This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2016.2531118, IEEE Transactions on Neural Systems and Rehabilitation Engineering

Discriminative Manifold Learning Based Detection of Movement-Related Cortical Potentials

Chuang Lin, *Member, IEEE*, Binghui Wang, Ning Jiang, *Senior Member, IEEE*, Ren Xu, *Student Member, IEEE*, Natalie Mrachacz-Kersting, and Dario Farina*, *Senior Member, IEEE*

***Abstract*— The detection of voluntary motor intention from EEG has been applied to closed-loop brain-computer interfacing (BCI). The movement-related cortical potential (MRCP) is a low frequency component of the EEG signal, which represents movement intention, preparation, and execution. In this study, we aim at detecting MRCPs from single-trial EEG traces. For this purpose, we propose a detector based on a discriminant manifold learning method, called locality sensitive discriminant analysis (LSDA), and we test it in both online and offline experiments with executed and imagined movements. The online and offline experimental results demonstrated that the proposed LSDA approach for MRCP detection outperformed the Locality Preserving Projection (LPP) approach, which was previously shown to be the most accurate algorithm so far tested for MRCP detection. For example, in the online tests, the performance of LSDA was superior than LPP in terms of a significant reduction in false positives (FP) (passive FP: 1.6 ± 0.9/min vs. 2.9 ± 1.0/min, p=0.002, active FP: 2.2 ± 0.8/min vs. 2.7 ± 0.6/min, p=0.03), for a similar rate of true positives. In conclusion, the proposed LSDA based MRCP detection method is superior to previous approaches and is promising for developing patient-driven BCI systems for motor function rehabilitation as well as for neuroscience research.**

***Index Terms*—Discriminant manifold learning, locality sensitive discriminant analysis (LSDA), movement-related cortical potential (MRCP), Brain–computer interface (BCI), electroencephalogram (EEG)**

This work was financially supported by the EU project DEMOVE (Contract Nr. 267888), Natural Science Foundation of China (Nr. 61272371), Fundamental Research Funds for the Central Universities (Nr.DUT14QY16).

C. Lin is with the Department of Neurorehabilitation Engineering, Bernstein Focus Neurotechnology, Bernstein Center for Computational Neuroscience, University Medical Center Göttingen, Georg-August University, 37075, Göttingen, Germany, (e-mail: lin.chuang@bccn.uni-goettingen.de).

N. Jiang, R. Xu, and D. Farina are with the Department of Neurorehabilitation Engineering, Bernstein Focus Neurotechnology, Bernstein Center for Computational Neuroscience, University Medical Center Göttingen, Georg-August University, 37075, Göttingen, Germany, (e-mail: lin.chuang@bccn.uni-goettingen.de, ning.jiang@bccn.uni-goettingen.de ren.xu@bccn.uni-goettingen.de, dario.farina@bccn.uni-goettingen.de).

B. H. Wang is with the department of Electrical and Computer Engineering, College of Engineering, Iowa State University, Ames, U.S.A. (e-mail: binghui.wang89@gmail.com).

Natalie Mrachacz-Kersting is with the Center for Sensory-Motor Interaction, Department of Health Science and Technology, Aalborg University Denmark (email: nm@hst.aau.dk)

*D. Farina (Email: dario.farina@bccn.uni-goettingen.de)

I. Introduction

Detection of motor intention from electroencephalogram (EEG) signals may be used to provide an effective feedback in neurorehabilitation [1]. In order to promote plasticity, a causal relation between intended movement and the corresponding neural feedback should be established [2]. Detection of voluntary motor intention with a brain-switch is also the basis component for the control of external devices with brain-computer interface (BCI) systems. There are two main types of non-invasive motor-related cortical signals that can be used for detecting motor intention: sensorimotor rhythms (SMR) [3] and movement-related cortical potentials (MRCP) [4]. SMR-based systems can achieve a good detection accuracy, either for busy and idle state or for different movement type, but with a relatively long latency [5], [6]. Conversely, the detection of MRCPs may be done with very short latency (not exceeding 300 ms) [7].

The MRCPs have very low frequency components (0.05 to 3 Hz [8]–[10]) and are associated with movement imagination and execution [8]–[10]. Specifically, a negative deflection is observed in the EEG up to 2 s prior to the onset of the movement, and is considered as a cortical representation of motor preparation [11], [12]. It is then followed by a rebound after the peak of maximum negativity [11]. The negative deflection can be divided into readiness potential (RP) and motor potential (MP), which have been associated with planning and execution of a movement, respectively [9], [11]. The rebound after the peak of maximum negativity is also known as a movement-monitoring potential (MMP) [9], [11], which has been associated with the precision of the movement [13]. It is believed that RP reflects an intention to act, which remains unconscious for part of its time course [14], or an index of resource mobilization [15], and it has been observed even if the movement is not executed but only imagined.

Methods for MRCP detection in BCI applications have been developed in recent years [1], [2], [16]–[19]. These BCI systems detect the motor intention from MRCP and translate it into commands for external devices [15], [16]. The early detection of an intended action in relation to the task onset is crucial in systems where the brain signal is used to drive an external device for neuromodulation purposes [7], [16]. Specifically, Niazi et al. [1] have shown that it is possible to

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2016.2531118, IEEE Transactions on Neural Systems and Rehabilitation Engineering

detect a user's motor intention from MRCPs by a simple matched filter (MF) with supervised extraction of the MRCP template, and used it to drive an electrical stimulator [16]. Jiang et al. [19] have also shown that independent component analysis (ICA) [20] can significantly improve the MRCPs' detection performance.

The major drawback of previous methods for MRCP detection is that they did not adaptively select the features [1], [16], [17], [19], [21]–[23]. Moreover, they did not consider the important intrinsic structure of MRCPs, i.e., they did not use the local neighborhood affinity of MRCPs for designing the detector. For this reason, recently, Ren et al. [2], [18], [24] proposed to use a classical manifold learning method, locality preserving projections (LPP [25], [26]), to detect MRCPs. This approach achieved good performance. However, the false positive rate (FPR) has been shown to be still relatively high and this is a critical issue in BCI control.

Nowadays, many techniques based on discriminant manifold learning have been developed for computer vision and pattern recognition, with specific applications in face recognition [27]–[31]. All the manifold learning based methods assume that naturally generated data lie on a low dimensional sub-manifold of their ambient space. Additionally, different from the one graph based methods [25], [26], more recent techniques [27]–[31] are based on two graphs, i.e., the within-class graph and the between-class graph, to discover both the local geometric structure and the global discriminant structure of the data manifold. In this way, the data between different classes can be maximally separated, and the variance of data within the same class can be minimized.

In this paper, we propose a discriminant manifold learning method, the locality sensitive discriminant analysis (LSDA) [31], based detector to further enhance the performance of MRCPs' detection. The procedure is based on the construction of the within-class graph and the between-class graph from a training set and to find a linear transformation matrix to map the high-dimensional data to a low-feature space. With the transformation matrix, both the local neighborhood information and global discriminant information of the data are effectively preserved. Finally, a nearest-neighbor (NN) classifier is applied for detection. The dramatic advantage of this novel detection scheme is that it preserves a higher true positive rate (TPR) and at the same time suppresses the FPR with respect to previous methods proposed for MRCP detection. This characteristic is due to the following considerations. First, the discriminative intrinsic features are automatically extracted by discovering the global discriminant structure and preserving the local structure concurrently. Second, to satisfy the manifold assumption, we artificially reproduce more samples (one sample corresponding to one MRCP signal) by moving a time window associated with the MRCP (see details in the METHODS part). The reproduced samples increase the available numbers of MRCPs, which benefits the final MRCP detection. Additionally, our method is very efficient, for it is a linear method, which requires a simple implementation [31]. In comparison, the former time-frequency and SVM based methods are much complicated. In [22], the features are pre-defined, the adopted classifier SVM is also complicated and vulnerable to parameters' selection comparing with our method, although it is a kind of popular methods for classification and regression analysis [32], [33].

II. METHODS

In the method part we first introduce the fundamental theory of LSDA, and then the offline and online experimental protocols. Finally, we introduce the statistical analysis methods for the experimental data. Throughout this paper, boldface uppercase letters denote matrices, boldface lowercase letters denote vectors, while italics denote scalars.

A. LSDA

Manifold learning has attracted much attention in computer vision and pattern recognition[25]–[31]. These methods often assume that naturally generated data lies on a sub-manifold of its ambient space, and aim at discovering the intrinsic structure of the data. However, for the detection purpose, extra discriminant structure should be incorporated into the data, and thus the discriminant manifold learning methods arise. Among them, LSDA is particularly interesting because of its algorithmic simplicity and computational efficiency [31].

Suppose there are N samples $\mathbf{X}=\{\mathbf{x}_i\}_{i=1}^{N}\in \mathrm{R}^{m\times N}$ generated from the underlying sub-manifold. We can build a nearest neighbor graph associated with a weight matrix to model the local geometric structure of the sub-manifold. Specifically, the graph $\mathbf{G}$ is defined with N vertices, with each representing a sample, and an edge is put between vertices $\mathbf{x}_i$ and $\mathbf{x}_j$, if $\mathbf{x}_i$ and $\mathbf{x}_j$ are "close", i.e., $\mathbf{x}_i$ is among the k nearest neighbors of $\mathbf{x}_j$ or vice versa. Then, the corresponding $N\times N$ weight matrix $\mathbf{W}$ can be defined with each W_{ij} representing the affinity of the edge joining vertices $\mathbf{x}_i$ and $\mathbf{x}_j$. For simplicity, we set $W_{ij}=1$.

Furthermore, to incorporate discriminative information into the sub-manifold, LSDA constructs two graphs: the within-class graph $\mathbf{G}_w$ and the between-class graph $\mathbf{G}_b$. Let $l(\mathbf{x}_i)$ be the class label of sample $\mathbf{x}_i$. For each $\mathbf{x}_i$, its k nearest neighbor set $\mathbf{N}(\mathbf{x}_i)=\{\mathbf{x}_i^{(1)},\mathbf{x}_i^{(2)},\cdots,\mathbf{x}_i^{(k)}\}$ can then be split into two parts: $\mathbf{N}_w(\mathbf{x}_i)$ and $\mathbf{N}_b(\mathbf{x}_i)$. $\mathbf{N}_w(\mathbf{x}_i)$ contains the neighbors sharing the same label with $\mathbf{x}_i$, while $\mathbf{N}_b(\mathbf{x}_i)$ contains the neighbors having different labels with $\mathbf{x}_i$. That is,

$$\mathbf{N}_{\mathrm{w}}(\mathbf{x}_i)=\{\mathbf{x}_i^{(m)}\mid l(\mathbf{x}_i^{(m)})=l(\mathbf{x}_i)\},\quad \mathbf{N}_{\mathrm{b}}(\mathbf{x}_i)=\{\mathbf{x}_i^{(n)}\mid l(\mathbf{x}_i^{(n)})\neq l(\mathbf{x}_i)\}$$

where $\mathbf{N}_{\mathrm{w}}(\mathbf{x}_i)\cup\mathbf{N}_{\mathrm{b}}(\mathbf{x}_i)=\mathbf{N}(\mathbf{x}_i)$ and $|\mathbf{N}_{\mathrm{w}}(\mathbf{x}_i)|+|\mathbf{N}_{\mathrm{b}}(\mathbf{x}_i)|=k$. Let $\mathbf{W}_w$ and $\mathbf{W}_b$ be the weight matrices of $\mathbf{G}_w$ and $\mathbf{G}_b$, respectively. Similarly, we have $W_{w,ij}=1$ if $\mathbf{x}_i\in\mathbf{N}_{\mathrm{w}}(\mathbf{x}_j)$ or $\mathbf{x}_j\in\mathbf{N}_{\mathrm{w}}(\mathbf{x}_i)$, and $W_{b,ij}=1$ if $\mathbf{x}_i\in\mathbf{N}_b(\mathbf{x}_j)$ or $\mathbf{x}_j\in\mathbf{N}_b(\mathbf{x}_i)$.

Now, consider the problem of mapping samples onto $\mathbf{G}_w$ and $\mathbf{G}_b$ to a line such that the connected samples of $\mathbf{G}_w$ are as close as possible while those of $\mathbf{G}_b$ are as far as possible. Let $\mathbf{y}=[y_1,y_2,\cdots,y_N]^T$ be such a map. Then, we have the following two objective functions

$$\min\sum_{i,j}(y_i-y_j)^2W_{w,ij} \tag{1a}$$

$$\max\sum_{i,j}(y_i-y_j)^2W_{b,ij} \tag{1b}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2016.2531118, IEEE Transactions on Neural Systems and Rehabilitation Engineering

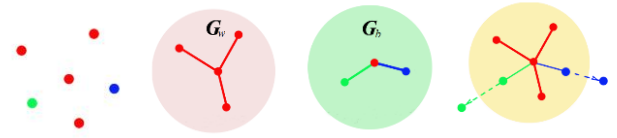

Fig.1. Illustration of how LSDA deals with neighbors with different labels. (a) The center point has five neighbors. The points with the red color belong to one class, the other neighbors belong to another class. (b) The within-class graph consists of points with the same label. (c) The between-class graph connects consists of points with different labels. (d) After LSDA, the margin between different classes is maximized.

under certain constraints, $W_{w,ij}=1$ if $\mathbf{x}_i \in \mathbf{N}_w(\mathbf{x}_j)$ or $\mathbf{x}_j \in \mathbf{N}_w(\mathbf{x}_i)$, and $W_{b,ij}=1$ if $\mathbf{x}_i \in \mathbf{N}_b(\mathbf{x}_j)$ or $\mathbf{x}_j \in \mathbf{N}_b(\mathbf{x}_i)$. Suppose $\mathbf{a}$ is the projection vector, that is, $\mathbf{y}^T=\mathbf{a}^T\mathbf{X}$. By simple algebraic formulations, formula (1a) and (1b) can be respectively reduced to:

$$\max_{\mathbf{a}} \mathbf{a}^T\mathbf{X}\mathbf{W}_w\mathbf{X}^T\mathbf{a} \tag{2a}$$

$$\max_{\mathbf{a}} \mathbf{a}^T\mathbf{X}\mathbf{L}_b\mathbf{X}^T\mathbf{a} \tag{2b}$$

where $\mathbf{L}_b=\mathbf{D}_b-\mathbf{W}_b$, and $\mathbf{D}_b$ is a diagonal matrix whose entries are $D_{b,ii}=\sum_j W_{b,ij}$.

Finally, the optimization problem of (2a) and (2b) becomes:

$$\arg\max_{\mathbf{a}} \mathbf{a}^T\mathbf{X}(\alpha\mathbf{L}_b+(1-\alpha)\mathbf{W}_w)\mathbf{X}^T\mathbf{a} \quad s.t.\quad \mathbf{a}^T\mathbf{X}\mathbf{D}_w\mathbf{X}^T\mathbf{a}=1 \tag{3}$$

where $\mathbf{D}_w$ is a diagonal matrix whose entries are $D_{w,ii}=\sum_j W_{w,ij}$ and the parameter α, ($0\le\alpha\le1$), controls the importance between the local intrinsic structure and the global discriminant structure.

The projection $\mathbf{a}$ that maximizes (3) can be obtained by solving the generalized eigenfunction ploblem:

$$\mathbf{X}(\alpha\mathbf{L}_b+(1-\alpha)\mathbf{W}_w)\mathbf{X}^T\mathbf{a}=\lambda\mathbf{X}\mathbf{D}_w\mathbf{X}^T\mathbf{a} \tag{4}$$

Let $\mathbf{A}=[\mathbf{a}_1,\mathbf{a}_2,...,\mathbf{a}_d]$ be the projection matrix composed of d eigenvectors corresponding to the first d largest eigenvalues $\lambda_1\ge\lambda_2\ge\cdots\ge\lambda_d$ solved by (4). Then, the optimal embedding is:

$$\mathbf{x}_i \rightarrow \mathbf{y}_i=\mathbf{A}^T\mathbf{x}_i \tag{5}$$

Finally, a nearest-neighbor (NN) classifier with Euclidean metric is applied to perform the detection.

In MRCP detection, the characteristics of the noise components may be similar to those of the MRCP signals to be detected. In the LSDA framework, as is shown in Fig. 1, although these noise components are neighbors to MRCP signals, are attributed to another class, and be described by G_b. After discriminant learning, such kind of noise signals are set as far as possible from the MRCP signals, which greatly declines their interference to the detection of MRCP signals. At the same time, by the construction of G_w, LSDA preserves the local structure. For this reason, it is theoretically expected that LSDA should achieve significantly lower FPR with similar TPR, with respect to the previously proposed LPP based MRCP detection [2], [18].

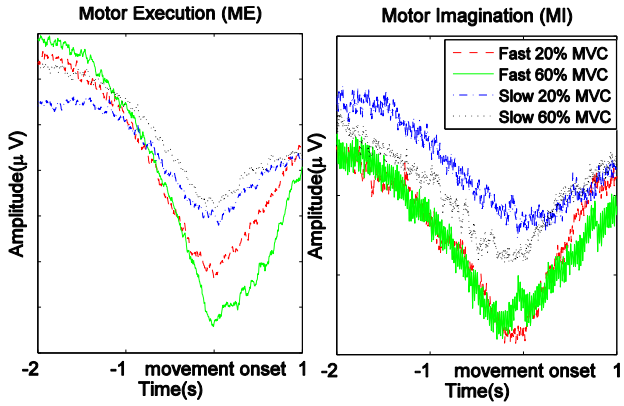

Fig. 2. Average EEG traces corresponding to the 4 tasks for motor execution (ME) and motor imagination (MI) across the 12 subjects.

B. *Offline Experimental Protocol*

We used data previously recorded [22] to test the offline performance of LSDA with respect to the current state-of-art method for MRCP detection, LPP [2], [18].

The dataset was collected from 12 healthy subjects (4 women and 8 men: 27±6 years old), who had no prior BCI experience. All procedures were approved by the local ethical committee (N-20100067), and the participants signed the informed consent before the experiment.

The subject was seated in a comfortable chair, in an electrically shielded room. Their right foot was fixed to a customized pedal. Force was recorded from a force transducer mounted to the pedal. At the beginning of the experiment, the maximum voluntary contraction (MVC) force was recorded according to 4×50 repetitions of cued movements. Each subject performed the following four tasks of ankle dorsiflexion: 1) 0.5s to reach 20% MVC ("Fast20"), 2) 0.5s to reach 60% MVC ("Fast60"), 3) 3s to reach 20% MVC ("Slow20"), and 4) 3s to reach 60% MVC ("Slow60") [22]. A custom made program (Knud Larsen, SMI, Aalborg University) was used to visually cue them to perform the movements correctly. Each movement was separated with 8-10 s and initiated by a digital trigger.

The EEG signals were band-pass filtered with a 2nd order Butterworth filter from 0.05-10 Hz in the forward and reverse direction (in our online experiment, the EEG signal were further filtered from 0.05-3Hz). To correct the poor spatial resolution of EEG, the signals were spatially filtered by a large Laplacian filter [34]. The signals corresponding to motor execution (ME) and motor imagination (MI) were used in the present study. MI is the most challenging for detection and the closest to the practical applications. The grand averages of the EEG traces for ME and MI across the 12 subjects are shown in Fig. 2.

The ME and MI EEG signals were segmented into two classes: 1) signal intervals that contained MRCPs during motor execution and imagination; and 2) noise signal intervals. The duration of the time windows was set to 2 s. The distance

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2016.2531118, IEEE Transactions on Neural Systems and Rehabilitation Engineering

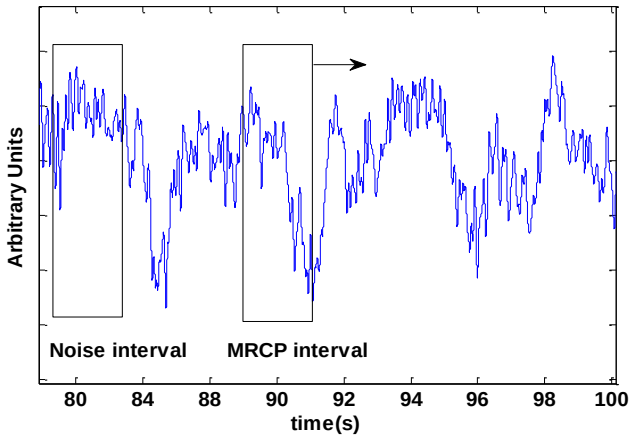

Fig. 3. Illustration of how to select MRCP and noise intervals. Rectangles represent the initial time windows with 2s duration. These time windows were shifted 10 times with the step of 0.05s for obtaining a number of intervals containing the MRCPs. Data comes from Fast20, subject 1.

between the center of the time window and peak of maximum negativity of EEG signal was set to 1 s. The time window was shifted 10 times towards the peak of the maximum negativity of EEG signal with a step of 0.05 s. All the shifted intervals corresponded to MRCP signal intervals while the remaining parts of the recording were considered as noise intervals (Fig. 3).

We extracted discriminative and locality preserving features of these two classes via LSDA [31] and LPP [25], respectively. A nearest-neighbor classifier was then applied to these features. The detection performance was compared between the two methods. Both in LSDA and LPP, the preserved dimension was set to 30. 30 percent of the intervals were used for training, the other parts were used for testing, and the results were 3-folds cross validated.

C. Online Experimental Protocol

The proposed detection method was also tested online with a specific experiment designed for this study.

1) Experimental Setup

Nine channels of EEG channels were collected using the actiCAP system and the g.USBamp amplifier (gTec, GmbH) from the standard positions Cz, Fz, FC1, FC2, C3, C4, CP1, CP2, and Pz. The ground electrode and reference electrode were placed on AFz and the left earlobe, respectively, according to the standard international 10-20 system. The EEG signals were low frequency filtered by a Butterworth filter and spatially filtered by a Laplacian filter [34] to obtain one EEG derivation.

Surface electromyography (sEMG) signals were also recorded in monopolar derivation from the right tibialis anterior (TA) muscle with disposable electrodes (Neuroline 720, Ambu) located at muscle belly. The ground electrode for EMG was placed at the right ankle.

2) Experimental procedures

The online experiment was executed on 8 healthy subjects (men, age: 30.1±3.8 yrs) and consisted of three sessions: preparation, training, and testing. The experiments were approved by the local ethical committee and the subjects signed an informed consent form before participation. The experimental procedures included the signal collection session, training session, and the testing session.

In the signal collection session, the subjects were asked to perform ballistic dorsiflexions of their right ankle. After each trial, the subject could see the low frequency filtered and spatially filtered EEG signal and modulate the movement to generate stable EEG signals.

In the training session, the subjects were asked to perform self-paced motor executions, *i.e.* isometric dorsiflexions, with resting intervals of 10-20 s in between. The EEG traces were segmented with windows starting 2.5 s before and ending 1.5 s after the movement onset, as estimated from the EMG signal. Corrupted trials were flagged by the experimenter, and excluded from further processing. These segments were used for the training of the detection methods. The reason for selecting executed movements rather than imagery as the training set is to obtain a precise alignment of the MRCPs in the training set by using the EMG signals, which is not possible for motor imagery. Moreover, the signal-to-noise ratio of MRCPs is greater for executed than imagined movements [4], [22]. In practice, this training approach would be limited to patients with residual muscle activity [18].

In the training session, the subjects were asked to perform 20 successful executions, which is a smaller number of trials with respect to previous studies [18].

After the training phase for LSDA and LPP, two projection matrices were obtained and used for dimensionality reduction in the testing trials. The nearest-neighbor classifier was used for classifying the MRCP signal intervals and noise intervals.

Two runs (continuous recording) were performed in the online session for testing LSDA and LPP. In each run, the subjects performed first motor executions and then imagery. The order of the two runs was randomized.

Both LSDA and LPP were applied on execution and imagery. During execution, the subject was asked to perform self-paced dorsiflexions 30 times. During imagery, the subjects were asked to perform self-paced dorsiflexion imagery also 30 times, and to lightly say ‘YES’ approximately 1 s after each imagery task in order to have reference labels for the motor imagery events. Preliminary investigations showed that light speech does not influence the detection of MRCP. A red bar appeared on the screen every time a detection occurred (either true or false detection) as an online feedback for the subject. The number of true positive and false positive were recorded.

Both executions and imagery in the testing session were referred to as active phases. There were 90 seconds passive phases before and after each active phase, in which the subjects were requested to totally relax. Eye movements, mouth movements, or slight trunk movements were allowed in this phase to mimic realistic condition. The number of passive false positive (pFP) was counted for the two intervals of 90 seconds.

For both executions and imagery, the true positive rate (TPR), active false positives (aFP), and passive false positives (pFP) were calculated offline. TPR was the ratio of true detections to the 30 attempts. The aFP was the number of false

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2016.2531118, IEEE Transactions on Neural Systems and Rehabilitation Engineering

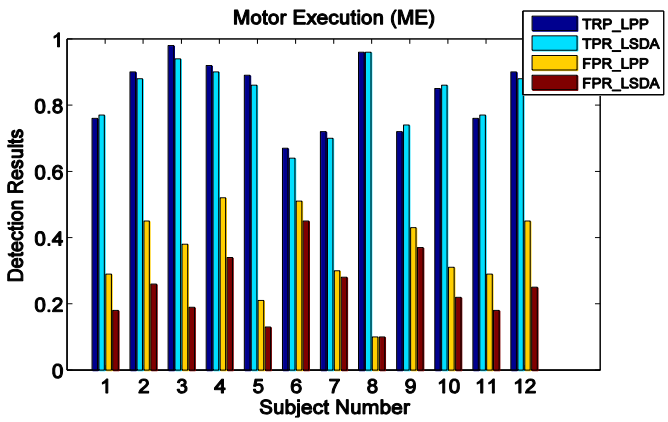

Fig. 4. Average detection accuracy with LSDA-NN and LPP-NN corresponding to the 4 tasks for motor execution (ME) across the 12 subjects. TPR and FPR is the ratio of the number of true and false detections to total attempts, respectively.

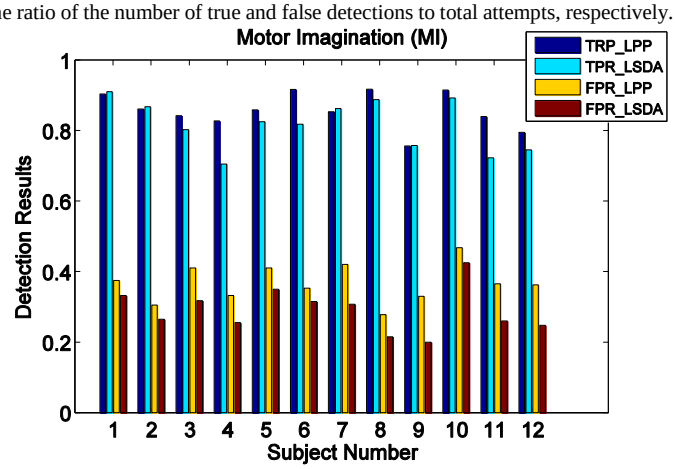

Fig. 5. Average detection accuracy with LSDA-NN and LPP-NN corresponding to the 4 tasks for motor imagination (MI) across the 12 subjects. TPR and FPR is the ratio of the number of true and false detections to total attempts, respectively.

detections per minute in the active phases and pFP was the number of false detections per minute in the passive phases.

D. Statistical Analysis

Due to the limited number of subjects, a nonparametric test (Mann–Whiteney U Test) was used to analyze significant differences among the method for detection and the type of task. The response variable of the Mann–Whiteney U Test included TPR, aFP, and pFP. The two factors of the test were: tasks (two levels: execution and imagery) and algorithms (two levels: LSDA-NN and LPP-NN). For all tests, the significance level was set to 0.05.

III. Experimental Results

A. Offline experimental results

For the offline data (Motor Exec., ME), as shown in Fig. 4, TPR was found similar for LSDA-NN and LPP-NN (82 ± 14% vs. 83 ± 14%, p=0.77). However, FPR (Exec.) with LSDA-NN was significantly lower than with LPP-NN (26 ± 12% vs. 34 ± 16%, p=0.028).

For the offline data (Motor Imag., MI), as shown in Fig. 5, TPR was found similar for LSDA-NN and LPP-NN (82 ± 7% vs. 85 ± 5%, p=0.17). However, FPR (Imag.) with LSDA-NN was significantly lower than with LPP-NN (29 ± 6% vs. 35 ± 7%, p=0.03).

The offline experimental results show the advantage of our scheme, which encourage us to investigate the online results, which are closer to real life situations.

B. Online experimental results

The results for the online analysis are summarized in Table I. The Mann–Whiteney U Test showed that there was no statistical significance between execution and imagery with LPP-NN for TPR, aFP, and pFP (p=0.90, 0.70, and 0.96, respectively). The same results were observed with LSDA-NN (p=0.78, 0.19, and 0.09, respectively). Following these result, the data from the two tasks were pooled together for comparison between methods.

aFP was the active false positive rate, i.e. false positives occurring during execution and imagination of the task. pFP was the passive false positive rate, i.e., false positives occurring in the idle state. pFP were divided into two groups: Prior90s and After90s, which represent the periods prior and after the execution/imagination trials in the experiment.

In the online experiments, the TPR was similar between LSDA-NN and LPP-NN (75.5 ± 12.0% vs. 73.5 ± 11.7%, p=0.57). However, both pFP and aFP with LSDA-NN was significantly lower than with LPP-NN (pFP: 1.6 ± 0.9/min vs. 2.9 ± 1.0/min, p=0.002; aFP: 2.2 ± 0.8/min vs. 2.7 ± 0.6/min, p=0.03).

Fig. 6 reports the improvement in FP when using the

TABLE I DETECTION ACCURACY WITH LSDA-NN AND LPP_NN.
TPR IS THE RATIO OF THE NUMBER OF TRUE DETECTIONS TO TOTAL ATTEMPTS; WHILE AFP AND PFP ARE THE NUMBER OF FALSE DETECTIONS IN ACTIVE PHASE AND PASSIVE PHASE, RESPECTIVELY. EXEC. AND IMAG. REPRESENT EXECUTION PHASE AND IMAGERY PHASE, RESPECTIVELY.

Subject	LPP-NN						LSDA-NN					
	TPR%		aFP/min		pFP		TPR%		aFP/min		pFP	
	Exec.	Imag.	Exec.	Imag.	Prior90s	After90s	Exec.	Imag.	Exec.	Imag.	Prior90s	After90s
1	81	90	2.9	2.0	2	2	91	90	1.7	2.2	2	0
2	80	80	2.7	1.6	4	3	81	80	0.7	1.2	4	0
3	70	60	2.3	2.8	3	4	71	80	0.9	3.0	2	3
4	65	73	2.6	2.1	6	3	63	67	2.5	2.8	5	3
5	83	90	2.4	2.7	4	5	86	90	2.1	2.2	3	2
6	82	81	2.8	3.0	5	6	84	81	1.8	2.0	2	1
7	60	58	3.5	4.2	6	7	61	53	2.8	3.3	4	3
8	71	52	3.2	3.1	4	5	69	61	2.9	2.7	3	2
Mean ± SD	**73.5 ± 11.7%**		**2.7 ± 0.6/min**		**2.9 ± 1.0/min**		**75.5 ± 12.0%**		**2.2 ± 0.8/min**		**1.6 ± 0.9/min**	

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2016.2531118, IEEE Transactions on Neural Systems and Rehabilitation Engineering

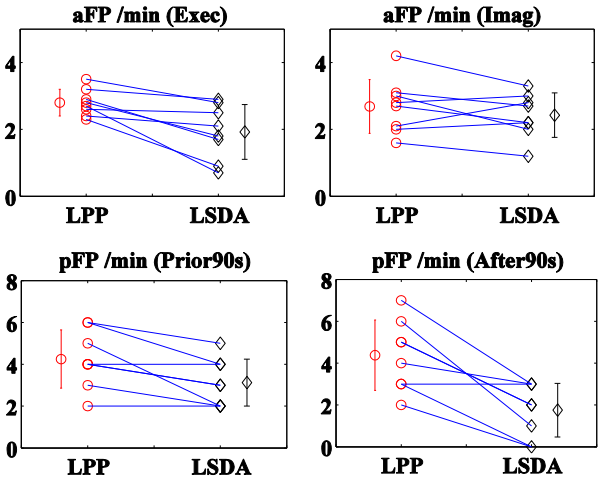

Fig. 6. aFP and pFP obtained with LSDA compared to LPP for all subjects.

proposed method for each subject. The LSDA corresponded to a decrease in aFP in 8 and 5 subjects out of 8 for execution and imagination respectively. Furthermore, pFP decreased with LSDA with respect to LPP in 6 and 7 subjects before and after the execution/imagination trials. In the cases in which LSDA was not effective in reducing FP, the difference with LPP was minimal, so that on average the decrease was significant.

IV. DISCUSSION AND CONCLUSION

The aim of this study was to improve the detection performance of MRCP in EEG traces for BCI applications. For this purpose, we applied a discriminative manifold learning method (LSDA).

We executed offline as well as online experiments for comparing the proposed method with the state of the art approach. The main characteristic of the new approach is the extraction of discriminative and locality preserving features for two classes. The nearest-neighbor classifier was then applied to these features. The detection performance was evaluated on the experimental data and compared with the previously proposed LPP method.

The results showed that the proposed application of LSDA for feature extraction outperformed LPP in MRCP detection because of a significantly reduction in FPR when maintaining a similar level of TPR. The proposed method can be a promising candidate for developing closed-loop BCI neurorehabilitation systems based on MRCP detection.

In the online experiment, the TPR of LPP (73.5 ± 11.7%) and LSDA (75.5 ± 12.0%) found in this study was slightly lower than that reported in [18]. This is due to differences between the two studies. Indeed, in the present study, we reduced the number of training trials to 20, which was only 67 % of the trials in [18]. This reduction was done for showing the possibility of reducing the training time. With limited training samples, the TPR decreased slightly with respect to the previous study but still acceptable for practical applications, as shown in [16].

It was observed that TPR in the online tests was lower than that of the offline tests. This is explained by the fact that we used different datasets in offline and online experiments. In the offline experiments, the dataset was taken from [22], whereas the MRCP signals were collected by cue-based activity of the subjects. The purpose of executing offline experiments is to find out whether the new algorithm is promising in suppressing FP. In the online experiments, we executed a new experiment and the MRCP signals were collected by self-paced activity of the subjects. The subjects involved in the two datasets are also different. The motor imagination experiments need the wholehearted cooperation of the subjects. Moreover, the calculation of FPR in the offline and online experiments was also slightly different, due to the different approaches in recording. In the offline experiment, FPR was the percentage of times in which noise segments were classified as MRCP signals. In the online experiment, the aFP was the number of false detections per minute in the active phases and pFP was the number of false detections per minute in the passive phases.

In conclusion, we proposed a discriminative manifold learning (LSDA) based scheme for MRCPs' detection. The detection performance was evaluated in both offline and online situation. Both the offline and online results showed that our scheme provided similar TPR as in state of the art methods but significantly reduced the FPR. Considering that reduction in FPR is of fundamental relevance in BCI applications, both for communication/control and for rehabilitation/training, the results shown are promising for the application of MRCP as source signal for BCI systems.

REFERENCES

[1] I. K. Niazi, N. Jiang, O. Tiberghien, J. F. Nielsen, K. Dremstrup, and D. Farina, "Detection of movement intention from single-trial movement-related cortical potentials," *Journal of Neural Engineering*, vol. 8, no. 6. p. 066009, Dec-2011.

[2] R. Xu, N. Jiang, N. Mrachacz-Kersting, C. Lin, G. Asin Prieto, J. C. Moreno, J. L. Pons, K. Dremstrup, and D. Farina, "A closed-loop brain-computer interface triggering an active ankle-foot orthosis for inducing cortical neural plasticity," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 7, pp. 2092–2101, 2014.

[3] G. Pfurtscheller and F. H. Lopes Da Silva, "Event-related EEG/MEG synchronization and desynchronization: Basic principles," *Clin. Neurophysiol.*, vol. 110, pp. 1842–1857, 1999.

[4] N. Birbaumer, "Slow cortical potentials: their origin, meaning, and clinical use," *Brain and behavior past, present, and future*. pp. 25–39, 1997.

[5] G. R. Müller-Putz, V. Kaiser, T. Solis-Escalante, and G. Pfurtscheller, "Fast set-up asynchronous brain-switch based on detection of foot motor imagery in 1-channel EEG," *Med. Biol. Eng. Comput.*, vol. 48, no. 3, pp. 229–233, Mar. 2010.

[6] G. Pfurtscheller and T. Solis-Escalante, "Could the beta rebound in the EEG be suitable to realize a 'brain switch'?," *Clin. Neurophysiol.*, vol. 120, no. 1, pp. 24–29, Jan. 2009.

[7] M. Grosse-Wentrup, D. Mattia, and K. Oweiss, "Using brain-computer interfaces to induce neural plasticity and restore function.," *J. Neural Eng.*, vol. 8, no. 2, p. 025004, 2011.

[8] O. F. do Nascimento, K. D. Nielsen, and M. Voigt, "Relationship between plantar-flexor torque generation and the magnitude of the movement-related potentials.," *Exp. brain Res.*, vol. 160, no. 2, pp. 154–165, Jan. 2005.

[9] S. M. Slobounov and W. J. Ray, "Movement-related potentials with reference to isometric force output in discrete and

repetitive tasks," *Exp. Brain Res.*, vol. 123, no. 4, pp. 461–473, Nov. 1998.

[10] N. Birbaumer, T. Elbert, A. G. Canavan, and B. Rockstroh, "Slow potentials of the cerebral cortex and behavior.," *Physiol. Rev.*, vol. 70, no. 1, pp. 1–41, Jan. 1990.

[11] H. Shibasaki and M. Hallett, "What is the Bereitschaftspotential?," *Clin. Neurophysiol.*, vol. 117, no. 11, pp. 2341–2356, Nov. 2006.

[12] O. F. do Nascimento, K. Dremstrup, and M. Voigt, "Movement-related parameters modulate cortical activity during imaginary isometric plantar-flexions," *Exp Brain Res*, vol. 171, pp. 70–90, 2006.

[13] H. Shibasaki, G. Barrett, E. Halliday, and A. M. Halliday, "Cortical potentials associated with voluntary foot movement in man," *Electroencephalogr. Clin. Neurophysiol.*, vol. 52, no. 6, pp. 507–516, Dec. 1981.

[14] B. Libet, C. Gleason, E. Wright, and D. Pearl, "Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential)," *Brain*, vol. 106, no. 3, pp. 623–642, 1983.

[15] W. C. McCallum and S. H. Curry, *Slow potential changes in the human brain*. NATO Advanced Research Workshop on Slow Potential Changes in the Human Brain, May, 1990, Il Ciocco, Italy. Plenum Press, 1993.

[16] I. K. Niazi, N. Mrachacz-Kersting, N. Jiang, K. Dremstrup, and D. Farina, "Peripheral Electrical Stimulation Triggered by Self-Pace Detection of Motor Intention Enhances Motor Evoked Potentials.," *IEEE Trans. Neural Rehabil. Syst. Eng.*, vol. 20, pp. 595–604, 2012.

[17] I. K. Niazi, N. Jiang, M. Jochumsen, J. F. Nielsen, K. Dremstrup, and D. Farina, "Detection of movement-related cortical potentials based on subject-independent training," *Med. Biol. Eng. Comput.*, vol. 51, pp. 507–512, 2013.

[18] R. Xu, N. Jiang, C. Lin, N. Mrachacz-Kersting, K. Dremstrup, and D. Farina, "Enhanced low-latency detection of motor intention from EEG for closed-loop brain-computer interface applications," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 2, pp. 288–296, 2014.

[19] N. Jiang, L. Gizzi, N. Mrachacz-Kersting, K. Dremstrup, and D. Farina, "A brain-computer interface for single-trial detection of gait initiation from movement related cortical potentials.," *Clin. Neurophysiol.*, vol. 126, no. 1, pp. 154–159, 2015.

[20] A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis.," *IEEE Trans. Neural Netw.*, vol. 10, no. 3, pp. 626–34, Jan. 1999.

[21] Y. Gu, O. F. do Nascimento, M.-F. Lucas, and D. Farina, "Identification of task parameters from movement-related cortical potentials.," *Med. Biol. Eng. Comput.*, vol. 47, no. 12, pp. 1257–1264, Dec. 2009.

[22] M. Jochumsen, I. K. Niazi, N. Mrachacz-Kersting, D. Farina, and K. Dremstrup, "Detection and classification of movement-related cortical potentials associated with task force and speed.," *J. Neural Eng.*, vol. 10, no. 5, p. 056015, 2013.

[23] D. Farina, O. F. do Nascimento, M.-F. Lucas, and C. Doncarli, "Optimization of wavelets for classification of movement-related cortical potentials generated by variation of force-related parameters.," *J. Neurosci. Methods*, vol. 162, no. 1, pp. 357–363, May 2007.

[24] T. C. Bulea, S. Prasad, A. Kilicarslan, and J. L. Contreras-Vidal, "Classification of stand-to-sit and sit-to-stand movement from low frequency EEG with locality preserving dimensionality reduction.," *Conf. Proc. ... Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.*, vol. 2013, pp. 6341–4, Jan. 2013.

[25] X. He and P. Niyogi, "Locality Preserving Projections," in *Proc. Conf. Advances in Neural Information Processing Systems*, pp. 153–160, 2003.

[26] X. He, S. Yan, Y. Hu, P. Niyogi, and H. J. Zhang, "Face recognition using Laplacianfaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 3, pp. 328–340, Mar. 2005.

[27] B.-H. Wang, C. Lin, X.-F. Zhao, and Z.-M. Lu, "Neighbourhood sensitive preserving embedding for pattern classification," *IET Image Process.*, vol. 8, no. 8, pp. 489–497, Aug. 2014.

[28] J. Gui, Z. Sun, W. Jia, R. Hu, Y. Lei, and S. Ji, "Discriminant sparse neighborhood preserving embedding for face recognition," *Pattern Recognit.*, vol. 45, no. 8, pp. 2884–2893, Aug. 2012.

[29] B. Raducanu and F. Dornaika, "A supervised non-linear dimensionality reduction approach for manifold learning," *Pattern Recognit.*, vol. 45, no. 6, pp. 2432–2444, Jun. 2012.

[30] S. Yan, D. Xu, B. Zhang, H. Zhang, Q. Yang, and S. Lin, "Graph Embedding and Extensions: A General Framework for Dimensionality Reduction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 1, pp. 40–51, Jan. 2007.

[31] D. Cai, X. He, K. Zhou, J. Han, and H. Bao, "Locality sensitive discriminant analysis," *Proc. 20th Int. Jt. Conf. Artifical Intell.*, pp. 708–713, Jan. 2007.

[32] M. Fernández-Delgado, E. Cernadas, S. Barro, and D. Amorim, "Do we need hundreds of classifiers to solve real world classification problems?," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 3133–3181, Jan. 2014.

[33] C.-C. Chang and C.-J. Lin, "LIBSVM," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 1–27, Apr. 2011.

[34] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K. Muller, "Optimizing Spatial filters for Robust EEG Single-Trial Analysis," *IEEE Signal Process. Mag.*, vol. 25, no. 1, pp. 41–56, 2008.