IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received July 10, 2019, accepted July 15, 2019, date of publication July 18, 2019, date of current version August 7, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2929545

Multiple Correlated Component Analysis for Identifying the Bilateral Location of Target in Visual Search Tasks

YIN TIAN[1,2], HAIYONG ZHANG[1], PEIYANG LI[1], AND YANG LI[2]
[1]Bio-information College, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
[2]College of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

Corresponding author: Yin Tian (tiany20032003@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61671097, and in part by the Chongqing Research Program of Basic Science and Frontier Technology under Grant cstc2017jcyjBX0007.

ABSTRACT N2pc is defined as a negative event-related potential component that appears after about 250 ms at posterior electrodes contralateral to a target's location in visual search, which can be used to measure attentional shifts between bilateral visual hemifields and locate the spatial location of lateral targets. However, the waves between the left and right hemispheres elicited by lateral targets usually exhibit a small amplitude difference and strong synchronicity, which may lead to low classification performance. Therefore, the present study explored the feasibility of a multiple correlated components analysis (MCORCA) methods to identify the lateral targets in visual search tasks with a single trial, which could weight the target signals by spatial filters to enlarge the amplitude difference between bilateral hemispheres and extract the linear combinations of multiple channels across trials with an optimal subset of correlated components to avoid the loss of efficient information. The classification rate achieved 82% with a single short-duration trial when using the proposed method with Leave-one-out-cross-validation (LOOCV). The findings demonstrated that the MCORCA-based methods could be used to improve the classification performance for the N2pc-based brain-computer interfaces (BCI) in visual search.

INDEX TERMS N2pc, MCORCA, visual search, brain–computer interfaces.

I. INTRODUCTION

A brain–computer interface (BCI) can convert brain signals into the required control signals for communication and the operation of devices without any motor muscular movements. Electroencephalography has been widely used for BCI due to its high temporal resolution, portability, ease of use, and low cost. There are four main types of EEG-based BCI: slow cortical potentials (SCP), sensorimotor rhythms, P300, and steady-state visual evoked potentials (SSVEP) [1]–[7]. These brain activities, related to specific internal or external events, which are called the event-related potential [8], have a high signal-to-noise ratio (SNR).

Traditionally, BCI systems have been used to assist people with severe motor disabilities to communicate or achieve motor control [9]–[12]. However, in the last decade, BCI research for healthy users has increasingly attracted attention, focusing on the augmentation of human functions [8] or the detection of users' mental states [13]. For example, observers were able to detect rare, attended targets, which could elicit a distinct positive waveform at approximately 300 ms after a rare or salient stimulus onset, i.e. P300 event-related potential [8], in the oddball paradigm [14]. The first P300-based BCI was a speller proposed by Farwell and Donchin [15]. Various P300 speller paradigms were developed and provided reasonable BCI control [5]–[7], [16]. One of these applications focused on augmenting the human search capabilities in the process of visual search [17]–[19]. The N2pc component usually serves as an indicator for attentional selection in the visual search paradigm [20]. N2pc was evoked at about 250 ms after stimulus onset over posterior electrodes contralateral to the visual hemifield of the targets' location in the visual search scene, which was related to the shift of attention towards a candidate target [21]–[23].

The associate editor coordinating the review of this manuscript and approving it for publication was Emrecan Demirors.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see http://creativecommons.org/licenses/by/4.0/

Visual search displays contain a unique singleton target among uniform distractor items, and the N2pc component may be triggered in response to the singleton item [21], [24]. Specifically, when a visual search target is clearly visible among a set of distractors, the ERP waveform becomes more negative at contralateral scalp sites compared to ipsilateral scalp sites. The voltage difference of the waveform between the contralateral and ipsilateral sites was defined as the N2pc component [25]. Generally, the N2pc appears during the time window from 180 to 300 ms after stimulus onset and the maximal amplitude of the N2pc is located at channels PO7/PO8 [21], which could represent attentional shifts to object locations. Since the N2pc is evoked by lateral visual targets in the left or right visual fields and presents a contralateral hemisphere, i.e. the left visual targets elicit the right hemisphere N2pc (mainly located at channel PO8) and the right visual targets elicit the left hemisphere N2pc (mainly located at channel PO7), this component may disappear for targets presented on the vertical meridian of the whole visual field [24]–[26].

Since the N2pc effectively locates the desired target in the hemifield of the visual field, the N2pc can be exploited as a control signal for BCI to identify a certain area of the visual field where the target is located. However, very few studies have considered the N2pc for controlling a BCI and the related classification performance still needs to be improved. For example, N2pc was utilized to localize the spatial location of targets in aerial images presented by a rapid serial visual presentation (RSVP) protocol using single-trial classification, achieving a median area under the receiver operating characteristic curve (AUC) of 0.76 with a support vector machine (SVM) [18]. Awni *et al.* also found that N2pc can help identify the left and right colored digits (i.e. pop-out targets) with linear discriminant analysis (LDA). When averaging three trials, the mean classification accuracy (CA) achieved 75% and one individual's CA reached nearly 90%, indicating that N2pc-based CA varied across participants. With single-trial classification, all of the subjects showed poor CA, which was only marginally better than chance [27]. Moreover, the combination of N2pc and P300 was used to manipulate a BCI for Internet browsing and a robotic arm with healthy users. And an application was also developed to assist disabled people with basic communication, achieving a CA of 80% after twelve repetitions of the stimuli [28].

For the N2pc-based BCI mentioned above, the amplitude of N2pc was chosen as a feature. However, the inter-trial variability in amplitude and latency impaired the detection of N2pc. Generally, in order to increase the signal-to-noise ratio and improve the classification performance in BCIs, the usual approach was to average multiple repetitions of each stimulus [27], [28]. The drawbacks of this approach were that repetitions could cause fatigue and a decrease in speed, which impairs the information transfer rate. Therefore, it was a good way to improve the task performance that single-trial extraction, combined information over multiple channels and enlarging the amplitude difference of N2pc between the left and right hemispheres, could achieve a high performance [29].

In this paper, we propose a novel method based on correlated component analysis (CORCA) for single-trial N2pc detection. In essence, EEG reflects thousands of synchronous neural oscillations in the brain [30]. When using EEG to measure signals from the same neurophysiological processes, it captures synchronous oscillations from channel recordings located in the task- related regions [31], which indicates that there are inherent correlations between EEG series from different channels in the spatial distributions determined by anatomy for the same cognitive tasks [32]. In visual search tasks, the activation region is mainly located around channel PO8 for the left visual-field-search target, while the activation regions are mainly located around channel PO7 for the right visual-search target. These contralateral activations further indicate that the neural oscillations related to the same visual-field-search tasks should hold strong synchronicity. Thus, preserving this inherent character will enhance the features' discriminability in EEG-based classifications. However, the canonical correlation analysis method requires the canonical projection vectors to be orthogonal, which drops this inherent correlation [32], [33]. Compared to CCA, CORCA proposed the linear combinations of multiple channels across subjects with maximal correlation so as to take advantage of the inherent correlations between channel recordings in the same cognitive tasks [32], which was introduced to recognize the SSVEP components and achieved satisfactory performance [4]. Zhang's group only extracted a maximally correlated component of the specific frequency in SSVEP recognition [4], ignoring the complementary relationships between multiple correlation components, which may lose the information with discriminative power. Thus, in this work, MCORCA is proposed to allow for the complementary correlations in multiple time series. Moreover, the MCORCA weighted the target signals by spatial filters to extract the spatial information for different visual search targets, which could increase the amplitude difference between bilateral hemispheres. The main goal of our study was to employ a novel method for feature extraction and classification in a visual search task that allows the efficient identification of the left visual field (LVF) and right visual field (RVF) lateral targets with respect to the horizontal midline. In addition, we will explore and comparatively analyze the performance for the subset of components with the proposed multiple correlated component analysis (MCORCA) and the classical SVM.

II. METHODS

A. PARTICIPANTS

Thirty healthy subjects (19–21 years of age, average age 20.4 years, 20 males and 10 females) participated. All subjects were right-handed with normal vision or corrected to normal vision. None of them had a history of any neurological or psychiatric disorders. All subjects signed a consent form before the experiments and received monetary

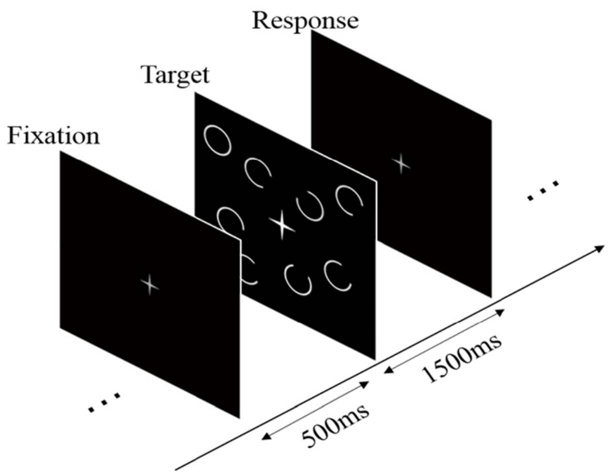


FIGURE 1. An example of the experiment. Each trial started with a fixation cross (0.5° × 0.5°) at the center of the display for 500 ms. The display randomly contained a variable number of items (4, 8 or 12) equally distributed in the four quadrants of fixation.

compensation after the experiments. The experimental procedures were approved by the Ethics Committee of Chongqing University of Posts and Telecommunications and conducted according to the Declaration of Helsinki.

B. EXPERIMENTAL DESIGN

The experiments were similar to our previous study [34]. A fixation cross ($0.5^\circ \times 0.5^\circ$) was presented at the center of the display throughout the entire block. Stimuli consisted of circles (O) and arcs (C) presented on a black background. The diameter of these items subtended 0.8°. One tenth of the circle was cut off to form a gap in each arc. The stimuli items were randomly placed in the four quadrants around the center cross, forming a square area (Fig. 1). In each trial, the display randomly contained a variable number of items (4, 8 or 12), equally distributed in the four quadrants of fixation. The visual search displays contained either an O among Cs, or all Cs, in which the circle was the target and the arcs were the distractors (serial search).

Each participant completed two sessions in the serial search task. Participants performed a series of 100 trials in each session, divided into 20 blocks of 5 trials each. Specifically, in each session, the 4 items condition was used for 20% of the trials, and the 8 and 12 items for 40% respectively. The target was absent in 10% of the trials (i.e. catch trials).

Each trial began with a central fixed cross flashing for 500 ms, which then changed to the stimulus array for 1500 ms or terminated when participants responded. A fixation cross was then presented for 500 ms and was immediately followed by the next trial. In the search task, participants were instructed to press the right mouse button quickly with their right hand when the target appeared in the right side of the center cross and press the left mouse button when the target appeared in the left. For no-target trials, participants were required to withhold their response. In all experiments, speed and accuracy were emphasized equally. No feedback was provided.

C. EEG RECORDING AND PROCESSING

Participants were required to sit comfortably in front of an LCD screen with their eyes approximately 57 cm away. EEG signals were recorded by a 64-channel Neuroscan EEG System (SynAmps2, bandpass filter: 0.05∼100 Hz, sampling rate: 1000 Hz), which was positioned according to the extended international 10-20 system. The vertex was used as the reference online and the impedances of all channels were maintained below 5 $k\Omega$.

The EEG was re-referenced offline into the infinity reference [35], [36]. After re-referencing, the data were down-sampled to 250 Hz and filtered with a band-pass of 0.1–30 Hz. From these filtered data, we further extracted segments corresponding to a time period of 200 ms before and 1000 ms following stimulus onset and utilized 200 ms prior to the stimulus for baseline correction. Moreover, we conducted linear-detrending and artifacts removal. For artifacts removal, we visually removed the EEG epochs containing common artifacts such as eye blinks, eye movements, head movement, and muscle activity. Automated rejection was further performed with the amplitude criteria of an absolute threshold ($< 60\ \mu$V).

D. CONTRALATERAL AND IPSILATERAL WAVEFORMS

In order to detect whether the N2pc component was elicited by the lateral target, we performed ERP analysis for all subjects. The data epochs were first separately averaged according to the appearance location of a target, i.e. the LVF and RVF target categories. Then, the averaged waveforms were divided into contralateral and ipsilateral waveforms across brain hemispheres. Thus, for LVF targets, the contralateral waveforms represented ERPs on the right hemisphere (i.e. right posterior-occipital region) while the ipsilateral waveforms represented ERPs on the left hemisphere, and vice versa [25].

E. DIFFERENCE AMPLITUDE-BASED FEATURE EXTRACTION

As mentioned above, the N2pc was closely related to the attention selection process, which was usually related to the lateral orienting of attention in the visual search tasks. The usual way to extract the amplitude of the N2pc was to compute the amplitude differences between paired posterior electrodes corresponding to symmetric positions with respect to the brain's median plane. Here, eight paired posterior symmetric electrodes were selected for further analysis, i.e., P1/2, P3/4, P5/6, P7/8, PO3/PO4, PO5/PO6, PO7/PO8, O1/2 [18].

For the position recognition of the visual search task, these amplitude differences are usually used as features for the input of classifiers such as SVM [18] and LDA [28]. For comparison, SVM with recursive feature elimination (SVM-RFE) using the difference amplitude feature has also been used to compare the performance with the proposed MCORCA method in the current study.

F. SUPPORT VECTOR MACHINE WITH RECURSIVE FEATURE ELIMINATION (SVM-RFE)

The support vector machine (SVM) [37], [38] has been widely used for data classification due to its excellent generalization performance. SVM implements the following idea: it maps the input vectors into a high dimensional feature space and finds a hyperplane which has the largest margin to split the data into two sets. The goal of the SVM classifier is to separate features $X \in \mathrm{R}^d$ into two classes by finding a decision function $f(x) = sign(w^T \phi(x) + b)$. SVM solves the following quadratic optimization problem:

$$\begin{aligned} &\min_{w,b,\xi} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\xi_i \\ &s.t.\ y_i(w^T\phi(x_i) + b) \geq 1 - \xi_i \end{aligned} \tag{1}$$

where $\boldsymbol{w}$ is the normal vector of the hyperplane, the function ϕ maps the original feature vector x_i into a higher dimensional space [39], the parameters ξ_i are slack variables that relax the constraints of the canonical hyperplane for non-separable cases, and C is a regularization parameter on the training error.

SVM-RFE was proposed by Guyon *et al.* [40], who eliminated recursively the features based on the concept of margin maximization. SVM-RFE used criteria from the coefficients in SVM models to assess features and iteratively discarded features that correspond to the smallest $\|w_i\|_2$ [41]. The present study adapted this method for feature ranking. At each iteration step, the electrode pair with the lowest score was eliminated. The SVM-RFE was performed based on the training data, leading to a ranked list of electrode pairs [41], [42].

G. MCORCA-BASED METHOD

The previous method, CORCA, proposed the linear combinations of multiple channels across subjects with maximal correlation [4], [32]. In other words, CORCA estimated a weight vector that maximized the inter-subject correlation with similar temporal activation among subjects. Given a set of N subjects data matrices $X_1, X_2, \ldots, X_N$, where $X_N \in \mathrm{R}^{N_c \times N_s}$, N_c and N_s denote the number of channels and sampling points, respectively. Then, the subject-aggregated data matrices can be defined as follows:

$$\overline{X}_1 = \left[X_{I_{11}}X_{I_{21}}\ldots X_{I_{P1}}\right] \tag{2}$$

$$\overline{X}_2 = \left[X_{I_{12}}X_{I_{22}}\ldots X_{I_{P2}}\right] \tag{3}$$

where $I_i = \{I_{i1}, I_{i2}\} = \{(1,2), (1,3), \ldots, (N-1, N)\}$ denote the set of all $P = N \times (N-1)/2$ subject pairs. With CORCA, a projection vector $\boldsymbol{\omega}$ was observed, which maximized the correlation between $y_1 = \overline{X}_1^T\boldsymbol{\omega}$ and $y_2 = \overline{X}_2^T\boldsymbol{\omega}$. The optimization problem was transformed to seek the Pearson Product Moment Correlation Coefficient between y_1y$_1$ and y$_2 y_2$:

$$\frac{y_1^T y_2}{\|y_1\|\,\|y_2\|} = \frac{\omega^T R_{12}\omega}{\left(\omega^T R_{11}\omega\right)^{1/2}\left(\omega^T R_{22}\omega\right)^{1/2}} \tag{4}$$

$$R_{ij} = \frac{1}{N_s \times P}\overline{X}_i\overline{X}_j^T \tag{5}$$

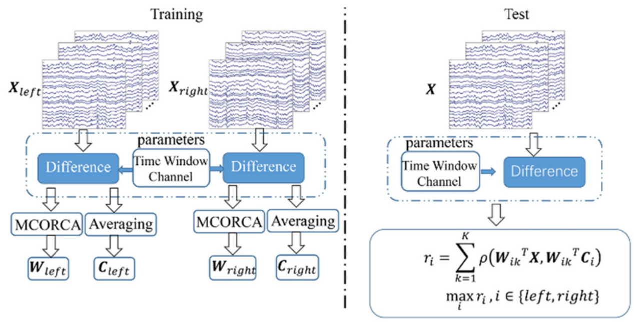


FIGURE 2. Illustration of the proposed MCORCA-based method.

where R_{ij} was the sample covariance matrix. Assuming $\omega^T R_{11}\omega = \omega^T R_{22}\omega$ [32], [43], the projection vector $\boldsymbol{\omega}$ which maximizes the inter-subject correlation between y_1 and y_2 can be estimated through the following single eigenvalue equation

$$(R_{21} + R_{12})\,\omega = \lambda\,(R_{11} + R_{22})\,\omega \tag{6}$$

where λ is the generalized eigenvalues, and the eigenvalues are arranged according to their absolute value. The optimal component projection is the eigenvector $\boldsymbol{\omega}$ of matrix with the maximal eigenvalues and captures the largest correlation between subjects.

In the current study, within-subject correlations were computed to obtain the spatial filters for LVF targets and RVF targets respectively. With the projection vector $\boldsymbol{\omega}$, we could obtain the projection signals of a single-trial test data and a reference signal and then calculate the correlation coefficient between the two signals as the classification feature [4].

With individual training data $\overline{X}_i \in \mathrm{R}^{N_c \times N_s \times N_t}, i \in \{left, right\}$ denoted as LVF targets or RVF targets, N_t denoting the number of trials, we can obtain the weight vector $\omega_{ik} \in \mathrm{R}^{N_c \times 1}$ with CORCA. ω_{ik} represents the k-th weight vector for the i-th class. Denote $C_i \in \mathrm{R}^{N_c \times N_s}$ as the reference signals which are obtained by averaging across multiple training trials $\overline{X}_i$. Next, the correlation coefficient between the test sample $X \in \mathrm{R}^{N_c \times N_s}$ and the reference signal C_i can be computed as follows:

$$r_i = \sum_{k=1}^{K}\rho\left(\omega_{ik}^T X, \omega_{ik}^T C_i\right), \quad i \in \{left, right\} \tag{7}$$

The class of the test data is then recognized as the class of the template signals with the maximal correlation coefficients:

$$\max_i r_i, \quad i \in \{left, right\} \tag{8}$$

Furthermore, we chose the k strongest components to be used for feature extraction and classification (named as multiple correlated component analysis (MCORCA)), and expected the optimal components subset to exhibit a high classification performance. Specifically, if k = 1, it represents the CORCA with a maximum correlation component. The schematic diagram of the proposed methods is shown in Fig. 2. The method includes the following steps:

Step 1, computing the amplitude differences between paired posterior electrodes for all data sets.

Step 2, in the training stage, the CORCA is used to calculate the spatial filters for LVF and RVF respectively. The reference signals are obtained by averaging across multiple training trials.

Step 3, in the test stage, the spatial filter is used to calculate the correlation coefficient between the test signal and all the reference signals by formula (7). The LVF and RVF lateral targets are recognized by formula (8).

H. PERFORMANCE EVALUATION

The classification accuracy (CA) and information transfer rate (ITR) were used to evaluate the performance of the methods mentioned above. We evaluated the influence of different component numbers ($k = 1, \ldots, 8$) on the CA and ITR. The SVM-RFE method was used to compare the performance with the MCORCA method.

For each participant, a leave-one-out cross validation was used to evaluate the classification performance. Specifically, one trial was selected as the test sample, and the rest of the trials were considered as the training samples to estimate the classifier. The procedure was repeated until each trial had been used once for testing. The feature selection was performed inside the cross-validation loop, which led to a specific ranking of the eight electrode pairs in each fold. For each fold, the classifier was trained on the best k ($k = 1, \ldots, 8$) components (electrode pairs) respectively, and tested on the corresponding components of the test set. The corresponding CAs over all folds were averaged to yield an overall accuracy. Finally, an estimate of the classification accuracy could be obtained for each number of selected components. Besides the CA, the ITR was also used to evaluate the performance [44]. The ITR can be quantified by the bit rate per minute:

$$B = \log_2 N + P \log_2 P + (1 - P) \log_2 [\frac{1-P}{N-1}] \tag{9}$$

$$ITR = B/T \tag{10}$$

where P is the classification accuracy and N is the number of classes. B is calculated in bits per trial according to Shannon's theorem and T is the time duration in minutes, which leads to the calculation of ITR in bits per minute. A paired t-test was implemented to investigate the statistical difference of the classification performance between MCORCA and SVM-RFE.

III. RESULTS

A. N2pc

The grand-averaged waveform for lateral targets at the paired channels of PO7/PO8 across all participants is shown in Fig. 3, in which the contralateral curve is defined as the average waveform at the channel PO7 for RVF targets and the waveform of the channel PO8 for LVF targets. Meanwhile, the ipsilateral curve is the average waveform at the channel PO7 for LVF targets and the waveform at the channel PO8 for

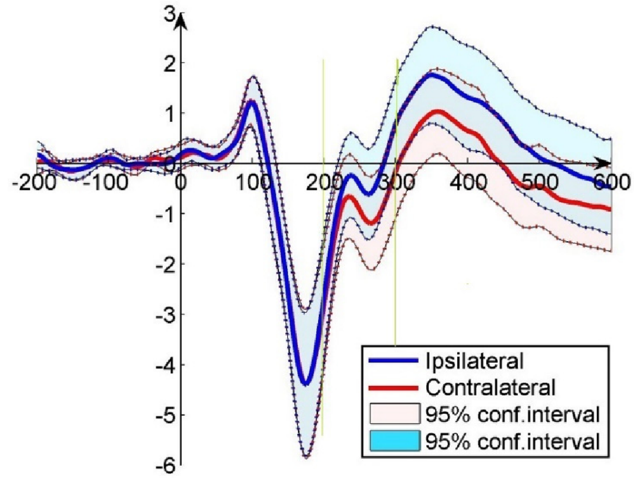


FIGURE 3. Grand-averaged ERPs over 30 participants with 95% confidence intervals at channels PO7 and PO8.

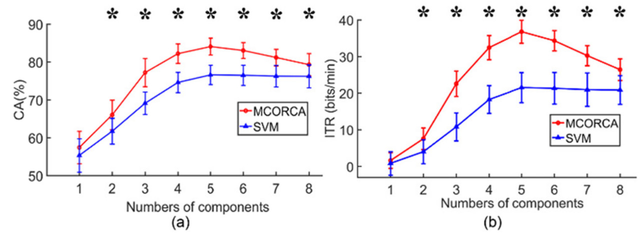


FIGURE 4. The average CAs (a) and ITRs (b) across all subjects of the two methods with different numbers of components. The asterisk (*) indicates significant difference between the two methods by paired t-tests ($p < 0.05$).

RVF targets. As shown in Fig. 3, N2pc appears at approximately 200–300 ms after stimulus onset. For clear, the 95% confidence intervals are also presented in Fig.3 [45]–[47]. For considering the variance between subjects, we use the averaged amplitude values in the time window from 200 to 300ms to test the difference between ipsilateral and contralateral waveforms rather than the peak value. The paired t-test showed that there was significant difference between the averaged ipsilateral and contralateral amplitude during the interval of 200-300ms ($t = 5.05$, $p < 0.05$).

B. CLASSIFICATION FOR LVF VERSUS RVF

As shown in the ERP analysis, significant amplitude differences between the contralateral and ipsilateral waveforms were obtained during the intervals of 200–300 ms after stimulus onset. Therefore, with the MCORCA-based method, the spatial filters of the LVF targets and RVF targets from individual training data were calculated in this interval respectively. The spatial filters (eigenvectors) corresponding to the eigenvalues were sorted in descending order, and the sorted spatial filters corresponding to the k best eigenvalues were used as the projection vector ω. As shown in Fig. 4, the classification accuracies and the ITRs presented a consistent tendency following the increased number of components of both MCORCA and SVM-RFE. From the statistical analysis with the paired t-test, the performance of the MCORCA-based method was better than that of the SVM-RFE method when the number of components selected was from 2 to 8 (all with $p < 0.05$).

To determine the proper number of components for high classification performance, we performed one-way ANOVA and multiple comparisons on the CA and ITR of different components across all subjects. For CA, there was a statistically significant main effect for the number of components (MCORCA: $F = 46.28$, $p < 0.001$; SVM-RFE: $F = 52.56$, $p < 0.001$). Post hoc multiple comparisons using the Bonferroni correction showed that there was significantly increased CA with the number of components when k < 5 (all with $p < 0.05$), but no significant increase of CA when k $\geq$ 5 (all with $p > 0.05$) for the two methods. However, there was a downward trend for the proposed MCORCA-based method. For the ITR, there was a significant increase in the ITR with the number of components at k < 5 for the two methods (all with $p < 0.05$), which was consistent with the CA results. No significant increase was observed at k $\geq$ 5 for the SVM-RFE method (all with $p > 0.05$). However, the ITR showed a significant decrease with the number of components for the MCORCA method. The results show that we can obtain the best number of components for classification and achieve good performance with the proposed method.

With the difference amplitude feature and recursive feature elimination, we trained the SVM classifier to distinguish LVF from RVF lateral targets for different numbers of components. The mean CA and ITR with amplitude and the SD of cross-validation were 75% $\pm$ 4.60% and 22 $\pm$ 4.20 bit/min, respectively, when the number of components was 5 (Fig. 4). This result showed that there was a poor performance for lateral targets detection based on single-trial classification using SVM-RFE, which was consistent with the previous study [27]. For the MCORCA-based method, the average CA and ITR respectively reached up to 82% $\pm$ 3.70% and 35 $\pm$ 3.31 bit/min when the number of components was 5. Overall, these results indicate that the MCORCA-based method for feature extraction and classification performed better than SVM-RFE in the visual search task with single-trial classification.

We also investigated the classification performance of the two methods across all the subjects at various time windows using a sliding window of 100 ms with a step of 50 ms (0–100 ms, 50–150 ms, 100–200 ms, 150–250 ms, 200–300 ms, 250–350 ms, 300–400 ms). The number of components was set as 5.

As shown in Fig. 5, the MCORCA-based method yields better performance than the SVM-RFE method in the time window of 200–300 ms after the stimulus onset. Consistent with the results of ERP analysis, the N2pc components were elicited by the target at approximately 250 ms post-stimulus (Fig. 3) and good performance was observed in this window. Furthermore, the time window between 200–300 ms was an optimal choice for CA and ITR, prior to the P300-based BCI (usually time window > 300 ms).

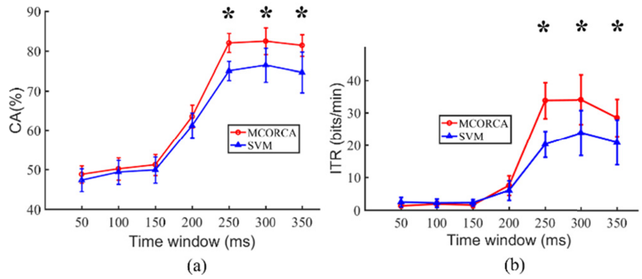


FIGURE 5. The average CAs (a) and ITRs (b) across all subjects of the two methods for a sliding window of 100 ms with a step of 50 ms. The asterisk (*) indicates significant difference between the two methods by paired t-tests ($p < 0.05$).

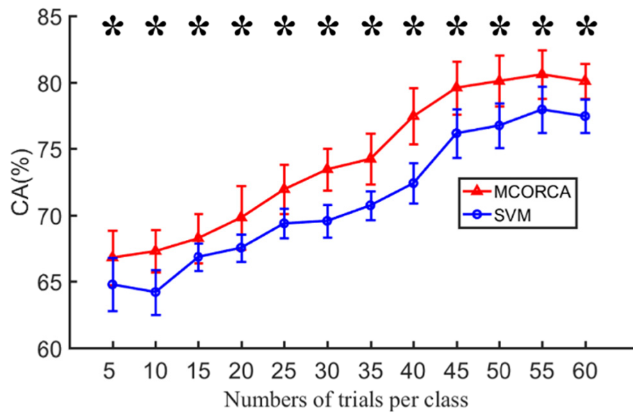


FIGURE 6. The averaged CAs across all subjects of the two methods with different numbers of training samples per class. The asterisk (*) indicates significant difference between the two methods by paired t-tests ($p < 0.05$).

To evaluate the influence of the sample numbers on the classification performance of MCORCA, the following 11 values (i.e. the number of training trials per class) were further tested: 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, and 60. For each class per subject, the training samples were randomly selected from the data set and the remaining trials served as the test set [48], [49]. The number of components was set as 5. For each size of the training set, this procedure was repeated 20 times [48]. Fig. 6 shows the average classification results for different numbers of training trials across all subjects. Using the paired t-test, we found that the proposed MCORCA method outperformed the SVM-RFE method in all conditions.

IV. DISCUSSION

In this study, the MCORCA method for N2pc-based BCIs was proposed to identify the LVF and RVF lateral targets with a single trial in visual search tasks. Our results demonstrate that the proposed MCORCA method can achieve good performance with a short duration (< 300 ms EEG data).

In our visual search paradigm, the N2pc component was elicited by responding to the lateral targets at about 200–300 ms post-stimulus (Fig. 3), which was consistent with the previous literature on the N2pc [18], [21], [25], [27]. During this time window, the MCORCA-based method was applied to learn the spatial filters of the LVF and RVF targets. The good performance of our method, which could detect the electrophysiological differences between LVF and RVF targets efficiently, can be attributed to the effective spatial filters learned during the training stage. Meanwhile, the EEG background artifacts were effectively removed and the SNR of the N2pc signals was improved by the spatial filters. With the spatial filters, the shared neuropsychophysiology components in response to the specific target locations were captured by calculating the optimal inter-trial correlative components instead of the maximal correlative components. In fact, traditional CCA methods require the canonical projection vectors to be orthogonal, which drops the inherent correlations among channel recordings that reflect the synchronous oscillations for a given cognitive task [32]. Compared with CCA, CORCA proposed the linear combinations of multiple channels across subjects with maximal correlation so as to take advantage the inherent correlations between channel recordings in the same cognitive tasks [4]. However, CORCA mainly utilized the component corresponding to maximal correlation [4], ignoring the complementary relationships between multiple correlation components, which may lose the information with discriminative power. Compared with this, MCORCA allows for complementary correlations in multiple time series, which enhances the SNR and discriminative power of features, resulting in good performance in a short duration (< 300 ms EEG data). As shown in Fig. 4, the highest classification performance was observed with the five strongest correlative components (5 paired electrodes), i.e. CA: 82% and ITR: 35 bit/min.

Moreover, the worst classification performance was observed when a maximal correlative component was selected, similar to the previous study [4]. In addition, when the number of the strongest correlative components was greater than 5 (i.e. $k > 5$), the CA and ITR showed a significant decline. Though the decline of the CA (when $k > 5$) did not have a significant effect, a declining trend of the CA could also be observed (Fig. 4a). However, the SVM-RFE classification based on difference amplitude features showed a poor performance with single-trial discrimination of the LVF and RVF lateral targets when compared to MCORCA, which resulted in inter-trial variability in the amplitude and latency of the ERPs. The findings revealed that our proposed MCORCA-based method could get the optimal subset on the number of electrodes to achieve high classification performance.

In the current study, there were variations in performance across subjects, and one reason for this difference might be the choice of time window of N2pc. It should be noted that determination of the time window was based on averaged ERP analysis across all subjects, ignoring individual differences. We did not evaluate the subject-specific N2pc latency, which may be a factor that led to induce these variations in performance. Moreover, it has been previously demonstrated that the latency and amplitude of N2pc is related to different visual search tasks [50], [51] and the interstimulus interval [52]. Because of the inter-subject variability, we trained the classification model with the same individual. In future, a transfer learning method could be applied to improve the subject-specific classification performance by transferring the general features learned from multiple subjects or experiments. A study that designs a proper experiment paradigm across subjects for the visual search task would be beneficial to BCI application. In addition, the performance of the MCORCA-based method for single-trial lateral targets identification was tested here on an offline experiment only, and online performance will be investigated in future work, where a larger number of training data could improve the performance. To some extent, the calibration stage is usually time-consuming, and may prove fatiguing for subjects. Therefore, reducing the calibration time is critical for actual BCI application. One of the most widely used approaches is the transfer learning method that exploits shared features between training data recorded from multiple subjects or sessions.

V. CONCLUSION

In this paper, a novel ERP identification method, the MCORCA-based method for N2pc-based BCIs, was proposed to identify the LVF and RVF lateral targets with single-trial classification in visual search tasks. The proposed method was designed to find the linear combinations of multichannel EEG signals that were the optimally correlated components subset in time instead of a maximally correlative component, and to take advantage of the inherent correlations between channel recordings in the same cognitive tasks. Moreover, the proposed method weights the target signals by spatial filters to extract the spatial information for different visual search targets, and improve the SNR of N2pc signals. In short, the MCORCA-based method could get an optimal subset of the correlated components and achieved a good classification performance on CA and ITR when compared to SVM-RFE.

REFERENCES

[1] L. F. Nicolas-Alonso and J. Gomez-Gil, "Brain computer interfaces, a review," *Sensors*, vol. 12, no. 2, pp. 1211–1279, Jan. 2012.
[2] Y. Zhang, D. Guo, P. Xu, Y. Zhang, and D. Yao, "Robust frequency recognition for SSVEP-based BCI with temporally local multivariate synchronization index," *Cognit. Neurodyn.*, vol. 10, no. 6, pp. 505–511, Dec. 2016.
[3] Y. Zhang, D. Guo, D. Yao, and P. Xu, "The extension of multivariate synchronization index method for SSVEP-based BCI," *Neurocomputing*, vol. 269, pp. 226–231, Dec. 2017.
[4] Y. Zhang, D. Guo, F. Li, E. Yin, Y. Zhang, P. Li, Q. Zhao, T. Tanaka, D. Yao, and P. Xu, "Correlated component analysis for enhancing the performance of SSVEP-based brain-computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 5, pp. 948–956, May 2018.
[5] C. Polprasert, P. Kukieattikool, T. Demeechai, J. A. Ritcey, and S. Siwamogsatham, "New stimulation pattern design to improve P300-based matrix speller performance at high flash rate," *J. Neural Eng.*, vol. 10, no. 10, Apr. 2013, Art. no. 036012.

[6] R. C. Panicker, S. Puthusserypady, and Y. Sun, "An asynchronous P300 BCI with SSVEP-based control state detection," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 6, pp. 1781–1788, Jun. 2011.

[7] S. T. Ahi, H. Kambara, and Y. Koike, "A dictionary-driven P300 speller with a modified interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 19, no. 1, pp. 6–14, Feb. 2011.

[8] E. B. J. Coffey, A.-M. Brouwer, E. S. Wilschut, and J. B. F. van Erp, "Brain–machine interfaces in space: Using spontaneous rather than intentionally generated brain signals," *Acta Astronautica*, vol. 67, no. 1, pp. 1–11, Jul./Aug. 2010.

[9] T. Elbert, B. Rockstroh, W. Lutzenberger, and N. Birbaumer, "Biofeedback of slow cortical potentials. I biofeedback à partir des potentiels corticaux lents. I," *Electroencephalogr. Clin. Neurophysiol.*, vol. 48, no. 3, pp. 293–301, Mar. 1980.

[10] A. Kübler, F. Nijboer, J. Mellinger, T. M. Vaughan, H. Pawelzik, G. Schalk, D. J. McFarland, N. Birbaumer, and J. R. Wolpaw, "Patients with ALS can use sensorimotor rhythms to operate a brain-computer interface," *Neurology*, vol. 64, no. 10, pp. 1775–1777, May 2005.

[11] J. J. Shih, D. J. Krusienski, and J. R. Wolpaw, "Brain-computer interfaces in medicine," *Mayo Clinic Proc.*, vol. 87, no. 3, pp. 268–279, Mar. 2012.

[12] S. R. Soekadar, N. Birbaumer, M. W. Slutzky, and L. G. Cohen, "Brain–machine interfaces in neurorehabilitation of stroke," *Neurobiol. Disease*, vol. 83, no. 1, pp. 172–179, Nov. 2015.

[13] A. Myrden and T. Chau, "A passive EEG-BCI for single-trial detection of changes in mental state," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 4, pp. 345–356, Apr. 2017.

[14] A. D. Gerson, L. C. Parra, and P. Sajda, "Cortically coupled computer vision for rapid image search," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 2, pp. 174–179, Jun. 2006.

[15] L. A. Farwell and E. Donchin, "Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalogr. Clin. Neurophys.*, vol. 70, no. 6, pp. 510–523, Dec. 1988.

[16] B. Rebsamen, C. Guan, H. Zhang, C. Wang, C. Teo, M. H. Ang, and E. Burdet, "A brain controlled wheelchair to navigate in familiar environments," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 18, no. 6, pp. 590–598, Dec. 2010.

[17] A. R. Marathe, A. J. Ries, V. J. Lawhern, B. J. Lance, J. Touryan, K. McDowell, and H. Cecotti, "The effect of target and non-target similarity on neural classification performance: A boost from confidence," *Frontiers Neurosci.*, vol. 9, p. 270, Aug. 2015.

[18] A. Matran-Fernandez and R. Poli, "Brain–computer interfaces for detection and localization of targets in aerial imagess," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 4, pp. 959–969, Apr. 2017.

[19] A. Matranfernandez and R. Poli, "Towards the automated localisation of targets in rapid image-sifting by collaborative brain-computer interfaces," *PLoS ONE*, vol. 12, no. 5, pp. 1–28, May 2017.

[20] M. J. Wieser, A. Hambach, and M. Weymar, "Neurophysiological correlates of attentional bias for emotional faces in socially anxious individuals—Evidence from a visual search task and N2pc," *Biol. Psychol.*, vol. 132, pp. 192–201, Feb. 2018.

[21] S. J. Luck and S. A. Hillyard, "Electrophysiological correlates of feature analysis during visual search," *Psychophysiology*, vol. 31, no. 3, pp. 291–308, May 1994.

[22] M. Eimer and A. Grubert, "Spatial attention can be allocated rapidly and in parallel to new visual objects," *Current Biol.*, vol. 24, no. 2, pp. 193–198, Jan. 2014.

[23] M. Jenkins, A. Grubert, and M. Eimer, "Category-based attentional guidance can operate in parallel for multiple target objects," *Biol. Psychol.*, vol. 135, pp. 211–219, May 2018.

[24] C. Hickey, J. J. Mcdonald, and J. Theeuwes, "Electrophysiological evidence of the capture of visual attention," *Cognit. Neurosci. J.*, vol. 39, no. 3, pp. 849–860, 2006.

[25] G. F. Woodman and S. J. Luck, "Serial deployment of attention during visual search," *J. Experim. Psychol. Human Perception Perform.*, vol. 29, no. 1, pp. 121–138, 2003.

[26] M. Eimer, M. Kiss, and S. Nicholas, "What top-down task sets do for us: An ERP study on the benefits of advance preparation in visual search," *J. Experim. Psychol. Hum. Perception Perform.*, vol. 37, no. 6, pp. 1758–1766, 2011.

[27] H. Awni, J. J. S. Norton, S. Umunna, and K. D. Federmeier, "Towards a brain computer interface based on the N2pc event-related potential," in *Proc. Int. IEEE/EMBS Conf. Neural Eng. (NER)*, Nov. 2013, pp. 1021–1024.

[28] J. L. S. Blasco, E. Iáñez, A. Úbeda, and J. M. Azorín, "Visual evoked potential-based brain–machine interface applications to assist disabled people," *Expert Syst. Appl.*, vol. 39, no. 9, pp. 7908–7918, Jul. 2012.

[29] Y. Huang, D. Erdogmus, M. Pavel, S. Mathan, and K. E. Hild, "A framework for rapid visual image search using single-trial brain evoked responses," *Neurocomputing*, vol. 74, no. 12, pp. 2041–2051, Jun. 2011.

[30] S. Yuval-Greenberg, O. Tomer, A. S. Keren, I. Nelken, and L. Y. Deouell, "Transient induced gamma-band response in EEG as a manifestation of miniature saccades," *Neuron*, vol. 58, no. 3, pp. 429–441, May 2008.

[31] J. A. Grahn, "Neuroscientific investigations of musical rhythm: Recent advances and future challenges," *Contemp. Music Rev.*, vol. 28, no. 3, pp. 251–277, Dec. 2009.

[32] J. P. Dmochowski, P. Sajda, J. Dias, and L. C. Parra, "Correlated components of ongoing EEG point to emotionally laden attention—A possible marker of engagement?" *Frontiers Hum. Neurosci.*, vol. 6, no. 2, p. 112, May 2012.

[33] S. Kotz *et al.*, Eds. *Breakthroughs in Statistics*. New York, NY, USA: Springer-Verlag, 1992

[34] Y. Tian, A. B. Chica, P. Xu, and D. Yao, "Differential consequences of orienting attention in parallel and serial search: An ERP study," *Brain Res.*, vol. 1391, no. 3, pp. 81–92, May 2011.

[35] D. Yao, "A method to standardize a reference of scalp EEG recordings to a point at infinity," *Physiol. Meas.*, vol. 22, no. 4, pp. 693–711, Oct. 2001.

[36] Y. Tian and D. Yao, "Why do we need to use a zero reference? Reference influences on the ERPs of audiovisual effects," *Psychophysiology*, vol. 50, no. 12, pp. 1282–1290, Dec. 2013.

[37] B. E. Boser, "A training algorithm for optimal margin classifiers," in *Proc. Workshop Comput. Learn. Theory*, Jul. 1992, pp. 144–152.

[38] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995.

[39] N. T. Renukadevi and P. Thangaraj, "Performance evaluation of SVM—RBF kernel for medical image classification," *Global J. Comput. Sci. Technol.*, vol. 13, no. 4, pp. 14–20, May 2013.

[40] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene Selection for Cancer Classification using Support Vector Machines," *Mach. Learn.*, vol. 46, nos. 1–3, pp. 389–422, Jan. 2002.

[41] T. N. Lal, M. Schroder, T. Hinterberger, J. Weston, M. Bogdan, N. Birbaumer, and B. Scholkopf, "Support vector channel selection in BCI," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1003–1010, Jun. 2004.

[42] K. Yan and D. Zhang, "Feature selection and analysis on correlated gas sensor data with recursive feature elimination," *Sens. Actuators B, Chem.*, vol. 212, pp. 353–363, Jun. 2015.

[43] J. P. Dmochowski, M. A. Bezdek, B. P. Abelson, J. S. Johnson, E. H. Schumacher, and L. C. Parra, "Audience preferences are predicted by temporal reliability of neural processing," *Nature Commun.*, vol. 5, no. 5, Jul. 2014, Art. no. 4567.

[44] Y. Zhang, P. Xu, K. Cheng, and D. Yao, "Multivariate synchronization index for frequency recognition of SSVEP-based brain–computer interface," *J. Neurosci. Methods*, vol. 221, pp. 32–40, Jan. 2014.

[45] G. A. Rousselet, J. J. Foxe, and J. P. Bolam, "A few simple steps to improve the description of group results in neuroscience," *Eur. J. Neurosci.*, vol. 44, no. 9, pp. 2647–2651, 2016.

[46] G. A. Rousselet, C. R. Pernet, and R. R. Wilcox, "Beyond differences in means: Robust graphical methods to compare two groups in neuroscience," *Eur. J. Neurosci.*, vol. 46, no. 2, pp. 1738–1748, Jul. 201s7.

[47] R. R. Wilcox and G. A. Rousselet, "A guide to robust statistical methods in neuroscience," *Current protocols Neurosci.*, vol. 82, no. 1, pp. 8.42.1–8.42.30, Jan. 2018.

[48] H. Lu, H.-L. Eng, C. Guan, K. N. Plataniotis, and A. N. Venetsanopoulos, "Regularized common spatial pattern with aggregation for EEG classification in small-sample setting," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 12, pp. 2936–2946, Dec. 2010.

[49] M. Grosse-Wentrup, C. Liefhold, K. Gramann, and M. Buss, "Beamforming in noninvasive brain–computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 4, pp. 1209–1219, Apr. 2009.

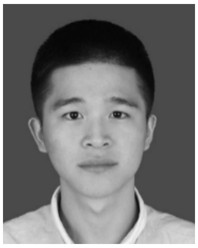
[50] R. Nako, R. Wu, T. J. Smith, and M. Eimer, "Item and category-based attentional control during search for real-world objects: Can you find the pants among the pans?" *J. Exp. Psychol. Hum. Perception Perform.*, vol. 40, no. 4, pp. 1283–1288, 2014.

[51] R. Nako, T. J. Smith, and M. Eimer, "Activation of New Attentional Templates for Real-world Objects in Visual Search," *J. Cognit. Neurosci.*, vol. 27, no. 5, pp. 902–912, 2015.

[52] B. Brisson and P. Jolicoeur, "The N2pc component and stimulus duration," *Neuroreport*, vol. 18, no. 11, pp. 1163–1166, Jul. 2007.

YIN TIAN received the Ph.D. degree in biomedical engineering from the School of Life Science and Technology, University of Electronic Science and Technology of China, in 2009. She is currently a Professor with the Bio-information College, Chongqing University of Posts and Telecommunications. Her research interests include cognitive neuroscience, EEG/fMRI data processing, machine learning, brain network technology, and brain–computer interface.

PEIYANG LI received the Ph.D. degree in biomedical engineering from the University of electronic science and technology of china, Chengdu, Sichuan, China, in 2018. He is currently with the Chongqing University of Posts and Telecommunications, Chongqing, China. His research interests include brain–computer interaction, convex optimization, machine learning, and pattern recognition.



HAIYONG ZHANG received the master's degree in biomedical engineering from the Bio-information College, Chongqing University of Posts and Telecommunications, in 2019. His research interests include biomedical signal processing, visual search analysis, machine learning, and brain–computer interface.

YANG LI received the master's degree in electronics and communication engineering from the Information Science and Technology College, Chengdu University of Technology, in 2017. He is currently pursuing the Ph.D. degree with the Chongqing University of Posts and Telecommunications. His research interests include signal processing, machine learning, and brain–computer interface.