Spatiotemporal Beamforming: A Transparent and Unified Decoding Approach to Synchronous Visual Brain-Computer Interfacing

Benjamin Wittevrongel* and Marc M. Van Hulle

Laboratory for Neuro- and Psychophysiology, Department of Neurosciences, KU Leuven, Leuven, Belgium

Brain-Computer Interfaces (BCIs) decode brain activity with the aim to establish a direct communication channel with an external device. Albeit they have been hailed to (re-)establish communication in persons suffering from severe motor- and/or communication disabilities, only recently BCI applications have been challenging other assistive technologies. Owing to their considerably increased performance and the advent of affordable technological solutions, BCI technology is expected to trigger a paradigm shift not only in assistive technology but also in the way we will interface with technology. However, the flipside of the quest for accuracy and speed is most evident in EEG-based visual BCI where it has led to a gamut of increasingly complex classifiers, tailored to the needs of specific stimulation paradigms and use contexts. In this contribution, we argue that spatiotemporal beamforming can serve several synchronous visual BCI paradigms. We demonstrate this for three popular visual paradigms even without attempting to optimizing their electrode sets. For each selectable target, a spatiotemporal beamformer is applied to assess whether the corresponding signal-of-interest is present in the preprocessed multichannel EEG signals. The target with the highest beamformer output is then selected by the decoder (maximum selection). In addition to this simple selection rule, we also investigated whether interactions between beamformer outputs could be employed to increase accuracy by combining the outputs for all targets into a feature vector and applying three common classification algorithms. The results show that the accuracy of spatiotemporal beamforming with maximum selection is at par with that of the classification algorithms and interactions between beamformer outputs do not further improve that accuracy.

Keywords: event-related potential, steady-state visual evoked potential, code-modulated visual evoked potential, BCI, P300, spatiotemporal filter

1. INTRODUCTION

In a Brain-Computer Interface (BCI), signals are recorded from the brain with the aim to enable users to interact with an external device, without the need for muscular, vocal or other means of communication (Vidal, 1973). Typically, the subject is presented with a number of targets, each one representing a specific action (e.g., a letter in a spelling interface, a movement of a prosthetic limb or a wheelchair, etc.) that can be selected either by redirecting gaze or attention to the intended

target (**Figure 1**) or by performing a mental action (e.g., imagined movement; Pfurtscheller and Neuper, 2001). Among the most performant BCI paradigms, in terms of accuracy and speed, are the ones where targets are visually stimulated and the user directs his/her gaze to target to be selected (for review, see Bin et al., 2009; Nicolas-Alonso and Gomez-Gil, 2012). Each stimulation paradigm elicits specific brain potentials, which are recorded from the scalp using electroencephalography (EEG) (Lotte et al., 2007), from the cortical surface using electrocorticography (ECoG) (Schalk and Leuthardt, 2011) or from implants such as microelectrode arrays and depth electrodes (Maynard et al., 1997; Kennedy et al., 2004). For each paradigm, analysis- and decoding techniques have been proposed tailored to maximize target prediction speed and accuracy.

One popular paradigm is based on the P300 event-related potential (ERP), an EEG component recorded over the centro-parietal region that exhibits a positive deflection in amplitude, peaking around 300 ms (whence its name), in sync with the onset of an infrequent stimulus (called oddball) to which the subject pays attention. In a P300-based BCI, the targets are overlaid with the oddball stimulus in a non-overlapping fashion (serial stimulation). The P300 can be elicited in response to a variety of sensory stimuli, including visual- (Farwell and Donchin, 1988; Sellers and Donchin, 2006; Manyakov et al., 2010a), auditory- (Sellers and Donchin, 2006; Furdea et al., 2009; Schreuder et al., 2010), and tactile stimuli (Brouwer and Van Erp, 2010). EEG recordings are typically cut into epochs, in sync with possible oddball events. However, in order to address the relatively low signal-to-noise ratio (SNR), several oddball events need to be generated and epochs averaged before P300 responses can be detected in them, resulting in a reduction

in communication speed. P300-based BCIs have been most extensively studied, resulting in a wide range of techniques for EEG feature extraction and classification (Lotte et al., 2007). The most simple ones consider template matching in which a (spatio)temporal template is first learned from P300-labeled responses and then used to identify the intended target as the one with the highest correlation coefficient (Krusienski et al., 2006) or consider peak picking in which the target having the largest difference between the lowest amplitude prior to and the highest amplitude in the P300 time window is selected as winner (Farwell and Donchin, 1988). The Linear Discriminant Analysis (LDA) (Jin et al., 2013) and its variants, including Fisher LDA (Krusienski et al., 2006; Manyakov et al., 2011b), Bayesian LDA (Hoffmann et al., 2008; Manyakov et al., 2011b) and Stepwise LDA (Farwell and Donchin, 1988; Krusienski et al., 2006; Manyakov et al., 2011b), have been applied to pre-processed epochs or after feature extraction by Common Spatio-Temporal Pattern (Krusienski et al., 2007), xDAWN (Rivet et al., 2009) and linear predictive coding (Momennezhad et al., 2014). More complex methods include multilayer feed-forward neural networks (Gulcar et al., 1998; Manyakov et al., 2011b), Nuclear Norm on tensor representation (Hunyadi et al., 2013), Bayesian statistical classifiers (Pires et al., 2008), ensemble methods with SVMs and a genetic algorithm (Ghoggali et al., 2013), and a fuzzy fusion of the outcome of template matching and peak picking (Salimi Khorshidi et al., 2007; Salimi-Khorshidi et al., 2008). As far as we are aware, state-of-the-art P300 classification for BCI is achieved by the linear Support Vector Machine (SVM) (Krusienski et al., 2006; Combaz et al., 2009, 2013; Manyakov et al., 2011b) as it outperforms its non-linear (Krusienski et al., 2006; Manyakov et al., 2011b) counterparts. The recently

FIGURE 1 | Graphical representation of a visual BCI. The subject gazes at the target on a display ("interface"), to which the desired action is associated, while his/her brain activity is being recorded. Each target is stimulated visually, in a specific manner, which leaves a trace in the recorded brain activity. From the preprocessed recordings, signal features are extracted and the classification algorithm predicts the intended target, after which the display is updated accordingly (visual feedback) and the corresponding action performed.

introduced spatiotemporal beamforming filter has been shown to be at par with an optimized linear SVM for P300 detection (Wittevrongel and Van Hulle, 2016a).

Unlike the P300 paradigm, which is inherently serial in nature, the SSVEP paradigm encodes the targets using periodically and simultaneously flickering stimuli, and in which each target adopts a unique frequency (Regan, 1979; Middendorf et al., 2000), unique phase (Lee et al., 2010; Lopez-Gordo et al., 2010; Manyakov et al., 2012) or unique frequency-phase combination (Jia et al., 2011). EEG signals are recorded over the occipital cortex and are expected to carry the same signal characteristics as the target being gazed at. Hence, SSVEP decoding methods are traditional relying on frequency domain methods such as the Fourier transform (Gao et al., 2003) and the Continuous Wavelet Transform (Zhang et al., 2010). Other frequency-detection algorithms include the Minimum Energy Combination (Friman et al., 2007; Chumerin et al., 2013; Combaz et al., 2013), the Multivariate Synchronization Index (MSI) (Zhang et al., 2014a, 2017) and time-domain analysis (Luo and Sullivan, 2010; Manyakov et al., 2010b). Among the machine learning algorithms that have been used for SSVEP target identification are Bayesian LDA (Cecotti, 2010), SVM (Singla and Haseena, 2014), k-Nearest Neighbor (kNN) (Kwak et al., 2015), and several neural network architectures (Manyakov et al., 2011a, 2013; Singla and Haseena, 2014; Kwak et al., 2017). However, the most accurate SSVEP decoding methods are based on the Canonical Correlation Analysis (CCA) (Lin et al., 2007), and its numerous extensions (Pan et al., 2011; Zhang et al., 2011; Chen et al., 2014, 2015; Nakanishi et al., 2014; Zhang et al., 2014b; Vu et al., 2016), and spatiotemporal beamforming (Wittevrongel and Van Hulle, 2016b,c).

Whereas the P300 paradigm encodes targets in terms of ERP responses and SSVEP in terms of frequency and/or phase combinations, the code-modulated VEP (cVEP) paradigm encodes the targets using lagged versions of a binary coding sequence. While record-breaking communication speed was achieved when it was introduced, it has been considerably less studied (Gao et al., 2014), and fewer decoding algorithms have been proposed for this paradigm. Originally, template matching (Bin et al., 2009) was used to identify the gazed target from averaged epochs, but also the CCA algorithm has been applied (Bin et al., 2011; Aminaka et al., 2015a; Wei et al., 2016) as well as SVMs (Aminaka et al., 2015b,c) and spatiotemporal beamforming (Wittevrongel et al., 2017). The latter two have shown state-of-the-art performance with shorter stimulation times.

Importantly, what the majority of the reported decoding algorithms have in common is that they have been construed to maximize performance of a given paradigm, even when used in a particular context. We argue that one of these algorithms, the spatiotemporal beamformer, can achieve state-of-the-art performance in all three paradigms. The beamformer is a filter that estimates the contribution of an a-priori defined activation pattern (i.e., a template, a signal-of-interest) into new data. It can appear as a spatial, temporal or spatiotemporal filter. Unlike in our previous work (Wittevrongel and Van Hulle, 2016a,b; Wittevrongel et al., 2017), which was aimed at investigating

the beamformer in the context of a specific paradigm, the first goal of this manuscript is to describe a generalized and unified approach to spatiotemporal beamforming for decoding synchronous EEG responses by demonstrating its versatility in the context of synchronous visual BCI. Secondly, in addition to a simple maximum-beamformer output classification, we also investigate whether interactions between beamformer outputs can increase decoder accuracies. We end by indicating future directions for spatiotemporal beamforming in BCI.

2. METHODS

2.1. Data Acquisition and Processing

The data used in this study has been collected in previous offline BCI studies. For more details about the experimental setting, interface and recordings, we refer the reader to the corresponding publications. All three studies were carried out in accordance with the recommendations and approval of the ethics committee of our university hospital UZ Leuven with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki.

Data for the P300 paradigm was collected using an interface consisting of 9 targets (Wittevrongel and Van Hulle, 2016a). In each trial, subjects were asked to direct their gaze to a cued target and continue to focus as all targets were individually highlighted 15 times for 100 ms (with 100 ms inter-stimulus interval) in pseudorandom order. In response to the cued target, when highlighted, a P300 response will be elicited. All targets were used as cue 4 times, leading to a total of 36 trials, each consisting of 15 "P300 epochs" (i.e., where the cued target was highlighted) and 120 epochs where another target was highlighted. Data from 21 subjects was collected using an ActiveTwo system (BioSemi, The Netherlands) with 32 active Ag/AgCl electrodes (and 4 additional electrodes around the eyes) at a sampling rate of 2,048 Hz. The raw data was offline referenced from the Common-Mode Sense (CMS) reference to average of the mastoid signals, corrected for eye artifacts using the RAAA-method (Croft and Barry, 2000), filtered between 0.5 and 15 Hz and cut in 0.6-s epochs time-locked to the onset of the stimulation. Each epoch was baselined by subtracting the average of the 100 ms pre-onset activity and downsampled to 64 Hz.

SSVEP data was obtained from an interface consisting of 4 targets, each adopting an unique frequency-phase combination of 12 or 15 Hz and 0 or π radians (Wittevrongel and Van Hulle, 2016b). During a trial, subject were asked to direct their gaze to a cued target and continue to focus when all targets were simultaneously flickering at their unique frequency-phase combinations for 5 s. All targets were cued 15 times, leading to a total of 60 5-s trials. Data from 20 subjects was recorded using an ActiveTwo system (BioSemi, The Netherlands) with 32 active Ag/AgCl electrodes (and 4 additional electrodes around the eyes) at a sampling rate of 2048 Hz. The raw data was offline referenced from the CMS reference to the average of the mastoid signals, corrected for eye artifacts using the RAAA-method (Croft and Barry, 2000), filtered between 4 and 20 Hz, cut in 5-s trials time-locked to the onset of the flickering, and downsampled to 512 Hz.

For the cVEP study, a 32-target interface was used (Wittevrongel et al., 2017). Targets were encoded using lagged versions of a binary m-sequence of length 63, presented at a stimulation rate of 120 Hz (1 m-sequence = 0.525 s). Each trial started with a cue, to which the subjects need to direct their gaze and to continue focusing while all targets were simultaneously presented with their unique lagged m-sequences. The sequences were repeated 10 times. Each target was cued 5 times, leading to a total of 160 trials. Data from 17 subjects was recorded using a Neuroscan SynampsRT device (Compumedics, Australia) with 32 active Ag/AgCl electrodes at a sampling rate of 1,000 Hz with ground and reference electrodes at FPz and AFz, respectively. The data was offline referenced to the average of the mastoid signals, filtered between 4 and 31 Hz, cut into trials that are time-locked to the onset of the stimulation, and finally downsampled to 120 Hz. Since no ocular electrodes were included in the recording, the data was not corrected for eye artifacts.

2.2. Spatiotemporal Beamforming

The beamforming principle was originally introduced as a spatial filter for radar, sonar and seismic data analysis (Van Veen and Buckley, 1988), but has also been employed in EEG analysis to isolate the signal originating from a predefined brain location (Van Veen et al., 1997), for ERP analysis (Treder et al., 2016), and to build a BCI application based on imagined movement detection (Grosse-Wentrup et al., 2009).

van Vliet et al. (2015) introduced a spatiotemporal extension to beamforming for single-trial N400 ERP detection in the context of processing semantic stimuli. To calculate the spatiotemporal filter, a spatiotemporal activation pattern $\mathbf{A} \in \mathbb{R}^{m \times n}$ needs to be defined, where m represents the number of channels (here, EEG electrodes) and n the number of samples. This activation pattern represents the signal-of-interest (or template) and can be obtained a-priori (van Vliet et al., 2015) or based on actual training data (Wittevrongel and Van Hulle, 2016a,b; Wittevrongel et al., 2017). The resulting beamformer $\mathbf{w} \in \mathbb{R}^{1 \times mn}$ is a multivariate filter that optimally isolates the targeted response from noise and possible other unrelated activity, by taking into account the information contained in the covariance matrix $\mathbf{\Sigma} \in \mathbb{R}^{(mn) \times (mn)}$ estimated from the available (training) data.

The original formulation of the Linearly Constrained Minimum Variance (LCMV) spatial beamformer $\mathbf{w}_{sp} \in \mathbb{R}^{m \times 1}$ minimizes the variance of the beamformer output $\mathbf{w}_{sp}^{\mathsf{T}}\mathbf{S}$:

$$\arg\min_{\mathbf{w}_{sp}} \mathbf{w}_{sp}^{\mathsf{T}}\mathbf{S}(\mathbf{w}_{sp}^{\mathsf{T}}\mathbf{S})^{\mathsf{T}} \Rightarrow \arg\min_{\mathbf{w}_{sp}} \mathbf{w}_{sp}^{\mathsf{T}}\mathbf{\Sigma}_{sp}\mathbf{w}_{sp}, \quad (1)$$

where $\mathbf{\Sigma}_{sp} \in \mathbb{R}^{m \times m}$ is the spatial covariance matrix of an EEG trial $\mathbf{S} \in \mathbb{R}^{m \times n}$. By adding the linear constraint:

$$\mathbf{a}_{sp}^{\mathsf{T}}\mathbf{w}_{sp} = 1, \quad (2)$$

where $\mathbf{a}_{sp} \in \mathbb{R}^{m \times 1}$ is the spatial activation pattern, trivial solutions of (1) are avoided, and signals that are similar to $\mathbf{a}_{sp}$ will be mapped to a value close to 1, allowing for an easy measure of

similarity. The solution of (1) under constraint (2) can be found using the method of Lagrange multipliers (Van Veen et al., 1997):

$$\mathbf{w}_{sp} = \frac{\mathbf{\Sigma}_{sp}^{-1}\mathbf{a}_{sp}}{\mathbf{a}_{sp}^{\mathsf{T}}\mathbf{\Sigma}_{sp}^{-1}\mathbf{a}_{sp}}. \quad (3)$$

The spatial beamformer can be expanded to a spatiotemporal variant as follows. Let $S \in \mathbb{R}^{m \times n \times r}$ be r trials, $\mathbf{X} \in \mathbb{R}^{r \times (mn)}$ a matrix where each row l is obtained by concatenating the rows of a corresponding trial $S[*, *, l]$ with $l \in [1..r]$, $\mathbf{\Sigma} \in \mathbb{R}^{(mn) \times (mn)}$ the covariance matrix of $\mathbf{X}^{\mathsf{T}}$, and $\mathbf{a}^{\mathsf{T}} \in \mathbb{R}^{1 \times (mn)}$ a vector containing the concatenated rows of the spatiotemporal activation pattern $\mathbf{A}$. The spatiotemporal LCMV beamformer $\mathbf{w} \in \mathbb{R}^{(mn) \times 1}$ with the linear constraint $\mathbf{a}^{\mathsf{T}}\mathbf{w} = 1$ can now be calculated as:

$$\mathbf{w} = \frac{\mathbf{\Sigma}^{-1}\mathbf{a}}{\mathbf{a}^{\mathsf{T}}\mathbf{\Sigma}^{-1}\mathbf{a}}, \quad (4)$$

and applied to the data as a simple weighted sum:

$$y = \mathbf{sw}, \quad (5)$$

where $\mathbf{s} \in \mathbb{R}^{1 \times (mn)}$ indicates the concatenated rows of a segment $\mathbf{S}$ and y represents the contribution of the activation pattern $\mathbf{A}$ in $\mathbf{S}$.

2.3. Beamformer Construction and Feature Extraction

As the activation pattern captures the signal-of-interest, each paradigm (P300, SSVEP, cVEP) will have unique activation patterns, and since EEG responses might different across subjects, beamformers should be constructed for each subject individually (Wittevrongel and Van Hulle, 2016a). In the next few sections, we will describe for each paradigm how one or more beamformers (BF) are constructed and used to extract feature vectors. A visual depiction of this process is given in **Figure 2**.
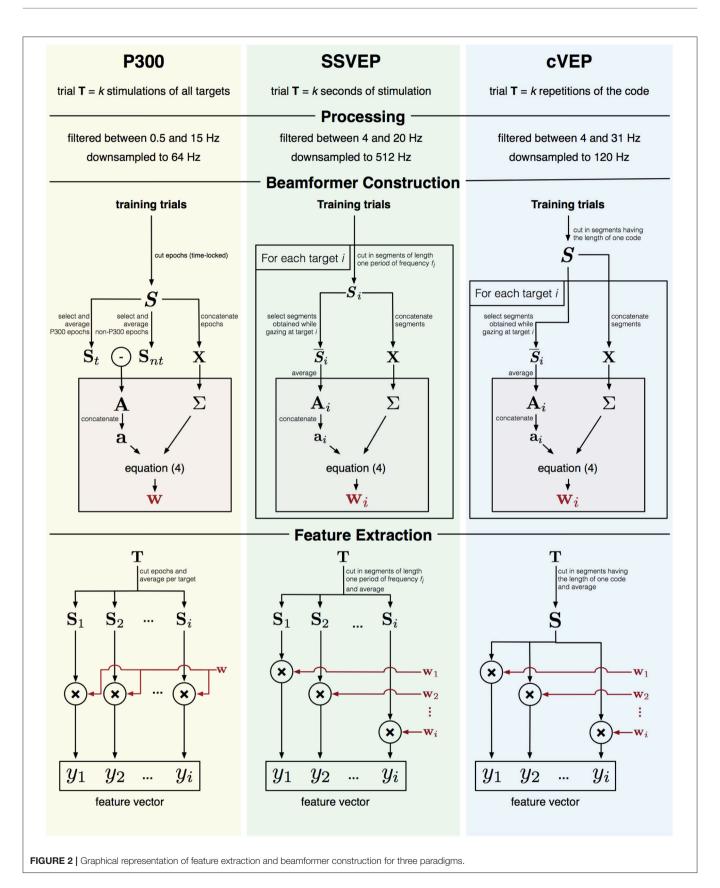
2.3.1. P300

For a P300-based BCI the activation pattern $\mathbf{A}$ represents the P300 ERP. One trial $\mathbf{T}$ in the P300 paradigm consists of k stimulations (i.e., highlights) of all targets. To obtain the activation pattern $\mathbf{A}$, all training trials are first cut into time-locked epochs $\mathbf{S}$ from 0 to 600 ms post-stimulus onset and their baseline subtracted. The activation pattern $\mathbf{A}$ can then be obtained as the difference between the average ($\mathbf{S}_t$) of the P300 epochs (i.e., when the target being gazed at is highlighted) and the average ($\mathbf{S}_{nt}$) of the non-P300 epochs. The covariance matrix $\mathbf{\Sigma}$ is calculated from $\mathbf{S}$ (using the concatenation approach described above), and a single beamformer $\mathbf{w}$ constructed.

To obtain the feature vector for a trial $\mathbf{T}$, the trial is first cut into time-locked epochs and all epochs corresponding to the same target are averaged. Each averaged epoch $\mathbf{S}_i \in \mathbb{R}^{m \times n}$ is then filtered by the beamformer $\mathbf{w}$ to obtain an value y_i, which represents the P300 contribution for target i.

2.3.2. SSVEP

Targets in the SSVEP paradigm are encoded with unique frequency-phase combinations, and one trial $\mathbf{T}$ consists of

FIGURE 2 | Graphical representation of feature extraction and beamformer construction for three paradigms.

k seconds of SSVEP stimulation, during which all targets are simultaneously flickering at their assigned frequency-phase combinations. Unlike for the P300 case, each target elicits a specific frequency and phase response, and it is therefore not possible to define a single activation pattern to discriminate targets, but rather multiple activation patterns, each one tailored to one target, are needed. To capture the frequency-phase response to target *i*, a time-domain analysis (Luo and Sullivan, 2010; Manyakov et al., 2010b) is adopted: each training trial $\mathbf{T}$ is cut into consecutive, non-overlapping segments $\mathbf{S}_i$ with a length *n* equal to a single period of stimulus frequency f_i. The activation pattern $\mathbf{A}_i$ for target *i* is then given by averaging the segments ($\overline{\mathbf{S}}_i \subset \mathbf{S}_i$) extracted from the trials during which target *i* was cued. The covariance matrix $\mathbf{\Sigma}$ was calculated from $\mathbf{S}_i$ and beamformer $\mathbf{w}_i$ from Equation (4).

The feature vector for a trial $\mathbf{T}$ contains the beamformer outputs y_i for each target *i*. To obtain y_i, the trial is cut into segments of length one period of frequency f_i, the segments are averaged and filtered using the corresponding beamformer $\mathbf{w}_i$.

To improve classification accuracy, the first 120 ms of each trial were removed prior to cutting the segments, as the SSVEP response is not stable during this time. (Nakanishi et al., 2014; Wittevrongel and Van Hulle, 2016b).

2.3.3. cVEP
In the cVEP paradigm, targets were encoded using lagged versions of a binary m-sequence, which was repeated *k* times over the course of one trial $\mathbf{T}$. Similar to the SSVEP paradigm, each target elicited a unique response and several activation patterns (one for each target) were defined. Each training trial $\mathbf{T}$ was cut into a maximal number of 0.525-s consecutive, non-overlapping segments $\mathbf{S}$. Note that the 0.525 s corresponds to the time needed to display one full m-sequence. The activation pattern $\mathbf{A}_i$ for target *i* was then obtained as the average of the segments $\overline{\mathbf{S}}_i \subset \mathbf{S}$ in response to target *i*. As for SSVEP, the covariance matrix $\mathbf{\Sigma}$ was calculated from $\mathbf{S}$, and beamformer $\mathbf{w}_i$ from Equation (4).

The feature vector for a trial is constructed by cutting the trial into segments (using the same method as above), averaging the segments ($\mathbf{S}$) and independently applying the beamformer $\mathbf{w}_i$ to obtain the corresponding y_i for each target *i*.

2.4. Target Identification
Based on the feature vector, a prediction for the intended target was made. We compared several prediction strategies. As the feature vectors contain estimates of the degree to which the activation pattern of each target is present in the trial, the most naive prediction is to select the target with the highest score (i.e., prediction = max(y_i)). However, as alternatives, we also applied three common classifiers (Nearest Neighbor (NN), LDA and SVM) to the features vectors. These classifiers can take into account possible interactions between beamformer outputs of different targets (i.e., features in the feature vector). The feature vectors of the training trials were used to train these classifiers. The complete procedure is depicted in **Figure 3**. The three classifiers were implemented by using the corresponding built-in functions in Matlab (2015) (*fitcdiscr, fitcknn, fitcecoc*).

2.5. Performance Estimation
Classification accuracy was estimated using cross-validation (4-fold for the P300 data and 5-fold for SSVEP and cVEP data), for increasing signal lengths (i.e., stimulus repetitions for P300 and cVEP and stimulation length for SSVEP).

For all paradigms, the same electrode set was included in the analysis, consisting of *Fz, Cz, Pz, Oz, O1, O2, PO3, PO4, P3,* and *P4*. The locations of the included electrodes (**Figure 4**) cover the scalp areas where routinely P300 (centro-parietal), SSVEP (occipital) and cVEP (parieto-occipital) responses are detected. Note that we do not attempt to optimize the electrode sets of the paradigms, but rather consider the same assembly for all three paradigms, which adds to the genericness of our approach.

2.6. Statistics
Since the distributions do not consistently follow a Gaussian distribution, we adopted the non-parametric (two-tailed) Wilcoxon signed rank test. We used this test to compare the accuracies of the different classifiers. The significance threshold was set to 0.0083 ($= \frac{0.05}{6}$) after applying the Bonferroni correction for multiple comparisons.
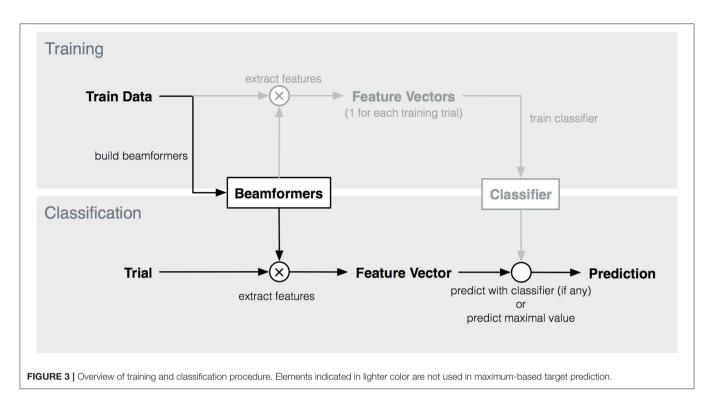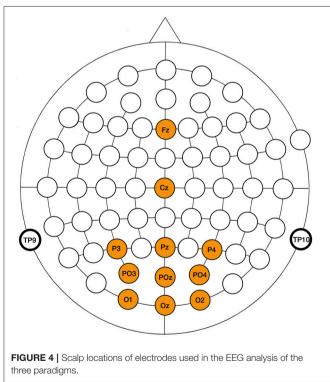
2.7. Data Availability
All data and analysis scripts are available at: https://kuleuven.box.com/v/SpatiotemporalBeamforming.

3. RESULTS
Figure 5A shows the classification accuracies of the four classifiers for the P300 paradigm with increasing stimulus repetitions (1 repetition = 1.8 s). For all stimulus repetitions in the P300 case, there is no significant difference between maximal beamformer output prediction and NN- or SVM-based classification. LDA-based prediction consistently has the lowest median accuracy and is significantly different up to 5 (compared to SVM, $p < 0.003$) and 10 (compared to max and NN, $p < 0.008$) stimulus repetitions. All four prediction approaches require at least two stimulus repetitions to surpass the 70% accuracy threshold (median accuracy of 86.1% for maximum-, 83.3% for kNN- and SVM- and 75.0% for LDA-based prediction) deemed necessary for establishing reliable communication (Kübler et al., 2004; Kübler and Birbaumer, 2008; Brunner et al., 2011; Combaz et al., 2013).

With the SSVEP paradigm (**Figure 5B**), the target identification accuracies of the four classification approaches are not significantly different for any of the stimulation lengths. The accuracies of all classifiers increase considerably up until 1.12 s stimulation, after which the maximal median accuracy of 100% is reached (for maximum-, kNN- and SVM-based prediction) and remains consistent for longer stimulation lengths. Note that the interquartile ranges (as well as the number of outliers) are considerably smaller compared to the P300 paradigm, indicating a higher consistency across subjects, which may be due to the fact that SSVEP is an automatic visual (sensory) response while P300 is a cognitive potential and thus more influenced by ongoing mental activity (e.g., attention).

FIGURE 3 | Overview of training and classification procedure. Elements indicated in lighter color are not used in maximum-based target prediction.

FIGURE 4 | Scalp locations of electrodes used in the EEG analysis of the three paradigms.

Similar to the P300-case, for the cVEP paradigm (**Figure 5C**), LDA-based prediction consistently exhibits the lowest median accuracy for all signal lengths. While the maximal-based prediction is significantly different from NN for a stimulation length of 1.05 s ($p = 0.006$, two-tailed Wilcoxon signed rank
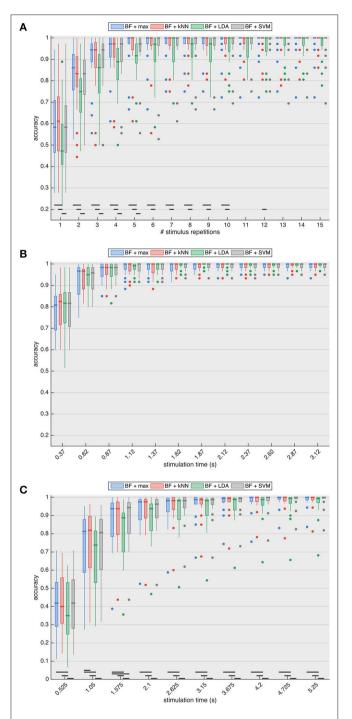
test) and from SVM for 1.625 s stimulation ($p = 0.002$, two-tailed Wilcoxon signed rank test), maximal-, NN- and SVM-based predictions are not significantly different for any other stimulation lengths. Two repetitions of the m-sequence (i.e., stimulation length of 1.05 s) are necessary for the median to surpass the 70% accuracy threshold (median accuracy of 81.3% for maximum-, 81.9% for kNN-, 80.6% SVM-, and 73.7% for LDA-based prediction). For all classifiers, the accuracies steadily increase with increasing stimulation length.

4. DISCUSSION

In this manuscript, we described a transparent and unified decoding approach to synchronous visual BCI, and showed its feasibility for three popular paradigms: P300, SSVEP, and cVEP. The feature vector of a trial was determined as the output of spatiotemporal beamformer(s), and we compared four methods for predicting the intended target.

The accuracies obtained by the NN- and SVM-based predictions are, with two exceptions for the cVEP paradigm, not significantly different from maximum-based prediction. The additional information given by the interaction between the beamformer outputs is not beneficial for target identification, and is even detrimental when adopting an LDA-based classifier. Furthermore, maximum-based prediction has a computational advantage as no feature extraction and classifier training is required after building the spatiotemporal beamformer(s) (see **Figure 3**).

For all three tested paradigms, an identical electrode set was included in the analysis to demonstrate the versatility and robustness of the spatiotemporal beamforming approach. As the
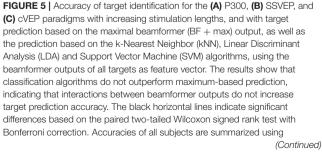
FIGURE 5 | Accuracy of target identification for the **(A)** P300, **(B)** SSVEP, and **(C)** cVEP paradigms with increasing stimulation lengths, and with target prediction based on the maximal beamformer (BF + max) output, as well as the prediction based on the k-Nearest Neighbor (kNN), Linear Discriminant Analysis (LDA) and Support Vector Machine (SVM) algorithms, using the beamformer outputs of all targets as feature vector. The results show that classification algorithms do not outperform maximum-based prediction, indicating that interactions between beamformer outputs do not increase target prediction accuracy. The black horizontal lines indicate significant differences based on the paired two-tailed Wilcoxon signed rank test with Bonferroni correction. Accuracies of all subjects are summarized using

(Continued)

aim of this study was to describe a unified methodology for synchronous BCI, the accuracies reported in this manuscript are likely not optimal and could be improved by optimizing the electrode set (Lal et al., 2004; Schröder et al., 2005; Lv and Liu, 2008; Arvaneh et al., 2011; Barachant and Bonnet, 2011; Wittevrongel et al., 2017) or the filtering range (Song and Epps, 2007; Manyakov et al., 2010a) for each subject individually. In our previous work (Wittevrongel and Van Hulle, 2016a,b; Wittevrongel et al., 2017), we have shown that the accuracies obtained by the spatiotemporal beamforming approach are at par of even outperform the state-of-the-art classifiers of the respective paradigms, and that training beamformers is often faster than alternative classification methods such as SVM, which would allow experimenters to explore the aforementioned optimizations within a reasonable amount of time.

Both spatial and spatiotemporal filtering serve to improve the SNR, which allows for more accurate ERP analysis or classification performance. While the LCMV beamformer approach with a data-driven activation pattern has also been applied to spatial filtering (Treder et al., 2016), the spatiotemporal filter used in this study has the advantage to jointly model both the spatial and temporal characteristics of the signal-of-interest. For the P300 ERP, for example, it has been shown that target prediction is more accurate when the early visual ERP components (e.g., N200 elicited over the occipital cortex) are included in the analysis (Bianchi et al., 2010; Kaufmann et al., 2011a), and many studies have therefore investigated stimulation paradigms that elicit additional ERP components (Kaufmann et al., 2011b; Jin et al., 2012, 2015). A spatial filter tuned to the centro-parietal P300 component would filter out these additional components, causing a reduction in prediction accuracy.

In recent years, various complex algorithms have been proposed that often rely on (multiple) extensions of an existing algorithm and that are tailored to a specific paradigm (e.g., the extensions for MSI Zhang et al., 2017 and CCA Lin et al., 2007; Pan et al., 2011; Zhang et al., 2011; Chen et al., 2014, 2015; Nakanishi et al., 2014; Zhang et al., 2014b; Vu et al., 2016 for SSVEP detection). Other proposed approaches apply a black-box classification algorithm (SVM, Neural Networks), where the experimenter does not have control over the pattern on which target discrimination is based. However, in the context of BCI, this pattern is in most cases well known, and in previous work we have shown that for example the pattern learned by a state-of-the-art black-box SVM in the context of a P300-BCI is very similar to the activation pattern obtained by simple averaging such as in the spatiotemporal beamformer approach (Wittevrongel and Van Hulle, 2016a). Unlike these algorithms, the beamforming approach described in this manuscript is less complex (as desired,

cfr. Occam's razor), fully transparent and applicable to various synchronous stimulation paradigms.

In future work, we will further develop the spatiotemporal beamformer. At this point, the main bottleneck is to estimate the high-dimensional covariance matrix, which, to be accurate, requires a significant amount of training data (cfr. the curse of dimensionality Pruzek, 1994; Schoukens and Pintelon, 2014). Compared to the spatial beamformer, the covariance matrices in the spatiotemporal variant consider the complete temporal dimension of the signal-of-interest for each evaluated channel (hence its dimension of $(mn) \times (mn)$, with m the number of channel and n the number of (time) samples). For the 10 channels selected in our study, the covariance matrices had the following dimensions: 380 ($=$ 10 channels $\times$ 38 samples) for P300; 420 or 340 (depending on the frequency) for SSVEP; and 630 for cVEP. Furthermore, since our approach requires spatiotemporal samples, training data is considerably harder to obtain. For a spatial beamformer, one typically can collect hundreds of samples in one second, while for the spatiotemporal variant the number of samples that can be collected per second depends on the temporal characteristics of the signal-of-interest. In our study, the P300 epochs were 0.6 s long (albeit that they are partially overlapping), SSVEP segments required 1/12 or 1/15 s of stimulation (depending on the frequency), and one signal-of-interest in the cVEP paradigm was 0.525 seconds long, leading to a mere 3 ($=$ 0.6 s with 0.2 s stimulus onset asynchrony), 12 or 15, and 1.9 samples extracted from the first second of stimulation for P300, SSVEP, and cVEP, respectively. In addition to reducing the number of channels included, alternative methods for the estimation of the covariance matrix can be considered. In this and previous studies, we estimated a pooled covariance matrix by including non-targeted segments for the covariance matrix estimation, which is often used in cases when the amount of samples is smaller than the dimensions of the covariance matrix (e.g., facial recognition, Edwards et al., 1998, BCI, Vidaurre and Blankertz, 2010). However, several methods for covariance matrix estimation with a small number of samples have been described (e.g., shrinkage, Ledoit and Wolf, 2004, regularization, Friedman, 1989, among others) and can be assessed in terms of improving the estimation of the covariance matrix.

At this time, we applied the beamformer for synchronous signals, but it can likely be extended for asynchronous detection of spatiotemporal signals of interest. As the beamformer outputs can be interpreted, the spatiotemporal beamformer in combination with a sliding window could be used as transformation of multi-electrode EEG to a single-trace estimate of the contribution of a spatiotemporal template in ongoing EEG recordings. An important aspect of this development will be to limit to amount of false positive detections.

5. CONCLUSION

In this study, we described spatiotemporal beamforming as a unified approach for target identification in BCI applications that rely on EEG responses synchronized to the stimulation onset. We have shown the versatility of our approach by applying it to three popular visual stimulation paradigms (P300, SSVEP, and cVEP), and demonstrated its robustness since high accuracies were reached with an identical (thus, non-optimized) electrode set for all paradigms. Unlike other classification algorithms, the spatiotemporal beamformer is fairly straightforward and fully transparent as the experimenter has full control over the activation pattern used for target discrimination. Finally, we showed that selecting the target with the highest beamformer output suffices to achieve a competitive accuracy, and that by accounting for possible interactions between beamformer outputs, accuracy does not further improve.

AUTHOR CONTRIBUTIONS

BW conducted the analysis. Both authors wrote and reviewed the manuscript.

FUNDING

REFERENCES

Aminaka, D., Makino, S., and Rutkowski, T. M. (2015a). "Chromatic and high-frequency cvep-based bci paradigm," in *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE* (Milan: IEEE), 1906–1909.

Aminaka, D., Makino, S., and Rutkowski, T. M. (2015b). "Classification accuracy improvement of chromatic and high–frequency code–modulated visual evoked potential–based bci," in *International Conference on Brain Informatics and Health* (London: Springer), 232–241.

Aminaka, D., Makino, S., and Rutkowski, T. M. (2015c). "Svm classification study of code-modulated visual evoked potentials," in *Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2015 Asia-Pacific* (Hong-Kong: IEEE), 1065–1070.

Arvaneh, M., Guan, C., Ang, K. K., and Quek, C. (2011). Optimizing the channel selection and classification accuracy in eeg-based bci. *IEEE Trans. Biomed. Eng.* 58, 1865–1873. doi: 10.1109/TBME.2011.2131142

Barachant, A., and Bonnet, S. (2011). "Channel selection procedure using riemannian distance for bci applications," in *2011 5th International IEEE/EMBS Conference on Neural Engineering (NER)* (Cancun: IEEE), 348–351.

Bianchi, L., Sami, S., Hillebrand, A., Fawcett, I. P., Quitadamo, L. R., and Seri, S. (2010). Which physiological components are more suitable for visual erp based brain–computer interface? a preliminary meg/eeg study. *Brain Topogr.* 23, 180–185. doi: 10.1007/s10548-010-0143-0

Bin, G., Gao, X., Wang, Y., Hong, B., and Gao, S. (2009). VEP-based brain-computer interfaces: time, frequency, and code modulations [research frontier]. *IEEE Comput. Intell. Mag.* 4, 22–26. doi: 10.1109/MCI.2009. 934562

Bin, G., Gao, X., Wang, Y., Li, Y., Hong, B., and Gao, S. (2011). A high-speed bci based on code modulation vep. *J. Neural Eng.* 8:025015. doi: 10.1088/1741-2560/8/2/025015

Brouwer, A.-M., and Van Erp, J. B. (2010). A tactile p300 brain-computer interface. *Front. Neurosci.* 4:19. doi: 10.3389/fnins.2010.00019

Brunner, C., Allison, B., Altstätter, C., and Neuper, C. (2011). A comparison of three brain-computer interfaces based on event-related desynchronization, steady state visual evoked potentials, or a hybrid approach using both signals. *J. Neural Eng.* 8:025010. doi: 10.1088/1741-2560/8/2/025010

Cecotti, H. (2010). "Classification of steady-state visual evoked potentials based on the visual stimuli duty cycle," in *2010 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL)* (Rome: IEEE), 1–5.

Chen, X., Wang, Y., Gao, S., Jung, T.-P., and Gao, X. (2015). Filter bank canonical correlation analysis for implementing a high-speed ssvep-based brain–computer interface. *J. Neural Eng.* 12:046008. doi: 10.1088/1741-2560/12/4/046008

Chen, X., Wang, Y., Nakanishi, M., Jung, T.-P., and Gao, X. (2014). "Hybrid frequency and phase coding for a high-speed ssvep-based bci speller," in *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE* (Chicago, IL: IEEE), 3993–3996.

Chumerin, N., Manyakov, N. V., van Vliet, M., Robben, A., Combaz, A., and Van Hulle, M. M. (2013). Steady-state visual evoked potential-based computer gaming on a consumer-grade eeg device. *IEEE Tran. Comput. Intell. AI Games* 5, 100–110. doi: 10.1109/TCIAIG.2012.2225623

Combaz, A., Chatelle, C., Robben, A., Vanhoof, G., Goeleven, A., Thijs, V., et al. (2013). A comparison of two spelling brain-computer interfaces based on visual p3 and ssvep in locked-in syndrome. *PLoS ONE* 8:e73691. doi: 10.1371/journal.pone.0073691

Combaz, A., Manyakov, N. V., Chumerin, N., Suykens, J. A., and Van Hulle, M. M. (2009). "Feature extraction and classification of eeg signals for rapid p300 mind spelling," in *International Conference on Machine Learning and Applications, 2009. ICMLA'09* (Miami Beach, FL: IEEE), 386–391.

Croft, R. J., and Barry, R. J. (2000). Removal of ocular artifact from the eeg: a review. *Neurophysiol. Clin.* 30, 5–19. doi: 10.1016/S0987-7053(00)00055-1

Edwards, G. J., Cootes, T. F., and Taylor, C. J. (1998). "Face recognition using active appearance models," in *European Conference on Computer Vision* (Freiburg: Springer), 581–595.

Farwell, L. A., and Donchin, E. (1988). Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalogr. Clin. Neurophysiol.* 70, 510–523. doi: 10.1016/0013-4694(88)90149-6

Friedman, J. H. (1989). Regularized discriminant analysis. *J. Am. Stat. Assoc.* 84, 165–175. doi: 10.1080/01621459.1989.10478752

Friman, O., Volosyak, I., and Graser, A. (2007). Multiple channel detection of steady-state visual evoked potentials for brain-computer interfaces. *IEEE Trans. Biomed. Eng.* 54, 742–750. doi: 10.1109/TBME.2006.889160

Furdea, A., Halder, S., Krusienski, D., Bross, D., Nijboer, F., Birbaumer, N., et al. (2009). An auditory oddball (p300) spelling system for brain-computer interfaces. *Psychophysiology* 46, 617–625. doi: 10.1111/j.1469-8986.2008.00783.x

Gao, S., Wang, Y., Gao, X., and Hong, B. (2014). Visual and auditory brain–computer interfaces. *IEEE Trans. Biomed. Eng.* 61, 1436–1447. doi: 10.1109/TBME.2014.2300164

Gao, X., Xu, D., Cheng, M., and Gao, S. (2003). A bci-based environmental controller for the motion-disabled. *IEEE Trans. Neural Syst. Rehabil. Eng.* 11, 137–140. doi: 10.1109/TNSRE.2003.814449

Ghoggali, S., Ungureanu, M., and Strungaru, R. (2013). A genetic algorithm and ensemble of svm approach for bci p300 classification problem. *Sci. Bull. Ser. C* 75, 129–142.

Grosse-Wentrup, M., Liefhold, C., Gramann, K., and Buss, M. (2009). Beamforming in noninvasive brain–computer interfaces. *Biomed. Eng. IEEE Trans.* 56, 1209–1219. doi: 10.1109/TBME.2008.2009768

Gulcar, H., Yilmaz, Y., and Demiralp, T. (1998). "Classification of p300 component in single trial event related potentials," in *Biomedical Engineering Days, 1998. Proceedings of the 1998 2nd International Conference* (Istanbul: IEEE), 48–50.

Hoffmann, U., Vesin, J.-M., Ebrahimi, T., and Diserens, K. (2008). An efficient p300-based brain–computer interface for disabled subjects. *J .Neurosci. Methods* 167, 115–125. doi: 10.1016/j.jneumeth.2007.03.005

Hunyadi, B., Signoretto, M., Debener, S., Van Huffel, S., and De Vos, M. (2013). "Classification of structured eeg tensors using nuclear norm regularization: improving p300 classification," in *2013 International Workshop on Pattern Recognition in Neuroimaging (PRNI)* (Philadelphia, PA: IEEE), 98–101.

Jia, C., Gao, X., Hong, B., and Gao, S. (2011). Frequency and phase mixed coding in ssvep-based brain–computer interface. *IEEE Trans. Biomed. Eng.* 58, 200–206. doi: 10.1109/TBME.2010.2068571

Jin, J., Allison, B. Z., Kaufmann, T., Kübler, A., Zhang, Y., Wang, X., et al. (2012). The changing face of p300 bcis: a comparison of stimulus changes in a p300 bci involving faces, emotion, and movement. *PLoS ONE* 7:e49688. doi: 10.1371/journal.pone.0049688

Jin, J., Sellers, E. W., Zhang, Y., Daly, I., Wang, X., and Cichocki, A. (2013). Whether generic model works for rapid erp-based bci calibration. *J. Neurosci. Methods* 212, 94–99. doi: 10.1016/j.jneumeth.2012.09.020

Jin, J., Sellers, E. W., Zhou, S., Zhang, Y., Wang, X., and Cichocki, A. (2015). A p300 brain–computer interface based on a modification of the mismatch negativity paradigm. *Int. J. Neural Syst.* 25:1550011. doi: 10.1142/S0129065715500112

Kaufmann, T., Hammer, E. M., and Kübler, A. (2011a). "ERPs contributing to classification in the P300 BCI," in *Proceedings of the Fifth International BCI Conference* (Graz: Graz University of Technology), 136–139.

Kaufmann, T., Schulz, S., Grünzinger, C., and Kübler, A. (2011b). Flashing characters with famous faces improves erp-based brain–computer interface performance. *J. Neural Eng.* 8:056016. doi: 10.1088/1741-2560/8/5/056016

Kennedy, P. R., Kirby, M. T., Moore, M. M., King, B., and Mallory, A. (2004). Computer control using human intracortical local field potentials. *IEEE Trans. Neural Syst. Rehabil. Eng.* 12, 339–344. doi: 10.1109/TNSRE.2004.834629

Krusienski, D., Sellers, E., and Vaughan, T. (2007). "Common spatio-temporal patterns for the p300 speller," in *3rd International IEEE/EMBS Conference on Neural Engineering, 2007. CNE'07* (Kohala Coast, HI: IEEE), 421–424.

Krusienski, D. J., Sellers, E. W., Cabestaing, F., Bayoudh, S., McFarland, D. J., Vaughan, T. M., et al. (2006). A comparison of classification techniques for the p300 speller. *J. Neural Eng.* 3:299. doi: 10.1088/1741-2560/3/4/007

Kübler, A., and Birbaumer, N. (2008). Brain-computer interfaces and communication in paralysis: extinction of goal directed thinking in completely paralysed patients? *Clin. Neurophysiol.* 119, 2658–2666. doi: 10.1016/j.clinph.2008.06.019

Kübler, A., Neumann, N., Wilhelm, B., Hinterberger, T., and Birbaumer, N. (2004). Predictability of brain-computer communication. *J. Psychophysiol.* 18, 121–129. doi: 10.1027/0269-8803.18.23.121

Kwak, N.-S., Müller, K.-R., and Lee, S.-W. (2015). A lower limb exoskeleton control system based on steady state visual evoked potentials. *J. Neural Eng.* 12:056009. doi: 10.1088/1741-2560/12/5/056009

Kwak, N.-S., Müller, K.-R., and Lee, S.-W. (2017). A convolutional neural network for steady state visual evoked potential classification under ambulatory environment. *PLoS ONE* 12:e0172578. doi: 10.1371/journal.pone.0172578

Lal, T. N., Schroder, M., Hinterberger, T., Weston, J., Bogdan, M., Birbaumer, N., et al. (2004). Support vector channel selection in bci. *IEEE Tran. Biomed. Eng.* 51, 1003–1010. doi: 10.1109/TBME.2004.827827

Ledoit, O., and Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *J. Multiv. Anal.* 88, 365–411. doi: 10.1016/S0047-259X(03)00096-4

Lee, P.-L., Sie, J.-J., Liu, Y.-J., Wu, C.-H., Lee, M.-H., Shu, C.-H., et al. (2010). An ssvep-actuated brain computer interface using phase-tagged flickering sequences: a cursor system. *Anna. Biomed. Eng.* 38, 2383–2397. doi: 10.1007/s10439-010-9964-y

Lin, Z., Zhang, C., Wu, W., and Gao, X. (2007). Frequency recognition based on canonical correlation analysis for ssvep-based bcis. *IEEE Trans. Biomed. Eng.* 54, 1172–1176. doi: 10.1109/TBME.2006.889197

Lopez-Gordo, M., Prieto, A., Pelayo, F., and Morillas, C. (2010). Use of phase in brain–computer interfaces based on steady-state visual evoked potentials. *Neural Proces. Lett.* 32, 1–9. doi: 10.1007/s11063-010-9139-8

Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F., and Arnaldi, B. (2007). A review of classification algorithms for eeg-based brain–computer interfaces. *J. Neural Eng.* 4:R1. doi: 10.1088/1741-2560/4/2/R01

Luo, A., and Sullivan, T. J. (2010). A user-friendly ssvep-based brain–computer interface using a time-domain classifier. *J. Neural Eng.* 7:026010. doi: 10.1088/1741-2560/7/2/026010

Lv, J., and Liu, M. (2008). "Common spatial pattern and particle swarm optimization for channel selection in bci," in *3rd International Conference on Innovative Computing Information and Control, 2008. ICICIC'08* (Dalian: IEEE), 457.

Manyakov, N. V., Chumerin, N., Combaz, A., and Hulle, M. (2010a). "On the selection of time interval and frequency range of eeg signal preprocessing for p300 brain-computer interfacing," in *XII Mediterranean Conference on Medical and Biological Engineering and Computing 2010* (Chalkidiki: Springer), 57–60.

Manyakov, N. V., Chumerin, N., Combaz, A., Robben, A., and Van Hulle, M. M. (2010b). "Decoding ssvep responses using time domain classification," in *IJCCI (ICFC-ICNC)* (Valencia), 376–380.

Manyakov, N. V., Chumerin, N., Combaz, A., Robben, A., van Vliet, M., and Van Hulle, M. M. (2011a). "Decoding phase-based information from ssvep recordings: A comparative study," in *2011 IEEE International Workshop on Machine Learning for Signal Processing (MLSP)* (Santander: IEEE), 1–6.

Manyakov, N. V., Chumerin, N., Combaz, A., and Van Hulle, M. M. (2011b). Comparison of classification methods for p300 brain-computer interface on disabled subjects. *Computat. Intell. Neurosci.* 2011:2. doi: 10.1155/2011/519868

Manyakov, N. V., Chumerin, N., Robben, A., Combaz, A., van Vliet, M., and Van Hulle, M. M. (2013). Sampled sinusoidal stimulation profile and multichannel fuzzy logic classification for monitor-based phase-coded ssvep brain–computer interfacing. *J. Neural Eng.* 10:036011. doi: 10.1088/1741-2560/10/3/036011

Manyakov, N. V., Chumerin, N., and Van Hulle, M. M. (2012). Multichannel decoding for phase-coded ssvep brain–computer interface. *Int. J. Neural Syst.* 22:1250022. doi: 10.1142/S0129065712500220

Matlab (2015). *Matlab Version 8.5.0 (R2015a)*. Natick, MA: The MathWorks Inc.

Maynard, E. M., Nordhausen, C. T., and Normann, R. A. (1997). The utah intracortical electrode array: a recording structure for potential brain-computer interfaces. *Electroencephalogr. Clin. Neurophysiol.* 102, 228–239. doi: 10.1016/S0013-4694(96)95176-0

Middendorf, M., McMillan, G., Calhoun, G., and Jones, K. S. (2000). Brain-computer interfaces based on the steady-state visual-evoked response. *IEEE Trans. Rehabil. Eng.* 8, 211–214. doi: 10.1109/86.847819

Momennezhad, A., Shamsi, M., Ebrahimnezhad, H., and Saberkari, H. (2014). Classification of eeg-p300 signals extracted from brain activities in bci systems using ν-svm and blda algorithms. *Appl. Med. Inform.* 34, 23–35.

Nakanishi, M., Wang, Y., Wang, Y.-T., Mitsukura, Y., and Jung, T.-P. (2014). A high-speed brain speller using steady-state visual evoked potentials. *Int. Jo. Neural Syst.* 24:1450019. doi: 10.1142/S0129065714500191

Nicolas-Alonso, L. F., and Gomez-Gil, J. (2012). Brain computer interfaces, a review. *Sensors* 12, 1211–1279. doi: 10.3390/s120201211

Pan, J., Gao, X., Duan, F., Yan, Z., and Gao, S. (2011). Enhancing the classification accuracy of steady-state visual evoked potential-based brain–computer interfaces using phase constrained canonical correlation analysis. *J. Neural Eng.* 8:036027. doi: 10.1088/1741-2560/8/3/036027

Pfurtscheller, G., and Neuper, C. (2001). Motor imagery and direct brain-computer communication. *Proc. IEEE* 89, 1123–1134. doi: 10.1109/5.939829

Pires, G., Castelo-Branco, M., and Nunes, U. (2008). "Visual p300-based bci to steer a wheelchair: a bayesian approach," in *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE* (Vancouver, BC: IEEE), 658–661.

Pruzek, R. M. (1994). "High dimensional covariance estimation: Avoiding the 'curse of dimensionality,'" in *Proceedings of the First US/Japan Conference on the Frontiers of Statistical Modeling: An Informational Approach* (Springer), 233–253.

Regan, D. (1979). Electrical responses evoked from the human brain. *Sci. Am.* 241, 134–146. doi: 10.1038/scientificamerican1279-134

Rivet, B., Souloumiac, A., Attina, V., and Gibert, G. (2009). xdawn algorithm to enhance evoked potentials: application to brain–computer interface. *IEEE Trans. Biomed. Eng.* 56, 2035–2043. doi: 10.1109/TBME.2009.2012869

Salimi-Khorshidi, G., Nasrabadi, A. M., and Golpayegani, M. H. (2008). Fusion of classic p300 detection methods' inferences in a framework of fuzzy labels. *Artif. Intell. Medicine* 44, 247–259. doi: 10.1016/j.artmed.2008.06.002

Salimi Khorshidi, G., Nasrabadi, A. M., and Hashemi Golpayegani, M. (2007). "Modifying the classic template matching technique using a fuzzy multi agent to have an accurate p300 detection," in *3rd Kuala Lumpur International Conference on Biomedical Engineering 2006* (Kuala Lumpur: Springer), 410–414.

Schalk, G., and Leuthardt, E. C. (2011). Brain-computer interfaces using electrocorticographic signals. *IEEE Rev. Biomed. Eng.* 4, 140–154. doi: 10.1109/RBME.2011.2172408

Schoukens, J., and Pintelon, R. (2014). *Identification of Linear Systems: A Practical Guideline to Accurate Modeling.* Elsevier.

Schreuder, M., Blankertz, B., and Tangermann, M. (2010). A new auditory multi-class brain-computer interface paradigm: spatial hearing as an informative cue. *PLoS ONE* 5:e9813. doi: 10.1371/journal.pone.0009813

Schröder, M., Lal, T. N., Hinterberger, T., Bogdan, M., Hill, N. J., Birbaumer, N., et al. (2005). Robust eeg channel selection across subjects for brain-computer interfaces. *EURASIP J. Appl. Signal Process.* 2005, 3103–3112. doi: 10.1155/ASP.2005.3103

Sellers, E. W., and Donchin, E. (2006). A p300-based brain–computer interface: initial tests by als patients. *Clin. Neurophysiol.* 117, 538–548. doi: 10.1016/j.clinph.2005.06.027

Singla, R., and Haseena, B. (2014). Comparison of ssvep signal classification techniques using svm and ann models for bci applications. *Int. J. Inform. Electron. Eng.* 4:6. doi: 10.7763/IJIEE.2014.V4.398

Song, L., and Epps, J. (2007). Classifying eeg for brain-computer interface: learning optimal filters for dynamical system features. *Comput. Intell. Neurosci.* 2007:57180. doi: 10.1155/2007/57180

Treder, M. S., Porbadnigk, A. K., Avarvand, F. S., Müller, K.-R., and Blankertz, B. (2016). The lda beamformer: optimal estimation of erp source time series using linear discriminant analysis. *Neuroimage* 129, 279–291. doi: 10.1016/j.neuroimage.2016.01.019

Van Veen, B. D., and Buckley, K. M. (1988). Beamforming: a versatile approach to spatial filtering. *IEEE Assp. Mag.* 5, 4–24. doi: 10.1109/53.665

Van Veen, B. D., Van Drongelen, W., Yuchtman, M., and Suzuki, A. (1997). Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *Biomed. Eng. IEEE Trans.* 44, 867–880.

van Vliet, M., Chumerin, N., De Deyne, S., Wiersema, J., Fias, W., Storms, G., et al. (2015). Single-trial ERP component analysis using a spatio-temporal lCMV beamformer. *Biomed. Eng. IEEE Trans.* 63, 55–66. doi: 10.1109/TBME.2015.2468588

Vidal, J. J. (1973). Toward direct brain-computer communication. *Annu. Rev. Biophys. Bioeng.* 2, 157–180. doi: 10.1146/annurev.bb.02.060173.001105

Vidaurre, C., and Blankertz, B. (2010). Towards a cure for bci illiteracy. *Brain Topogra.* 23, 194–198. doi: 10.1007/s10548-009-0121-6

Vu, H., Koo, B., and Choi, S. (2016). "Frequency detection for ssvep-based bci using deep canonical correlation analysis," in *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (Budapest: IEEE), 001983–001987.

Wei, Q., Feng, S., and Lu, Z. (2016). Stimulus specificity of brain-computer interfaces based on code modulation visual evoked potentials. *PLoS ONE* 11:e0156416. doi: 10.1371/journal.pone.0156416

Wittevrongel, B., and Van Hulle, M. M. (2016a). Faster p300 classifier training using spatiotemporal beamforming. *Int. J. Neural Syst.* 26:1650014. doi: 10.1142/S0129065716500143

Wittevrongel, B., and Van Hulle, M. M. (2016b). Frequency- and phase encoded ssvep using spatiotemporal beamforming. *PLoS ONE* 11:e0159988. doi: 10.1371/journal.pone.0159988

Wittevrongel, B., and Van Hulle, M. M. (2016c). "Hierarchical online ssvep spelling achieved with spatiotemporal beamforming," in *Statistical Signal Processing Workshop (SSP), 2016 IEEE* (Palma de Mallorca: IEEE), 1–5.

Wittevrongel, B., Van Wolputte, E., and Van Hulle, M. M. (2017). Code-modulated visual evoked potentials using fast stimulus presentation and spatiotemporal beamformer decoding. *Sci. Rep.* 7:15037. doi: 10.1038/s41598-017-15373-x

Zhang, Y., Guo, D., Yao, D., and Xu, P. (2017). The extension of multivariate synchronization index method for ssvep-based bci. *Neurocomputing* 269, 226–231. doi: 10.1016/j.neucom.2017.03.082

Zhang, Y., Xu, P., Cheng, K., and Yao, D. (2014a). Multivariate synchronization index for frequency recognition of ssvep-based brain–computer interface. *J. Neurosci. Methods* 221, 32–40. doi: 10.1016/j.jneumeth.2013.07.018

Zhang, Y., Zhou, G., Jin, J., Wang, X., and Cichocki, A. (2014b). Frequency recognition in ssvep-based bci using multiset canonical correlation analysis. *Int. J. Neural Syst.* 24:1450013. doi: 10.1142/S0129065714500130

Zhang, Y., Zhou, G., Zhao, Q., Onishi, A., Jin, J., Wang, X., et al. (2011). "Multiway canonical correlation analysis for frequency components recognition in SSVEP-based BCIS," in *Neural Information Processing: 18th International Conference, ICONIP 2011, Proceedings, Part I* (Shanghai: Springer), 287–295. doi: 10.1007/978-3-642-24955-6_35

Zhang, Z., Li, X., and Deng, Z. (2010). "A cwt-based ssvep classification method for brain-computer interface system," in *2010 International Conference on Intelligent Control and Information Processing (ICICIP)* (Dalian: IEEE), 43–48.