Utilizing Retinotopic Mapping for a Multi-Target SSVEP BCI With a Single Flicker Frequency

Alexander Maye, Dan Zhang, and Andreas K. Engel

***Abstract*— In brain–computer interfaces (BCIs) that use the steady-state visual evoked response (SSVEP), the user selects a control command by directing attention overtly or covertly to one out of several flicker stimuli. The different control channels are encoded in the frequency, phase, or time domain of the flicker signals. Here, we present a new type of SSVEP BCI, which uses only a single flicker stimulus and yet affords controlling multiple channels. The approach rests on the observation that the relative position between the stimulus and the foci of overt attention result in distinct topographies of the SSVEP response on the scalp. By classifying these topographies, the computer can determine at which position the user is gazing. Offline data analysis in a study on 12 healthy volunteers revealed that 9 targets can be recognized with about 95 ± 3% accuracy, corresponding to an information transfer rate (ITR) of 40.8 ± 3.3 b/min on average. We explored how the classification accuracy is affected by the number of control channels, the trial length, and the number of EEG channels. Our findings suggest that the EEG data from five channels over parieto-occipital brain areas are sufficient for reliably classifying the topographies and that there is a large potential to improve the ITR by optimizing the trial length. The robust performance and the simple stimulation setup suggest that this approach is a prime candidate for applications on desktop and tablet computers.**

***Index Terms*— Brain-computer interfaces, Neurocontrollers, Spatial filters, Pattern recognition.**

I. Introduction

THE steady-state visual evoked potential (SSVEP) is a brain response which can be measured by an electroencephalogram (EEG) when a subject looks at periodic luminance- or contrast-modulated stimuli [1]. It has the same frequency as the stimulation sequence and a phase delay that is fixed for a given stimulation frequency [2]. These properties, together with the high signal-to-noise ratio of SSVEPs, are frequently exploited for brain-computer interfaces (BCIs) by presenting a set of flicker stimuli which differ in their frequency, phase, temporal pattern, or combinations thereof and decoding the SSVEP response while the user pays attention to the target stimulus [3].

Most SSVEP-based BCIs employ a set of fixed frequencies to encode different control channels. Directing overt or covert attention to a stimulus which flickers at one of these frequencies increases the power at this frequency in the EEG, which can be used as a feature for automatic classification [4], [5]. Using this approach, up to 45 channels have been encoded for a spelling application [6], and ITRs up to 320 bits/min have been achieved [7]. Various approaches to optimize the performance of such frequency-coded SSVEP BCIs, which are also referred to as f-VEP BCIs [8] or FDMA (frequency division multiple access [9]) BCIs have been developed [10], [11].

The approach we present in this article is fundamentally different from these frequency-coded SSVEP BCIs. Instead of decoding targets from power spectra, our approach classifies the spatial distributions of the SSVEP power across the scalp (topographies) that are elicited by a single flicker stimulus which appears at different positions in the visual field of the user. Hence our approach encodes channels in the spatial domain and therefore belongs to the class of spatial or SDMA (spatial division multiple access) SSVEP BCIs.

Spatial properties of the SSVEP have rarely been exploited for encoding channels in a BCI. However, there is a substantial body of electrophysiological findings about the lateralization of VEPs and SSVEPs which support the concept of an SDMA BCI. Common textbook knowledge holds that extrafoveal light stimulation elicits strong responses over visual cortical areas of the contralateral hemisphere [12]. However, an ipsilateral activation maximum has also been observed for a pattern reversal stimulus at 2 Hz and termed paradoxical lateralization [13]. This phenomenon seems to depend on the presentation of the stimulus above or below the horizontal meridian in some subjects, but may be absent in others [14]. These differences may result from variations in the individual anatomical convolutions of the cortex in relation to the electrode placement [15]. Differences of the lateralization in response to parafoveal stimuli have been observed even at the level of the various VEP components. For example, the N160 was found to be stronger at contralateral electrodes, while the P120 was larger at ipsilateral electrodes [16]. Together with the amplitude, also the polarity of VEPs can change between upper and lower

Manuscript received August 26, 2015; revised February 8, 2016 and August 3, 2016; accepted February 5, 2017. Date of publication April 25, 2017; date of current version August 6, 2017. This work was supported by in part EU under Grant ERC-2010-AdG-269716 and Grant FP7-ICT-270212, in part by DFG under Grants GRK 1247/1/2 and SFB TRR169/B1/Z2, in part by the Landesforschungsfoerderung Hamburg, and in part by the Natural Science Foundation of China under Grant 61101151. *(Corresponding author: Alexander Maye.)*

A. Maye and A. K. Engel are with the Department of Neurophysiology and Pathophysiology, University Medical Center Hamburg-Eppendorf, D-20246 Hamburg, Germany (e-mail: a.maye@uke.de; ak.engel@uke.de).

D. Zhang is with the Department of Psychology, School of Social Sciences, Tsinghua University, Beijing 100084, China (e-mail: dzhang@mail.tsinghua.edu.cn).

Digital Object Identifier 10.1109/TNSRE.2017.2666479

field stimuli [17]. Analysis of the visual evoked magnetic response using MEG suggests that the dissociation between the location of sources and their actual projection on the scalp may be responsible for the observed inconsistencies of power distributions [18], and that adequate models of the spatial SSVEP patterns should include multiple local and distributed sources [19]. As different stimulation frequencies can activate different cortical networks, SSVEP topographies also depend on the stimulation frequency [20].

These studies show that the retinotopic mapping, i.e. the mapping of visual input from the retina to neurons in the visual areas of the cortex, causes distinct EEG patterns which depend on the position of a flicker stimulus in the visual field. This led us to the hypothesis that at least the SSVEP topographies for stimuli above and below the horizontal meridian and to the left and right of the vertical meridian should be sufficiently distinct from each other to facilitate their automatic recognition. The critical link between the electrophysiological studies and the application of their findings in a BCI is the idea that a presentation of stimuli at various eccentricities when the subject fixates a central location is equivalent to a central stimulus presentation with the subject gazing at targets in the periphery. This was confirmed in a study where the spatial distribution of VEPs elicited by a flashed checkerboard pattern was used to recognize at which corner of a rhombic stimulus subjects were gazing [21]. As explained above, the different parafoveal positions of the flicker associated with the four gaze directions caused distinct ERP waveforms at the five electrodes over occipital and parietal brain areas (Pz, O1, O2, Oz, inion). A consistent classification accuracy of over 90% was achieved. More recently, a VEP BCI study employed a central low-speed reversal stimulation with subjects gazing at targets to the left and right [22]. This resulted in a characteristic change of the VEP waveform in the interval from 50 to 200 ms post-stimulus between corresponding electrodes over the left and right hemispheres at occipital, parietal and central regions. Classification accuracies of 90% for two classes and 84.2% for three classes were achieved.

The same idea using SSVEPs instead of VEPs was realized by [23]. The study used a central LED flicker which was flanked at the left and right side by black patches. A comparison of the SSVEP power at occipital electrodes O1 and O2 allowed the method to recognize at which of the three targets subjects were looking. A fourth control channel was added by utilizing the Berger effect, i.e. by determining from the occipital alpha power if subjects opened or closed their eyes. An average classification accuracy of 78% was achieved in four subjects. A preliminary study suggested that not only the position of single stimuli in the visual field but even more complex shapes can be reconstructed from the SSVEP topography [24].

In addition to the application for SDMA BCIs, the retinotopic mapping is also employed for increasing the signal-to-noise ratio in conventional FDMA BCIs. In this setup, two lights, flickering to the left and right of a fixation spot at the same frequency but opposite phase, elicit SSVEPs over the visual cortex in the respective contralateral hemispheres. Taking the difference of the EEG signals from both locations (e.g. O1-O2) still yields the SSVEP, but at the same time reduces the noise waveforms that are common to both hemispheres [25]. This visual half-field stimulation can also help with reducing the number of stimulation frequencies in FDMA BCIs by using different frequency combinations for the left and right half-field stimuli [26], [27]. Two fundamental differences distinguish visual half-field stimulation from our approach. First, in BCIs using visual half-field stimulation, a set of frequencies or combinations thereof is used to encode different channels, whereas different locations in the visual field of the user encode channels in the proposed approach. Second, data classification is based on spectral features of the EEG signals elicited by the visual half-field stimuli, while our approach is based on the differences in the SSVEP topographies when the user gazes at different targets.

In our study we implemented a nine-target SSVEP BCI using only one centrally presented steady-state flicker stimulus. Different targets were placed in different spatial locations relative to the flicker stimulus. Overt attention to different targets resulted in distinct spatial topographies of the elicited SSVEP responses, which constituted the basis for BCI classification. As the very first study of this new type of SDMA BCI, the contribution of the present study is threefold: first, we designed and implemented a nine-target SDMA BCI paradigm; second, we developed a Canonical Correlation Analysis (CCA) based algorithm for efficient extraction of the SSVEP spatial topographies under multiple overt attention conditions; third, we investigated the possible number of targets for reliable BCI control based on the offline classification performances. The offline results show that the proposed BCI is a promising candidate for simple but efficient BCI communication.

II. Methods

A. Stimulation

The main function of the stimulation setup is to generate a single visual flicker and to define fixation targets for the subjects which bring the flicker stimulus in different positions on the retina in relation to the fovea. This functionality is provided by a stimulus, which resembles a clock-face (Fig. 1), but has only 8 digits evenly distributed along the perimeter and an additional digit in the center. The dial is presented at a visual angle of 27° (∼26 cm) on an LCD computer monitor (24 in, DELL 2405FPW, 55 cm viewing distance) and flickers at 15 Hz, a quarter of the refresh rate. In humans, the amplitude of the SSVEP response peaks at this frequency [28]–[30], and at the size of the stimulus SSVEPs are reliably evoked in all subjects.

For obtaining training data for the classifier, the gaze direction of the subjects is cued by a small black disc (2°). The target digit and the cue disc in the background are not flickering; therefore, subjects always fixate stationary spots on the screen.

B. Subjects and Experimental Procedure

Twelve subjects aged between 23 and 45 (6 females, average: 29 years) participated voluntarily in this study. They were

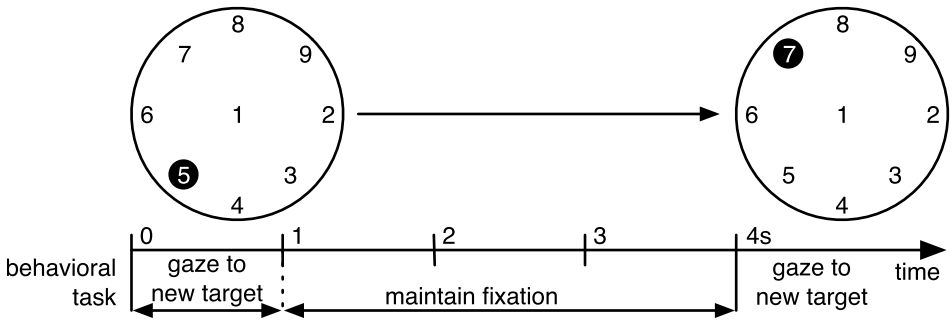

Fig. 1. Stimulus (top) and timing (bottom) of the experiment. The large white disc (clock-face) flickers at the stimulation frequency. Targets are indicated by a steady black background. The example shows the target sequence "5", "7".

mostly university students who had previously participated in electrophysiological studies but not studies using visual flicker or other BCI paradigms. They gave written informed consent and received financial compensation. All of them had normal vision and were free of neurological and ophtalmological disorders. The study was approved by the ethics committee of the medical association of the city of Hamburg, Germany.

Subjects were cued to fixate all 9 targets in random order. They were instructed to pay overt attention to each new target as soon as possible and then maintain fixation until the next target was cued. They were requested to avoid head movements, eye blinks, swallowing or any other muscular activity apart from moving the eyes to the target. Each target was shown for 4 s, and 1 s at the beginning of this epoch allowed for the time needed to adjust the gaze direction. After the presentation of all 9 targets, the screen was blanked for 2 s before the next random sequence of targets was presented. Subjects could blink or swallow in this interval. Ten sequences were presented in a block. Then a short break of 1-2 min was administered. The first block was used to familiarize subjects with the stimulation and the task. Then, another six or seven blocks were recorded in total, depending on the ability of the subject to avoid eye blinks. This resulted in 60 repetitions (trials) per class in 9 of the twelve subjects, and 70 repetitions per class in the remaining 3. An experiment lasted between 1 and 1.5 h. Like in most SSVEP BCI paradigms, subjects did not perform any training prior to the data recording. As the spatial attention task was relatively easy and straightforward, we only asked the subjects to gaze at the cued digit with no additional behavioral measurements such as eye-tracking or EOGs. Nevertheless, compliance with the task was checked intermittently using a web-cam at the top of the screen that was looking at the subject's face. The high classification accuracies suggest that all subjects reliably carried out the task.

The stimulation software was written in Matlab (The Mathworks, Natick, MA, USA) using the Psychophysics Toolbox extensions [31]–[33].

C. EEG Recording and Data Preprocessing

A 32-channel EEG was recorded using BioSemi's ActiveTwo AD-box (BioSemi instrumentation, Amsterdam, The Netherlands). It uses active electrodes referenced to an active common-mode-sense electrode and a passive driven-right-leg electrode. Electrodes were placed according to the 10-20 international system. During stimulation, the data buffer of the Fieldtrip toolbox [34] was running and storing EEG data on a PC's hard disk at a sample rate of 2048 Hz. Recordings took place in a regular lab environment with ambient illumination from ceiling lights and without any electrical or acoustic shielding. Subjects were seated in a reclining chair and encouraged to find a comfortable position.

Data were preprocessed offline by a notch filter to remove line noise and a 1–60 Hz bandpass filter. They were then downsampled to 512 Hz. In order to arrive at a conservative estimate of the performance and to simulate the data processing in a hypothetical online version of the paradigm, we did not perform any artifact rejection.

D. Data Analysis and Classification

In principle, the distribution of the power and phases of the SSVEP across all channels could be used for classification. A differential weighting of the channels can facilitate the classification process though, because channels have a distinct sensitivity to power modulations caused by different positions of the stimulus in the visual field. Moreover, it can be expected that only the power modulations in very narrow frequency bands around the stimulation frequency and its harmonics carry information for target recognition. These two aspects make our data a use case for Canonical Correlation Analysis (CCA), which was shown to exceed the frequency recognition accuracy of traditional power spectral density approaches in the context of SSVEP-based BCIs [35]. The main idea of this approach is to find a set of linear combinations between channels to maximize the correlations between two multivariate signals. For frequency recognition, a sine wave at the stimulation frequency is correlated with the EEG signal observed in N channels. Technically, instead of a single sine wave, a set consisting of a sine and a cosine wave as well as their harmonics is used since the phase relation between the template and the EEG signal is not known, and considering harmonics can improve the recognition accuracy.

In formal terms, CCA determines the canonical coefficients in matrices $\mathbf{A}$ and $\mathbf{B}$ such that the canonical correlations $\mathbf{r} = [\rho_1, \ldots, \rho_M]$ between $\mathbf{AX}$ and $\mathbf{BY}$ of two multivariate signals $\mathbf{X}$ and $\mathbf{Y}$ become maximal. Here $\rho_i = \rho\,(\mathbf{a}(i)\mathbf{X}, \mathbf{b}(i)\mathbf{Y})$ is the ith canonical correlation ($\mathbf{a}_i$ and $\mathbf{b}_i$ being the ith row

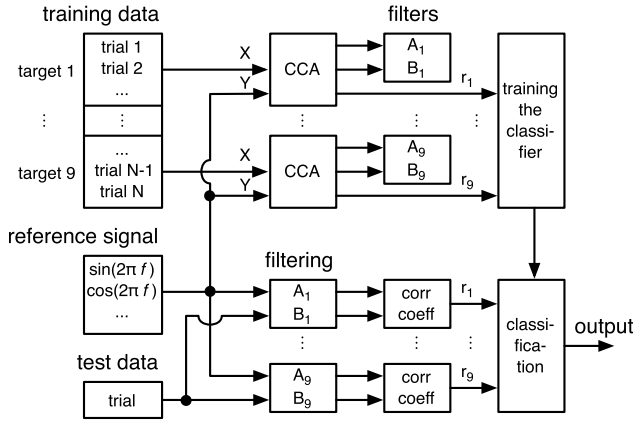

Fig. 2. Schematic of the data analysis for training (upper part) and testing (lower part) the BCI.

of **A** and **B** respectively), and $M = \min(\text{rank}(\mathbf{X}), \text{rank}(\mathbf{Y}))$ is the number of canonical variables. Typically only the first canonical correlation is used in subsequent analyses, but here we apply the full vector **r** in the classification.

We employ CCA for determining the canonical correlations **r** between the EEG signal $\mathbf{X} = [x_1(t)', \ldots, x_L(t)']$, L the number of channels, and a set of sine and cosine functions at the stimulation frequency f_{stim} and the second and third harmonics:

$$\mathbf{Y} = \begin{bmatrix} \sin(2\pi f_{stim} t) \\ \cos(2\pi f_{stim} t) \\ \sin(4\pi f_{stim} t) \\ \cos(4\pi f_{stim} t) \\ \sin(6\pi f_{stim} t) \\ \cos(6\pi f_{stim} t) \end{bmatrix}' \quad (1)$$

We chose to include the second and third harmonics as we had observed strong responses at these frequencies in almost all subjects. All 32 channels were used for the CCA, i.e. $L = 32$.

The training phase of the BCI system comprises two steps, computation of the filters and training of the classifier. First, coefficient matrices **A** and **B** are computed for each of the $C = 9$ target classes separately. Trials with EEG data when the subject attended to target c are concatenated and submitted to a CCA, yielding matrices $\mathbf{A}_c$ and $\mathbf{B}_c$ for each of the C targets. Then, the coefficients of correlation between each trial filtered with all C filters $\mathbf{A}_c\mathbf{X}$ ($c = 1 \ldots C$) and the corresponding reference signals $\mathbf{B}_c\mathbf{Y}$ are computed. The resulting $C \times M$ correlation coefficients constitute the feature vector for this trial, which together with the trial's label is used to train the classifier. We use multi-class linear discriminant analysis [36]. This method projects multi-dimensional feature vectors in a space in which the classes are maximally separated.

In the application phase, data from the trial to be classified is filtered with all C filters $\mathbf{A}_c$, and the correlation coefficients with the corresponding reference signals $\mathbf{B}_c\mathbf{Y}$ are computed. The resulting correlations constitute the feature vector that is forwarded to the classifier, which then outputs the recognized target. Fig. 2 summarizes the training and testing procedures in a flow diagram.

The proposed correlation-based feature extraction method accounts for both, amplitude and phase information: given a fixed pair of coefficient matrices **A** and **B**, data from two trials with similar SSVEP amplitudes but opposite SSVEPs phases result in canonical correlation coefficients of opposite signs and similar absolute values. Thereby the distinct amplitude-phase response patterns for different BCI commands (as observed in Fig. 3) can be effectively captured.

The focus of the current study is to get an initial estimate of the number of output channels that could be controlled by a spatial SSVEP BCI. To derive this estimate, we grouped the data in all possible combinations of 2, 3, 4 etc. out of 9 classes and assessed the classification accuracy for each combination by leave-one-sample-out cross-validation. The combination that achieved the lowest validation error was selected as the representative for the corresponding group.

In a similar manner, we analyzed the effect of the number of EEG channels on the classification accuracy. For each channel, we ran a cross validation on the data with this channel removed. Afterwards, the channel whose removal retained the highest classification accuracy was dropped, and the procedure was iterated on the remaining channels until 3 channels were left.

In order to estimate the calibration time in a potential online application of the proposed paradigm, we determined the relation between the number of training trials and the resulting classification accuracies. From the full data set of each subject, we generated a series of subsets of increasing size by randomly selecting trials until the required number for each class was reached. Then leave-one-sample-out cross-validation was used to assess the classification accuracy for this subset as described before. This process of randomly subsampling the training data and cross-validating the accuracy was repeated 100 times and yielded an average classification accuracy for a training set of the given size.

ITR was computed as

$$ITR = \frac{60}{T}\left(\log_2 C + P\log_2 P + (1-P)\log_2 \frac{1-P}{C-1}\right) \quad (2)$$

where P is classification accuracy, C the number of classes and T the time required for computing the output [37]–[39].

III. Results

In order to verify our central hypothesis, that presenting visual flicker at different positions in the visual field generates distinct distributions of SSVEPs, we analyzed the SSVEP response when subjects fixated the different targets. Figure 3 shows examples from the three subjects with the highest classification accuracies.

The SSVEP power topographies (see Fig. 3a–c) show in general the typical power maximum over occipital and parietal areas. A systematic shift of the power maximum between right and left occipital areas when gazing at the targets in a clockwise sequence can be observed. When subjects gazed at target 6, for example, the flicker stimulus appeared in their right visual hemifield, hence an increase in the activity over the visual cortex of the contralateral, i.e. left, hemisphere was expected. The opposite relation holds for target 2. However, subject b) shows activation maxima over the

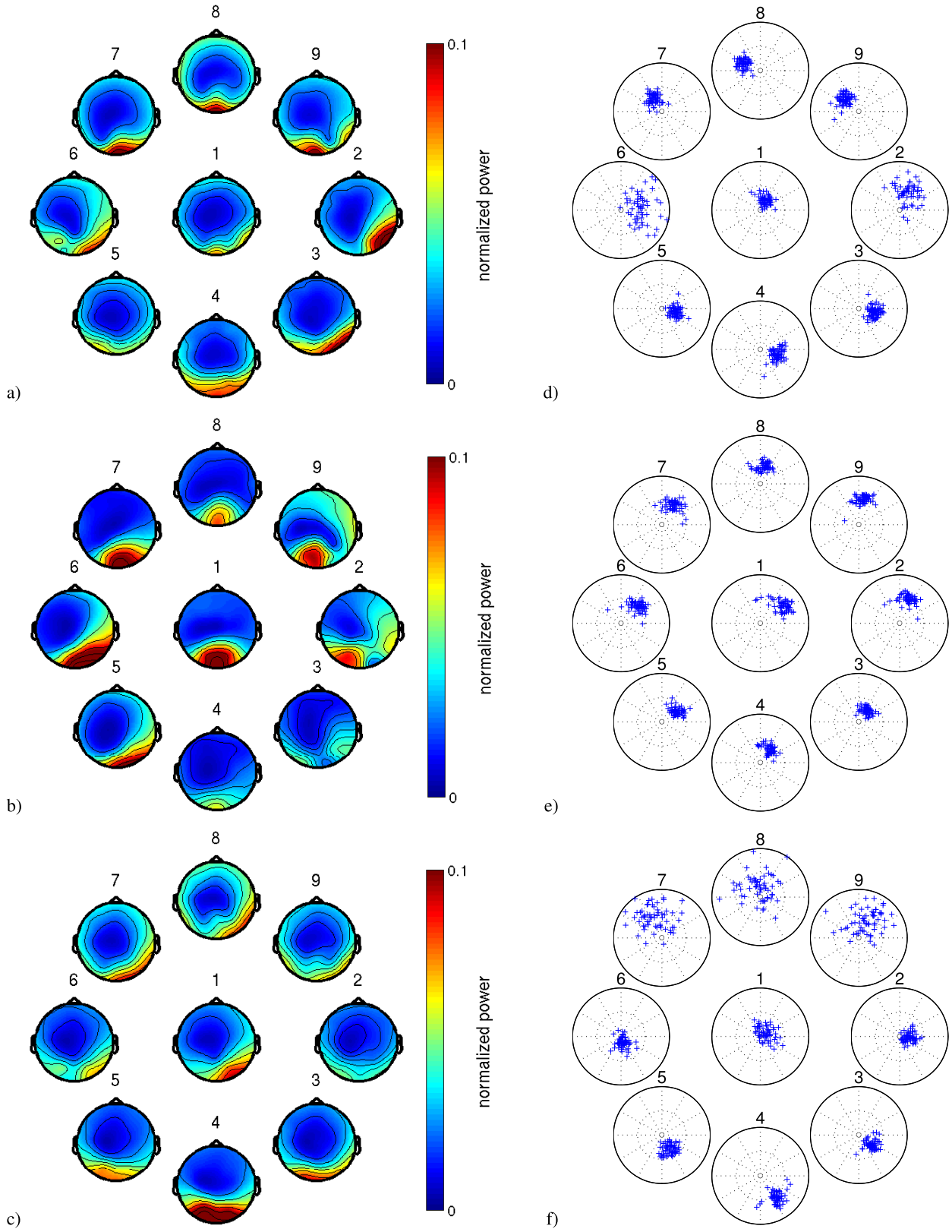

Fig. 3. **a-c)** Individual topographies of normalized power in the three subjects with the highest accuracy at 9 classes. Color represents power at the stimulation frequency relative to the neighboring frequency bins (signal-to-noise ratio). **d-e)** Phase-amplitude plots of the single-trial SSVEP responses in the same three subjects.

ipsilateral hemisphere, a paradoxical lateralization as described by [13]. A similar trend can be seen for subject c) although less clearly.

A comparison of the topographies along the vertical midline reveals a shift of the power distribution towards the central region. Another interesting comparison in subjects a) and c) is

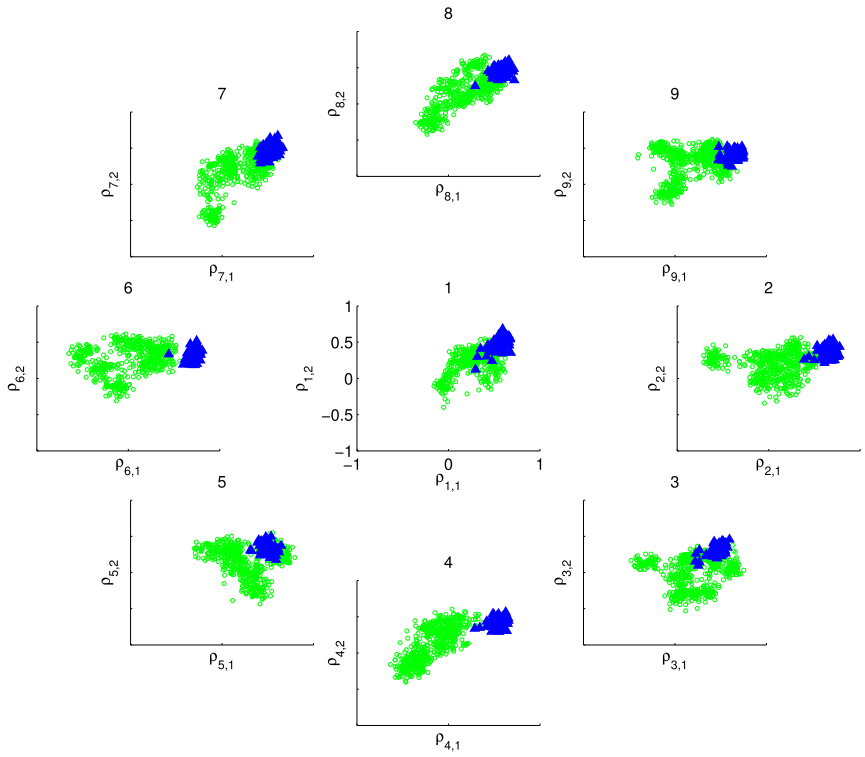

Fig. 4. Scatter plot of the first two correlation coefficients (ρ_1 and ρ_2) for trials when the respective target was attended (blue triangles) or unattended (green circles). Data from the subject with the best performance.

between corresponding targets above and below the horizontal midline, i.e. 7 vs. 5 and 9 vs. 3. It suggests a flip of the power maximum to the opposite hemisphere when the stimulus crosses the horizontal midline, a finding that was also described by [14].

Next we analyzed the amplitude and phase response. Figure 3d–e shows polar plots of the single-trial SSVEP responses at occipital electrode O1. Distinct amplitude-phase patterns were observed when the subjects attended to different targets. Like in the power topographies, substantial variability across subjects could be observed.

Figure 3 demonstrates distinct patterns when subjects gazed at the different targets, suggesting that a BCI which is based on the automatic classification of spatial SSVEP amplitude and phase responses is feasible. To substantiate this hypothesis, we show the distribution of the features that are input to the classifier for the three subjects with the best performance in Fig. 4. Only the first two correlations of the filtered data with the class-specific reference signal are shown in the two axes of the scatter plots, i.e. $\rho_{c,1}$ and $\rho_{c,2}$ of the correlation vector $\mathbf{r_c}$ between $\mathbf{A}_c\mathbf{X}$ and $\mathbf{B}_c\mathbf{Y}$. The distribution shows that the CCA successfully determined filters that maximize the correlation for trials when the subject attended the respective class; data from trials when other targets were attended, and that were filtered with the same filter, yielded considerably smaller correlations. The good separation between attended and unattended trials when only the first two correlations are used as features lets expect a high accuracy when all canonical correlations are used for classification.

We show these classification accuracies as a function of the number of targets to be distinguished in Fig. 5a. The best binary classification accuracy reached 100% for most subjects, although the corresponding target combinations were subject specific. As the number of control channels increases, the recognition accuracy decreases and reaches an average of 95% (±3.2%, range 88–98%) when all nine targets are used. The shown values result from a leave-one-sample-out cross-validation of classifying the recorded data and can be considered as an upper bound for the corresponding values that could be expected in an online version of the paradigm.

As Fig. 5a shows only the highest classification accuracy for the corresponding number of classes, we analyzed in more detail the distribution of these accuracies for the subject with the best performance. The size of the dots in Fig. 5b reflects the number of combinations of two, three etc. targets that were classified with the accuracy given on the ordinate. When the BCI employs only two or three targets, the majority of combinations yields a classification accuracy of 100%. When more classes (4–9) are used, classification accuracies seem to follow a normal distribution with a mean at around 98%. In a real application, naturally the combination yielding the highest accuracy would be selected.

Looking at the maximum accuracies for different numbers of control channels, we were wondering if there are target combinations with good separability that are common to all subjects, and whether our hypothesis about the the prevalence of the top-bottom combinations over left-right combinations can be confirmed. We therefore counted for each combination

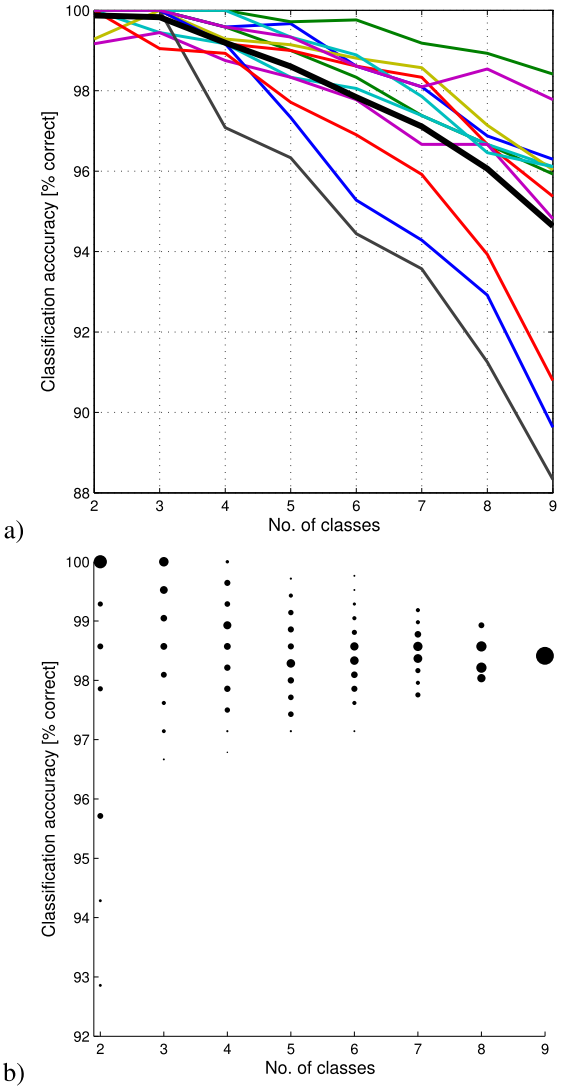

Fig. 5. **a)** Maximum classification accuracy when subsets of the nine classes were used. Each line corresponds to one subject; the thick black line is the average accuracy. **b)** Distribution of classification accuracies across the different combinations of the number of classes given on the abscissa. The size of each dot represents the number of such combinations that afford the accuracy given on the ordinate. Dot sizes are normalized to sum up to the same value for each number of classes. Data are from a single subject (best performance). Chance levels decrease monotonically from 50% for 2 classes to 11% for 9 classes.

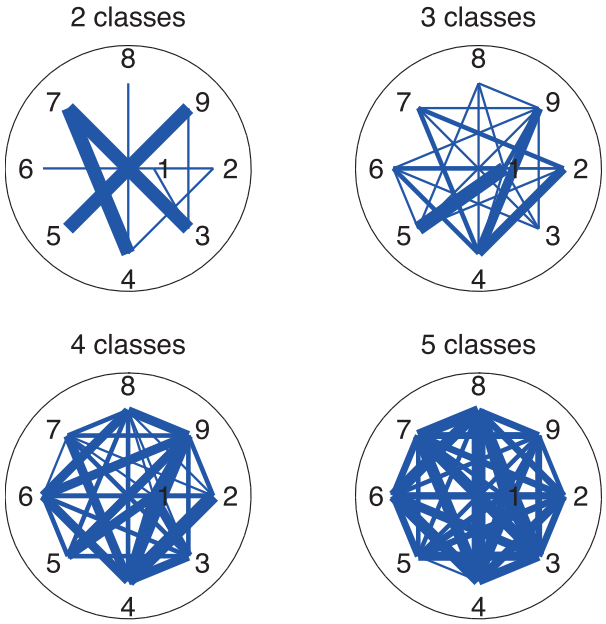

Fig. 6. Lines connect class sets of highest discriminability when 2, 3, 4 or 5 classes have to be distinguished. Line thickness shows how frequently the corresponding class pair was selected for best discriminability over all subjects. Target 1 was slightly set off the center to minimize visual overlap.

of targets the number of subjects that achieved their best accuracy with this combination. These counts are visualized in Fig. 6. For two classes, there is a clear advantage for target combinations in the most distant quadrants of the visual field. When three classes are used, the central target was frequently among the best combinations. The set of optimal combinations of four targets resembles the one for three targets, with a prominent appearing of the target at the bottom of the stimulus (4). The set of five targets is dominated again by targets in upper and lower hemifield.

We computed the ITR resulting from the obtained classification accuracies and analyzed the influence of the trial length (Fig. 7). The ITR is rising with the number of classes and reaches an average of about 40.8±3.3 bits/min across subjects for nine classes (Fig. 7a). This estimation is based on the whole duration of 4 s for each target. Fig. 7b) shows that for all subjects the ITR could also be improved by using shorter trials. The average ITR increases to 51.4±5.5 bits/min when trials of 3 s duration are used. Even shorter trials of 2s can further increase the average ITR to 65.2±15.4 bits/min at the expense of increasing the standard deviation. Hence performance across users is more robust for longer trials. Note that for all trial lengths, the first second, during which the user adjust the gaze direction, is always discarded; therefore, in a trial length of 2 s, for example, only 1 s of data is actually used for the classification.

The former results were obtained by including all 32 channels in the classification procedure. Frequency-coded SSVEP BCIs, in contrast, often use the signal from only a single or a few electrode(s). An interesting question therefore is if the proposed spatially-coded SSVEP BCI requires the EEG signal from all over the scalp, or if local signals from a few electrodes could be sufficient for achieving the same performance. We addressed this question empirically by another off-line analysis of the data to determine the relevance of each channel. We devised an iterative schema which in each step determined the electrode with the least negative impact on the classification accuracy (estimated by leave-one-sample-out cross-validation, cf. Fig. 5). This electrode was dropped from further analyzes and the next iteration was started.

The effect of the gradual reduction in the number channels that were employed in the classification is shown in Fig. 8a. For all subjects this curve exhibits an inverted U-shape, with a slight increase in classification accuracy when reducing the number of used channels to about 20, a plateau until the channel number is further reduced to about 8 and a decrease for a reduction down to 3 channels. It is interesting

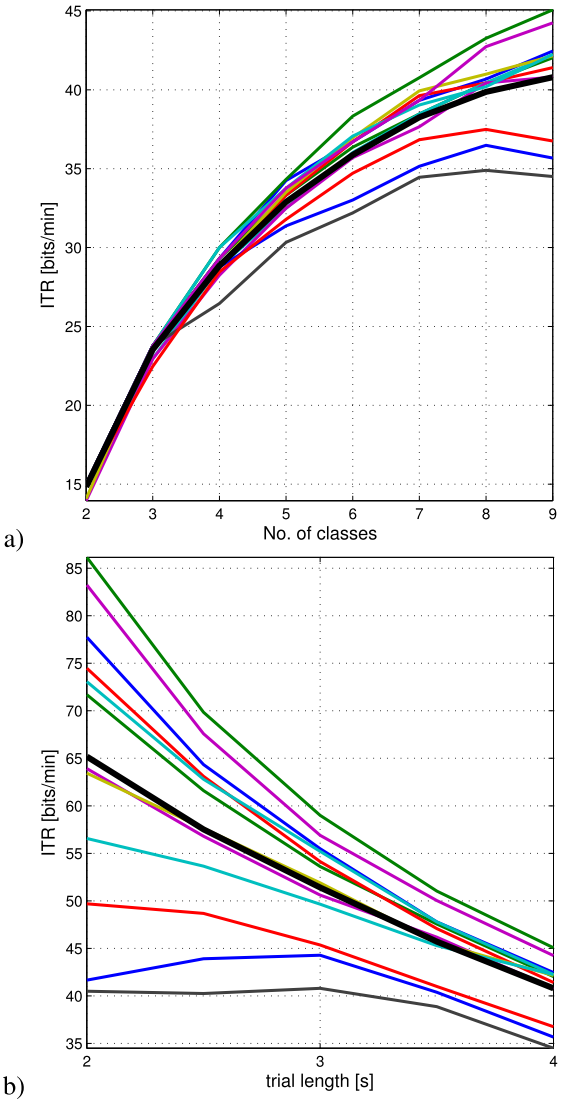

Fig. 7. **a)** ITR for different numbers of classes. **b)** ITR as a function of the trial length. Each line corresponds to one subject; the thick black line is the average.

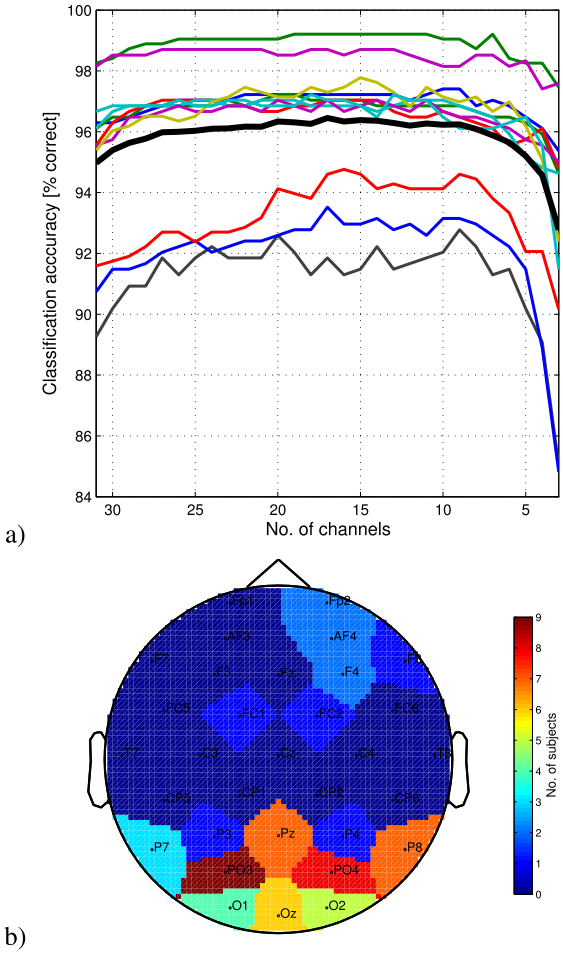

Fig. 8. **a)** Maximum classification accuracy (estimated by leave-one-sample-out cross-validation) depending on the number of EEG channels. Each line corresponds to one subject; the thick black line is the average accuracy. **b)** Number of subjects for which the respective electrode contributed to the highest accuracy when the classification was based on only 5 EEG channels (out of 32).

to note that the accuracies for operating the classification with only 5 channels is the same or slightly better than with all 32 channels in all subjects.

To find out which the most informative channels in this BCI approach are, we counted for each electrode in how many subjects this electrode was among the set of 5 that "survived" the iterative dropping schema. Figure 8b shows that in most subjects the classification relied on channels at positions Pz, PO3/4, P7/8, Oz and O1/2. The signals from these electrodes seem to have the highest relevance for the high classification accuracies that we achieved with the proposed BCI approach.

The effect that the size of the training set has on the classification accuracy can be seen in Fig. 9. On average (black curve), accuracy decreases by about 5% when only 20 instead of 60 trials per class comprise the training data. Thus training time could be reduced from about 50 minutes (like in the current study) to about 17 minutes with only a minor drop in performance. For the best two subjects, even a single training session with 10 trials per class (taking about 8 minutes) would be enough to achieve a classification accuracy of more than 90%. Subjects with a poorer performance have the steepest learning curves, and with 60 training trials per class they can achieve a similar performance.

IV. Discussion

In this study, we presented a single visual flicker stimulus of fixed frequency and position to subjects. By directing their attention overtly to different target positions on the screen, the stimulus was projected to different positions on their retinas. We then investigated the retinotopic mapping to the visual brain areas by analyzing power and phase distributions of the SSVEP response. We found that these distributions are sufficiently distinct to allow a classifier to recognize which target the subject was attending to from the EEG data.

Whereas the SSVEP responses were highly idiosyncratic regarding the spatial topographies of the amplitude and phase responses, they were consistent enough within each subject

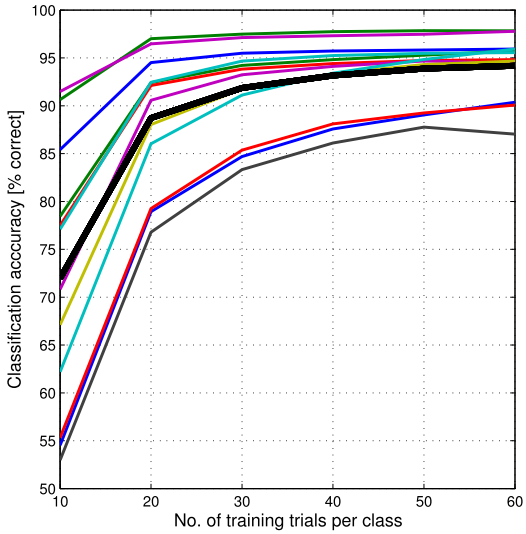

Fig. 9. Classification accuracy (estimated by leave-one-sample-out cross-validation) as a function of the number of training trials per class. Each line corresponds to one subject; the thick black line is the average accuracy.

to provide an average classification accuracy of 95±3% for recognizing nine targets. Our method constructs target-specific spatial filters by CCA and uses the canonical correlations between the filtered data and a reference signal for classification. EOG and other possible muscle artifacts are not likely to induce a systematic bias for the observed high classification accuracies, because the CCA method targets specifically the stimulation frequency and locations over occipital and parietal brain regions. Our results are in line with previous neurophysiological findings on retinotopic mapping in the human brain [13], [14], [17]. We made a further step to show that this mapping can be accurately identified at the single-trial level using EEG.

Conventional SSVEP BCIs usually employ one stimulation frequency for each output command. Different output commands are coded by the SSVEP frequency when the subjects are directly attending to one of the steady-state flicker stimuli. Therefore, the number of targets available on the computer screen is restricted by factors such as the screen's refresh rate and the responsive frequency range of the human visual system. The proposed BCI system significantly extends the conventional SSVEP BCIs and achieves multiple output commands using only one flicker stimulus. In our design, different output commands are coded by different spatial locations relative to this stimulus, and the elicited SSVEP topographies serve as the carrier of information about the target of attention. As the proposed SSVEP BCI relies on a spatial coding strategy rather than the frequency coding strategy, the number of BCI output commands is determined by the spatial "resolution" of the retinotopic mapping as revealed by human EEG recordings. Our results suggest the possibility of operating the proposed SDMA BCI with up to nine simultaneous output commands. Whether more output commands can be included will be investigated in our future studies. Our findings about the retinotopic mapping of flicker stimuli can potentially be incorporated in frequency-coded SSVEP BCIs by extending each frequency-coded target to at least up to nine spatial targets, which may greatly increase the number of output commands.

Our paradigm has a number of advantages over conventional SSVEP BCIs: First, it uses only a single stimulation frequency, yet several control channels can be encoded. This has a number of implications. Adjusting stimulation frequencies to the subject's individual response profile is simplified, because the proposed paradigm requires only a single frequency with a suitable SSVEP response. A single stimulation frequency may also turn out advantageous for improving the system's ITR because cortical areas need not be entrained by different frequencies with each new target selected. We plan to verify these hypotheses with an online version of the proposed method in a subsequent study. Parameters like the number of electrodes or the size of the analysis window will also be optimized in a follow-up study.

Second, the proposed interface is very economical with screen space. It does not require a subdivision of the screen into areas that generate the stimulation for the respective control channel. It uses instead a single stimulation area, and control channels figure on the screen just as points. Actually the targets need not appear at all on the screen; they could be likewise marked at the screen's frame. Moreover, the screen contents itself could be used as the SSVEP stimulus, be it by blanking it or by reversing its contrast at the stimulation frequency. A straightforward application for the present system could be a virtual phone with digits on a dial like in the now historic analog telephones.

And third, the simplicity of the overall system together with the previously noted advantages allows us to anticipate mobile applications using tablet PCs and even smart phones. The precious screen space on mobile devices needs not be traded off against the number of control channels. Moreover, our explorative study revealed that high BCI performances could be achieved with only 5 electrodes (plus reference), which indicates that the proposed system can be implemented with both simple, portable stimulation as well as recording devices. Of course such applications require further investigations of the dependency between the size of the stimulus and the classification performance. The aim of this study was to investigate the feasibility of a spatial SSVEP BCI. We did not attempt to optimize neither the performance (e.g., number of targets, trial duration, classification accuracy) nor user comfort (e.g., number of electrodes). For instance, although subjects do not have to gaze directly at a flickering screen area (as in f-VEP BCIs), the relatively large size of the surrounding flicker is likely to induce visual fatigue or discomfort. Nonetheless our study introduces a BCI method as an alternative to existing, frequency-coded SSVEP BCIs. We conjecture that it shares many of the advantages with these existing approaches, like, for example, the modest training requirements, the resilience against EEG artifacts or the simple stimulation setup, but that more research is needed to reveal the most powerful approach.

We conclude by contemplating about the question why spatial SSVEP BCIs have attracted relatively little attention so far. We think that a major reason may lie in the quite

diverse and at times inconsistent findings on the lateralization of SSVEPs for extrafoveally presented stimuli as outlined in the introduction. We admit that the current study does not contribute much to clarify the causes of the considerable variability of lateralized SSVEP topographies across subjects or to systematize them. However, we demonstrate that within each subject, the change with stimulus position is systematic enough to enable an automatic recognition of the target of overt attention. Notwithstanding that many details of our method require further elaboration in follow-up studies, we think that we have demonstrated the potential of the lateralization effect for BCI applications.

REFERENCES

[1] D. Regan, "Recent advances in electrical recording from the human brain," *Nature*, vol. 253, no. 5491, pp. 401–407, Feb. 1975.

[2] F. Di Russo, W. A. Teder-Sälejärvi, and S. A. Hillyard, "Steady-state VEP and attentional visual processing," in *The Cognitive Electrophysiology of Mind and Brain*. A. Zani and A. M. Proverbio, Eds. 2002, pp. 259–274.

[3] F.-B. Vialatte, M. Maurice, J. Dauwels, and A. Cichocki, "Steady-state visually evoked potentials: Focus on essential paradigms and future perspectives," *Prog. Neurobiol.*, vol. 90, no. 4, pp. 418–438, Apr. 2010.

[4] X. Gao, D. Xu, M. Cheng, and S. Gao, "A BCI-based environmental controller for the motion-disabled," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 11, no. 2, pp. 137–140, Jun. 2003.

[5] D. Zhang, A. Maye, X. Gao, B. Hong, A. K. Engel, and S. Gao, "An independent brain–computer interface using covert non-spatial visual selective attention," *J. Neural Eng.*, vol. 7, no. 1, p. 016010, 2010.

[6] X. Chen, Z. Chen, S. Gao, and X. Gao, "A high-ITR SSVEP-based BCI speller," *Brain-Comput. Interfaces*, vol. 1, nos. 3–4, pp. 181–191, 2014.

[7] X. Chen, Y. Wang, M. Nakanishi, X. Gao, T.-P. Jung, and S. Gao, "High-speed spelling with a noninvasive brain–computer interface," *Proc. Nat. Acad. Sci. USA*, vol. 112, no. 44, pp. E6058–E6067, 2015. [Online]. Available: http://www.pnas.org/content/112/44/E6058.short and http://www.pnas.org/content/112/44/E6058.full

[8] G. Bin, X. Gao, Y. Wang, B. Hong, and S. Gao, "VEP-based brain-computer interfaces: Time, frequency, and code modulations [research frontier]," *IEEE Comput. Intell. Mag.*, vol. 4, no. 4, pp. 22–26, Nov. 2009.

[9] S. Gao, Y. Wang, X. Gao, and B. Hong, "Visual and auditory brain–computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 5, pp. 1436–1447, May 2014.

[10] I. Volosyak, "SSVEP-based Bremen–BCI interface—Boosting information transfer rates," *J. Neural Eng.*, vol. 8, no. 3, p. 036020, 2011.

[11] E. Yin, Z. Zhou, J. Jiang, Y. Yu, and D. Hu, "A dynamically optimized SSVEP brain–computer interface (BCI) speller," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 6, pp. 1447–1456, Jun. 2015.

[12] R. H. Wurtz and E. R. Kandel, "Central visual pathways," in *Principles of Neural Science*, vol. 4, E. R. Kandel, J. H. Schwartz, and T. M. Jessell, Eds. New York, NY, USA: McGraw-Hill, 2000, pp. 523–547.

[13] G. Barrett, L. Blumhardt, A. M. Halliday, E. Halliday, and A. Kriss, "A paradox in the lateralisation of the visual evoked response," *Nature*, vol. 261, no. 5557, pp. 253–255, May 1976.

[14] H. A. Baseler, E. E. Sutter, S. A. Klein, and T. Carney, "The topography of visual evoked response properties across the visual field," *Electroencephalogr. Clin. Neurophysiol.*, vol. 90, no. 1, pp. 65–81, Jan. 1994.

[15] H. Baseler and E. E. Sutter, "M and P components of the VEP and their visual field distribution," *Vis. Res.*, vol. 37, no. 6, pp. 675–690, Mar. 1997.

[16] M. D. Rugg, C. R. Lines, and A. D. Milner, "Further investigation of visual evoked potentials elicited by lateralized stimuli: Effects of stimulus eccentricity and reference site," *Electroencephalogr. Clin. Neurophysiol./Evoked Potentials Sec.*, vol. 62, no. 2, pp. 81–87, Mar. 1985.

[17] F. Di Russo *et al.*, "Identification of the neural sources of the pattern-reversal VEP," *NeuroImage*, vol. 24, no. 3, pp. 874–886, Feb. 2005.

[18] G. Harding, B. Janday, and R. Armstrong, "The topographic distribution of the magnetic P100M to full- and half-field stimulation," *Documenta Ophthalmol.*, vol. 80, no. 1, pp. 63–73, Mar. 1992.

[19] S. G. Thorpe, P. L. Nunez, and R. Srinivasan, "Identification of wave-like spatial structure in the SSVEP: Comparison of simultaneous EEG and MEG," *Statist. Med.*, vol. 26, no. 21, pp. 3911–3926, Sep. 2007.

[20] J. Ding, G. Sperling, and R. Srinivasan, "Attentional modulation of SSVEP power depends on the network tagged by the flicker frequency," *Cerebral Cortex*, vol. 16, no. 7, pp. 1016–1029, Jul. 2006.

[21] J. J. Vidal, "Real-time detection of brain events in EEG," *Proc. IEEE*, vol. 65, no. 5, pp. 633–641, May 1977.

[22] N. Yoshimura and N. Itakura, "Usability of transient VEPs in BCIs," in *Recent Advances in Brain-Computer Interface Systems*, R. Fazel, Ed. Rijeka, Croatia: InTech, 2011, ch. 6, pp. 119–134.

[23] Y. Punsawad and Y. Wongsawat, "Multi-command SSVEP-based BCI system via single flickering frequency half-field stimulation pattern," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug./Sep. 2011, pp. 1101–1104.

[24] C.-H. Han, H.-J. Hwang, and C.-H. Im, "Classification of visual stimuli with different spatial patterns for single-frequency, multi-class SSVEP BCI," *Electron. Lett.*, vol. 49, no. 22, pp. 1374–1376, Oct. 2013.

[25] A. Materka and M. Byczuk, "Alternate half-field stimulation technique for SSVEP-based brain-computer interfaces," *Electron. Lett.*, vol. 42, no. 6, pp. 321–322, Mar. 2006.

[26] Z. Yan, X. Gao, G. Bin, B. Hong, and S. Gao, "A half-field stimulation pattern for SSVEP-based brain-computer interface," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Sep. 2009, pp. 6461–6464.

[27] Z. Yan, X. Gao, and S. Gao, "Right-and-left visual field stimulation: A frequency and space mixed coding method for SSVEP based brain-computer interface," *Sci. China Inf. Sci.*, vol. 54, no. 12, pp. 2492–2498, Dec. 2011.

[28] C. S. Herrmann, "Human EEG responses to 1–100 Hz flicker: Resonance phenomena in visual cortex and their potential correlation to cognitive phenomena," *Experim. Brain Res.*, vol. 137, nos. 3–4, pp. 346–353, 2001.

[29] M. A. Pastor, J. Artieda, J. Arbizu, M. Valencia, and J. C. Masdeu, "Human cerebral activation during steady-state visual-evoked responses," *J. Neurosci.*, vol. 23, no. 37, pp. 11621–11627, 2003.

[30] Y. Wang, R. Wang, X. Gao, B. Hong, and S. Gao, "A practical VEP-based brain-computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 2, pp. 234–240, Jun. 2006.

[31] D. G. Pelli, "The VideoToolbox software for visual psychophysics: Transforming numbers into movies," *Spatial Vis.*, vol. 10, no. 4, pp. 437–442, 1997.

[32] D. H. Brainard, "The psychophysics toolbox," *Spatial Vis.*, vol. 10, no. 4, pp. 433–436, 1997.

[33] M. Kleiner, D. Brainard, D. Pelli, A. Ingling, R. Murray, and C. Broussard, "What's new in psychtoolbox-3?" *Perception*, vol. 36, no. 14, pp. 1–16, 2007.

[34] R. Oostenveld, P. Fries, E. Maris, and J.-M. Schoffelen, "FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data," *Comput. Intell. Neurosci.*, vol. 2011, Oct. 2011, Art. no. 156869.

[35] Z. Lin, C. Zhang, W. Wu, and X. Gao, "Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pp. 2610–2614, Dec. 2006.

[36] R. C. Rao, "The utilization of multiple measurements in problems of biological classification," *J. Roy. Statist. Soc.*, vol. 10, no. 2, pp. 159–203, 1948.

[37] J. R. Wolpaw, H. Ramoser, D. J. McFarland, and G. Pfurtscheller, "EEG-based communication: Improved accuracy by response verification," *IEEE Trans. Rehabil. Eng.*, vol. 6, no. 3, pp. 326–333, Sep. 1998.

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

[39] P. Yuan, X. Gao, B. Allison, Y. Wang, G. Bin, and S. Gao, "A study of the existing problems of estimating the information transfer rate in online brain–computer interfaces," *J. Neural Eng.*, vol. 10, no. 2, p. 026014, 2013.

Alexander Maye received the degree in computer science in Dresden and Berlin, with a focus on the dynamics of models of neuronal oscillations, and the Ph.D. degree from the Institute of Technology in Berlin in 2002. He was with the Zuse Institute Berlin, Berlin, where he was involved in brain atlases. He was with the University Medical Center Hamburg-Eppendorf, where he was involved in combine computational modeling and electrophysiological studies of the human brain.

Dan Zhang received the B.E. degree in automation and the Ph.D. degree in biomedical engineering from Tsinghua University, Beijing, China, in 2005 and 2011, respectively. He was a Post-Doctoral Fellow with the School of Medicine, Tsinghua University, from 2011 to 2013. He is currently an Associate Professor with the Department of Psychology, Tsinghua University. His research interests include social neuroscience, engineering psychology, and brain-computer interfaces.

Andreas K. Engel received the degree in medicine and philosophy from the Technical University of Munich in 1987. From 1987 to 1995, he was a Staff Scientist with the Max Planck Institute for Brain Research, Frankfurt, where he developed a long-standing interest in the dynamics of sensory processing, intermodal and sensorimotor integration, and theories of perception, action, attention, and consciousness. In 1996, he established an independent group with the Max Planck Institute for Brain Research, funded by the Heisenberg Program of the German Research Foundation. In 2000, he was with the Research Center Jülich to set up the newly established Cellular Neurobiology Group, Institute for Medicine. In 2002, he was the Chair of Neurophysiology with the University Medical Center, Hamburg-Eppendorf. He is currently a Professor of Physiology and the Head of the Department of Neurophysiology and Pathophysiology with the University Medical Center Hamburg-Eppendorf.