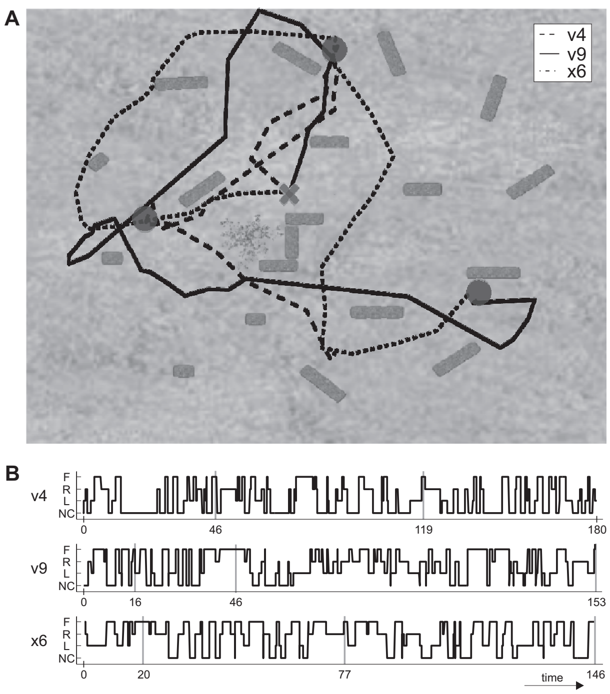

Fig. 4. A. Map of the freeSpace virtual environment showing the best performance (route) for each subject. The rectangles indicate hedges, the circles the pickup areas (collision detection) and the ”x“ marks the starting point. Subject v9 (continuous line) and x6 (dotted line) successfully collected the 3 coins. Subject v4 (dashed line) succeeded in picking up only 2 coins in 3 minutes. B. Classification output as function of time (L=left hand, R=right hand, F=foot, NC=non-control).

TABLE IV

DISTRIBUTION OF THE CLASSIFICATION RESULTS OF THE FREESPACE EXPERIMENT. FOR EACH SUBJECT (ID), RUN (R) AND SESSION THE FREQUENCY OF OCCURRENCE IN PERCENT OF DETECTED LEFT HAND (L), RIGHT HAND (R), FOOT OR TONGUE (F/T) MOTOR IMAGERY AND NON-CONTROL (NC) IS REPORTED.

		Session 1				Session 2			
ID	Rn	L	R	F/T	NC	L	R	F/T	NC
v4	1	18	18	6	58	**14**	**19**	**21**	**46**
	2	14	17	10	59	14	20	15	51
	3	7	4	4	85	16	22	18	44
v9	1	16	12	27	45	**25**	**19**	**36**	**20**
	2	20	18	29	33	17	26	33	24
	3	20	26	30	24	14	13	30	43
x6	1	30	24	39	7	4	8	10	78
	2	27	23	43	7	5	6	14	75
	3	**20**	**22**	**38**	**20**	4	4	9	83

Online EOG reduction and EMG detection was used for the first time in our feedback experiments. The muscle detection algorithm has been used to identify possible muscle activity in realtime. Accordingly, it is possible to use this information to avoid the classification of artifact data. A threshold value can be used to modify the sensitivity and specificity of the detector. An open and interesting issue is the discussion on the desired system response. One can think e.g. of a ”system freeze” or ”pause-mode”. In any case, this approach provides new possibilities for patients which are unable to control their own muscle activity (e.g. spasms).

The feedback training results show that 3 bipolar channel contains enough information to control a 3-class BCI (synchronized protocol) with an accuracy of 80%. Reducing the number of EEG channels is important because of (i) an increase of the usability (less time needed for electrode placement) and (ii) a minimization of electrode failures (e.g. exact position on the scalp, impedance, ...). Adaptation to subject-specific parameters is crucial to obtain a reliable classification in a short space of time. By default DSLVQ was applied to the data after each session. In the future this part can be replaced by a on-line adaptation method presented recently [23].

A new type of feedback was presented to subjects during the feedback experiments. The smiley was introduced because of the ”richer“ visual feedback (colors, position, shape of the mouth) compared to the bargraph or basket feedback [12], [14]. The expectation was a increased motivation for the subjects resulting in an improved performance. Interviews with the subject confirmed that the motivation to make the smiley laugh was high.

One very important issue for self-paced (asynchronous) BCIs is the evaluation criteria or measure of performance. We presented TP and FP rates computed on a sample-by-sample basis from data collected using a synchronized protocol with longer inter-trial intervals. For each subject the very first attempts of self-paced control were used for evaluation. The achieved average FP rates of 16.9%/19% (18.9/21.3 min out of 112 min of NC) were to high and the mean TP rates of 28.4%/25.1% to low. During 18.6 out of 30 trials (62%), however, subjects suceeded in switching from NC to IC. One can assume that longer feedback training period helps to increase the performance. TP/FP rates, however, depend strongly on the definition of the TP and FP intervals. The fact that MI-induced changes in EEG activity are not time-locked (delay) and have a variable duration make this definition difficult. One problem emerging from the cue-based design might be the expectation of the next cue to come. This expectation can unintentionally induce subjects to change the brain activity and produce FP. Nevertheless we are confident that sample-by-sample based TP/FP rates are most suited to characterize self-paced BCI performance. To compute the right TP/FP rates it is necessary to access the subjects ”real“ intend and compare it with the BCI output. This information, however, is not directly accessible. One option to obtain this information might be an interactive experimental design, where subjects autonomous decide on timing and type of MI and give immediate feedback, e.g. by interview or by pressing a button, on the correctness of the BCI output. When working with severely paralyzed people, however, motor interaction may be impossible.

The freeSpace paradigm was introduced because no instructions, except on how to navigate and the aim to collect coins, had to be given to the subjects. The paradigm is motivating, entertaining and most important, it gives an ample scope on how to achieve the goal. Each subject succeeded in navigating through the VE and collecting coins. As can be seen from the distribution of the BCI classification result (Tab.IV) and emerged from the interviews, for navigation both MI and NC were used. The paths in Fig.4.A show that each subject choose his own way through the freeSpace. Subject v4 and v9 could improve the performance from the first to the second session. This was not possible for subject x6. Also during the training x6 had a high variability of the performance. The overall trend, however, was towards higher classification accuracies. At this stage the NC state was not explicitly tested. However, as stated by the subjects, periods of NC were important. For further experiments, the paradigm can easily be enhanced by e.g. adding predefined periods of NC.

Although the “freeSpace” VE was implemented for three-dimensional (3-D), stereoscopic representation, at this stage only a conventional computer screen was used for visualization. One possible option for the future is to train users to operate BCI-based devices (e.g. wheelchair) in the Virtual Reality [24].

LDA and band power features are a good choice for the discrimination between different MI tasks. The question is whether this classifier/feature is best suited to identify MI patterns in the ongoing EEG. Finding proper features and classifier is one important task for future research. Wavelet-packet analysis [25] or phase relationships [26] may contribute to solve this problem. Compared to CFR1, CFR2 was sensitive to the non-stationarity of EEG. By adapting the detection threshold TH_{zero} this was taken into account. Higher values of TH_{zero} cause a decrease of FP, however, also the motivation of the subjects might decrease because generating TP is much more difficult. On the other hand small values cause lots of TP but FP as well. The varying TP/FP rates in Tab. II reflect this relationship. For the training this implies to start with lower values which can be increased when subjects achieved reliable control. These procedure, however, results in in poor TP/FP performance values. Gaining BCI-control is not only depending from machine learning, but also psychological aspects play an important role.

One drawback of defining IC by merging data from three MI tasks was that CFR2 had a ”preference“ (bias) for a certain MI patterns. This behavior was not visible during the evaluation experiments. After the first freeSpace experiments, however, subjects stated that switching into the IC state was easier for certain MI patterns. Therefore one strategy developed by the subjects was to switch into IC by performing the preferred MI first and thereafter switching to the desired one. In Tab.IV this effect is visible for subject v4 in session 2. Also in Fig.4.B for subject v4 the preference of right MI is visible. One solution to overcome this problem is to redesign the classifiers from ”1 vs. 1“ classification to ”1 vs. rest“. Given the classification accuracy does not decrease the benefit for a higher number of classes would be a reduction of the computational effort to N_C (number of classes) LDAs.

V. Conclusion

The methods and training procedure presented in this work enabled users to gain control of a self-paced BCI. Three bipolar EEG channels were analyzed and motor imagery was used as experimental strategy. In order to assure that no muscle activity was used for control, EMG was detected and reported to subjects online; furthermore EOG artifacts were reduced. Finding a proper evaluation method or performance measure is still an open issue. Actually, the BCI community, however, is addressing this important topic [4].

The study show that subjects successfully navigated through the freeSpace VE and collected coins by autonomously switching between different mental states. In doing so, each subject chose the way independently. These are further steps which help BCIs become a real alternative to standard communication channels.

Acknowledgment

This work was supported by the “Fonds zur Förderung der Wissenschaftlichen Forschung” in Austria, project P16326-B02 and in part by the EU research project PRESENCCIA IST 2006-27731. The authors also would like to thank Hewlett-Packard for their support.

References

[1] J. Wolpaw, N. Birbaumer, D. McFarland, G. Pfurtscheller, and T. Vaughan, “Brain-computer interfaces for communication and control.” *Clin Neurophysiol*, vol. 113, no. 6, pp. 767–791, Jun 2002.

[2] G. Pfurtscheller, C. Neuper, and N. Birbaumer, *”Human Brain-Computer Interface,” in Motor cortex in voluntary movements: a distributed system for distributed functions*, ser. Methods and New Frontiers in Neuroscience, E. Vaadia and E. C. P. A. Riehle, Eds. CRC Press, 2005.

[3] N. Birbaumer, “Brain-computer-interface research: Coming of age.” *Clin Neurophysiol*, vol. 117, no. 3, pp. 479–483, Mar 2006. [Online]. Available: http://dx.doi.org/10.1016/j.clinph.2005.11.002

[4] S. Mason, J. Kronegg, J. Huggins, M. Fatourechi, and A. Schlögl, “Evaluating the performance of self-paced brain computer interface technology,” Tech. Rep., 2006.

[5] S. Mason, J. Kronegg, J. Huggins, M. Fatourechi, and A. Schleogl, “A brain-controlled switch for asynchronous control applications.” *IEEE Trans Biomed Eng*, vol. 47, no. 10, pp. 1297–1307, Oct 2000.

[6] J. del R. Millàn and J. Mourino, “Asynchronous BCI and local neural classifiers: an overview of the Adaptive Brain Interface project.” *IEEE Trans Neural Syst Rehabil Eng*, vol. 11, no. 2, pp. 159–161, Jun 2003.

[7] J. Borisoff, S. Mason, A. Bashashati, and G. Birch, “Brain-computer interface design for asynchronous control applications: improvements to the LF-ASD asynchronous brain switch.” *IEEE Trans Biomed Eng*, vol. 51, no. 6, pp. 985–992, Jun 2004.

[8] R. Scherer, G. Müller, C. Neuper, B. Graimann, and G. Pfurtscheller, “An asynchronously controlled EEG-based virtual keyboard: improvement of the spelling rate.” *IEEE Trans Biomed Eng*, vol. 51, no. 6, pp. 979–984, Jun 2004.

[9] G. Müller-Putz, R. Scherer, G. Pfurtscheller, and R. Rupp, “EEG-based neuroprosthesis control: a step towards clinical practice.” *Neurosci Lett*, vol. 382, no. 1-2, pp. 169–174, 2005. [Online]. Available: http://dx.doi.org/10.1016/j.neulet.2005.03.021

[10] B. Blankertz, G. Dornhege, M. Krauledat, K.-R. Mller, V. Kunzmann, F. Losch, and G. Curio, “The Berlin Brain-Computer Interface: EEG-based communication without subject training.” *IEEE Trans Neural Syst Rehabil Eng*, vol. 14, no. 2, pp. 147–152, Jun 2006.

[11] D. J. McFarland, W. A. Sarnacki, T. M. Vaughan, and J. R. Wolpaw, “Brain-computer interface (BCI) operation: signal and noise during early training sessions.” *Clin Neurophysiol*, vol. 116, no. 1, pp. 56–62, Jan 2005. [Online]. Available: http://dx.doi.org/10.1016/j.clinph.2004.07.004

[12] G. Pfurtscheller and C. Neuper, "Motor Imagery and Direct Brain-Computer Communication," *Proceedings of the IEEE*, vol. 89, pp. 1123–1134, 2001.

[13] C. Vidaurre, A. Schlögl, R. Cabeza, R. Scherer, and G. Pfurtscheller, "Study of on-line adaptive discriminant analysis for eeg-based brain computer interfaces," *IEEE Trans. Biomed. Eng.*, p. in press.

[14] G. Krausz, R. Scherer, G. Korisek, and G. Pfurtscheller, "Critical decision-speed and information transfer in the "Graz Brain-Computer Interface"." *Appl Psychophysiol Biofeedback*, vol. 28, no. 3, pp. 233–240, Sep 2003.

[15] R. Scherer, A. Schlögl, G. Müller-Putz, and G. Pfurtscheller, "Inside the Graz-BCI: rtsBCI," in *Proceedings of the 2nd International Brain-Computer Interface Workshop and Training Course 2004*, ser. Biomedizinische Technik, vol. 49, no. 1. Schiele & Schön, 2004, pp. 81–82.

[16] A. Schlögl, C. Keinrath, D. Zimmermann, R. Scherer, R. Leeb, and G. Pfurtscheller, "A fully automated correction method of EOG artifacts in EEG recordings." *Clinical Neurophysiology, in press*.

[17] R. Scherer, A. Schlögl, and G. Pfurtscheller, "Online detection and reduction of electrooculographic (eog) and electromyographic (emg) artifacts," in *Proceedings of the Gemeinsame Jahrestagung der Deutschen, der Österreichischen und der Schweizerischen Gesellschaften für Biomedizinische Technik*, 2006.

[18] F. H. L. da Silva, K. van Hulten, J. G. Lommen, W. S. van Leeuwen, C. W. van Veelen, and W. Vliegenthart, "Automatic detection and localization of epileptic foci." *Electroencephalogr Clin Neurophysiol*, vol. 43, no. 1, pp. 1–13, Jul 1977.

[19] M. Pregenzer, "Frequency component selection for an EEG-based brain to computer interface." *IEEE Trans Rehabil Eng*, vol. 7, no. 4, pp. 413–419, Dec 1999.

[20] ——, "Dslvq," Ph.D. dissertation, Graz University of Technology, 1997.

[21] C. Neuper, R. Scherer, M. Reiner, and G. Pfurtscheller, "Imagery of motor actions: differential effects of kinesthetic and visual-motor mode of imagery in single-trial EEG." *Cogn Brain Res*, vol. 25, no. 3, pp. 668–677, Dec 2005. [Online]. Available: http://dx.doi.org/10.1016/j.cogbrainres.2005.08.014

[22] G. Pfurtscheller, C. Brunner, A. Schlögl, and F. H. L. da Silva, "Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks." *Neuroimage*, vol. 31(1), pp. 205–10, Jan 2006. [Online]. Available: http://dx.doi.org/10.1016/j.neuroimage.2005.12.003

[23] C. Vidaurre, A. Schlögl, R. Cabeza, R. Scherer, and G. Pfurtscheller, "A Fully On-line Adaptive BCI," *IEEE Trans Rehab Eng*, p. in press, 2006.

[24] J. Tarr and W. Warren, "Virtual reality in behavioral neuroscience and beyond," *Nat Rev Neurosci*, vol. 5, pp. 1089–92, 2002.

[25] B. Graimann, J. E. Huggins, S. P. Levine, and G. Pfurtscheller, "Toward a direct brain interface based on human subdural recordings and wavelet-packet analysis." *IEEE Trans Biomed Eng*, vol. 51, no. 6, pp. 954–962, Jun 2004.

[26] C. Brunner, R. Scherer, B. Graimann, G. Supp, and G. Pfurtscheller, "Online control of a brain-computer interface using phase synchronization," *IEEE Trans. Biomed. Eng.*, p. in press, 2006.