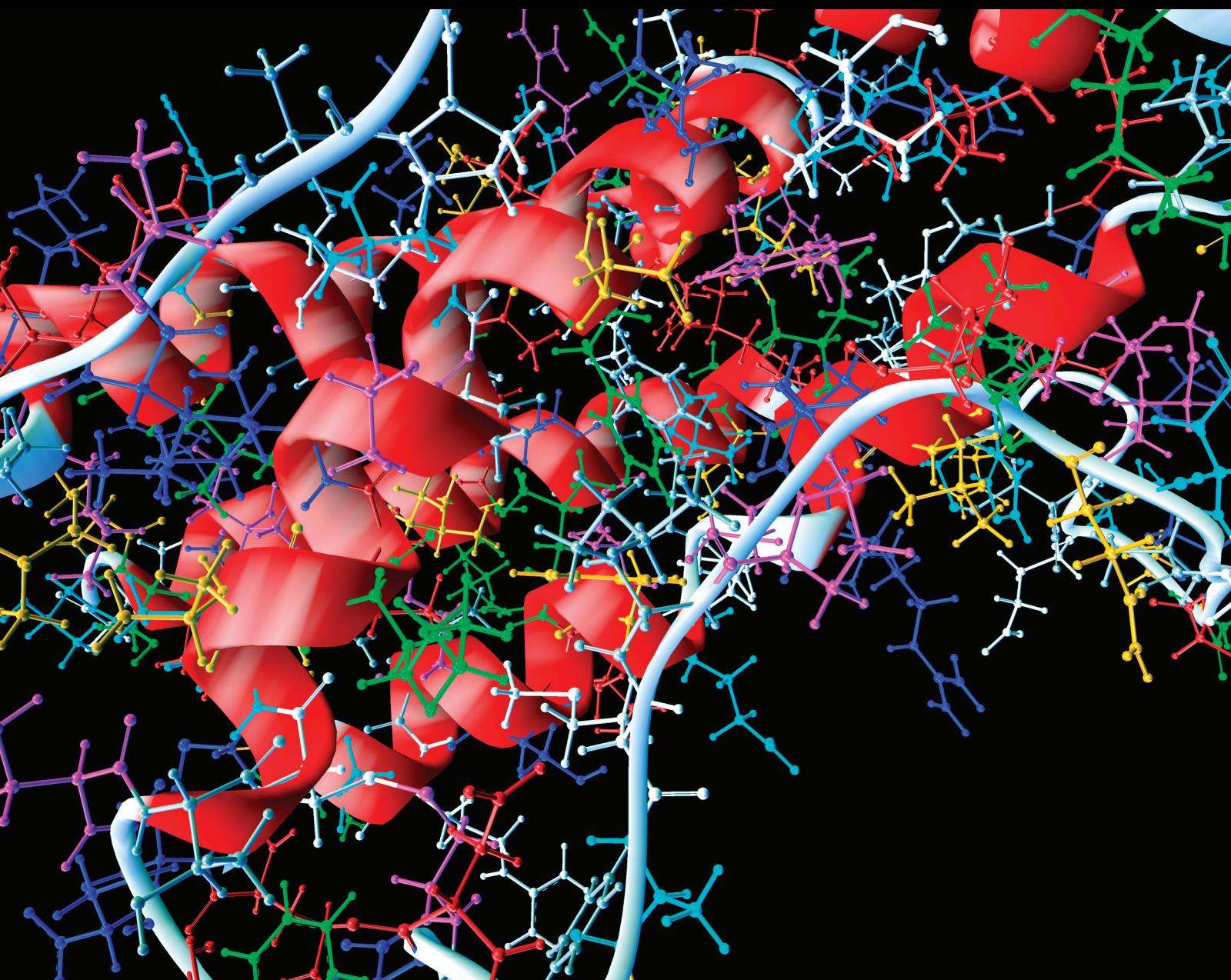


Advanced Signal Processing for Cardiovascular and Neurological Diseases

Lead Guest Editor: Dingchang Zheng

Guest Editors: Fei Chen, Peng Li, and Sheng-Yu Peng



Advanced Signal Processing for Cardiovascular and Neurological Diseases

Advanced Signal Processing for Cardiovascular and Neurological Diseases

Lead Guest Editor: Dingchang Zheng

Guest Editors: Fei Chen, Peng Li, and Sheng-Yu Peng



Copyright © 2018 Hindawi. All rights reserved.

This is a special issue published in "Computational and Mathematical Methods in Medicine." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Editorial Board

Raul Alcaraz, Spain	Luca Faes, Italy	Giuseppe Pontrelli, Italy
Emil Alexov, USA	Maria E. Fantacci, Italy	Jesús Poza, Spain
Konstantin G. Arbeev, USA	Giancarlo Ferrigno, Italy	Christopher Pretty, New Zealand
Georgios Archontis, Cyprus	Marc Thilo Figge, Germany	Mihai V. Putz, Romania
Enrique Berjano, Spain	Alfonso T. García-Sosa, Estonia	Jose Joaquin Rieta, Spain
Junguo Bian, USA	Humberto González-Díaz, Spain	Jan Rychtar, USA
Elia Biganzoli, Italy	Igor I. Goryanin, Japan	Vinod Scaria, India
Konstantin Blyuss, UK	Marko Gosak, Slovenia	Xu Shen, China
Hans A. Braun, Germany	Damien Hall, Australia	Simon A. Sherman, USA
Zoran Bursac, USA	Roberto Hornero, Spain	Dong Song, USA
Thierry Busso, France	Tingjun Hou, China	Xinyuan Song, Hong Kong
Guy Carrault, France	Seiya Imoto, Japan	João M. R. S. Tavares, Portugal
Filippo Castiglione, Italy	Hsueh-Fen Juan, Taiwan	Jlenia Toppi, Italy
Carlos Castillo-Chavez, USA	Martti Juhola, Finland	Nelson J. Trujillo-Barreto, UK
Prem Chapagain, USA	Rafik Karaman, Palestine	Anna Tsantili-Kakoulidou, Greece
Hsiu-Hsi Chen, Taiwan	Andrzej Kloczkowski, USA	Markos G. Tsipouras, Greece
Wai-Ki Ching, Hong Kong	Chung-Min Liao, Taiwan	Po-Hsiang Tsui, Taiwan
Nadia A. Chuzhanova, UK	Ezequiel López-Rubio, Spain	Gabriel Turinici, France
Maria N. D.S. Cordeiro, Portugal	Reinoud Maex, Belgium	Raoul van Loon, UK
Cristiana Corsi, Italy	Valeri Makarov, Spain	Liangjiang Wang, USA
Irena Cosic, Australia	Kostas Marias, Greece	Ruisheng Wang, USA
Qi Dai, China	Juan Pablo Martínez, Spain	David A. Winkler, Australia
Chuangyin Dang, Hong Kong	Richard J. Maude, Thailand	Gabriel Wittum, Germany
Didier Delignières, France	Michele Migliore, Italy	Yu Xue, China
Jun Deng, USA	John Mitchell, UK	Yongqing Yang, China
Thomas Desaive, Belgium	Luminita Moraru, Romania	Chen Yanover, Israel
David Diller, USA	Chee M. Ng, USA	Xiaojun Yao, China
Michel Dojat, France	Michele Nichelatti, Italy	Kaan Yetilmezsoy, Turkey
Irini Doytchinova, Bulgaria	Kazuhisa Nishizawa, Japan	Hiro Yoshida, USA
Esmaeil Ebrahimie, Australia	Francesco Pappalardo, Italy	Henggui Zhang, UK
Georges El Fakhri, USA	Manuel F. G. Penedo, Spain	Yuhai Zhao, China
Issam El Naqa, USA	Jesús Picó, Spain	Xiaoqi Zheng, China
Angelo Facchiano, Italy	Alberto Policriti, Italy	

Contents

Advanced Signal Processing for Cardiovascular and Neurological Diseases

Dingchang Zheng , Fei Chen , Peng Li , and Sheng-Yu Peng
Editorial (2 pages), Article ID 3416540, Volume 2018 (2018)

An Intelligent Parkinson's Disease Diagnostic System Based on a Chaotic Bacterial Foraging Optimization Enhanced Fuzzy KNN Approach

Zhennao Cai , Jianhua Gu, Caiyun Wen, Dong Zhao, Chunyu Huang, Hui Huang , Changfei Tong, Jun Li, and Huiling Chen 
Research Article (24 pages), Article ID 2396952, Volume 2018 (2018)

Common Interferences Removal from Dense Multichannel EEG Using Independent Component Decomposition

Weifeng Li , Yuxiaotong Shen , Jie Zhang, Xiaolin Huang , Ying Chen , and Yun Ge 
Research Article (9 pages), Article ID 1482874, Volume 2018 (2018)

Single-Frequency Ultrasound-Based Respiration Rate Estimation with Smartphones

Linfei Ge , Jin Zhang , and Jing Wei
Research Article (8 pages), Article ID 3675974, Volume 2018 (2018)

Does the Temporal Asymmetry of Short-Term Heart Rate Variability Change during Regular Walking? A Pilot Study of Healthy Young Subjects

Xinpei Wang, Chang Yan, Bo Shi, Changchun Liu, Chandan Karmakar, and Peng Li 
Research Article (9 pages), Article ID 3543048, Volume 2018 (2018)

Kernel Principal Component Analysis of Coil Compression in Parallel Imaging

Yuchou Chang  and Haifeng Wang 
Research Article (9 pages), Article ID 4254189, Volume 2018 (2018)

Scenario Screen: A Dynamic and Context Dependent P300 Stimulator Screen Aimed at Wheelchair Navigation Control

Omar Piña-Ramirez , Raquel Valdes-Cristerna , and Oscar Yanez-Suarez 
Research Article (16 pages), Article ID 7108906, Volume 2018 (2018)

Novel Signal Noise Reduction Method through Cluster Analysis, Applied to Photoplethysmography

William Waugh, John Allen , James Wightman, Andrew J. Sims, and Thomas A. W. Beale 
Research Article (8 pages), Article ID 6812404, Volume 2018 (2018)

Linearized and Kernelized Sparse Multitask Learning for Predicting Cognitive Outcomes in Alzheimer's Disease

Xiaoli Liu, Peng Cao , Jinzhu Yang, and Dazhe Zhao
Research Article (13 pages), Article ID 7429782, Volume 2018 (2018)

Sequential Probability Ratio Testing with Power Projective Base Method Improves Decision-Making for BCI

Rong Liu, Yongxuan Wang, Xinyu Wu, and Jun Cheng
Research Article (10 pages), Article ID 2948742, Volume 2017 (2018)

Editorial

Advanced Signal Processing for Cardiovascular and Neurological Diseases

Dingchang Zheng ,¹ Fei Chen ,² Peng Li ,³ and Sheng-Yu Peng⁴

¹Faculty of Medical Science, Anglia Ruskin University, Chelmsford CM1 1SQ, UK

²Department of Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen, China

³Division of Sleep and Circadian Disorders, Brigham & Women's Hospital; Division of Sleep Medicine, Harvard Medical School, Boston, MA 02115, USA

⁴Department of Electrical Engineering, National Taiwan University of Science and Technology, Taiwan

Correspondence should be addressed to Dingchang Zheng; dingchang.zheng@anglia.ac.uk

Received 24 June 2018; Accepted 24 June 2018; Published 8 July 2018

Copyright © 2018 Dingchang Zheng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Advanced signal processing and computing techniques have been consistently playing a significant role in the field of biomedical engineering research. This special issue focused on the use and elaboration of latest advanced techniques for biomedical data analysis, including but not limited to deep machine learning, compressed sensing, and nonlinear dynamical approaches. Nine out of twenty-one submitted manuscripts in response to this special issue were finally accepted for publication, ranging from (i) noise suppression and removal in EEG and arterial photoplethysmography (PPG) signals; (ii) nonlinear dynamical approaches and multivariate and multiscale techniques for cardiovascular and neurophysiological imaging and signal processing; (iii) machine learning and deep neural network applications of cognitive outcome prediction for Alzheimer's diseases and Parkinson's diseases diagnosis; (iv) advanced signal processing to improve decision-making in brain-computer interface (BCI); and (v) acquisition and analysis of respiratory signals and rates using smartphones.

Since physiological signals are prone to measurement noise and capricious artefact, advanced signal processing is necessary to improve signal quality. W. Li et al. proposed a method based on independent component decomposition for common interference removal in a multichannel EEG recording system. W. Waugh et al. developed an algorithm to reduce sporadic noise in a continuous periodic signal, which has been validated in noise removal in arterial PPG signals.

This technique can be generalised to be applicable to a wide range of other physiological signals. This special issue also accepted a study from Y. Chang and H. Wang who proposed a new channel compression technique in parallel MRI for imaging acceleration via kernel principal component analysis (KPCA).

Heart rate asymmetry (HRA) reflects the balancing regulation of the sympathetic and parasympathetic nervous systems. X. Wang et al. applied short-term HRA analysis to examine whether and how HRA changes during low intensity daily exercises. Once implemented into medical devices, this technique could be used for disease prediction to identify health-related alterations during daily life.

In the category of applying advanced biomedical signal analysis for diseases diagnosis and prediction, Z. Cai et al. developed an enhanced fuzzy k-nearest neighbor (FKNN) method by coupling the chaotic bacterial foraging optimization with Gauss mutation (CBFO) approach with FKNN for the early detection of Parkinson's disease. X. Liu et al. developed and evaluated linearized and kernelized sparse multitask learning for predicting cognitive outcomes in Alzheimer's disease. Their results showed that multitask learning methods not only achieved better prediction performance than the state-of-the-art competitive methods but also effectively fused the multimodality data.

Fast and reliable decision-making is important for real time BCI applications. R. Liu et al. applied the sequential

probability ratio testing (SPRT) with power projective based method to speed up decision-making while trading off errors in BCI. O. Piña-Ramirez et al. proposed a "scenario" stimulation screen, which can be used for commanding a wheelchair even when users have no previous experience on the BCI spelling task.

Regarding the application of advanced signal processing in medical device development, this special issue accepted one study, which analysed cardiovascular signals utilizing mobile devices. L. Ge et al. designed and evaluated a smartphone-based respiration rate detection system utilizing single-frequency ultrasonic signals, where advanced respiration rate detection algorithm was also proposed.

In summary, we believe that this special issue has succeeded to collect papers covering many key areas of advanced signal processing for cardiovascular and neurological diseases. The special issue has provided an international forum for researchers working in the fields of biomedical engineering, medical physics, computational neuroscience, and integrative physiology to present the most recent ideas for understanding, diagnosing, and treatment of cardiovascular and neurological diseases. Ultimately, it will improve scientific debate in this interdisciplinary field.

Acknowledgments

We thank the authors for their excellent contributions to this special issue and the reviewers for their critical comments on the manuscripts that significantly helped to clarify some crucial points and improved the quality.

*Dingchang Zheng
Fei Chen
Peng Li
Sheng-Yu Peng*

Research Article

An Intelligent Parkinson's Disease Diagnostic System Based on a Chaotic Bacterial Foraging Optimization Enhanced Fuzzy KNN Approach

Zhennao Cai ,¹ Jianhua Gu,¹ Caiyun Wen,² Dong Zhao,³ Chunyu Huang,⁴ Hui Huang ,⁵ Changfei Tong,⁵ Jun Li,⁵ and Huiling Chen ⁵

¹School of Computer Science and Engineering, Northwestern Polytechnical University, Xian 710072, China

²Department of Radiology, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, Zhejiang 325035, China

³College of Computer Science and Technology, Changchun Normal University, Changchun 130032, China

⁴College of Computer Science and Technology, Changchun University of Science Technology, Changchun 130032, China

⁵College of Mathematics, Physics and Electronic Information Engineering, Wenzhou University, Wenzhou, Zhejiang 325035, China

Correspondence should be addressed to Huiling Chen; chenhuiling.jlu@gmail.com

Received 29 November 2017; Revised 2 April 2018; Accepted 21 May 2018; Published 21 June 2018

Academic Editor: Dingchang Zheng

Copyright © 2018 Zhennao Cai et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Parkinson's disease (PD) is a common neurodegenerative disease, which has attracted more and more attention. Many artificial intelligence methods have been used for the diagnosis of PD. In this study, an enhanced fuzzy k -nearest neighbor (FKNN) method for the early detection of PD based upon vocal measurements was developed. The proposed method, an evolutionary instance-based learning approach termed CBFO-FKNN, was developed by coupling the chaotic bacterial foraging optimization with Gauss mutation (CBFO) approach with FKNN. The integration of the CBFO technique efficiently resolved the parameter tuning issues of the FKNN. The effectiveness of the proposed CBFO-FKNN was rigorously compared to those of the PD datasets in terms of classification accuracy, sensitivity, specificity, and AUC (area under the receiver operating characteristic curve). The simulation results indicated the proposed approach outperformed the other five FKNN models based on BFO, particle swarm optimization, Genetic algorithms, fruit fly optimization, and firefly algorithm, as well as three advanced machine learning methods including support vector machine (SVM), SVM with local learning-based feature selection, and kernel extreme learning machine in a 10-fold cross-validation scheme. The method presented in this paper has a very good prospect, which will bring great convenience to the clinicians to make a better decision in the clinical diagnosis.

1. Introduction

Parkinson's disease (PD), a degenerative disorder of the central nervous system, is the second most common neurodegenerative disease [1]. The number of people suffering from PD has increased rapidly worldwide [2], especially in developing countries in Asia [3]. Although its underlying cause is unknown, the symptoms associated with PD can be significantly alleviated if detected in the early stages of illness [4–6]. PD is characterized by tremors, rigidity, slowed movement, motor symptom asymmetry, and impaired posture [7, 8]. Research has shown phonation and speech disorders are also common among PD patients [9]. In fact, phonation and

speech disorders can appear in PD patients as many as five years before being clinically diagnosed with the illness [10]. The voice disorders associated with PD include dysphonia, impairment in vocal fold vibration, and dysarthria, disability in correctly articulating speech phonemes [11, 12]. Little et al. [13] first attempted to identify PD patients with dysphonic indicators using a combination of support vector machines (SVM), efficient learning machines, and the feature selection approach. The study results indicated that the proposed method efficiently identified PD patients with only four dysphonic features.

Inspired by the results obtained by Little et al. [13], many other researchers conducted studies on the use of

machine learning techniques to diagnose PD patients on the same dataset (hereafter Oxford dataset). In [14], Das made a comparison of classification score for diagnosis of PD between artificial neural networks (ANN), DMneural, and Regression and Decision Trees. The ANN classifier yielded the best results of 92.9%. In [15], AStröm et al. designed a parallel feed-forward neural network system and yielded an improvement of 8.4% on PD classification. In [16], Sakar et al. proposed a method that combined SVM and feature selection using mutual information to detect PD and obtained a classification accuracy of 92.75%. In [17], a PD detection method developed by Li et al. using an SVM and a fuzzy-based nonlinear transformation method yielded a maximum classification accuracy of 93.47%. In another study, Shahbaba et al. [18] compared the classification accuracies of a nonlinear model based on a combination of the Dirichlet processes, multinomial logit models, decision trees, and support vector machines, which yielded the highest classification score of 87.7%. In [19], Psarakis et al. put forward novel convergence methods and model improvements for multiclass mRVMs. The improved model achieved an accuracy of 89.47%. In [20], Guo et al. proposed a PD detection method with a maximum classification accuracy of 93.1% by combination of genetic programming and the expectation maximization algorithm (GP-EM). In [21], Luukka used a similarity classifier and a feature selection method using fuzzy entropy measures to detect PD, and a mean classification accuracy of 85.03% is achieved. In [22], Ozciit et al. presented rotation forest ensemble classifiers with feature selection using the correlation method to identify PD patients; the proposed model yielded a highest classification accuracy of 87.13%. In [23], Spadoto et al. used a combination of evolutionary-based techniques and the Optimum-Path Forest (OPF) classifier to detect PD with a maximum classification accuracy of 84.01%. In [24], Polat integrated fuzzy C-means clustering-based feature weighting (FCMFW) into a KNN classifier, which yielded a PD classification accuracy of 97.93%. In [25], Chen et al. combined a fuzzy k -nearest neighbor classifier (FKNN) with the principle component analysis (PCA-FKNN) method to detect PD; the proposed diagnostic system yielded a maximum classification accuracy of 96.07%. In [26], Zuo et al. developed an PSO-enhanced FKNN based PD diagnostic system with a mean classification accuracy of 97.47%. In [27–29], Babu et al. proposed a ‘projection based learning meta-cognitive radial basis function network (PBL-McRBFN)’ approach for the prediction of PD, which obtained an testing accuracy of 96.87% on the gene expression data sets, 99.35% on standard vocal data sets, 84.36% on gait PD data sets, and 82.32% on magnetic resonance images. In [30], the hybrid intelligent system for PD detection was proposed which included several feature preprocessing methods and classification techniques using three supervised classifiers such as least-square SVM, probabilistic neural networks, and general regression neural network; the experimental results gives a maximum classification accuracy of 100% for the PD detection. Furthermore, in [31], Gök et al. developed a rotation forest ensemble KNN classifier with a classification accuracy of 98.46%. In [32], Shen et al. proposed an enhanced SVM based on fruit fly optimization algorithm, and have

achieved 96.90% classification accuracy for diagnosis of PD. In [33], Peker designed a minimum redundancy maximum relevance (mRMR) feature selection algorithm with the complex-valued artificial neural network to diagnosis of PD, and obtained a classification accuracy of 98.12%. In [34], Chen et al. proposed an efficient hybrid kernel extreme learning machine with feature selection approach. The experimental results showed that the proposed method can achieve the highest classification accuracy of 96.47% and mean accuracy of 95.97% over 10 runs of 10-fold CV. In [35], Cai et al. have proposed an optimal support vector machine (SVM) based on bacterial foraging optimization (BFO) combined with the relief feature selection to predict PD, the experimental results have demonstrated that the proposed framework exhibited excellent classification performance with a superior classification accuracy of 97.42%.

Different from the work of Little et al., Sakar et al. [36] designed voice experiments with sustained vowels, words, and sentences from PD patients and controls. The paper reported that sustained vowels had more PD-discriminative power than the isolated words and short sentences. The study result achieved 77.5% accuracy by using SVM classifier. From then on, several works have been proposed to detect PD using this PD dataset (hereafter Istanbul dataset). Zhang et al. [37] proposed a PD classification algorithm that integrated a multi-edit-nearest-neighbor algorithm with an ensemble learning algorithm. The algorithm achieved higher classification accuracy and stability compared with the other algorithms. Abrol et al. [38] proposed a kernel sparse greedy dictionary algorithm for classification tasks, comparing with kernel K-singular value decomposition algorithm and kernel multilevel dictionary learning algorithm. The method achieved an average classification accuracy of 98.2% and the best accuracy of 99.4% on the Istanbul PD dataset with multiple types of sound recordings. In [39], the authors investigated six classification algorithms, including Adaboost, support vector machines, neural network with multilayer perceptron (MLP) structure, ensemble classifier, K-nearest neighbor, naive Bayes, and presented feature selection algorithms including LASSO, minimal redundancy maximal relevance, relief, and local learning-based feature selection on the Istanbul PD dataset. The paper indicated that applying feature selection methods greatly increased the accuracy of classification. The SVM and KNN classifiers with local learning-based feature selection obtained the optimum prediction ability and execution times.

As shown above, ANN and SVM have been extensively applied to the detection of PD. However, understanding the underlying decision-making processes of ANN and SVM is difficult due to their black-box characteristics. Compared to ANN and SVM, FKNN is much simpler and yield more easily interpretable results. FKNN [40, 41] classifiers, improved versions of traditional k -nearest neighbor (KNN) classifiers, have been studied extensively since first proposed for the use of diagnostic purposes. In recent years, many variant versions of KNNs based on fuzzy sets theory and several extensions have been developed, such as fuzzy rough sets, intuitionistic fuzzy sets, type 2 fuzzy sets, and possibilistic theory based KNN [42]. FKNN allows for the representation of imprecise

knowledge via the introduction of fuzzy measures, providing a powerful method of similarity description among instances. In FKNN methods, fuzzy set theories are introduced into KNNs, which assign membership degrees to different classes instead of the distances to their k -nearest neighbors. Thus, each of the instances is assigned a class membership value rather than binary values. When it comes to the voting stage, the highest class membership function value is selected. Then based on these properties, FKNN has been applied to numerous practical problems, such as medical diagnosis problems [25, 43], protein identification and prediction problems [44, 45], bankruptcy prediction problems [46], slope collapse prediction problems [47], and grouting activity prediction problems [48].

The classification performance of an FKNN greatly relies on its tuning parameters, neighborhood size (k), and fuzzy strength (m). Therefore, the two parameters should be precisely determined before applying FKNN to practical problems. Several studies concerning parameter tuning in FKNN have been conducted. In [46], Chen et al. presented the particle swarm optimization (PSO) based method to automatically search for the two tuning parameters of an FKNN. According to the results of the study, the proposed method could be effectively and efficiently applied to bankruptcy prediction problems. More recently, Cheng et al. [48] developed a differential evolution optimization approach to determine the most appropriate tuning parameters of an FKNN and successfully applied to grouting activity prediction problems in the construction industry. Later, Cheng et al. [47] proposed using firefly algorithm to tune the hyperparameters of the FKNN model. The FKNN model was then applied to slope collapse prediction problems. The experiment results indicated that the developed method outperformed other common algorithms. The bacterial foraging optimization (BFO) method [49], a relatively new swarm-intelligence algorithm, mimics the cooperative foraging behavior of several bacteria on a multidimensional continuous search space and, therefore, effectively balances exploration and exploitation events. Since its introduction, BFO has been subtly introduced to real-world optimization problems [50–55], such as optimal controller design problems [49], stock market index prediction problems [56], automatic circle detection problems involving digital images [57], harmonic estimation problems [58], active power filter design problems [59], and especially the parameter optimization of machine learning methods [60–63]. In [60], BFO was introduced to wavelet neural network training and applied successfully to load forecasting. In [61], an improved BFO algorithm was proposed to fine-tune the parameters of fuzzy support vector machines to identify the fatigue status of the electromyography signal. The experimental results have shown that the proposed method is an effective tool for diagnosis of fatigue status. In [62], BFO was proposed to learn the structure of Bayesian networks. The experimental results verify that the proposed BFO algorithm is a viable alternative to learn the structures of Bayesian networks and is also highly competitive compared to state-of-the-art algorithms. In [63], BFO was employed to optimize the

training parameters appeared in adaptive neuro-fuzzy inference system for speed control of matrix converter- (MC-) fed brushless direct current (BLDC) motor. The simulation results have reported that the BFO approach is much superior to the other nature-inspired algorithms. In [64], a chaotic local search based BFO (CLS-BFO) was proposed, which introduced the DE operator and the chaotic search operator into the chemotaxis step of the original BFO.

Inspired from the above works, in this paper, the BFO method was integrated with FKNN for the maximum classification performance. In order to further improve the diversity of the bacteria swarm, chaos theory combination with the Gaussian mutation was introduced in BFO. Then, the resulting CBFO-FKNN model was applied to the detection of PD. In our previous work, we have applied BFO in the classification of speech signals for PD diagnosis [35]. In this work, we have further improved the BFO by embedding the chaotic theory and Gauss mutation and combined with the effective FKNN classifier. In order to validate the effectiveness of the proposed CBFO-FKNN approach, FKNN based on five other meta-heuristic algorithms including original BFO, particle swarm optimization (PSO), genetic algorithms (GA), fruit fly optimization (FOA), and firefly algorithm (FA) was implemented for strict comparison. In addition, advanced machine learning methods, including the support vector machine (SVM), kernel based extreme learning machine (KELM) methods, and SVM with local learning-based feature selection (LOGO) [65] (LOGO-SVM), were compared with the proposed CBFO-FKNN model in terms of classification accuracy (ACC), area under the receiver operating characteristic curve (AUC), sensitivity, and specificity. The experimental results show that the proposed CBFO-FKNN approach has exhibited high ACC, AUC, sensitivity, and specificity on both datasets. This work is a fully extended version of our previously published conference paper [66] and that further improved method has been provided.

The main contributions of this study are as follows:

- (a) First, we introduce chaos theory and Gaussian mutation enhanced BFO to adaptively determine the two key parameters of FKNN, which aided the FKNN classifier in more efficiently achieving the maximum classification performance, more stable and robust when compared to five other bio-inspired algorithms-based FKNN models and other advanced machine learning methods such as SVM and KELM.
- (b) The resulting model, CBFO-FKNN, is introduced to discriminate the persons with PD from the healthy ones on the two PD datasets of UCI machine learning repository. It is promising to serve as a computer-aided decision-making tool for early detection of PD.

The remainder of this paper is structured as follows. In Section 2, background information regarding FKNN, BFO, chaos theory, and Gaussian mutation is presented. The implementation of the proposed methodology is explained in Section 3. In Section 4, the experimental design is described in detail. The experimental results and a discussion are presented in Section 5. Finally, Section 6 concludes the paper.

2. Background Information

2.1. Fuzzy k-Nearest Neighbor (FKNN). In this section, a brief description of FKNN is provided. A detailed description of FKNN can be referred to in [41]. In FKNN, the fuzzy membership values of samples are assigned to different categories as follows:

$$u_i(x) = \frac{\sum_{j=1}^K u_{ij} \left(1 / \|x - x_j\|^{2/(m-1)} \right)}{\sum_{j=1}^K \left(1 / \|x - x_j\|^{2/(m-1)} \right)} \quad (1)$$

where $i=1,2,\dots,C$, $j=1,2,\dots,K$, C represents the number of classes, and K means the number of nearest neighbors. The fuzzy strength parameter (m) is used to determine how heavily the distance is weighted when calculating each neighbor's contribution to the membership value. $m \in (1, \infty)$. $\|x - x_j\|$ is usually selected as the value of m . In addition, the Euclidean distance, the distance between x and its j th nearest neighbor x_j , is usually selected as the distance metric. Furthermore, u_{ij} denotes the degree of membership of the pattern x_j from the training set to class i among the k -nearest neighbors of x . In this study, the constrained fuzzy membership approach was adopted in that the k -nearest neighbors of each training pattern (i.e., x_k) were determined, and the membership of x_k in each class was assigned as

$$u_{ij}(x_k) = \begin{cases} 0.51 + \left(\frac{n_j}{K} \right) * 0.49, & \text{if } j = i \\ \left(\frac{n_j}{K} \right) * 0.49, & \text{if } j \neq i. \end{cases} \quad (2)$$

The value of n_j denotes the number of neighbors belonging to j^{th} class. The membership values calculated using (2) should satisfy the following equations:

$$\sum_{I=1}^C \mu_{ij} = 1, \quad j = 1, 2, \dots, n, \quad C \text{ is the number of classes} \quad (3)$$

$$0 < \sum_{j=1}^n u_{ij} < n, \quad u_{ij} \in [0, 1].$$

After calculating all of the membership values of a query sample, it is assigned to the class with which it has the highest degree of membership, i.e.,

$$C(x) = \arg \max_{i=1}^C (u_i(x)) \quad (4)$$

2.2. Bacterial Foraging Optimization (BFO). The bacterial foraging algorithm (BFO) is a novel nature-inspired optimization algorithm proposed by Passino in 2002 [49]. The BFO simulates the mechanism of approaching or moving away while sensing the concentration of peripheral substances in bacterial foraging process. This method contains four basic behaviors: chemotaxis, swarming, reproduction, and elimination-dispersal.

2.2.1. Chemotaxis. The chemotaxis behavior simulates two different positional shifts of *E. coli* bacterium that depend on the rotation of the flagellum, namely, tumbling and moving. The tumbling refers to looking for new directions and the moving refers to keeping the direction going. The specific operation is as follows: first, a unit step is moved in a certain random direction. If the fitness value of the new position is more suitable than the previous one, it will continue to move in that direction; if the fitness value of the new position is not better than before, the tumble operation is performed and moves in another random direction. When the maximum number of attempts is reached, the chemotaxis step is stopped. The chemotaxis step to operate is indicated by the following:

$$\theta^i(j+1, k, l) = \theta^i(j, k, l) + C(i) * dct_i \quad (5)$$

$$dct_i = \frac{\Delta(i)}{\sqrt{\Delta^T(i) \Delta(i)}}$$

where $\theta^i(j, k, l)$ is the position of the i th bacterium. The j , k , and l , respectively, indicate the number of bacterial individuals to complete the chemotaxis, reproduction, and elimination-dispersal. $C(i)$ is the chemotaxis step length for the i th bacteria to move. Δ is the random vector between [-1, 1].

2.2.2. Swarming. In the process of foraging, the bacterial community can adjust the gravitation and repulsion between the cell and the cell, so that the bacteria in the case of aggregation characteristics and maintain their relatively independent position. The gravitation causes the bacteria to clump together, and the repulsion forces the bacteria to disperse in a relatively independent position to obtain food.

2.2.3. Reproduction. In the reproduction operation of BFO algorithm, the algorithm accumulates the fitness values of all the positions that the bacterial individual passes through in the chemotaxis operation and arranges the bacteria in descending order. Then the first half of the bacteria divides themselves into two bacteria by binary fission, and the other half die. As a result, the new reproduced bacterial individual has the same foraging ability as the original individual, and the population size of bacterial is always constant.

2.2.4. Elimination-Dispersal. After the algorithm has been reproduced for several generations, the bacteria will undergo elimination-dispersal at a given probability Ped , and the selected bacteria will be randomly redistributed to new positions. Specifically, if a bacterial individual in the bacterial community satisfies the probability Ped of elimination-dispersal, the individual loses the original position of foraging and randomly selects a new position in the solution space, thereby promoting the search of the global optimal solution.

2.3. Chaotic Mapping. Chaos, as a widespread nonlinear phenomenon in nature, has the characteristics of randomness, ergodicity, sensitivity to initial conditions and so on [67]. Due to the characteristics of ergodicity and randomness, chaotic

motions can traverse all the states in a certain range according to their own laws without repetition. Therefore, if we use chaos variables to search optimally, we will undoubtedly have more advantages than random search. Chaos ergodicity features can be used to optimize the search and avoid falling into the local minima; therefore, chaos optimization search method has become a novel optimization technique. Chaotic sequences generated by different mappings can be used such as logistic map, sine map, singer map, sinusoidal map, and tent map. In this paper, several chaotic maps were tried and the best one was chosen to combine with the BFO algorithm. According to the preliminary experiment, logistic map has achieved the best results. Thus, the chaotic sequences are generated by using logistic map as follows:

$$x_{i+1} = ux_i(1 - x_i) \quad (6)$$

u is the control parameter and let $u = 4$. When $u = 4$, the logistic mapping comes into a thorough chaotic state. Let $x_i \in (0, 1)$ and $x_i \neq 0.25, 0.5, 0.75$.

The initial bacterial population θ is mapped to the chaotic sequence that has been generated according to (6), resulting in a corresponding chaotic bacterial population pch .

$$pch = x_i * \theta \quad (7)$$

2.4. Gaussian Mutation. The Gaussian mutation operation has been derived from the Gaussian normal distribution and has demonstrated its effectiveness with application to evolutionary search [68]. This theory was referred to as classical evolutionary programming (CEP). The Gaussian mutations have been used to exploit the searching capabilities of ABC [69], PSO [70], and DE [71]. Also, Gaussian mutation is more likely to create a new offspring near the original parent because of its narrow tail. Due to this, the search equation will take smaller steps allowing for every corner of the search space to be explored in a much better way. Hence it is expected to provide relatively faster convergence. The Gaussian density function is given by

$$f_{gaussian(0,\sigma^2)}(\alpha) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\alpha^2/2\sigma^2} \quad (8)$$

where σ^2 is the variance for each member of the population.

3. Proposed CBFO-FKNN Model

In this section, we described the new evolutionary FKNN model based on the CBFO strategy. The two key parameters of FKNN were automatically tuned based on the CBFO strategy. As shown in Figure 1, the proposed methodology has two main parts, including the inner parameter optimization procedure and outer performance evaluation procedure. The main objective of the inner parameter optimization procedure was to optimize the parameter neighborhood size (k) and fuzzy strength parameter (m) by using the CBFO technique via a 5-fold cross-validation (CV). Then, the obtained best values of (k, m) were input into the FKNN prediction model in order to perform the PD diagnostic

TABLE 1: Description of the Oxford PD data set.

Label	Feature
S1	MDVP:Fo(Hz)
S2	MDVP:Fhi(Hz)
S3	MDVP:Flo(Hz)
S4	MDVP:Jitter(%)
S5	MDVP:Jitter(Abs)
S6	MDVP:RAP
S7	MDVP:PPQ
S8	Jitter:DDP
S9	MDVP:Shimmer
S10	MDVP:Shimmer(dB)
S11	Shimmer:APQ3
S12	Shimmer:APQ5
S13	MDVP:APQ
S14	Shimmer:DDA
S15	NHR
S16	HNR
S17	RPDE
S18	D2
S19	DFA
S20	Spread1
S21	Spread2
S22	PPE

classification task in the outer loop via the 10-fold CV. The classification error rate was used as the fitness function.

$$fitness = \frac{(\sum_{i=1}^K testError_i)}{k} \quad (9)$$

where $testError_i$ means the average test error of the FKNN classifier.

The main steps conducted by the CBFO strategy are described in detail as shown in Algorithm 1.

4. Experimental Design

4.1. Oxford Parkinson's Disease Data. The Oxford Parkinson's disease data set was donated by Little et al. [13], abbreviation as Oxford dataset. The data set was used to discriminate patients with PD from healthy controls via the detection of differences in vowel sounds. Various biomedical voice measurements were collected from 31 subjects. 23 of them are patients with PD, and 8 of them are healthy controls. The subjects ranged from 46 to 85 years of age. Each subject provided an average of six sustained vowel "ahh..." phonations, ranging from 1 to 36 seconds in length [13], yielding 195 total samples. Each recording was subjected to different measurements, yielding 22 real-value features. Table 1 lists these 22 vocal features and their statistical parameters.

4.2. Istanbul Parkinson's Disease Data. The second data set in this study was deposited by Sakar et al. [36] from Istanbul, Turkey, abbreviation as Istanbul dataset. It contained multiple types of sound recordings, including sustained vowels,

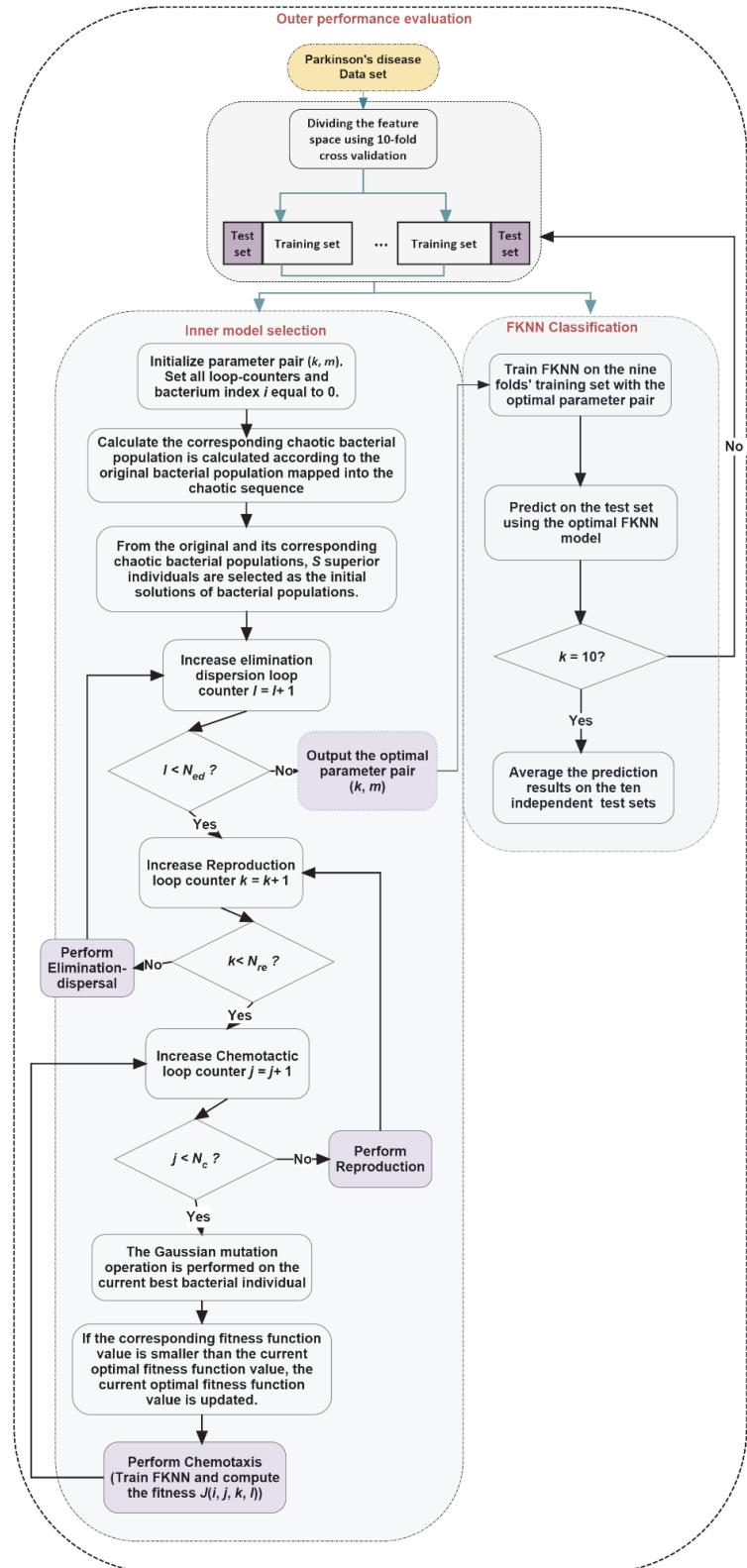


FIGURE 1: Flowchart of the proposed CBFO-FKNN diagnostic system.

```

Begin
Step 1: Parameter Initialization. Initialize the number of dimensions in the search space p, the
swarm size of the population S, the number of chemotactic steps Nc, the swimming length Ns, the
number of reproduction steps Nre, the number of elimination-dispersal events Ned, the
elimination-dispersal probability Ped, the size of the step C(i) taken in the random direction
specified by the tumble.
Step 2: Population Initialization. Calculate chaotic sequence according to Eq. (6). The
corresponding chaotic bacterial population is calculated according to the original bacterial
population mapped into the chaotic sequence according to Eq. (7). From the original and its
corresponding chaotic bacterial populations, S superior individuals are selected as the initial
solutions of bacterial populations.
Step 3: for ell=1:Ned /*Elimination and dispersal loop*/
    for K=1:Nre /*Reproduction loop*/
        for j=1:Nc /* chemotaxis loop*/
            Intertime=Intertime+1; /* represent the number of iterations*/
            for i=1:
                /*fobj represents calculating the fitness of the ith bacterium at the jth
                chemotactic, Kth reproductive, and lth elimination-dispersal steps.*/
                J(i,j,K,ell)=fobj(P(:,i,j,K,ell));
                /* Jlast stores this value since a cost better than a run may be identified.*/
                Jlast=J(i,j,K,ell);
                /* gbest(1,:) stores the current optimal bacterial individual.*/
                gbest(1,:)=P(:,i,j,K,ell);
                Tumble according to Eq.(5)
                /*Swim (for bacteria that seem to be headed in the right direction)*/
                m=0; /* Initialize counter for swim length*/
                while m<Ns
                    m=m+1;
                    if J(i,j+1,K,ell)<Jlast
                        /* Jlast stores this value since a cost better than a run may be identified.*/
                        Jlast=J(i,j+1,K,ell);
                        Tumble according Eq.(5)
                        if Jlast<Gbest
                            /* Gbest stores the current optimal fitness function value.*/
                            Gbest = Jlast;
                            gbest(1,:)=P(:,i,j+1,K,ell);
                        End
                        else
                            m=Ns;
                        End
                Gaussian mutation operation
                Moth_pos_m_gaus=gbest(1,:)*(1+randn(1));
                Moth_fitness_m_gaus=fobj(Moth_pos_m_gaus);
                Moth_fitness_s=fobj(gbest(1,:));
                Moth_fitness_comb=[Moth_fitness_m_gaus,Moth_fitness_s];
                [~,mm]=min(Moth_fitness_comb);
                if mm==1
                    gbest(1,:)=Moth_pos_m_gaus;
                end
                fitnessGbest = fobj(gbest(1,:));
                if fitnessGbest<Gbest
                    Gbest = fitnessGbest;
                end
            End
        End /*Go to next bacterium*/
    End /*Go to the next chemotactic*/
/*Reproduction*/
Jhealth=sum(J(:, :, K, ell), 2); /* Set the health of each of the S bacteria*/
[Jhealth, sortind]=sort(Jhealth); /*Sorts the nutrient concentration in order of ascending*/
/* Rearrange the bacterial population*/
P(:, :, 1, K+1, ell)=P(:, sortind, Nc+1, K, ell);
/*Split the bacteria (reproduction)*/
for i=1:Sr

```

```

/*The least fit do not reproduce, the most fit ones split into two identical copies*/
P(:,i+Sr,1,K+1,ell)=P(:,i,1,K+1,ell);
End
End /*Go to next reproduction*/
/*Elimination-Dispersal*/
for m=1:s
    if Ped>rand /*randomly generates a new individual anywhere in the solution space.*/
        Reinitialize bacteria m
    End
End
End /*Go to next Elimination-Dispersal*/
End

```

ALGORITHM 1: The steps of CBFO.

numbers, words, and short sentences from 68 subjects. Specifically, the training data collected from 40 persons including 20 patients with PD ranging from 43 to 77 and 20 healthy persons ranging from 45 to 83, while testing data was collected from 28 different patients with PD ranging 39 and 79. In this study, we selected only 3 types of sustained vowel recordings /a/, /o/, and /u/, with similar data type to the Oxford PD dataset. We merged them together and produced a database which contains total 288 sustained vowels samples and the analyses were made on these samples. As shown in Table 2, a group of 26 linear and time-frequency based features are extracted for each voice sample.

4.3. Experimental Setup. The experiment was performed on a platform of Windows 7 operating system with an Intel (R) Xeon (R) CPU E5-2660 v3 @ 2.6 GHz and 16GB of RAM. The CBFO-FKNN, BFO-FKNN, PSO-FKNN, GA-FKNN, FOA-FKNN, FA-FKNN, SVM, and KELM classification models were implemented with MATLAB 2014b. The LIBSVM package [72] was used for the SVM classification. The algorithm available at <http://www3.ntu.edu.sg/home/egbhuang> was used for the KELM classification. The CBFO-FKNN method was implemented from scratch. The data was scaled into a range of [0, 1] before each classification was conducted.

The parameters C and γ in $K(x, x_i) = \exp(-\gamma \|x - x_i\|^2)$ used during the SVM and KELM classifications were determined via the grid search method; the search ranges were defined as $C \in \{2^{-5}, 2^{-3}, \dots, 2^{15}\}$ and $\gamma \in \{2^{-15}, 2^{-13}, \dots, 2^5\}$. A population swarm size of 8, chemotactic step number of 25, swimming length of 4, reproduction step number of 3, elimination-dispersal event number of 2, and elimination-dispersal probability of 0.25 were selected for the CBFO-FKNN. The chemotaxis step value was established through trial and error, as shown in the experimental results section. The initial parameters of the other four meta-heuristic algorithms involved in training FKNN are chosen by trial and error as reported in Table 3.

4.4. Data Classification. A stratified k -fold CV [73] was used to validate the performance of the proposed approach and other comparative models. In most studies, k is given the value of 10. During each step, 90% of the samples are used

TABLE 2: Description of the Istanbul PD data set.

Label	Feature
S1	Jitter(local)
S2	Jitter(local, absolute)
S3	Jitter(rap)
S4	Jitter(ppq5)
S5	Jitter(ddp)
S6	Number of pulses
S7	Number of periods
S8	Mean period
S9	Standard dev. of period
S10	Shimmer(local)
S11	Shimmer(local, dB)
S12	Shimmer(apq3)
S13	Shimmer(apq5)
S14	Shimmer(apql1)
S15	Shimmer(dda)
S16	Fraction of locally unvoiced frames
S17	Number of voice breaks
S18	Degree of voice breaks
S19	Median pitch
S20	Mean pitch
S21	Standard deviation
S22	Minimum pitch
S23	Maximum pitch
S24	Autocorrelation
S25	Noise-to-Harmonic
S26	Harmonic-to-Noise

to form a training set, and the remaining samples are used as the test set. Then, the average of the results of all 10 trials is computed. The advantage of this method is that all of the test sets remain independent, ensuring reliable results.

A nested stratified 10-fold CV, which has been widely used in previous research, was used for the purposes of this study [74]. The classification performance evaluation was conducted in the outer loop. Since a 10-fold CV was used in the outer loop, the classifiers were evaluated in one

TABLE 3: Parameter setting of other optimizers involved in training FKNN.

Parameters	GA	PSO	FA	FOA
Population size	8	8	8	8
Max iteration	250	250	250	250
Search space	[2 ⁻⁸ , 2 ⁸]			
Crossover rate	0.8	-	-	-
Mutation rate	0.05	-	-	-
Acceleration constants	-	2	-	-
Inertia weight	-	1	-	-
Differential weight	-	-	-	-
Alpha	-	-	0.5	-
Beta	-	-	0.2	-
Gamma	-	-	1	-
ax	-	-	-	20
bx	-	-	-	10
ay	-	-	-	20
by	-	-	-	10

independent fold of data, and the other nine folds of data were left for training. The parameter optimization process was performed in the inner loop. Since a 5-fold CV was used in the inner loop, the CBFO-FKNN searched for the optimal values of k and m , and the SVM and KELM searched for the optimal values of C and γ in the remaining nine folds of data. The nine folds of data were further split into one fold of data for the performance evaluation, and four folds of data were left for training.

4.5. Evaluation Criteria. ACC, AUC, sensitivity, and specificity were taken to evaluate the performance of different models. These measurements are defined as

$$ACC = \frac{TP + TN}{(TP + FP + FN + TN)} \times 100\% \quad (10)$$

$$Sensitivity = \frac{TP}{(TP + FN)} \times 100\% \quad (11)$$

$$Specificity = \frac{TN}{(FP + TN)} \times 100\% \quad (12)$$

where TP is the number of true positives, FN means the number of false negatives, TN represents the true negatives, and FP is the false positives. AUC [75] is the area under the ROC curve.

5. Experimental Results and Discussion

5.1. Benchmark Function Validation. In order to test the performance of the proposed algorithm CBFO, 23 benchmark functions which include unimodal, multimodal, and fixed-dimension multimodal were used to do experiments. These functions are listed in Tables 4–6 where Dim represents the dimension, Range is the search space, and f_{min} is the best value.

In order to verify the validity of the proposed algorithm, the original BFO, Firefly Algorithm(FA)[76], Flower Pollination Algorithm (FPA)[77], Bat Algorithm (BA)[78], Dragonfly Algorithm (DA)[79], Particle Swarm Optimization (PSO)[80], and the improved BFO called PSOBFO were compared on these issues. The parameters of the above algorithm are set according to their original papers, and the specific parameter values are set as shown in Table 7. In order to ensure that the results obtained are not biased, 30 independent experiments are performed. In all experiments, the number of population size is set to 50 and the maximum number of iterations is set to 500.

Tables 8–10 show average results (Avg), standard deviation (Stdv), and overall ranks for different algorithms dealing with F1–23 issues. It should be noted that the ranking is based on the average result (Avg) of 30 independent experiments for each problem. In order to visually compare the convergence performance of our proposed algorithm and other algorithms, Figures 2–4 use the logarithmic scale diagram to reflect the convergence behaviors. In Figures 2–4, we only select typical function convergence curves from unimodal functions, multimodal functions, and fixed-dimension multimodal functions, respectively. The results of the unimodal F1–F7 are shown in Table 8. As shown, the optimization effect of CBFO in F1, F2, F3, and F4 is the same as the improved PSOBFO, but the performance is improved compared with the original BFO. Moreover, From the ranking results, it can be concluded that, compared with other algorithms, CBFO is the best solution to solve the problems of F1–F7.

With respect to the convergence trends described in Figure 2, it can be observed that the proposed CBFO is capable of testifying a very fast convergence and it can be superior to all other methods in dealing with F1, F2, F3, F4, F5, and F7. For F1, F2, F3, and F4, the CBFO has converged so fast during few searching steps compared to other algorithms. In particular, when dealing with cases F1, F2, F3, and F4, the trend converges rapidly after 250 iterations.

The calculated results for multimodal F8–F13 are tabulated in Table 9. It is observed that CBFO has attained the exact optimal solutions for 30-dimension problems F8 and F12 in all 30 runs. From the results for F9, F10, F11, and F13 problems, it can be agreed that the CBFO yields very competitive solutions compared to the PSOBFO. However, based on rankings, the CBFO is the best overall technique and the overall ranks show that the BFO, FA, BA, PSO, FPA, and DA algorithms are in the next places, respectively.

According to the corresponding convergence trend recorded in Figure 3, the relative superiority of the proposed CBFO in settling F8, F11, and F12 test problems can be recognized. In tackling F11, the CBFO can dominate all its competitors in tackling F11 only during few iterations. On the other hand, methods such as FPA, BA, DA, and PSO still cannot improve the quality of solutions in solving F11 throughout more steps.

The results for F14 to F23 are tabulated in Table 10. The results in Table 10 reveal that the CBFO is the best algorithm and can outperform all other methods in dealing with F15 problems. In F16, F17, and F19, it can be seen that

TABLE 4: Unimodal benchmark functions.

Function	Dim	Range	f_{\min}
$f_1(x) = \sum_{i=1}^n x_i^2$	30	[-100, 100]	0
$f_2(x) = \sum_{i=1}^n x_i + \prod_{i=1}^n x_i $	30	[-10, 10]	0
$f_3(x) = \sum_{i=1}^n \left(\sum_{j=1}^i x_j \right)^2$	30	[-100, 100]	0
$f_4(x) = \max_i \{ x_i , 1 \leq i \leq n\}$	30	[-100, 100]	0
$f_5(x) = \sum_{i=1}^{n-1} [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$	30	[-30, 30]	0
$f_6(x) = \sum_{i=1}^n ([x_i + 0.5])^2$	30	[-100, 100]	0
$f_7(x) = \sum_{i=1}^n i x_i^4 + \text{random}[0, 1)$	30	[-1.28, 1.28]	0

TABLE 5: Multimodal benchmark functions.

Function	Dim	Range	f_{\min}
$f_8(x) = \sum_{i=1}^n -x_i \sin(\sqrt{ x_i })$	30	[-500, 500]	-418.9829*5
$f_9(x) = \sum_{i=1}^n [x_i^2 - 10 \cos(2\pi x_i) + 10]$	30	[-5.12, 5.12]	0
$f_{10}(x) = -20 \exp\left(-0.2 \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}\right) - \exp\left(\frac{1}{n} \sum_{i=1}^n \cos(2\pi x_i)\right) + 20 + e$	30	[-32, 32]	0
$f_{11}(x) = \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	30	[-600, 600]	0
$f_{12}(x) = \frac{\pi}{n} \left\{ 10 \sin(\pi y_1) + \sum_{i=1}^{n-1} (y_i - 1)^2 [1 + 10 \sin^2(\pi y_{i+1})] + (y_n - 1)^2 \right\} + \sum_{i=1}^n u(x_i, 10, 100, 4)$	30	[-50, 50]	0
$y_i = 1 + \frac{x_i + 1}{4} u(x_i, a, k, m) \begin{cases} k(x_i - a) & x_i > a \\ 0 & -a < x_i < a \\ k(x_i - a) & x_i < -a \end{cases}$			
$f_{13}(x) = 0.1 \left\{ \sin^2(3\pi x_1) + \sum_{i=1}^n (x_i - 1)^2 [1 + \sin^2(3\pi x_i + 1)] + (x_n - 1)^2 [1 + \sin^2(2\pi x_n)] \right\} + \sum_{i=1}^n u(x_i, 5, 100, 4)$	30	[-50, 50]	0

the optimization effect of all the algorithms is not much different. In dealing with F20 case, the CBFO's performance is improved compared to original BFO and the improved PSOBFO. Especially in solving F18, the proposed algorithm is much better than the improved PSOBFO. From Figure 4, we can see that the convergence speed of the CBFO is better than other algorithms in dealing with F15, F18, F19, and F20. For F15, it surpasses all methods.

In order to investigate significant differences of obtained results for the CBFO over other competitors, the Wilcoxon rank-sum test [81] at 5% significance level was also employed in this paper. The p values of comparisons are reported in Tables 11–13. In each table, each p value which is not lower

than 0.05 is shown in bold face. It shows that the differences are not significant.

The p values are also provided in Table 11 for F1–F7. Referring to the p values of the Wilcoxon test in Table 11, it is verified that the proposed algorithm is statistically meaningful. The reason is that all p values are less than 0.05 except PSOBFO in F1, F2, F3, and F4. According to the p values in Table 12, all values are less than 0.05 except PSOBFO in F11 problem. Hence, it can be approved that the results of the CBFO are statistically improved compared to the other methods. As can be seen from the p value in Table 13, the CBFO algorithm is significantly better than the PSOBFO, FPA, BA, and PSO for F14–F23.

TABLE 6: Fixed-dimension multimodal benchmark functions.

Function	Dim	Range	f_{\min}
$f_{14}(x) = \left(\frac{1}{500} + \sum_{j=1}^{25} \frac{1}{j + \sum_{i=1}^2 (x_i - a_{ij})^6} \right)^{-1}$	2	[-65,65]	1
$f_{15}(x) = \sum_{i=1}^{11} \left[a_i - \frac{x_1(b_i^2 + b_i x_2)}{b_i^2 + b_i x_3 + x_4} \right]^2$	4	[-5, 5]	0.00030
$f_{16}(x) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1 x_2 - 4x_2^2 + 4x_2^4$	2	[-5,5]	-1.0316
$f_{17}(x) = \left(x_2 - \frac{5.1}{4\pi^2} x_1^2 + \frac{5}{\pi} x_1 - 6 \right)^2 + 10 \left(1 - \frac{1}{8\pi} \right) \cos x_1 + 10$	2	[-5,5]	0.398
$f_{18}(x) = [1 + (x_1 + x_2 + 1)^2 (19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1 x_2 + 3x_2^2)] \times [30 + (2x_1 - 3x_2)^2 \text{ times } (18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1 x_2 + 27x_2^2)]$	2	[-2,2]	3
$f_{19}(x) = -\sum_{i=1}^4 c_i \exp \left(-\sum_{j=1}^3 a_{ij} (x_j - p_{ij})^2 \right)$	3	[1,3]	-3.86
$f_{20}(x) = -\sum_{i=1}^4 c_i \exp \left(-\sum_{j=1}^6 a_{ij} (x_j - p_{ij})^2 \right)$	6	[0,1]	-3.32
$f_{21}(x) = -\sum_{i=1}^5 \left[(X - a_i) (X - a_i)^T + c_i \right]^{-1}$	4	[0,10]	-10.1532
$f_{22}(x) = -\sum_{i=1}^7 \left[(X - a_i) (X - a_i)^T + c_i \right]^{-1}$	4	[0,10]	-10.4028
$f_{23}(x) = -\sum_{i=1}^{10} \left[(X - a_i) (X - a_i)^T + c_i \right]^{-1}$	4	[0,10]	-10.5363

TABLE 7: Parameters setting for the involved algorithms.

Method	Population size	Maximum generation	Other parameters
BFO	50	500	$\Delta \in [-1, 1]$
BA	50	500	Q Frequency $\in [0, 2]$; A Loudness: 0.5; r Pulse rate: 0.5
DA	50	500	$w \in [0.9, 0.2]$; $s = 0.1$; $a = 0.1$; $c = 0.7$; $f = 1$; $e = 1$
FA	50	500	$\beta_0 = 1$; $\alpha \in [0, 1]$; $\gamma = 1$
FPA	50	500	switch probability $p = 0.8$; $\lambda = 1.5$
PSO	50	500	inertial weight = 1; $c_1 = 2$; $c_2 = 2$
PSOBFO	50	500	inertial weight = 1; $c_1 = 1.2$; $c_2 = 0.5$; $\Delta \in [-1, 1]$

The results demonstrate that the utilized chaotic mapping strategy and Gaussian mutation in the CBFO technique have improved the efficacy of the classical BFO, in a significant manner. On the one hand, applying the chaotic mapping strategy to the bacterial population initialization process can speed up the initial exploration of the algorithm. On the other hand, adding Gaussian mutation to the current best bacterial individual in the iterative process helps to jump out of the local optimum. In conclusion, the proposed CBFO can make a better balance between explorative and exploitative trends using the embedded strategies.

5.2. Results on the Parkinson's Disease. Many studies have demonstrated that the performance of BFO can be affected heavily by the chemotaxis step size $C(i)$. Therefore, we have also investigated the effects of $C(i)$ on the performance of the CBFO-FKNN. Table 14 displays the detailed results of CBFO-FKNN model with different values of $C(i)$ on the two datasets. In the table, the mean results and their standard deviations (in parentheses) are listed. As shown, the CBFO-FKNN

model performed best with an average accuracy of 96.97%, an AUC of 0.9781, a sensitivity of 96.87%, and a specificity of 98.75% when $C(i) = 0.1$ on the Oxford dataset and an average accuracy of 83.68%, an AUC of 0.6513, a sensitivity of 96.92%, and a specificity of 33.33% when $C(i) = 0.2$ on the Istanbul dataset. Furthermore, the CBFO-FKNN approach also yielded the most reliable results with the minimum standard deviation when $C(i) = 0.1$ and $C(i) = 0.2$ on the Oxford dataset and Istanbul dataset, respectively. Therefore, values of 0.1 and 0.2 were selected as the parameter value of $C(i)$ for CBFO-FKNN on the two datasets, respectively, in the subsequent experimental analysis.

The ACC, AUC, sensitivity, specificity, and optimal (k, m) pair values of each fold obtained via the CBFO-FKNN model with $C(i) = 0.1$ and $C(i) = 0.2$ on the Oxford dataset and Istanbul dataset are shown in Tables 15 and 16, respectively. As shown, each fold possessed a different parameter pair (k, m) since the parameters for each set of fold data were automatically determined via the CBFO method. With the optimal parameter pair, the FKNN yielded different optimal

TABLE 8: Results of unimodal benchmark functions (F1-F7).

F	CBFO	PSOBFO	BFO	FA	FPA	BA	DA	PSO
F1	Avg	0	0	8.73E-03	9.84E-03	1.45E+03	1.70E+01	2.15E+03
	Stdv	0	0	3.85E-03	3.20E-03	4.07E+02	2.09E+00	1.13E+03
	Rank	1	1	3	4	7	5	8
F2	Avg	0	0	3.55E-01	3.88E-01	4.59E+01	3.32E+01	1.53E+01
	Stdv	0	0	7.44E-02	8.27E-02	1.49E+01	3.35E+01	6.54E+00
	Rank	1	1	3	4	7	6	5
F3	Avg	0	0	4.96E-12	2.59E+03	1.99E+03	1.15E+02	1.46E+04
	Stdv	0	0	8.97E-12	8.38E+02	4.84E+02	3.68E+01	8.91E+03
	Rank	1	1	3	7	6	4	8
F4	Avg	0	0	3.24E-02	8.43E-02	2.58E+01	3.78E+00	2.95E+01
	Stdv	0	0	5.99E-03	1.60E-02	3.96E+00	3.02E+00	8.22E+00
	Rank	1	1	3	4	7	5	8
F5	Avg	2.90E+01	0	6.55E+04	2.33E+02	2.57E+05	4.48E+03	4.96E+05
	Stdv	2.62E-02	0	NA	4.30E+02	1.88E+05	1.24E+03	6.46E+05
	Rank	2	1	5	3	7	4	8
F6	Avg	1.34E-01	3.71E-01	2.11E+03	1.14E-02	1.53E+03	1.70E+01	2.06E+03
	Stdv	1.76E-02	5.99E-02	1.15E+04	4.71E-03	4.23E+02	2.51E+00	1.52E+03
	Rank	2	3	8	1	6	4	7
F7	Avg	3.62E-04	4.88E-03	3.77E-03	1.08E-02	4.60E-01	1.89E+01	6.92E-01
	Stdv	3.21E-04	3.44E-03	3.33E-03	2.79E-03	1.42E-01	2.00E+01	3.79E-01
	Rank	1	3	2	4	5	7	6
Sum of ranks		9	11	27	27	45	35	50
Average rank		1.2857	1.5714	3.8571	3.8571	6.4286	5	7.1429
Overall rank		1	2	3	3	7	5	8

TABLE 9: Results of multimodal benchmark functions (F8-F13).

F	CBFO	PSOBFO	BFO	FA	FPA	BA	DA	PSO
F8	Avg	-3.47E+04	-2.55E+03	-2.47E+03	-6.55E+03	-7.58E+03	-7.45E+03	-5.44E+03
	Stdv	1.79E+04	5.80E+02	5.25E+02	6.70E+02	2.12E+02	6.56E+02	5.55E+02
	Rank	1	7	8	5	2	3	6
F9	Avg	-2.89E+02	-2.90E+02	-2.88E+02	3.37E+01	1.44E+02	2.73E+02	1.71E+02
	Stdv	2.98E-01	0	8.61E-01	1.13E+01	1.68E+01	3.08E+01	4.15E+01
	Rank	2	1	3	4	5	7	6
F10	Avg	-9.66E+12	-1.07E+13	-9.08E+12	5.47E-02	1.31E+01	5.56E+00	1.02E+01
	Stdv	3.21E+11	3.97E-03	7.34E+11	1.31E-02	1.59E+00	3.77E+00	2.15E+00
	Rank	2	1	3	4	8	5	7
F11	Avg	0	0	4.99E-03	6.53E-03	1.49E+01	6.35E-01	1.65E+01
	Stdv	0	0	3.18E-03	2.63E-03	3.38E+00	6.31E-02	8.41E+00
	Rank	1	1	3	4	7	5	8
F12	Avg	1.34E-11	1.27E-08	3.04E-10	2.49E-04	1.16E+02	1.33E+01	7.90E+04
	Stdv	3.46E-11	2.02E-08	5.97E-10	1.06E-04	4.75E+02	4.93E+00	4.26E+05
	Rank	1	3	2	4	7	6	8
F13	Avg	4.20E-02	9.92E-02	9.92E-02	3.18E-03	6.18E+04	2.77E+00	4.46E+05
	Stdv	4.64E-02	2.52E-08	4.17E-10	2.53E-03	9.34E+04	4.37E-01	7.19E+05
	Rank	2	3	3	1	7	5	8
Sum of ranks		9	16	22	22	36	31	43
Average rank		1.5000	2.6667	3.6667	3.6667	6.0000	5.1667	7.1667
Overall rank		1	2	3	3	7	5	8

TABLE 10: Results of fixed-dimension multimodal benchmark functions (F14-F23).

F		CBFO	PSOBFO	BFO	FA	FPA	BA	DA	PSO
F14	Avg	9.83E+00	3.11E+00	2.96E+00	1.82E+00	1.04E+00	4.53E+00	1.30E+00	4.41E+00
	Stdv	4.51E+00	1.71E+00	2.22E+00	8.42E-01	1.56E-01	3.91E+00	6.96E-01	3.20E+00
	Rank	8	5	4	3	1	7	2	6
F15	Avg	4.33E-04	9.49E-04	6.24E-04	2.85E-03	7.44E-04	8.29E-03	3.73E-03	1.41E-03
	Stdv	1.65E-04	3.00E-04	2.25E-04	4.71E-03	1.41E-04	1.35E-02	5.95E-03	4.04E-04
	Rank	1	4	2	6	3	8	7	5
F16	Avg	-1.03E+00							
	Stdv	5.23E-06	1.60E-04	7.96E-06	3.36E-09	2.55E-08	8.94E-04	3.47E-06	2.49E-03
	Rank	1	1	1	1	1	1	1	1
F17	Avg	3.98E-01	3.99E-01						
	Stdv	2.24E-06	4.80E-05	2.02E-06	1.76E-09	6.28E-09	5.45E-04	1.84E-07	1.65E-03
	Rank	1	1	1	1	1	1	1	1
F18	Avg	3.00E+00	3.01E+00	3.00E+00	3.00E+00	3.00E+00	3.10E+00	3.00E+00	3.24E+00
	Stdv	2.00E-04	6.53E-03	3.13E-04	2.59E-08	1.60E-06	8.65E-02	6.09E-07	3.61E-01
	Rank	1	6	1	1	1	6	1	8
F19	Avg	-3.86E+00	-3.86E+00	-3.86E+00	-3.86E+00	-3.86E+00	-3.83E+00	-3.86E+00	-3.84E+00
	Stdv	4.86E-04	4.56E-03	5.87E-04	1.03E-09	2.38E-06	2.49E-02	1.16E-03	2.10E-02
	Rank	1	1	1	1	1	8	1	7
F20	Avg	-3.29E+00	-3.24E+00	-3.27E+00	-3.28E+00	-3.31E+00	-2.89E+00	-3.25E+00	-2.71E+00
	Stdv	2.41E-02	2.38E-02	2.48E-02	6.10E-02	6.06E-03	1.31E-01	1.01E-01	3.54E-01
	Rank	2	6	4	3	1	7	5	8
F21	Avg	-6.03E+00	-1.01E+01	-9.80E+00	-7.92E+00	-1.01E+01	-4.64E+00	-6.61E+00	-3.67E+00
	Stdv	9.74E-01	4.27E-02	1.28E+00	3.47E+00	1.30E-01	2.43E+00	2.62E+00	1.31E+00
	Rank	6	1	3	4	1	7	5	8
F22	Avg	-6.45E+00	-1.01E+01	-1.02E+01	-9.89E+00	-1.02E+01	-5.03E+00	-7.35E+00	-4.33E+00
	Stdv	1.22E+00	9.60E-01	9.61E-01	1.94E+00	4.87E-01	2.93E+00	2.98E+00	1.67E+00
	Rank	6	3	1	4	1	7	5	8
F23	Avg	-6.91E+00	-9.73E+00	-9.98E+00	-1.05E+01	-1.02E+01	-5.36E+00	-6.35E+00	-4.42E+00
	Stdv	1.30E+00	1.82E+00	1.63E+00	1.07E-06	4.94E-01	2.90E+00	3.36E+00	1.33E+00
	Rank	5	4	3	1	2	7	6	8
Sum of ranks		32	32	21	25	13	59	34	60
Average rank		3.2	3.2	2.1	2.5	1.3	5.9	3.4	6
Overall rank		4	4	2	3	1	7	6	8

TABLE 11: The calculated p-values from the functions (F1-F7) for the CBFO versus other optimizers.

Problem	PSOBFO	BFO	FA	FPA	BA	DA	PSO
F1	1	1.73E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06
F2	1	1.73E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06
F3	1	1.73E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06
F4	1	1.73E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06
F5	1.73E-06	1.73E-06	6.04E-03	1.73E-06	1.73E-06	1.73E-06	1.73E-06
F6	1.73E-06						
F7	1.92E-06	3.52E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06

classification performance values in each fold. This was attributed to the adaptive tuning of the two parameters by the CBFO based on the specific distribution of each data set.

In order to investigate the convergence behavior of the proposed CBFO-FKNN method, the classification error rate versus the number of iterations was recorded. For simplicity,

herein we take the Oxford dataset for example. Figures 5(a)–5(d) display the learning curves of the CBFO-FKNN for folds 1, 3, 5, and 7 in the 10-fold CV, respectively. As shown, all four fitness curves of CBFO converged into a global optimum in fewer than 20 iterations. The fitness curves gradually improved from iterations 1 through 20 but

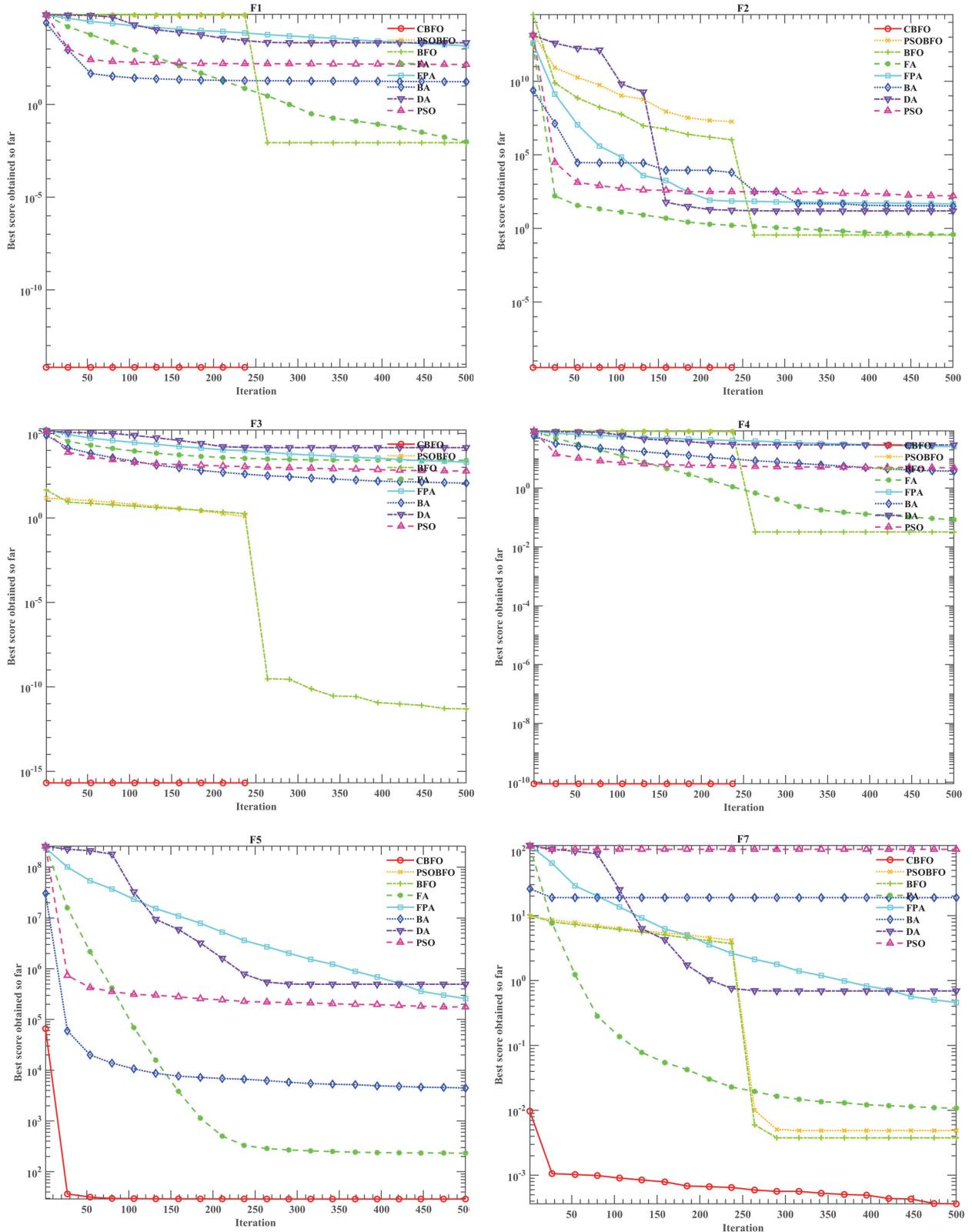


FIGURE 2: Convergence curves of unimodal functions.

TABLE 12: The calculated p-values from the functions (F8-F13) for the CBFO versus other optimizers.

Problem	PSOBFO	BFO	FA	FPA	BA	DA	PSO
F8	1.73E-06	1.73E-06	1.73E-06	1.92E-06	1.73E-06	1.73E-06	1.92E-06
F9	1.73E-06	6.89E-05	1.73E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06
F10	1.73E-06	4.90E-04	1.73E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06
F11	1	1.73E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06
F12	3.52E-06	5.79E-05	1.73E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06
F13	1.73E-06	1.92E-06	3.61E-03	1.73E-06	1.73E-06	1.73E-06	1.73E-06

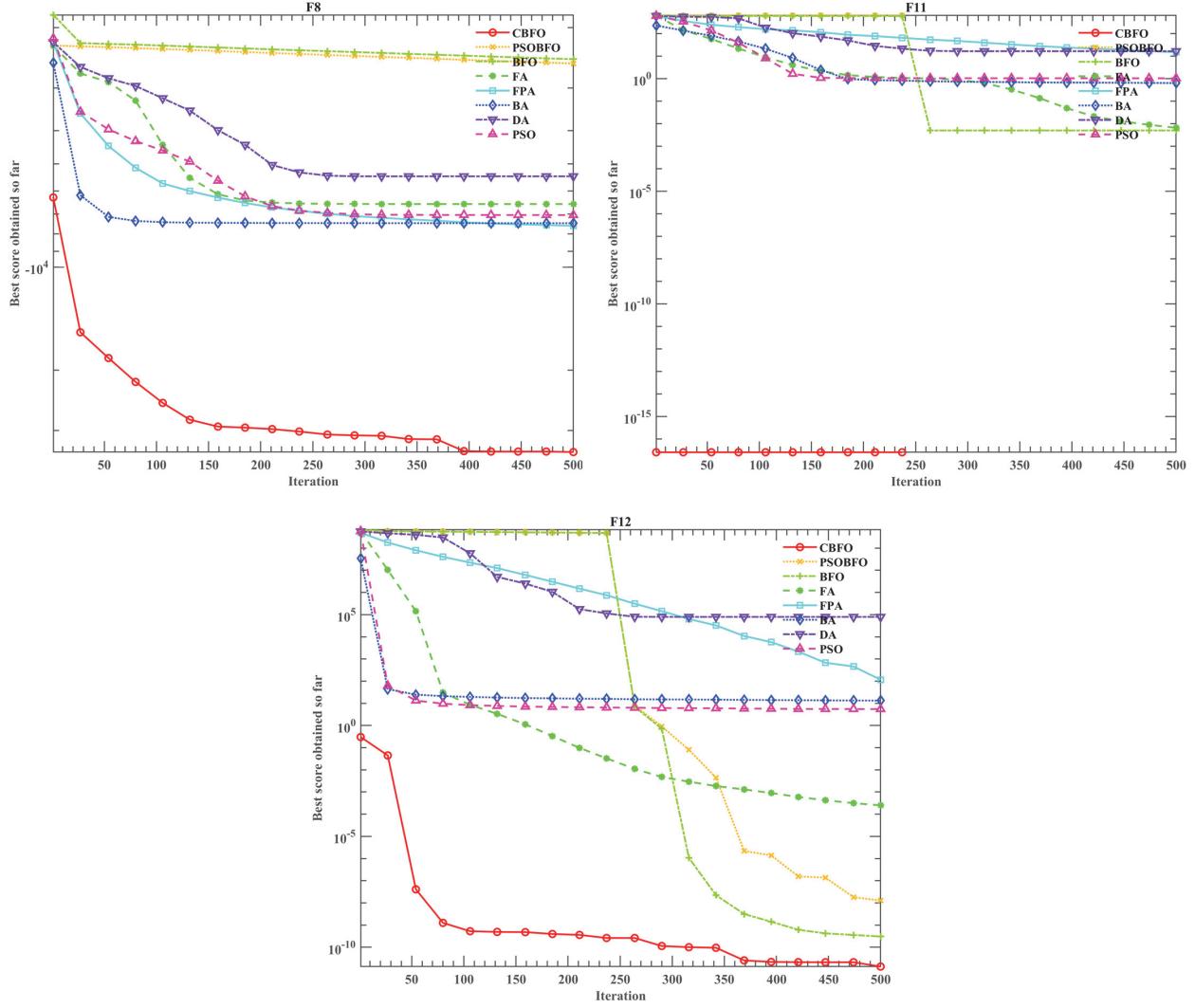


FIGURE 3: Convergence curves of multimodal functions.

exhibited no significant improvements after iteration 20. The fitness curves ceased after 50 iterations (the maximum number of iterations). The error rates of the fitness curves decreased rapidly at the beginning of the evolutionary process and continued to decrease slowly after a certain number of iterations. During the latter part of the evolutionary process, the fitness curves remained stable until the stopping criteria,

the maximum number of iterations, were satisfied. Thus, the proposed CBFO-FKNN model efficiently converged toward the global optima.

To validate the effectiveness of the proposed method, the CBFO-FKNN model was compared to five other meta-heuristic algorithms-based FKNN models as well as three other advanced machine learning approaches including

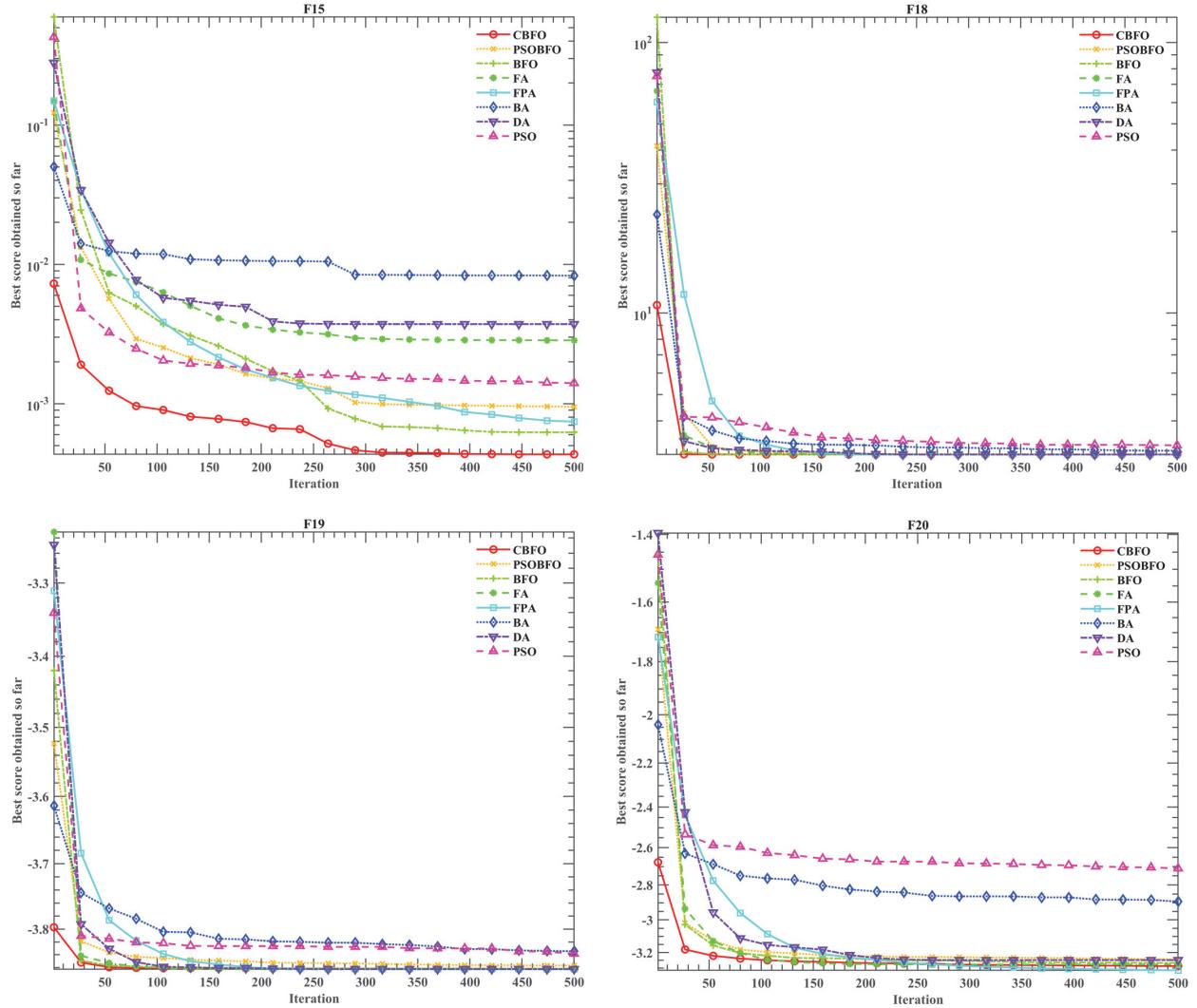


FIGURE 4: Convergence curves based on fixed-dimension multimodal functions.

TABLE 13: The calculated p-values from the functions (F14-F23) for the CBFO versus other optimizers.

Problem	PSOBFO	BFO	FA	FPA	BA	DA	PSO
F14	6.34E-06	6.98E-06	5.22E-06	3.18E-06	2.22E-04	1.73E-06	1.06E-04
F15	3.88E-06	8.31E-04	1.73E-06	1.24E-05	1.92E-06	4.29E-06	1.92E-06
F16	2.35E-06	7.50E-01	1.73E-06	1.73E-06	1.73E-06	1.97E-05	1.73E-06
F17	1.92E-06	3.60E-01	1.73E-06	1.73E-06	1.73E-06	2.60E-06	1.73E-06
F18	2.35E-06	8.45E-01	1.73E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06
F19	1.92E-06	8.22E-02	1.73E-06	1.73E-06	1.73E-06	8.19E-05	1.73E-06
F20	3.41E-05	8.94E-04	6.44E-01	2.60E-06	1.73E-06	3.82E-01	1.73E-06
F21	1.73E-06	2.13E-06	4.99E-03	1.73E-06	8.22E-03	7.04E-01	6.34E-06
F22	3.88E-06	2.60E-06	1.64E-05	1.73E-06	1.85E-02	1.85E-01	8.92E-05
F23	1.24E-05	2.60E-05	1.73E-06	1.73E-06	1.96E-02	4.05E-01	1.02E-05

TABLE 14: Detailed results of CBFO-FKNN with different values of $C(i)$ on the two datasets.

$C(i)$	Oxford dataset				Istanbul dataset			
	ACC	AUC	Sen	Spec	ACC	AUC	Sen	Spec
0.05	0.9542 (0.0370)	0.9417 (0.0774)	0.9666 (0.0356)	0.9167 (0.1620)	0.8230 (0.0636)	0.6180 (0.1150)	0.9694 (0.0413)	0.2667 (0.2108)
	0.9697 (0.0351)	0.9781 (0.0253)	0.9687 (0.0432)	0.9875 (0.0395)	0.8054 (0.0414)	0.5946 (0.0746)	0.9559 (0.0297)	0.2333 (0.1405)
0.1	0.9489 (0.0629)	0.9479 (0.0609)	0.9358 (0.1158)	0.9600 (0.0843)	0.8155 (0.0669)	0.6074 (0.1204)	0.9648 (0.0450)	0.2500 (0.2257)
	0.9589 (0.0469)	0.9466 (0.0860)	0.9600 (0.0555)	0.9333 (0.1610)	0.8368 (0.0283)	0.6512 (0.0698)	0.9691 (0.0360)	0.3333 (0.1571)
0.25	0.9587 (0.0536)	0.9459 (0.0901)	0.9669 (0.0459)	0.9250 (0.1687)	0.8257 (0.0770)	0.6385 (0.1560)	0.9603 (0.0328)	0.3167 (0.2987)
	0.9639 (0.0352)	0.9689 (0.0308)	0.9670 (0.0454)	0.9708 (0.0623)	0.8090 (0.0439)	0.6165 (0.1112)	0.9478 (0.0534)	0.2833 (0.2491)

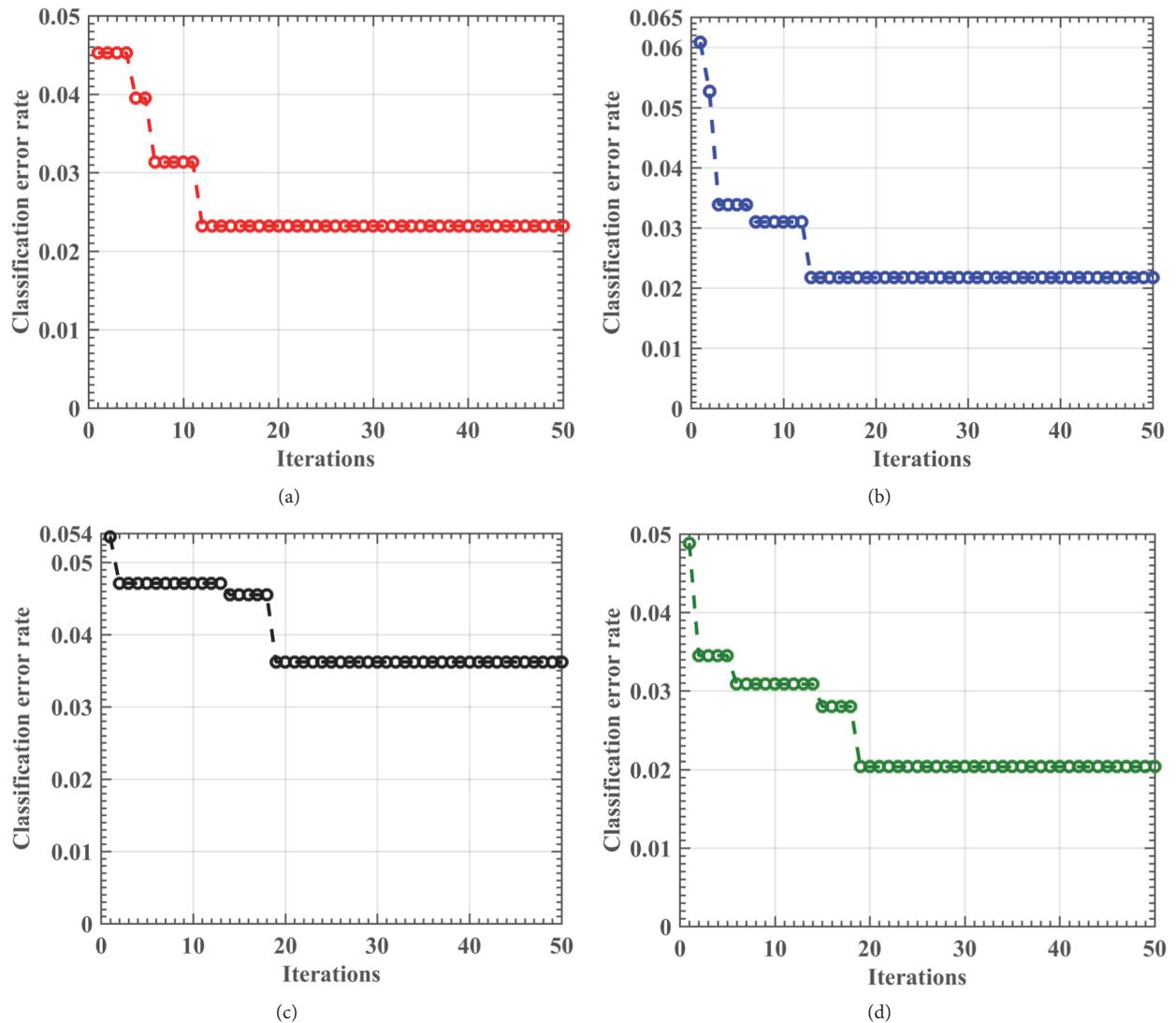


FIGURE 5: Learning curves of CBFO for fold 2 (a), fold 4 (b), fold 6 (c), and fold 8 (d) during the training stage.

TABLE 15: Detailed classification results of CBFO-FKNN on the Oxford dataset.

Fold	CBFO-FKNN						
	No.	ACC	AUC	Sen	Spec	k	m
1		0.9474	0.9667	0.9333	1.0000	1	1.77
2		1.0000	1.0000	1.0000	1.0000	1	2.94
3		0.9500	0.9688	0.9375	1.0000	1	3.92
4		0.9500	0.9375	1.0000	0.8750	1	6.89
5		0.9500	0.9667	0.9333	1.0000	1	9.33
6		0.9000	0.9412	0.8824	1.0000	1	7.26
7		1.0000	1.0000	1.0000	1.0000	1	9.21
8		1.0000	1.0000	1.0000	1.0000	1	7.61
9		1.0000	1.0000	1.0000	1.0000	1	8.95
10		1.0000	1.0000	1.0000	1.0000	1	7.25
Mean		0.9697	0.9781	0.9687	0.9875	1	6.51

TABLE 16: Detailed classification results of CBFO-FKNN on the Istanbul dataset.

Fold	CBFO-FKNN						
	No.	ACC	AUC	Sen	Spec	k	m
1		0.8571	0.7273	0.9545	0.5000	3	4.80
2		0.8276	0.5833	1.0000	0.1667	3	3.70
3		0.8276	0.7065	0.9130	0.5000	3	7.30
4		0.8276	0.5833	1.0000	0.1667	3	4.16
5		0.8966	0.7500	1.0000	0.5000	3	9.40
6		0.7931	0.6232	0.9130	0.3333	3	2.50
7		0.8621	0.7283	0.9565	0.5000	3	9.70
8		0.8276	0.5833	1.0000	0.1667	3	4.30
9		0.8276	0.5833	1.0000	0.1667	3	8.20
10		0.8214	0.6439	0.9545	0.3333	3	7.04
Mean		0.8368	0.6513	0.9692	0.3333	3	6.11

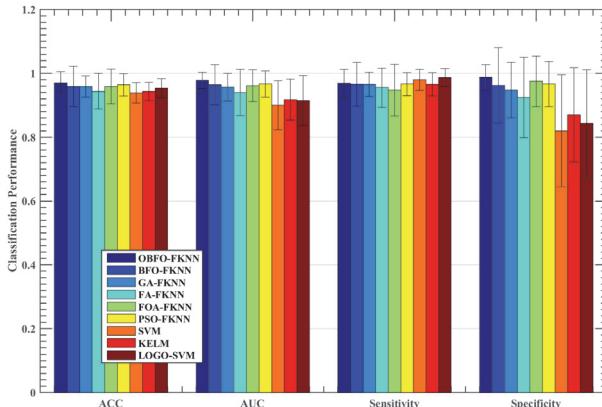


FIGURE 6: Comparison results obtained on the Oxford dataset by the nine methods.

SVM, KELM, and SVM with local learning-based feature selection (LOGO-SVM). As shown in Figure 6, the CBFO-FKNN method performed better than other competitors in terms of ACC, AUC, and sensitivity on the Oxford

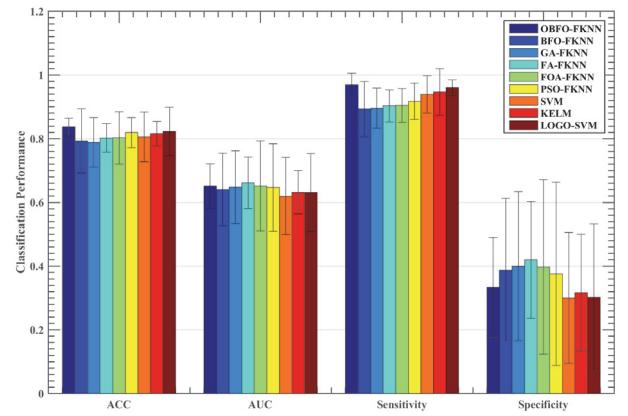


FIGURE 7: Comparison results obtained on the Istanbul dataset by the nine methods.

dataset. We can see that the CBFO-FKNN method yields the highest average ACC value of 96.97%, followed by PSO-FKNN, LOGO-SVM, KELM, SVM, FOA-FKNN, FA-FKNN, and BFO-FKNN. GA-FKNN has got the worst result among the all methods. On the AUC metric, OBFO-FKNN obtained similar results with FA-FKNN, followed by FOA-FKNN, GA-FKNN, PSO-FKNN, BFO-FKNN, KELM, and LOGO-SVM, and SVM has got the worst result. On the sensitivity metric, CBFO-FKNN has achieved obvious advantages, LOGO-FKNN ranked second, followed by KELM, SVM, PSO-FKNN, FOA-FKNN, FA-FKNN, and GA-FKNN. BFO-FKNN has got the worst performance. On the specificity metric, FA-FKNN achieved the maximum results, GA-FKNN and FOA-FKNN have achieved similar results, which ranked second, followed by BFO-FKNN, PSO-FKNN, CBFO-FKNN, and SVM. KELM and LOGO-SVM have obtained similar results, both of which got the worst performance. Regarding the Istanbul dataset, CBFO-FKNN produced the highest result with the ACC of 83.68%, while the LOGO-SVM and PSO-FKNN method yields the second best average ACC value as shown in Figure 7, followed by

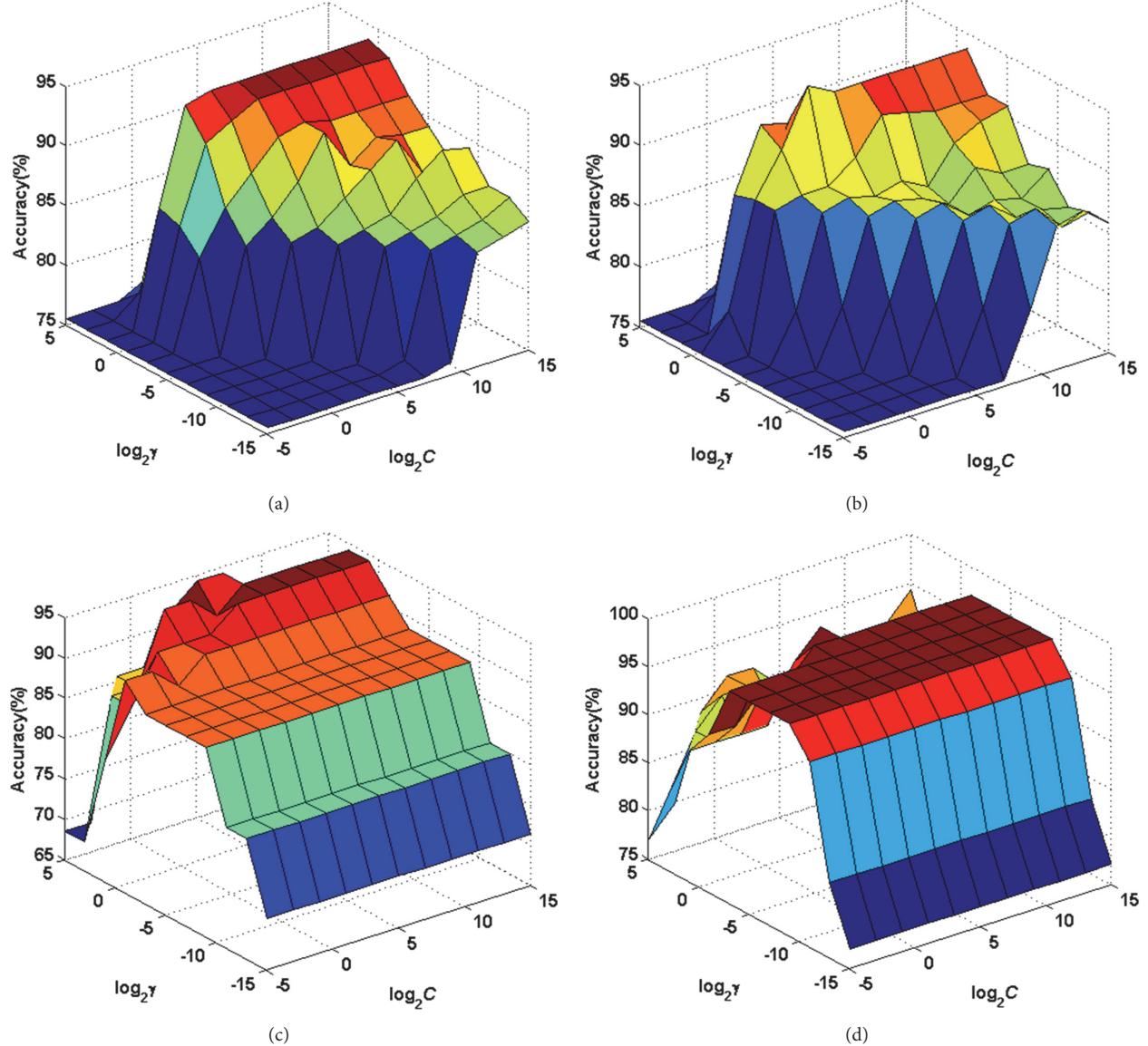


FIGURE 8: Training accuracy surfaces of SVM and KELM via the grid search method on the Oxford dataset. (a) Fold 2 for SVM. (b) Fold 4 for SVM on the data. (c) Fold 6 for KELM on the data. (d) Fold 8 for SVM on the data.

KELM, SVM, FOA-FKNN, FA-FKNN, BFO-FKNN, and GA-FKNN. From Figures 6 and 7, we can also find that the CBFO-FKNN can yield a smaller or comparative standard deviation than the other counterparts in terms of the four performance metrics on the both datasets. Additionally, we can find that the SVM with local learning-based feature selection can improve the performance of the two datasets. It indicates that there are some irrelevant features or redundant features in these two datasets. It should be noted that the LOGO method was used for feature selection, all the features were ranked by the LOGO, then all the feature subsets were evaluated incrementally, and finally the feature subset achieved the best accuracy was chosen as the one in the experiment.

According to the results, the superior performance of the proposed CBFO-FKNN indicates that the proposed method was the most robust tool for detection of PD among the nine methods. The main reason may lie in that the OBL mechanism greatly improves the diversity of the population and increases the probability of BFO escaping from the local optimum. Thus, it gets more chances to find the optimal neighborhood size and fuzzy strength values by the CBFO, which aided the FKNN classifier in more efficiently achieving the maximum classification performance. Figure 8 displays the surface of training classification accuracies achieved by the SVM and KELM methods for several folds of the training data via the grid search strategy on the Oxford dataset.

TABLE 17: The confusion matrix obtained by CBFO-FKNN via 10-fold CV for each group.

Male	Predicted PD	Predicted health
Actual PD	97	3
Actual health	2	16
Female	Predicted PD	Predicted health
Actual PD	44	3
Actual health	2	28
Old	Predicted PD	Predicted health
Actual PD	87	4
Actual health	0	18
Young	Predicted PD	Predicted health
Actual PD	56	0
Actual health	0	30

Through the experimental process, we can find the original BFO is more prone to overfitting; this paper introduces chaotic initialization, enriches the diversity of the initial population, and improves the convergence speed of the population as well; in addition, this paper also introduced Gaussian mutation strategy for enhancing the ability of the algorithm to jump out of local optimum, so as to alleviate the overfitting problem of FKNN in the process of classification.

We have also investigated whether the diagnosis was affected by age and gender. Herein, we have taken the Oxford dataset for example. The dataset was divided by the age (old or young) and gender (male or female), respectively. Regarding the age, we have chosen the mean age of 65.8 years as the dividing point. The samples in the old group are more than 65.8, and the samples in the young group are less than 65.8. Therefore, we can obtain four groups of data including male group, female group, old group, and young group. The classification results of the four groups in terms of confusion matrix are displayed in Table 17. As shown, we can find that either in the male group or in the female group 3 PD samples were wrongly classified as healthy ones, and 2 healthy samples were misjudged as PD ones. It indicates that the gender has little impact on the diagnostic results. In the old group, we can find that 4 PD samples were wrongly identified as healthy ones. However, none of the samples were misjudged in the young group. It suggests that the speech samples in the old group are much easier to be wrongly predicted than those in the young group.

To further investigate the impact of gender and age on the diagnosis results. We have further divided the samples into male group and female group on the premise of young and old age and old group and young group on the premise of male and female, respectively. So we can obtain 8 groups as shown in Table 18, and the detailed classification results are displayed in terms of confusion matrix. As shown, we can find that the probability of the sample being misclassified is closer in the old group and young group on the premise of male and female. It can be also observed that there was no sample being wrongly predicted in male and female groups on the premise

of young persons, while there was one sample being wrongly predicted in male and female groups on the premise of old persons, respectively. We can arrive at the conclusion that the presbyphonic may play a confounding role in the female and male dysphonic set, and the results of diagnosis were less affected by gender.

The classification accuracies of other methods applied to the diagnosis of PD are presented for comparison in Table 19. As shown, the proposed CBFO-FKNN method achieved relatively high classification accuracy and, therefore, it could be used as an effective diagnostic tool.

6. Conclusions and Future Work

In this study, we have proposed a novel evolutionary instance-based approach based on a chaotic BFO and applied it to differentiating the PD from the healthy people. In the proposed methodology, the chaos theory enhanced BFO strategy was used to automatically determine the two key parameters, thereby utilizing the FKNN to its fullest potential. The results suggested that the proposed CBFO-FKNN approach outperformed five other FKNN models based on nature-inspired methods and three commonly used advanced machine learning methods including SVM, LOGO-SVM, and KELM, in terms of various performance metrics. In addition, the simulation results indicated that the proposed CBFO-FKNN could be used as an efficient computer-aided diagnostic tool for clinical decision-making. Through the experimental analysis, we can arrive at the conclusion that the presbyphonic may play a confounding role in the female and male dysphonic set, and the results of diagnosis were less affected by gender. Additionally, the speech samples in the old group are much easier to be wrongly predicted than those in the young group.

In future studies, the proposed method will be implemented in a distributed environment in order to further boost its PD diagnostic efficacy. Additionally, implementing the feature selection using CBFO strategy to further boost the performance of the proposed method is another future work. Finally, due to the small vocal datasets of PD, we will generalize the proposed method to much larger datasets in the future.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of article.

Acknowledgments

This research is supported by the National Natural Science Foundation of China (61702376 and 61402337). This research is also funded by Zhejiang Provincial Natural Science Foundation of China (LY17F020012, LY14F020035, LQ13F020011, and LQ13G010007) and Science and Technology Plan Project of Wenzhou of China (ZG2017019, H20110003, Y20160070, and Y20160469).

TABLE 18: The confusion matrix obtained by CBFO-FKNN for each group with precondition.

		Predicted PD	Predicted health
Male			
Actual PD		62	1
Actual health		0	6
Female		Predicted PD	Predicted health
Actual PD		27	1
Actual health		0	12
		Predicted PD	Predicted health
Young			
Actual PD		37	0
Actual health		0	12
Old		Predicted PD	Predicted health
Actual PD		61	2
Actual health		0	6
Male		Predicted PD	Predicted health
Actual PD		35	2
Actual health		0	12
		Predicted PD	Predicted health
Female			
Actual PD		27	1
Actual health		0	12
Young		Predicted PD	Predicted health
Actual PD		19	0
Actual health		0	18

TABLE 19: Comparison of the classification accuracies of various methods.

Study	Method	Accuracy (%)
Little et al. (2009)	Pre-selection filter + Exhaustive search + SVM	91.4(bootstrap with 50 replicates)
Shahbaba et al. (2009)	Dirichlet process mixtures	87.7(5-fold CV)
Das (2010)	ANN	92. (hold-out)
Sakar et al. (2010)	Mutual information based feature selection + SVM	92.75(bootstrap with 50 replicates)
Psorakis et al. (2010)	Improved mRVMs	89.47(10-fold CV)
Guo et al. (2010)	GP-EM	93.1(10-fold CV)
Ozciit et al. (2011)	CFS-RF	87.1(10-fold CV)
Li et al. (2011)	Fuzzy-based non-linear transformation + SVM	93.47(hold-out)
Luukka (2011)	Fuzzy entropy measures + Similarity classifier	85.03(hold-out)
	Particle swarm optimization + OPF	73.53(hold-out)
Spadoto et al. (2011)	Harmony search + OPF	84.01(hold-out)
	Gravitational search algorithm + OPF	84.01(hold-out)
AStrom et al. (2011)	Parallel NN	91.20(hold-out)
Chen et al.(2013)	PCA-FKNN	96.07(10-fold CV)
Babu et al. (2013)	projection based learning for meta-cognitive radial basis function network (PBL-McRBFN)	99.35% (hold-out)
Hariharan et al. (2014)	integration of feature weighting method, feature selection method and classifiers	100%(10-fold CV)
Cai et al. (2017)	support vector machine (SVM) based on bacterial foraging optimization (BFO)	97.42%(10-fold CV)
This Study	CBFO-FKNN	97.89%(10-fold CV)

References

- [1] L. M. de Lau and M. M. Breteler, "Epidemiology of Parkinson's disease," *The Lancet Neurology*, vol. 5, no. 6, pp. 525–535, 2006.
- [2] S. K. van den Eeden, C. M. Tanner, A. L. Bernstein et al., "Incidence of Parkinson's disease: variation by age, gender, and race/ethnicity," *American Journal of Epidemiology*, vol. 157, no. 11, pp. 1015–1022, 2003.
- [3] E. R. Dorsey, R. Constantinescu, J. P. Thompson et al., "Projected number of people with Parkinson disease in the most populous nations, 2005 through 2030," *Neurology*, vol. 68, no. 5, pp. 384–386, 2007.
- [4] N. Singh, V. Pillay, and Y. E. Choonara, "Advances in the treatment of Parkinson's disease," *Progress in Neurobiology*, vol. 81, no. 1, pp. 29–44, 2007.
- [5] B. T. Harel, M. S. Cannizzaro, H. Cohen, N. Reilly, and P. J. Snyder, "Acoustic characteristics of Parkinsonian speech: A potential biomarker of early disease progression and treatment," *Journal of Neurolinguistics*, vol. 17, no. 6, pp. 439–453, 2004.
- [6] J. Rusz, R. Cmejla, H. Ruzickova, and E. Ruzicka, "Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease," *The Journal of the Acoustical Society of America*, vol. 129, no. 1, pp. 350–367, 2011.
- [7] J. Jankovic, "Parkinsons disease: clinical features and diagnosis," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 79, pp. 368–376, 2008.
- [8] J. Massano and K. P. Bhatia, "Clinical approach to Parkinson's disease: features, diagnosis, and principles of management," *Cold Spring Harbor Perspectives in Medicine*, vol. 2, no. 6, Article ID a008870, 2012.
- [9] A. K. Ho, R. Iansek, C. Marigliani, J. L. Bradshaw, and S. Gates, "Speech impairment in a large sample of patients with Parkinson's disease," *Behavioural Neurology*, vol. 11, no. 3, pp. 131–137, 1998.
- [10] B. Harel, M. Cannizzaro, and P. J. Snyder, "Variability in fundamental frequency during speech in prodromal and incipient Parkinson's disease: A longitudinal case study," *Brain and Cognition*, vol. 56, no. 1, pp. 24–29, 2004.
- [11] R. J. Baken and R. F. Orlikoff, *Clinical Measurement of Speech and Voice*, Singular Publishing Group, San Diego, CA, USA, 2nd edition, 2000.
- [12] L. Brabenec, J. Mekyska, Z. Galaz, and I. Rektorova, "Speech disorders in Parkinson's disease: early diagnostics and effects of medication and brain stimulation," *Journal of Neural Transmission*, vol. 124, no. 3, pp. 303–334, 2017.
- [13] M. A. Little, P. E. McSharry, E. J. Hunter, J. Spielman, and L. O. Ramig, "Suitability of dysphonia measurements for telemonitoring of Parkinson's disease," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 4, pp. 1015–1022, 2009.
- [14] R. Das, "A comparison of multiple classification methods for diagnosis of Parkinson disease," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1568–1572, 2010.
- [15] F. Åström and R. Koker, "A parallel neural network approach to prediction of Parkinson's Disease," *Expert Systems with Applications*, vol. 38, no. 10, pp. 12470–12474, 2011.
- [16] C. O. Sakar and O. Kursun, "Telediagnosis of parkinson's disease using measurements of dysphonia," *Journal of Medical Systems*, vol. 34, no. 4, pp. 591–599, 2010.
- [17] D.-C. Li, C.-W. Liu, and S. C. Hu, "A fuzzy-based data transformation for feature extraction to increase classification performance with small medical data sets," *Artificial Intelligence in Medicine*, vol. 52, no. 1, pp. 45–52, 2011.
- [18] B. Shahbaba and R. Neal, "Nonlinear models using Dirichlet process mixtures," *Journal of Machine Learning Research*, vol. 10, pp. 1829–1850, 2009.
- [19] I. Psorakis, T. Damoulas, and M. A. Girolami, "Multiclass relevance vector machines: sparsity and accuracy," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 21, no. 10, pp. 1588–1598, 2010.
- [20] P. F. Guo, P. Bhattacharya, and N. Kharma, "Advances in Detecting Parkinsons Disease," *Medical Biometrics*, pp. 306–314, 2010.
- [21] P. Luukka, "Feature selection using fuzzy entropy measures with similarity classifier," *Expert Systems with Applications*, vol. 38, no. 4, pp. 4600–4607, 2011.
- [22] A. Ozcift and A. Gulten, "Classifier ensemble construction with rotation forest to improve medical diagnosis performance of machine learning algorithms," *Computer Methods and Programs in Biomedicine*, vol. 104, no. 3, pp. 443–451, 2011.
- [23] A. A. Spadoto, R. C. Guido, F. L. Carnevali, A. F. Pagnin, A. X. Falcão, and J. P. Papa, "Improving Parkinson's disease identification through evolutionary-based feature selection," *Conference proceedings: IEEE Engineering in Medicine and Biology Society*, vol. 2011, pp. 7857–7860, 2011.
- [24] K. Polat, "Classification of Parkinson's disease using feature weighting method on the basis of fuzzy C-means clustering," *International Journal of Systems Science*, vol. 43, no. 4, pp. 597–609, 2012.
- [25] H.-L. Chen, C.-C. Huang, X.-G. Yu et al., "An efficient diagnosis system for detection of Parkinson's disease using fuzzy k-nearest neighbor approach," *Expert Systems with Applications*, vol. 40, no. 1, pp. 263–271, 2013.
- [26] W.-L. Zuo, Z.-Y. Wang, T. Liu, and H.-L. Chen, "Effective detection of Parkinson's disease using an adaptive fuzzy k-nearest neighbor approach," *Biomedical Signal Processing and Control*, vol. 8, no. 4, pp. 364–373, 2013.
- [27] G. Sateesh Babu and S. Suresh, "Parkinson's disease prediction using gene expression- A projection based learning meta-cognitive neural classifier approach," *Expert Systems with Applications*, vol. 40, no. 5, pp. 1519–1529, 2013.
- [28] G. S. Babu, S. Suresh, and B. S. Mahanand, "A novel PBL-McRBFN-RFE approach for identification of critical brain regions responsible for Parkinson's disease," *Expert Systems with Applications*, vol. 41, no. 2, pp. 478–488, 2014.
- [29] G. Sateesh Babu, S. Suresh, K. Uma Sangumathi, and H. J. Kim, "A Projection Based Learning Meta-cognitive RBF Network Classifier for Effective Diagnosis of Parkinson's Disease," in *Advances in Neural Networks - ISNN 2012*, vol. 7368 of *Lecture Notes in Computer Science*, pp. 611–620, Springer, Berlin, Germany, 2012.
- [30] M. Hariharan, K. Polat, and R. Sindhu, "A new hybrid intelligent system for accurate detection of Parkinson's disease," *Computer Methods and Programs in Biomedicine*, vol. 113, no. 3, pp. 904–913, 2014.
- [31] M. Gök, "An ensemble of k-nearest neighbours algorithm for detection of Parkinson's disease," *International Journal of Systems Science*, vol. 46, no. 6, pp. 1108–1112, 2015.
- [32] L. Shen, H. Chen, Z. Yu et al., "Evolving support vector machines using fruit fly optimization for medical data classification," *Knowledge-Based Systems*, vol. 96, pp. 61–75, 2016.

- [33] M. Peker, B. Şen, and D. Delen, "Computer-aided diagnosis of Parkinson's disease using complex-valued neural networks and mRMR feature selection algorithm," *Journal of Healthcare Engineering*, vol. 6, no. 3, pp. 281–302, 2015.
- [34] H.-L. Chen, G. Wang, C. Ma, Z.-N. Cai, W.-B. Liu, and S.-J. Wang, "An efficient hybrid kernel extreme learning machine approach for early diagnosis of Parkinson's disease," *Neurocomputing*, vol. 184, no. 4745, pp. 131–144, 2016.
- [35] Z. Cai, J. Gu, and H.-L. Chen, "A New Hybrid Intelligent Framework for Predicting Parkinson's Disease," *IEEE Access*, vol. 5, pp. 17188–17200, 2017.
- [36] B. E. Sakar, M. E. Isenkul, C. O. Sakar et al., "Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings," *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 4, pp. 828–834, 2013.
- [37] H. Zhang, L. Yang, Y. Liu et al., "Classification of Parkinson's disease utilizing multi-edit nearest-neighbor and ensemble learning algorithms with speech samples," *Biomedical Engineering Online*, vol. 15, no. 1, 2016.
- [38] V. Abrol, P. Sharma, and A. K. Sao, "Greedy dictionary learning for kernel sparse representation based classifier," *Pattern Recognition Letters*, vol. 78, pp. 64–69, 2016.
- [39] İ. Cantürk and F. Karabiber, "A Machine Learning System for the Diagnosis of Parkinson's Disease from Speech Signals and Its Application to Multiple Speech Signal Types," *Arabian Journal for Science and Engineering*, vol. 41, no. 12, pp. 5049–5059, 2016.
- [40] A. Józwik, "A learning scheme for a fuzzy k-NN rule," *Pattern Recognition Letters*, vol. 1, no. 5-6, pp. 287–289, 1983.
- [41] J. M. Keller, M. R. Gray, and J. A. Givens, "A fuzzy K-nearest neighbor algorithm," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-15, no. 4, pp. 580–585, 1985.
- [42] J. Derrac, S. García, and F. Herrera, "Fuzzy nearest neighbor algorithms: Taxonomy, experimental analysis and prospects," *Information Sciences*, vol. 260, pp. 98–119, 2014.
- [43] D.-Y. Liu, H.-L. Chen, B. Yang, X.-E. Lv, L.-N. Li, and J. Liu, "Design of an enhanced Fuzzy k-nearest neighbor classifier based computer aided diagnostic system for thyroid disease," *Journal of Medical Systems*, vol. 36, no. 5, pp. 3243–3254, 2012.
- [44] J. Sim, S.-Y. Kim, and J. Lee, "Prediction of protein solvent accessibility using fuzzy k-nearest neighbor method," *Bioinformatics*, vol. 21, no. 12, pp. 2844–2849, 2005.
- [45] Y. Huang and Y. Li, "Prediction of protein subcellular locations using fuzzy k-NN method," *Bioinformatics*, vol. 20, no. 1, pp. 21–28, 2004.
- [46] H.-L. Chen, B. Yang, G. Wang et al., "A novel bankruptcy prediction model based on an adaptive fuzzy k-nearest neighbor method," *Knowledge-Based Systems*, vol. 24, no. 8, pp. 1348–1359, 2011.
- [47] M.-Y. Cheng and N.-D. Hoang, "A Swarm-Optimized Fuzzy Instance-based Learning approach for predicting slope collapses in mountain roads," *Knowledge-Based Systems*, vol. 76, pp. 256–263, 2015.
- [48] M.-Y. Cheng and N.-D. Hoang, "Groutability estimation of grouting processes with microfine cements using an evolutionary instance-based learning approach," *Journal of Computing in Civil Engineering*, vol. 28, no. 4, Article ID 04014014, 2014.
- [49] K. M. Passino, "Biomimicry of bacterial foraging for distributed optimization and control," *IEEE Control Systems Magazine*, vol. 22, no. 3, pp. 52–67, 2002.
- [50] L. Liu, L. Shan, Y. Dai, C. Liu, and Z. Qi, "A Modified Quantum Bacterial Foraging Algorithm for Parameters Identification of Fractional-Order System," *IEEE Access*, vol. 6, pp. 6610–6619, 2018.
- [51] B. Hernandez-Ocana, O. Chavez-Bosquez, J. Hernandez-Torruco, J. Canul-Reich, and P. Pozos-Parra, "Bacterial Foraging Optimization Algorithm for Menu Planning," *IEEE Access*, vol. 6, pp. 8619–8629, 2018.
- [52] A. M. Othman and H. A. Gabbar, "Enhanced microgrid dynamic performance using a modulated power filter based on enhanced bacterial foraging optimization," *Energies*, vol. 10, no. 6, 2017.
- [53] S. S. Chouhan, A. Kaul, U. P. Singh, and S. Jain, "Bacterial Foraging Optimization Based Radial Basis Function Neural Network (BRBFNN) for Identification and Classification of Plant Leaf Diseases: An Automatic Approach Towards Plant Pathology," *IEEE Access*, vol. 6, pp. 8852–8863, 2018.
- [54] B. Turanoğlu and G. Akkaya, "A new hybrid heuristic algorithm based on bacterial foraging optimization for the dynamic facility layout problem," *Expert Systems with Applications*, vol. 98, pp. 93–104, 2018.
- [55] X. Lv, H. Chen, Q. Zhang, X. Li, H. Huang, and G. Wang, "An improved bacterial-foraging optimization-based machine learning framework for predicting the severity of somatization disorder," *Algorithms*, vol. 11, no. 2, article 17, 2018.
- [56] R. Majhi, G. Panda, B. Majhi, and G. Sahoo, "Efficient prediction of stock market indices using adaptive bacterial foraging optimization (ABFO) and BFO based techniques," *Expert Systems with Applications*, vol. 36, no. 6, pp. 10097–10104, 2009.
- [57] S. Dasgupta, S. Das, A. Biswas, and A. Abraham, "Automatic circle detection on digital images with an adaptive bacterial foraging algorithm," *Soft Computing*, vol. 14, no. 11, pp. 1151–1164, 2010.
- [58] S. Mishra, "A hybrid least square-fuzzy bacterial foraging strategy for harmonic estimation," *IEEE Transactions on Evolutionary Computation*, vol. 9, no. 1, pp. 61–73, 2005.
- [59] S. Mishra and C. N. Bhende, "Bacterial foraging technique-based optimized active power filter for load compensation," *IEEE Transactions on Power Delivery*, vol. 22, no. 1, pp. 457–465, 2007.
- [60] M. Ulagammai, P. Venkatesh, P. S. Kannan, and N. Prasad Padhy, "Application of bacterial foraging technique trained artificial and wavelet neural networks in load forecasting," *Neurocomputing*, vol. 70, no. 16–18, pp. 2659–2667, 2007.
- [61] Q. Wu, J. F. Mao, C. F. Wei et al., "Hybrid BF-PSO and fuzzy support vector machine for diagnosis of fatigue status using EMG signal features," *Neurocomputing*, vol. 173, pp. 483–500, 2016.
- [62] C. Yang, J. Ji, J. Liu, J. Liu, and B. Yin, "Structural learning of Bayesian networks by bacterial foraging optimization," *International Journal of Approximate Reasoning*, vol. 69, pp. 147–167, 2016.
- [63] T. S. Sivarani, S. Joseph Jawhar, C. Agees Kumar, and K. Prem Kumar, "Novel bacterial foraging-based ANFIS for speed control of matrix converter-fed industrial BLDC motors operated under low speed and high torque," *Neural Computing and Applications*, pp. 1–24, 2016.
- [64] F. Zhao, Y. Liu, Z. Shao, X. Jiang, C. Zhang, and J. Wang, "A chaotic local search based bacterial foraging algorithm and its application to a permutation flow-shop scheduling problem," *International Journal of Computer Integrated Manufacturing*, vol. 29, no. 9, pp. 962–981, 2016.
- [65] Y. Sun, S. Todorovic, and S. Goodison, "Local-learning-based feature selection for high-dimensional data analysis," *IEEE*

- Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1610–1626, 2010.
- [66] H. Chen, J. Lu, Q. Li, C. Lou, D. Pan, and Z. Yu, “A New Evolutionary Fuzzy Instance-Based Learning Approach: Application for Detection of Parkinson’s Disease,” in *Advances in Swarm and Computational Intelligence*, Y. Tan, Y. Shi, F. Buarque et al., Eds., vol. 9141 of *Lecture Notes in Computer Science*, pp. 42–50, Springer International Publishing, 2015.
 - [67] T. Kapitaniak, “Continuous control and synchronization in chaotic systems,” *Chaos, Solitons & Fractals*, vol. 6, no. C, pp. 237–244, 1995.
 - [68] T. Bäck and H.-P. Schwefel, “An overview of evolutionary algorithms for parameter optimization,” *Evolutionary Computation*, vol. 1, no. 1, pp. 1–23, 1993.
 - [69] X. Cheng and M. Jiang, “An improved artificial bee colony algorithm based on Gaussian mutation and chaos disturbance,” *Lecture Notes in Computer Science*, vol. 7331, no. 1, pp. 326–333, 2012.
 - [70] N. Higashi and H. Iba, “Particle swarm optimization with Gaussian mutation,” in *Proceedings of the 2003 IEEE Swarm Intelligence Symposium, SIS 2003*, pp. 72–79, April 2003.
 - [71] C. Jena, M. Basu, and C. K. Panigrahi, “Differential evolution with Gaussian mutation for combined heat and power economic dispatch,” *Soft Computing*, vol. 20, no. 2, pp. 681–688, 2016.
 - [72] C. Chang and C. Lin, “LIBSVM: a Library for support vector machines,” *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, article 27, 2011.
 - [73] S. L. Salzberg, “On comparing classifiers: pitfalls to avoid and a recommended approach,” *Data Mining and Knowledge Discovery*, vol. 1, no. 3, pp. 317–328, 1997.
 - [74] A. Statnikov, I. Tsamardinos, Y. Dosbayev, and C. F. Aliferis, “GEMS: a system for automated cancer diagnosis and biomarker discovery from microarray gene expression data,” *International Journal of Medical Informatics*, vol. 74, no. 7–8, pp. 491–503, 2005.
 - [75] T. Fawcett, “ROC graphs: Notes and practical considerations for researchers,” *Machine Learning*, vol. 31, pp. 1–38, 2004.
 - [76] A. H. Gandomi, X.-S. Yang, S. Talatahari, and A. H. Alavi, “Firefly algorithm with chaos,” *Communications in Nonlinear Science and Numerical Simulation*, vol. 18, no. 1, pp. 89–98, 2013.
 - [77] X. S. Yang, “Flower pollination algorithm for global optimization,” in *Unconventional Computation and Natural Computation*, vol. 7445 of *Lecture Notes in Computer Science*, pp. 240–249, Springer, Berlin, Germany, 2012.
 - [78] X.-S. Yang, “A new metaheuristic bat-inspired Algorithm,” *Studies in Computational Intelligence*, vol. 284, pp. 65–74, 2010.
 - [79] S. Mirjalili, “Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems,” *Neural Computing and Applications*, vol. 27, pp. 1053–1073, 2016.
 - [80] J. Kennedy and R. Eberhart, “Particle swarm optimization,” in *Proceedings of the IEEE International Conference on Neural Networks*, pp. 1942–1948, Perth, Australia, December 1995.
 - [81] J. Derrac, S. García, D. Molina, and F. Herrera, “A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms,” *Swarm and Evolutionary Computation*, vol. 1, no. 1, pp. 3–18, 2011.

Research Article

Common Interferences Removal from Dense Multichannel EEG Using Independent Component Decomposition

Weifeng Li , Yuxiaotong Shen , Jie Zhang, Xiaolin Huang , Ying Chen , and Yun Ge 

School of Electronic Science and Engineering, Nanjing University, Nanjing 210023, China

Correspondence should be addressed to Xiaolin Huang; xlhuang@nju.edu.cn and Yun Ge; geyun@nju.edu.cn

Received 26 October 2017; Revised 14 February 2018; Accepted 18 March 2018; Published 27 May 2018

Academic Editor: Fei Chen

Copyright © 2018 Weifeng Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To improve the spatial resolution, dense multichannel electroencephalogram with more than 32 leads has gained more and more applications. However, strong common interference will not only conceal the weak components generated from the specific isolated neural source, but also lead to severe spurious correlation between different brain regions, which results in great distortion on brain connectivity or brain network analysis. Starting from the fast independent component analysis algorithm, we first derive the mixing matrix of independent source components based on the baseline signals prior to tasks. Then, we identify the common interferences as those components whose mixing vectors span the minimum angles with respect to the unitary vector. By assuming that both the common interferences and their corresponding mixing vectors stay consistent during the entire experiment, we apply the demixing and mixing matrix to the task signals and remove the inferred common interferences. Subsequently, we validate the method using simulation. Finally, the index of global coherence is calculated for validation. It turns out that the proposed method can successfully remove the common interferences so that the prominent coherence of mu rhythms in motor imagery tasks is unmasked. The proposed method can gain wide applications because it reveals the true correlation between the local sources in spite of the low signal-to-noise ratio.

1. Introduction

Electroencephalogram (EEG) collected from the scalp is the integration of the electrical activities of amounts of cortex neurons blurred by the skull [1]. Although it is widely accepted that EEG has the advantage of high temporal resolution, the spatial resolution remains as a problem [2]. To improve the spatial resolution, dense multichannel EEG (with more than 32 channels) and high-density EEG (with more than 128 channels) have gained more and more applications. However, the more the channels are used, the more the redundant information is involved. It directly results in the fact that the weak components generated from the specific isolated neural source are deeply concealed by the common components from the surrounding sources [3]. Moreover, these redundancies can lead to a spurious correlation/coordination between different brain regions while in fact little or none is present. It will greatly distort the result of the brain connectivity or brain network analysis, which becomes more and

more popular [4–15]. Therefore, it is of great importance to unmask the isolated source-corresponding component from the originally collected signals with too much redundant information or common interferences.

Among multichannel EEG redundancy-removal methods, one representative is surface Laplacian reference scheme [16, 17]. After subtracting the average potential in the local neighborhood, the original signals referencing to one or two common locations are converted to referencing to the respective local one. Typically, the signal amplitude will greatly decrease, with the expected return of redundancy removal. The surface Laplacian reference scheme is theoretically simple and easy to implement. However, using the arithmetical mean within the neighborhood as the local reference may be a little bit rough, regardless of the conduction differences among the neighbor leads. In addition, great attention should be paid to the selection of the neighborhood.

Another representative is independent component analysis (ICA) [18, 19]. In fact, ICA has long been applied to EEG

preprocessing [20–27] including electrooculography artifacts removal. Recently, Whitmore and Lin have succeeded in removing distal electrical reference as well as volume-conducted noises from local field potentials using ICA [25]. It greatly motivates us to step further, trying a more general common interference removal.

In the presented manuscript, we do not identify the source or the frequency of the common interference. Instead, we only assume that the common interference will affect the different channel most evenly and the mixing vectors keep constant during the whole experiment, regardless of the mental activities. In addition, by regarding both the common interferences and their transfer vectors as identical in the entire experimental circumstance, we adopt the component extracted from the baseline data. We validate the proposed method on BCI competition dataset 1 [28, 29]. It turns out that the method can successfully unmask the coherence in mu rhythm during a motor imaginary task. Since high-density EEG and brain connectivity or brain network are the trends in neuroscience, the proposed method can gain wide applications.

In the manuscript, we first describe the method in Section 2, and then in Section 3 the method is validated using simulation series as well as experimental data provided in BCI Competition IV, and finally results are discussed in Section 4.

2. Methods

The method includes three steps in order: independent components decomposition, the common interference identification, and removal and inverse transformation.

2.1. Independent Component Decomposition. Mathematically, given the independent M sources as $\mathbf{S} = (s_{i,j})$, $i = 1, 2, \dots, M$, $j = 1, 2, \dots, L$, in which j represents the sampling time index, the N -channel ($N \geq M$) collected signal denoted as $\mathbf{X} = (x_{i,j})$, $i = 1, 2, \dots, N$, $j = 1, 2, \dots, L$, can be calculated as

$$\mathbf{X} = \mathbf{AS}, \quad (1)$$

in which \mathbf{A} is the N -by- M mixing matrix. Theoretically, each row of \mathbf{A} represents a set of combination weights of the M different sources on the specific channel, and each column of \mathbf{A} , denoted as $\bar{\mathbf{A}}_j$, reflects the relative impacts of the j th source on all the N different channels.

The independent component decomposition is to resolve (1) to obtain

$$\mathbf{S} = \mathbf{A}^{-1}\mathbf{X} = \mathbf{WX}, \quad (2)$$

where \mathbf{W} is called demixing matrix. Because neither \mathbf{W} nor \mathbf{S} is known a priori, the maximization of non-Gaussianity or minimization of mutual information principle is conventionally employed to approximate the \mathbf{W} as well as \mathbf{S} through iteration [12].

Herein, we adopt FastICA algorithm proposed by Hyvärinen [13] for independent component decomposition. The fixed-point iteration scheme as well as the maximum-negative entropy principle is employed to find the orthogonal

rotation matrix \mathbf{W} with the maximal non-Gaussian measure of the prewhitened data. And then the mixing matrix \mathbf{A} can be calculated as

$$\mathbf{A} = \mathbf{W}^{-1}. \quad (3)$$

2.2. Common Interference Identification and Removal. Subsequently, we try to identify and remove the common interference through analyzing the mixing matrix \mathbf{A} .

The putative common interference component is assumed as a distal signal that has approximately same effect on all electrodes. In order to obtain local brain activities more accurately, these distal common interference components should be removed. To do this, the vector angles are calculated between each $\bar{\mathbf{A}}_j$ and a unit vector, and the smaller the angle is, the more uniform the impacts of the corresponding source (independent component) across channels are and the more likely the corresponding source is a common interference. We delete this source through setting the corresponding k th independent source $s_{k,j}$, $j = 1, 2, \dots, L$ as 0, obtain the processed $\hat{\mathbf{S}}$, and finally derive the deabundance signals as

$$\hat{\mathbf{X}} = \mathbf{A}\hat{\mathbf{S}} \quad (4)$$

3. Experiments

3.1. Simulation. To validate the proposed method, we first applied it to simulation series. We define the three collected channel signals which are determined by three independent components, i.e., $s_1 = \sin(2\pi \times 10t)$, $s_2 = \cos(2\pi t)$, and random Gaussian noises with $\mu = 0$, $\sigma = 10$, and the mixing matrix $\mathbf{A} = \begin{bmatrix} 1 & -0.5 & 0.19 \\ 0.2 & 1 & 0.21 \\ -0.4 & 0.4 & 0.2 \end{bmatrix}$. As described in Section 2, the collected signals are derived by $\mathbf{X} = \mathbf{AS}$. The Gaussian component is deliberately set with great amplitude and is treated as the common interference. Theoretically, we can obtain the pure signal without common interference via setting the 3rd column elements as 0 s. We plot the pure signal of Channels 1 and 2 in Figure 1(a), and the collected contaminated signals in Figure 1(b). Then, we apply the proposed method to \mathbf{X} . After common interference removal, signals of Channels 1 and 2 are plotted in Figure 1(c). To quantitatively evaluate the signal quality, we also calculate the linear correlation coefficient between the collected signals and the pure signals, both before and after common interference removal.

Figure 1 shows that the proposed method nearly doubles the correlation coefficient with the pure signals, and wave form also indicates the signal quality is greatly improved, even in such low signal-to-noise ratios.

3.2. Application to Scalp EEG

3.2.1. Data Description. We apply the proposed method to the calibration data in dataset 1 of BCI Competition IV, provided by the Berlin BCI group [20, 21].

This dataset includes three artificial data (#c, #d, and #e) as well as four data pieces recorded from 4 healthy subjects (#a, #b, #f, and #g) in motor imagery experiments. Each data

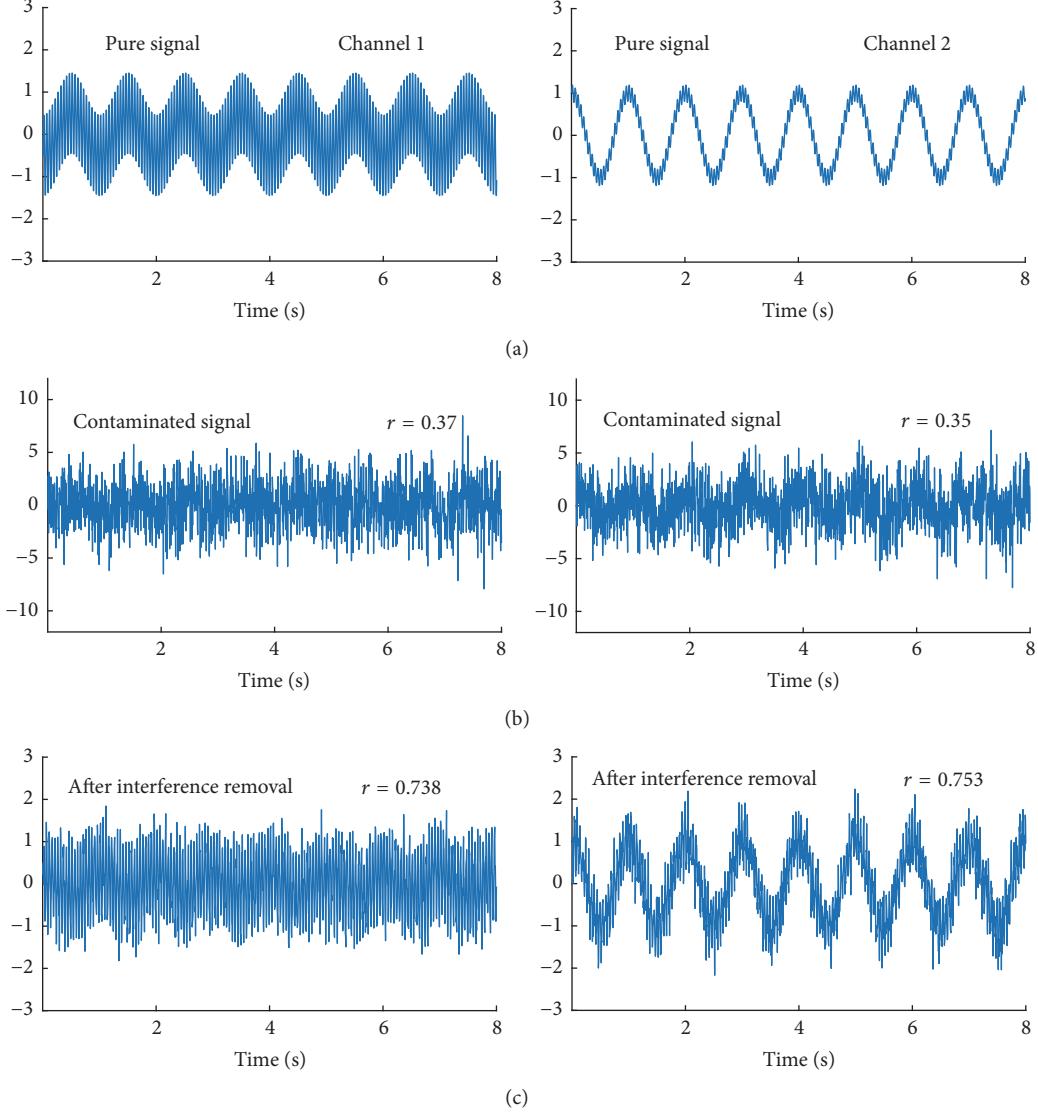


FIGURE 1: Simulation results ((a) is the pure signal, (b) is the contaminated signal, and (c) is the postprocessed signal. As we can see, although the noise is strong, the proposed method greatly improved the signal quality by doubling the correlation coefficient with the pure signal. Thus it validates the proposed method).

includes 59-channel continuous EEG or artificial simulated EEG, with a sampling rate of 1000 Hz and high cut-off frequency of 200 Hz. To compress the data size, the provider downsampled the data to 100 Hz after low-pass filtering them with stopband edge frequency 49 Hz [21]. According to the data information, we plot the lead locations in Figure 2.

In each experiment, before the first cue was given, the very first duration of 16 s can be considered the baseline signal, and then 200 trials of cue-response with 8 s duration were followed. Each trial consists of 4 s cue and motor imagery task, 2 s blank screen, and 2 s fixation. Motor imagery can be movement of left hand, right hand, or feet, and for each subject two classes of motor imagery were chosen. The first 2.56 s sections beginning with the cue are used for the following analysis.

3.2.2. Common Interference Removal. The baseline signals are firstly taken as original data to calculate the best orthogonal rotation matrix W and no more than 59 independent components S by FastICA [13]. The stopping criterion of FastICA is set as the minimum weight change of 10^{-5} .

Although the brain activities related independent sources might be different between the baseline and task trials, both the common interference signal itself and its corresponding transfer vector are assumed to be identical in the entire experiment. Therefore, the putative common interference components calculated by the baseline signals can be extended to the following task state of the EEG treatment. That is, W is applied to task trial signals:

$$\mathbf{S}^{\text{task}} = \mathbf{W}\mathbf{X}^{\text{task}} \quad (5)$$

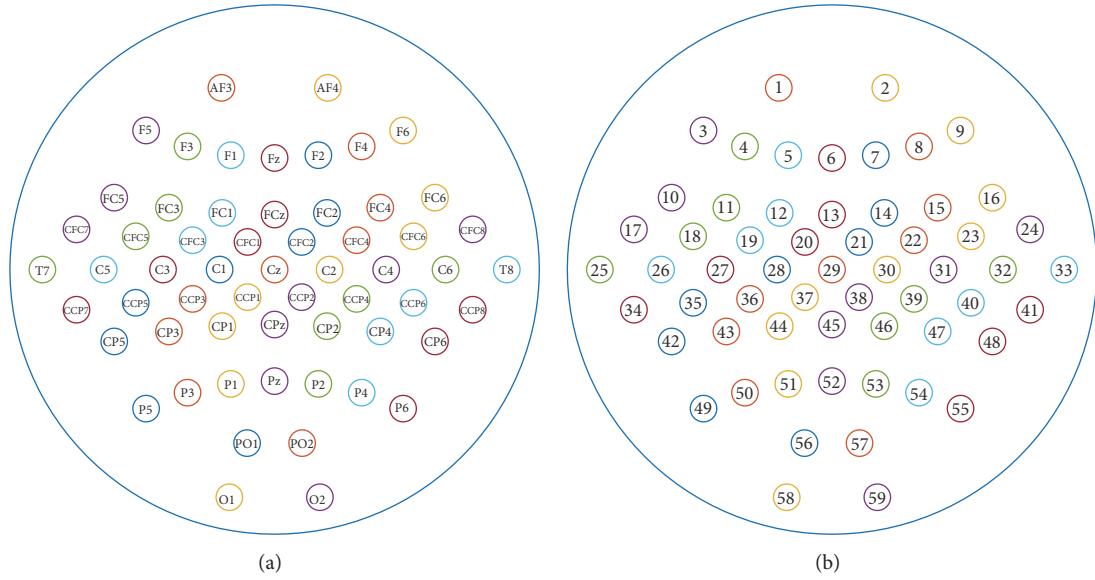


FIGURE 2: Lead locations for signals in dataset 1 ((a) presents the lead label, and (b) presents the lead number, in case we would refer to it in the manuscript).

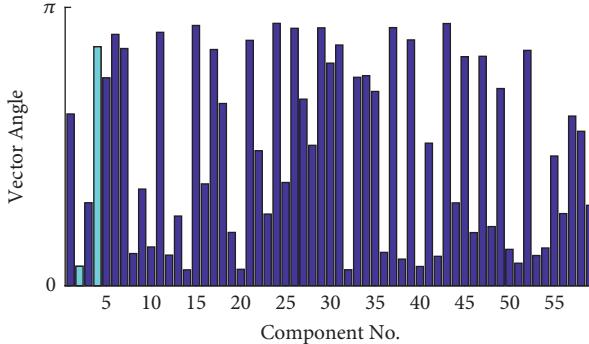


FIGURE 3: The vector angle of demixed independent component (the light blue marks the two components that are treated as common interference and removed).

After deleting the common interference, we obtain the processed signal as

$$\widehat{\mathbf{X}}^{\text{task}} = \mathbf{A}\widehat{\mathbf{s}}^{\text{task}} \quad (6)$$

3.2.3. EEG Results. We apply the proposed method to EEG. As a representative, we present the vector angle derived from #a in Figure 3, in which the light blue marks the two components treated as the common interference and then removed. As seen, these two components are not of the two smallest vector angles. However, we set an additional restriction that all elements in the mapping vector should be of the same sign. Therefore, in this case components 2 and 4 are determined as the common interference.

We also examined the EEG series before and after the processing. As a representative, we plot two leads of subject #a in Figure 4. As shown in Figure 4, the original signals collected from leads 5 (F1) and 7 (F2) are highly correlated. And the eye movement artifacts are obvious, occurring from

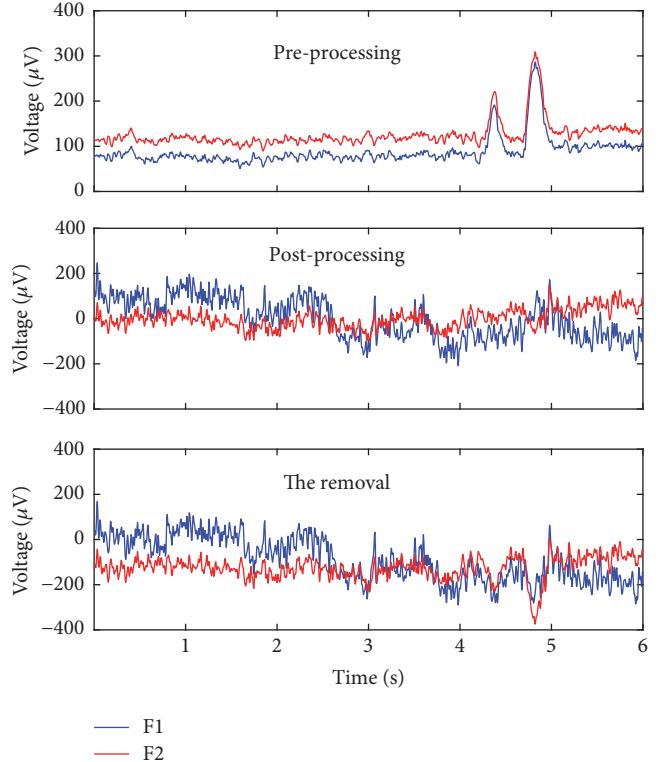


FIGURE 4: A representative EEG series result of subject #a. (Frontal EEG is usually contaminated by ocular and eye movement artifacts. This section includes two obvious ocular artifacts, occurring from 4.3 s to 5 s. After processing, the eye movement artifacts are successfully removed, and the correlation between F1 and F2 is alleviated)

4.3 s to 5 s. After processing, the correlation is alleviated, and the eye movement artifacts are removed.

It is difficult to provide an accurate signal quality evaluation, because we in fact do not know the “real” signal.

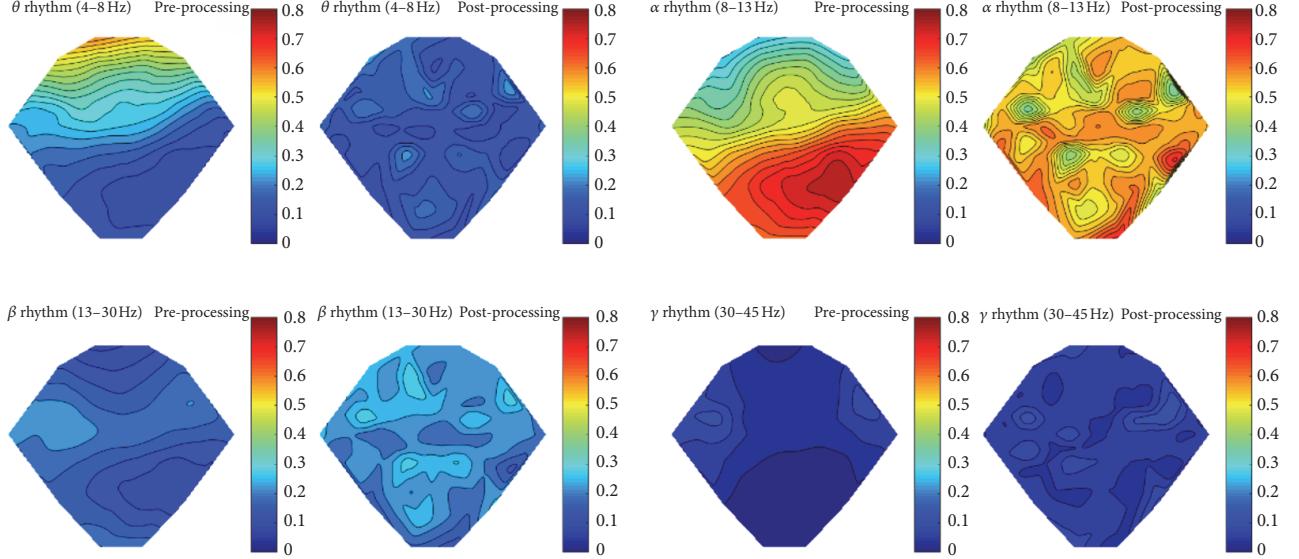


FIGURE 5: The relative power comparison between the original signal and the common interference removed signal. (Color represents the specific rhythm power relative to the power of the entire frequency band. As we see, after processing, the relative power reveals more distribution characteristics. It proves that we do uncover the intrinsic isolated neural activities, which were concealed by the strong common interference)

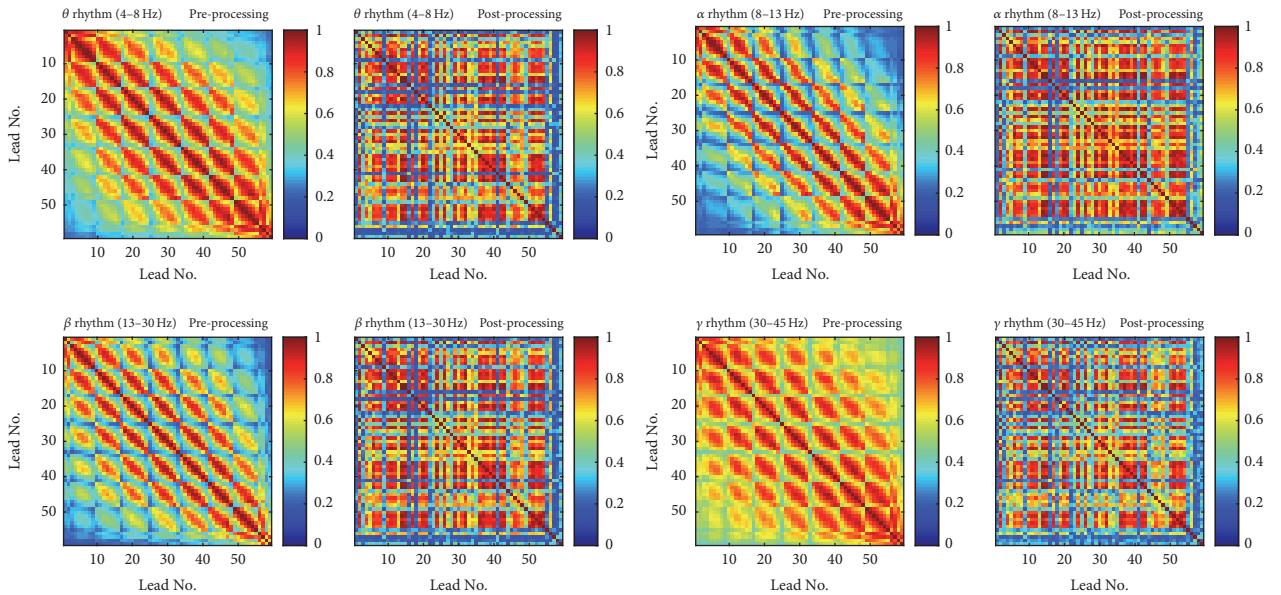


FIGURE 6: Coherence heatmap comparison between the original signal and the common interference removed signal. (As we can see, before processing, theta and gamma rhythm both present strong coherence for nearly all lead pairs. After processing, coherence differences among different pairs become obvious. In addition, the brighter lines parallel to the diagonal line diffuse to wider region after processing. It implies that after the common interference removal coherences between some far-away lead pairs become unconcealed and prominent)

However, we tried to calculate relative power as well as coherence and made comparison between the original signals (preprocessing) and processed signals with common interference removed (postprocessing). Taking the subject #a as an example, we present the relative power in Figure 5 and coherence heatmaps in Figure 6 for the commonly defined EEG rhythms.

As we can see from Figure 5, maps of the original signals have bigger connected regions, whereas after processing maps reveal more distribution characteristics. It proves that we do uncover the intrinsic isolated neural activities, which were concealed by strong common interference.

As we can see from Figure 6, for the original signal, the common interference imposes strong coherence on the

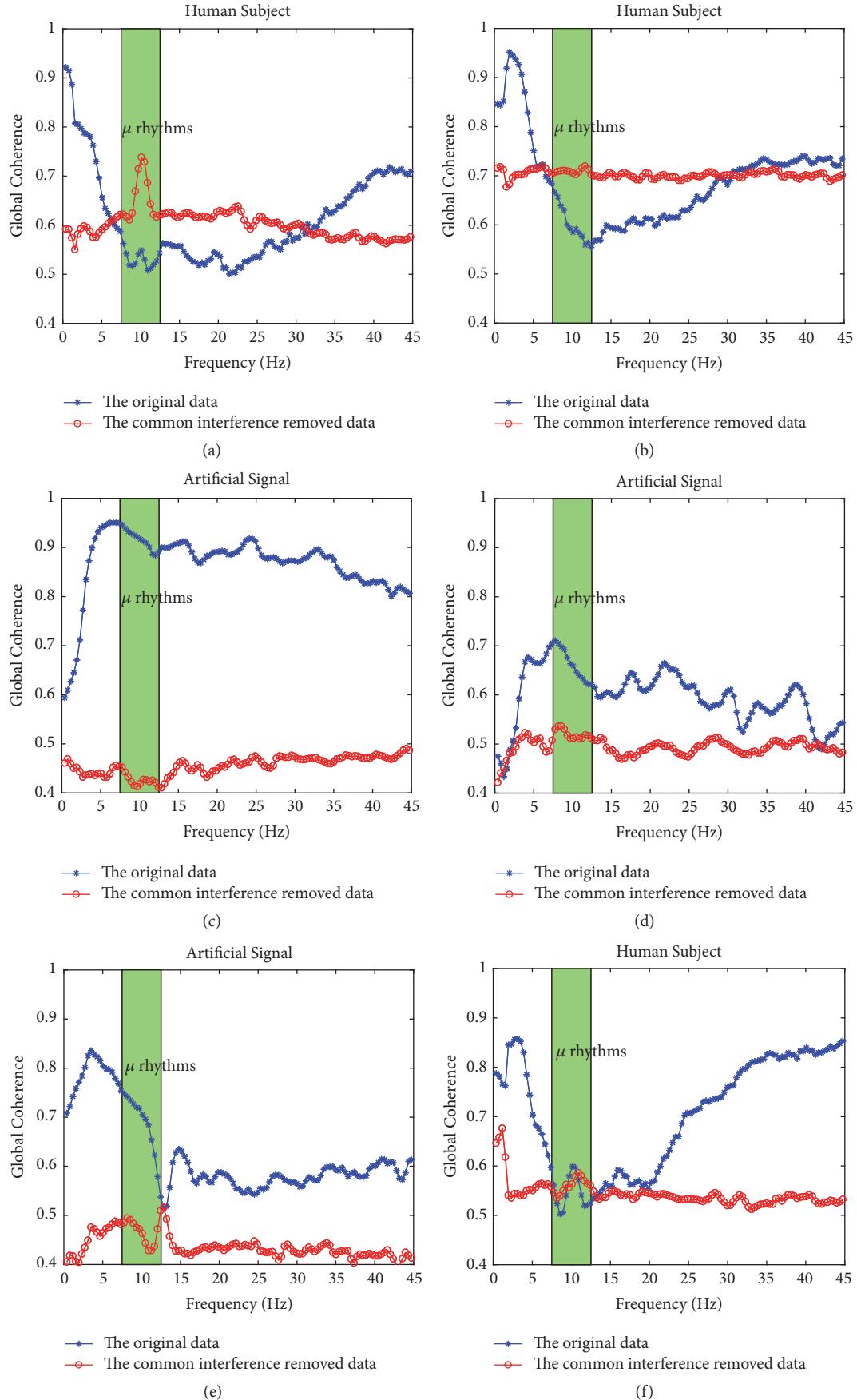


FIGURE 7: Continued.

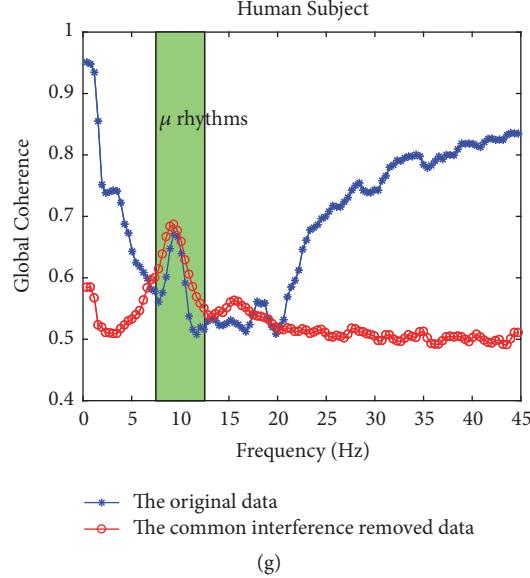


FIGURE 7: Comparison of the index of global coherence between the original and the processed signals. (The horizontal axis represents the frequency, and the vertical axis represents the index of global coherence. In the original data, all human subjects, i.e., (a), (b), (f), and (g), present high coherences in both low frequency band and high frequency band, which indicates a universal conductance induced consistency on scalp. However, as for the processed data, the high coherences in that two frequency bands are both suppressed while a coherence peak in mu rhythms (the green shade area) becomes prominent except for subject #b. It implies that although we did not mean to filter the specific frequency, the spurious high coherences caused by the common interferences are greatly alleviated. We cannot observe mu coherence in the artificial signals, i.e., (c), (d), and (e))

leads in the same neighborhood. That leads to the brighter lines parallel to the diagonal line in the heatmap, which may conceal coherence between leads that are not close in location. However, after being processed by the proposed method, the bright neighborhood diffused, and coherences between some far-away lead pairs become unconcealed and prominent.

We further calculated an interesting index, i.e., the global coherence [3, 22], and made comparison. All global coherence results for the data sets 1 are presented in Figure 7, in which the horizontal axis represents the frequency, and the vertical axis represents the index of global coherence.

As shown in Figures 7(a), 7(b), 7(f), and 7(g), in the original data, all human subjects present high coherences in both low frequency band and high frequency band, which indicates a universal conductance induced consistency on scalp. However, as for the processed data, the high coherences in that two frequency bands are both suppressed while a coherence peak in mu rhythms (the green shade area in Figure 7) becomes prominent except for subject #b. It implies that although we did not mean to filter the specific frequency, the spurious high coherences caused by the common interferences are greatly alleviated. Meanwhile, the coherence in mu rhythms, which are intrinsically related to the motor imaginary, is unmasked. And as to the artificial signals in Figures 7(c), 7(d), and 7(e), we cannot observe mu coherence. Since these signals are artificial, we consider it reasonable. Therefore, the above results demonstrate that the proposed method is successful.

4. Discussions and Conclusion

Coherence is the equivalence of correlation in frequency domain. In active brains, correlation analysis in time domain is difficult because the EEG amplitude is very weak for desynchronization. In these cases, coherence is the appropriate substitute. However, whether in time domain or in frequency domain, the spurious correlation brought by the common interference imposes a big problem on unmasking the true cooperation between the weak neural sources. In the presented work, we propose an independent component decomposition based method; the two most crucial innovations include the following: (1) the angle between the mixing vector and the unitary vector rather than the frequency or morphology is used to identify the common interference; and (2) the independent component source and the mixing vectors derived from the baseline signal are applied to the following task signals. As to (2), since most EEG experiments are implemented in stimulus-locking paradigm, the proposed method can gain wide applications. In brief, the proposed method presents successful application in the motor imaginary EEG of BCI Competition IV and reveals the coherence peak in motor related mu rhythms.

Appendix

The cross-spectral matrix \mathbf{C} is calculated as

$$\mathbf{C}_{ij}^X(f) = \frac{1}{K} \sum_{k=1}^K X_i^k(f) X_j^k(f)^*, \quad (\text{A.1})$$

in which $X_i^k(f)$ and $X_j^k(f)$ are the spectrums calculated from channels i and j , respectively, at frequency f . Then the cross-spectral matrix is singular value decomposed as $\mathbf{C} = \mathbf{USV}$, where \mathbf{S} is the diagonal matrix with each diagonal element, denoted as λ_i , being an eigenvalue, and $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$. Finally, the global coherence is calculated as

$$\text{Coh}_{\text{Global}} = \frac{\lambda_1}{\sum_{i=1}^N \lambda_i}. \quad (\text{A.2})$$

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported by Forward-Looking Project on the Integration of Industry, Education and Research of Jiangsu Province (Grant no. BY2015069-06) and Social Development Program of Primary Research & Development Plan in Jiangsu Province (Grant nos. BE2016733 and BE2017679).

References

- [1] X. Ma, X. Huang, Y. Shen et al., "EEG based topography analysis in string recognition task," *Physica A: Statistical Mechanics and its Applications*, vol. 469, pp. 531–539, 2017.
- [2] M. Gavaret, L. Maillard, and J. Jung, "High-resolution EEG (HR-EEG) and magnetoencephalography (MEG)," *Neurophysiologie Clinique / Clinical Neurophysiology*, vol. 45, no. 1, pp. 105–111, 2015.
- [3] M. I. Franco, L. Turin, A. Mershin, and E. M. C. Skoulakis, "Molecular vibration-sensing component in *Drosophila melanogaster* olfaction," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 9, pp. 3797–3802, 2011.
- [4] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nature Reviews Neuroscience*, vol. 10, no. 3, pp. 186–198, 2009.
- [5] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: Uses and interpretations," *NeuroImage*, vol. 52, no. 3, pp. 1059–1069, 2010.
- [6] X. Ma, X. Huang, Y. Ge et al., "Brain Connectivity Variation Topography Associated with Working Memory," *PLoS ONE*, vol. 11, no. 12, p. e0165168, 2016.
- [7] S. Palva and J. M. Palva, "Discovering oscillatory interaction networks with M/EEG: challenges and breakthroughs," *Trends in Cognitive Sciences*, vol. 16, no. 4, pp. 219–229, 2012.
- [8] K. K. L. Liu, R. P. Bartsch, A. Lin, R. N. Mantegna, and P. C. Ivanov, "Plasticity of brain wave network interactions and evolution across physiologic states," *Frontiers in Neural Circuits*, vol. 9, no. OCTOBER, pp. 1–15, 2015.
- [9] R. P. Bartsch, K. K. L. Liu, A. Bashan, and P. C. Ivanov, "Network physiology: How organ systems dynamically interact," *PLoS ONE*, vol. 10, no. 11, Article ID e0142143, 2015.
- [10] P. C. H. Ivanov, K. K. L. Liu, and R. P. Bartsch, "Focus on the emerging new fields of network physiology and network medicine," *New Journal of Physics*, vol. 18, no. 10, Article ID 100201, 2016.
- [11] A. Lin, K. K. L. Liu, R. P. Bartsch, and P. C. Ivanov, "Delay-correlation landscape reveals characteristic time delays of brain rhythms and heart interactions," *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences*, vol. 374, no. 2067, 2016.
- [12] J. R. Moorman, D. E. Lake, and P. C. Ivanov, "Early detection of sepsis - A role for network physiology?" *Critical Care Medicine*, vol. 44, no. 5, pp. e312–e313, 2016.
- [13] R. P. Bartsch and P. C. Ivanov, "Coexisting forms of coupling and Phase-Transitions in physiological networks," *Communications in Computer and Information Science*, vol. 438, pp. 270–287, 2014.
- [14] P. C. H. Ivanov and et al., "Physiologic networks: topological and functional transitions across sleep stages," *Sleep*, vol. 35, 2012, Supplement S: A52-A53.
- [15] A. Bashan, R. P. Bartsch, J. W. Kantelhardt, S. Havlin, and P. C. Ivanov, "Network physiology reveals relations between network topology and physiological function," *Nature Communications*, vol. 3, article no. 702, 2012.
- [16] P. L. Nunez and R. Srinivasan, "Electric Fields of the Brain: The neurophysics of EEG," *Electric Fields of the Brain: The neurophysics of EEG*, pp. 1–611, 2009.
- [17] G. Fein, J. Raz, F. F. Brown, and E. L. Merrin, "Common reference coherence data are confounded by power and phase effects," *Electroencephalography and Clinical Neurophysiology*, vol. 69, no. 6, pp. 581–584, 1988.
- [18] P. Comon, "Independent component analysis, A new concept?" *Signal Processing*, vol. 36, no. 3, pp. 287–314, 1994.
- [19] A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 10, no. 3, pp. 626–634, 1999.
- [20] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of Neuroscience Methods*, vol. 134, no. 1, pp. 9–21, 2004.
- [21] A. Delorme, T. Sejnowski, and S. Makeig, "Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis," *NeuroImage*, vol. 34, no. 4, pp. 1443–1449, 2007.
- [22] R. N. Vigário, "Extraction of ocular artefacts from EEG using independent component analysis," *Electroencephalography and Clinical Neurophysiology*, vol. 103, no. 3, pp. 395–404, 1997.
- [23] J. Onton, M. Westerfield, J. Townsend, and S. Makeig, "Imaging human EEG dynamics using independent component analysis," *Neuroscience & Biobehavioral Reviews*, vol. 30, no. 6, pp. 808–822, 2006.
- [24] G. G. Knyazev, J. Y. Slobodskoj-Plusnin, A. V. Bocharov, and L. V. Pylkova, "The default mode network and EEG alpha oscillations: An independent component analysis," *Brain Research*, vol. 1402, pp. 67–79, 2011.
- [25] N. W. Whitmore and S.-C. Lin, "Unmasking local activity within local field potentials (LFPs) by removing distal electrical signals using independent component analysis," *NeuroImage*, vol. 132, pp. 79–92, 2016.
- [26] B. Blankertz, G. Dornhege, M. Krauledat, K.-R. Müller, and G. Curio, "The non-invasive Berlin brain-computer interface: fast acquisition of effective performance in untrained subjects," *NeuroImage*, vol. 37, no. 2, pp. 539–550, 2007.
- [27] K. K. Liu, R. P. Bartsch, Q. D. Ma, and P. C. Ivanov, "Major component analysis of dynamic networks of physiologic organ interactions," *Journal of Physics: Conference Series*, vol. 640, p. 012013, 2015.

- [28] <http://www.bbci.de/competition/iv/#dataset1>.
- [29] P. P. Mitra and H. Bokil, *Observed Brain Dynamics*, Oxford Univ Press, New York, NY, USA, 2008.

Research Article

Single-Frequency Ultrasound-Based Respiration Rate Estimation with Smartphones

Linfei Ge ,¹ **Jin Zhang** ,¹ and **Jing Wei**²

¹Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen, Guangdong, China

²Department of Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen, Guangdong, China

Correspondence should be addressed to Jin Zhang; zhangj4@sustc.edu.cn

Received 2 December 2017; Revised 18 February 2018; Accepted 28 March 2018; Published 6 May 2018

Academic Editor: Dingchang Zheng

Copyright © 2018 Linfei Ge et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Respiration monitoring is helpful in disease prevention and diagnosis. Traditional respiration monitoring requires users to wear devices on their bodies, which is inconvenient for them. In this paper, we aim to design a noncontact respiration rate detection system utilizing off-the-shelf smartphones. We utilize the single-frequency ultrasound as the media to detect the respiration activity. By analyzing the ultrasound signals received by the built-in microphone sensor in a smartphone, our system can derive the respiration rate of the user. The advantage of our method is that the transmitted signal is easy to generate and the signal analysis is simple, which has lower power consumption and thus is suitable for long-term monitoring in daily life. The experimental result shows that our system can achieve accurate respiration rate estimation under various scenarios.

1. Introduction

Respiration is one of the most important vital signs for human beings. Many chronic diseases can be detected through respiration monitoring, such as asthma and chronic obstructive pulmonary diseases [1, 2]. Respiration monitoring can also reflect the users' sleep and emotion status. Respiration rate detection is a key function in respiration monitoring [3]. Most existing respiration rate detection devices are wearable, for example, respiration belts [4] and oronasal airflow meters [5]. Users need to wear these devices on their chests or faces during respiration monitoring, which is extremely inconvenient. Therefore, these devices are mainly applicable for medical examination for disease diagnosis, but they are not suitable for long-term everyday monitoring.

To overcome the disadvantage of the wearable devices, recently, researchers started investigating noncontact respiration monitoring methods. Some works use wireless signals to detect the respiration rate. For example, Abdelnasser et al. leveraged the WiFi signals [6–10], Lazaro et al. used the UWB signals [11, 12], and Rahman et al. used a microwave radar to detect respiration [13–15]. These systems require deploying extra wireless transceivers to transmit and receive

wireless signals, which makes the system expensive. Instead of specially deployed wireless transceivers, some researchers proposed using smartphones to detect vital signs, which is easier to access in daily life. Some works used built-in inertial sensors in smartphones to monitor vital signs [16–19], while others leveraged ultrasonic signals to conduct respiration detection and sleep monitoring [20, 21]; they utilized built-in speakers and microphones in smartphones to play and record ultrasound signals and extract useful information such as respiration patterns from them [22]. In our paper, we will also use ultrasound as the media to detect the respiration rate. Actually, ultrasound is a proper medium for vital sign detection, as it can be easily generated by smartphones, which are taken along by people all the time and are suitable for long-term monitoring. Besides, ultrasound is a mechanical wave; therefore, users do not need to worry about the electromagnetic radiation for long-term monitoring. Because of the above advantages, ultrasound is also considered to be used for other applications, such as activity recognition and human computer interaction [23–25]. However, in existing ultrasound-based monitoring systems, complicated signals such as frequency modulated continuous wave (FMCW) or orthogonal frequency division

multiplexing (OFDM) are used as the carrier to conduct object detection, which requires complicated modulation and demodulation signal processing modules. The complexity in signal generation and signal processing leads to large power consumption, which is not suitable for power-limited mobile devices such as smartphones.

In this paper, to overcome the above-mentioned limitation of existing solutions, we propose a smartphone-based respiration rate detection system using a single-frequency ultrasonic signal, which enables simple signal processing and low power consumption. We observed that when a single-frequency cosine ultrasonic signal is transmitted by the speaker and the signals reflected by the human being are collected using the built-in microphone, the signal strength of the received signal can reflect the breath pattern of the person being detected. The reason is that the received signal is a combination of the reflected signal and the line-of-sight signal. The chest movement of the respiration activity changes the distance between the speaker and the microphone and, therefore, changes the phase of the reflected signal, which finally results in the amplitude of the total received signal changing periodically. By analyzing the amplitude of the received signal, the respiration rate can be correctly detected. In our proposed system, to detect the respiration rate, we only need to calculate the amplitude of the signal, without complicated frequency domain analysis. Our method guarantees a high detection accuracy while keeping the analysis relatively simple.

The contributions of our paper are as follows. First, we make an observation that the received signal strength of the single-frequency ultrasound signal can reflect the respiration pattern of the user, by theoretical analysis. Second, we propose a smartphone-based respiration rate detection system utilizing single-frequency ultrasonic signals. We also design a rate detection algorithm to estimate the respiration rate based on the amplitude of the received signal. Third, we implement the system on an Android smartphone and the experimental result shows that our system can achieve accurate respiration rate estimation results under various testing scenarios.

2. Ultrasonic Signal Analysis

In this section, we will analyze the characteristics of the received ultrasonic signal when a single-frequency signal is transmitted by the speaker and reflected by the tester. The analysis result shows that the strength of the received signal reflects the respiration rate; thus, by detecting the signal strength, the respiration rate can be estimated.

2.1. Overview. In this paper, we use the built-in speaker of a smartphone to generate single-frequency ultrasound of 20 kHz. Most off-the-shelf smartphones can generate sound up to 22 kHz using their built-in speakers [20, 21]. The smartphone is placed in front of the tester. The ultrasound signal is reflected by the human body and captured by the built-in microphone of the smartphone. The received signal is mainly composed of two parts. One is the signal directly propagated from the speaker to the microphone. The other

is the part that is reflected by the user's moving chest. These two signals have a superposition at the receiving end. Because of the movements of the chest while breathing, the received signal varies. In the following subsection, we will derive the received signal strength of the composed signal. We observed that the signal strength varies along with the respiration. From the amplitude of the received signal, we can extract the respiration waveform and estimate the respiration rate.

2.2. Receiving Signal Analysis. In our system, we use the speaker and microphone of a smartphone to transmit and receive ultrasonic signals. At the transmitter end, the speaker emits a single-frequency cosine signal

$$S(t) = A \cos(2\pi f_s t), \quad (1)$$

where A is the amplitude and f_s is the frequency of the generated ultrasound signal.

The received signal is a superposition of two components: the static signal which propagates directly from the speaker to the microphone and is reflected by the static reflectors and the dynamic one which is caused by the movements of the tester's chest.

For the static component, it contains the line-of-sight signal directly transmitted from the speaker to the microphone and the signals reflected by static objects around. The static reflectors only change the phase of the signal without changing its frequency; therefore, the static component, which is the sum of all static rays, can be written as follows:

$$R_s(t) = A_1 \cos(2\pi f_s t + \phi_s), \quad (2)$$

where A_1 , f_s are the amplitude and frequency of the sound signal, respectively, and ϕ_s is a constant phase change.

For the dynamic component, the periodical movement of the chest causes a periodical distance change between the smartphone and the chest. Thus, the propagation delay varies because of the chest motion. Therefore, the phase also changes periodically. So, the dynamic component can be written as follows:

$$R_d(t) = A_2 \cos(2\pi f_s t + \phi_d(t)), \quad (3)$$

where A_2 , f_s are the amplitude and frequency of the sound signal. $\phi_d(t)$ is the periodical phase change caused by the chest movement, which is written as

$$\phi_d(t) = \frac{2\pi f_s (D_0 + D \cos(2\pi f_b t))}{v}, \quad (4)$$

where D_0 is the constant distance of the propagation path, D and f_b are the amplitude and frequency of the chest movement while breathing, and v is the speed of sound. Here, $D_0 + D \cos(2\pi f_b t)$ is the distance between the smartphone and the chest while breathing, and $(D_0 + D \cos(2\pi f_b t))/v$ is the propagation delay.

At the receiver end, the received signal is the superposition of the static component and dynamic component:

$$R(t) = R_s(t) + R_d(t) = A_r(t) \cos(2\pi f_s t + \varphi), \quad (5)$$

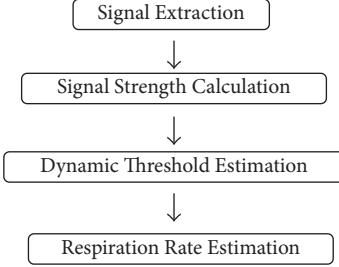


FIGURE 1: Algorithm overview.

where

$$A_r(t) = \sqrt{A_1^2 + A_2^2 + 2A_1A_2 \cos(\phi_s - \phi_d(t))}, \quad (6)$$

$$\tan \varphi = \frac{A_1 \cos \phi_s + A_2 \cos \phi_d(t)}{A_1 \sin \phi_s + A_2 \sin \phi_d(t)}. \quad (7)$$

Bringing (4) into (6), we have

$$\begin{aligned} A_r(t) &= \sqrt{A_1^2 + A_2^2 + 2A_1A_2 \cos\left(\phi_s - \frac{2\pi f_s(D_0 + D \cos(2\pi f_b t))}{\nu}\right)} \\ &= \sqrt{A_1^2 + A_2^2 + 2A_1A_2 \cos\left(\phi_s - \frac{2\pi f_s D_0}{\nu} - \frac{2\pi f_s D}{\nu} \cos(2\pi f_b t)\right)}. \end{aligned} \quad (8)$$

In the above expression of $A_r(t)$, $\cos(2\pi f_b t)$ varies with t at a frequency of f_b , which is the breathing rate. So, the amplitude of the received signal $A_r(t)$ varies periodically at the same frequency with breath. Therefore, the signal strength, which is the square of the amplitude, follows the same changing pattern. Therefore, by detecting the received signal strength, we can estimate the chest movement of a person.

3. Respiration Rate Estimation Algorithm

In this section, we present our respiration rate estimation algorithm based on single-frequency ultrasound signals. We aim to run our algorithm and make it work well on smartphones. Considering the limited resources and power supply of smartphones, we try our best to reduce the complexity of our algorithm. We are trying to find the most effective way instead of the most accurate way for signal processing. Our system is composed of four stages: signal extraction, signal strength calculation, dynamic threshold estimation, and respiration rate estimation, as shown in Figure 1.

In the first stage, the smartphone generates an inaudible 20 kHz ultrasound signal, plays it with the built-in speaker, and records the signal using the microphone. The recording process is to sample the received sound signal $R(t)$ with sampling rate F_s , which achieves a discrete signal:

$$X_n = R\left(\frac{n}{F_s}\right), \quad n = 1, 2, 3, \dots \quad (9)$$

In the second stage, we calculate the received signal strength. From the analysis in Section 2, we know that the

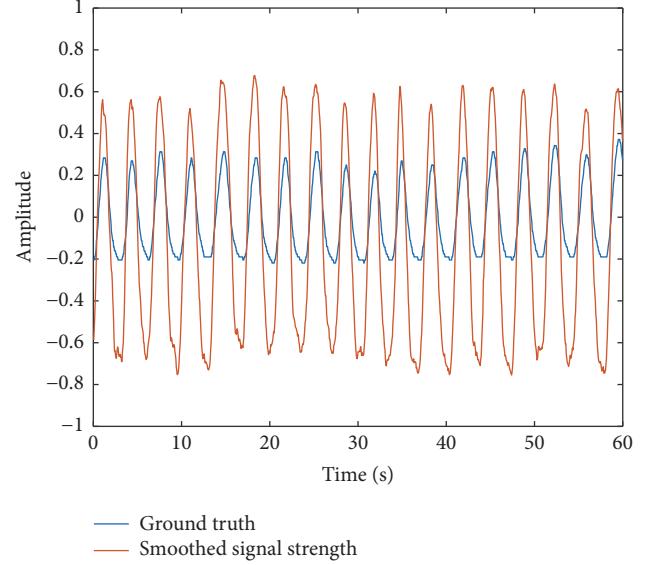


FIGURE 2: Smoothed signal strength compared with ground truth.

received signal strength changes at the same frequency as the breath. We can estimate the respiration rate based on the received signal strength P_m . Although the audio file is sampled with a high sampling frequency F_s , the received signal strength P_m can be calculated at relatively low frequency F_p , where $F_p = F_s/K$ and K is the coefficient that is used to reduce the sampling rate of signal strength. The signal strength P_m can be defined as the average of the signal strength of K samples in X_n :

$$P_m = \frac{1}{K} \sum_{n=(m-1)*K+1}^{m*K} X_n^2, \quad m = 1, 2, 3, \dots \quad (10)$$

Then, we smooth the received signal strength using a moving average filter with a window size of W points. The smoothed signal strength \bar{P}_m is given by

$$\bar{P}_m = \frac{1}{W} \sum_{i=\max[i-W, 1]}^m P_i \quad (m = 1, 2, 3, \dots). \quad (11)$$

Figure 2 shows the smoothed received signal strength \bar{P}_m as well as the ground truth achieved by respiration belt SCHUHFRIED Biofeedback Xpert [26]. The result shows that the smoothed signal strength matches the ground truth well.

With the estimated respiration waveform, we can derive the respiration rate of the tester. To simplify the algorithm, instead of frequency domain analysis, we want to use simple time domain analysis to detect the respiration rate by counting the number of peaks and valleys in the strength signal. To accurately count the peaks and valleys, a threshold is required, and by comparing the signal strength with the threshold, we can get the respiration period and respiration rate. However, in some cases, the signal strength may fluctuate severely because of the changing environment, just as Figure 3 shows. To solve this problem, we conduct dynamic threshold estimation at the third stage before calculating the

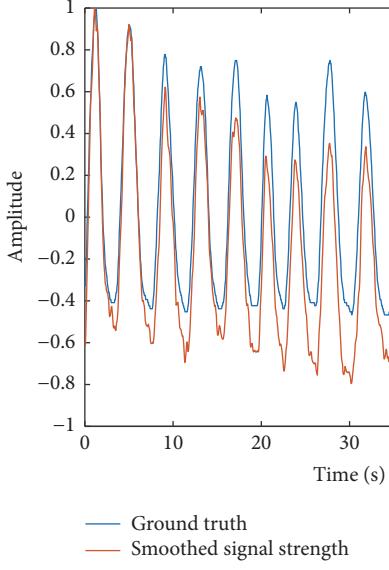


FIGURE 3: Signal strength fluctuation caused by the environment while testing.

respiration rate in the fourth stage. We calculate the dynamic threshold P_{th} by averaging W_{th} continuous points in \bar{P}_m .

$$P_{\text{th}}(m) = \frac{1}{W_{\text{th}}} \sum_{i=\lfloor m/W_{\text{th}} \rfloor * W_{\text{th}} + 1}^{\lfloor (m/W_{\text{th}}) + 1 \rfloor * W_{\text{th}}} \bar{P}_i, \quad m = 1, 2, 3, \dots \quad (12)$$

The dynamic threshold P_{th} varies as the signal strength \bar{P}_m changes.

At the last stage, we use the smoothed signal strength \bar{P}_m and the dynamic threshold P_{th} to estimate the respiration rate of the tester. We record the times that the estimated respiration waveform passes through the threshold, and then we get the period of the breathing. Using the average value of several recent periods, we can get the estimated respiration rate. Using this algorithm, we get Figure 4. This figure shows that our algorithm works well and gets a mean estimation error of 0.32 bpm in this example. We also test the case of deep breath; the algorithm still performs well as Figure 5 shows. The mean estimation error under deep breath is 0.35 bpm.

4. System Evaluation

4.1. System Implementation and Evaluation Setup. We develop an Android application to implement our algorithm on smartphones and evaluate the performance under various scenarios. We conduct experiments on two smartphones, a Xiaomi MI5 and a Samsung Galaxy S4, which are both based on the Android OS. Our application is developed with a minimum version of Android 4.0.4 and it works well on both smartphones. In our system, to generate ultrasound, we first generate a sound file in Pulse Code Modulation (PCM) format, and then we use AudioTrack in Android to play the generated sound file. The main speaker of the smartphone is used to transmit the ultrasound signal. In the receiving end, we use one microphone to receive the reflected signal, which

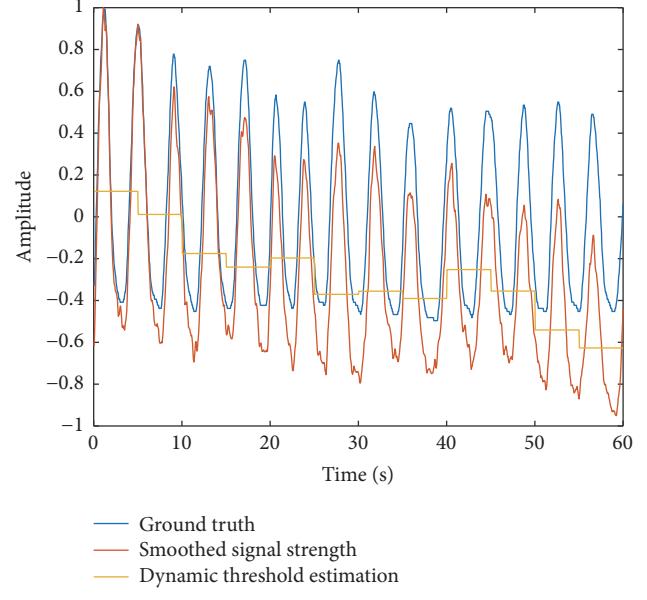


FIGURE 4: Dynamic threshold estimation.

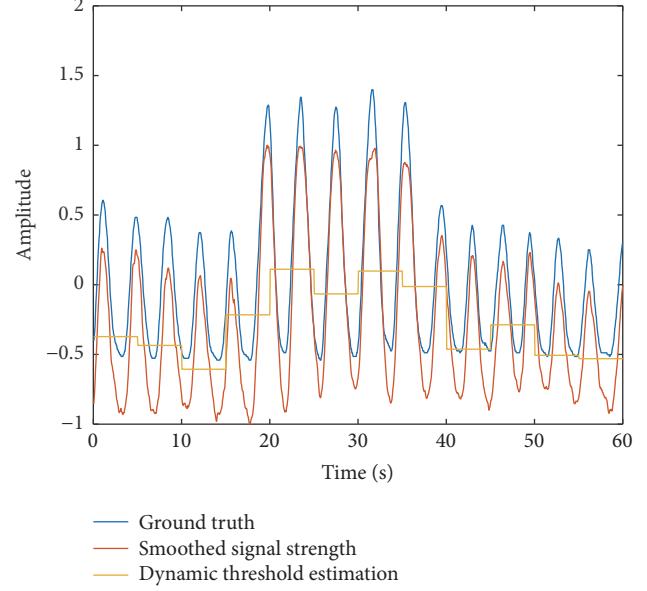


FIGURE 5: Dynamic threshold estimation under deep breath.

works on the mono record mode instead of the stereo mode to reduce the complexity of the computation. The ultrasound signal is at the frequency of 20 kHz. The sampling rate of the microphone is 48 kHz which is the maximum sampling rate that most smartphones can support. To achieve the ground truth, we use the SCHUHFRIED Biofeedback Xpert [26] to monitor the respiration rate of the testers. Figure 6 shows the experimental scenario in the office environment.

In the remaining part of this section, without specific instruction, the parameters are set as follows: sampling rate of the smartphone $F_s = 48$ kHz and coefficient $K = 2400$. Thus, the sampling rate of P_m is at a frequency of



FIGURE 6: Experimental scenario in the office.

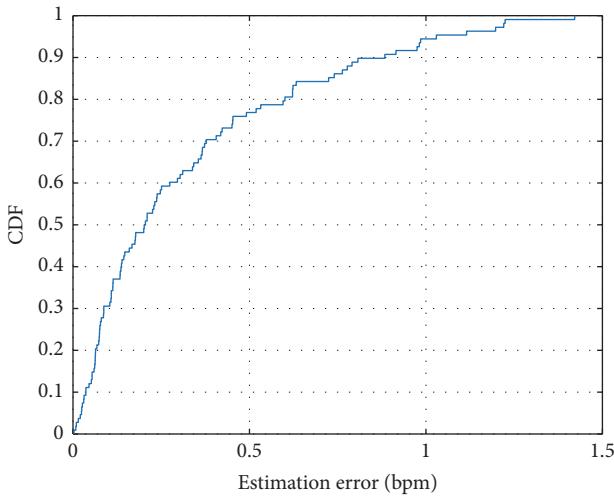


FIGURE 7: CDF of the estimation error.

$F_p = F_s/K = 20$ Hz. The window size of moving average filter $W = 5$. The window size of dynamic threshold estimation $W_{\text{th}} = 100$.

4.2. Evaluation Results. We evaluate the overall performance of the system and calculate the cumulative distribution functions (CDF) of estimation errors on the respiration rate, when the distance between the smartphone and the tester is 15 cm. As Figure 7 shows, over 90% of the results have an estimation error under 0.8 bpm. The medium estimation error is 0.2101 bpm while the mean estimation error is 0.4137 bpm. Traditional medical respiration monitoring devices usually allow an error of 5%, considering that a normal breath rate is usually 15–20 bpm, and our estimation error is less than 4%, which is accurate enough for daily use.

Figure 8 shows the respiration rate comparison for a test that lasts for 30 minutes. From this figure, we find that the estimated breathing rate follows the ground truth well. Due to the convenience of the smartphones, they are suitable for our algorithm to do a long-term monitoring.

Compared with Wang et al.'s work [21], we get similar results in estimation error with a much simpler algorithm.

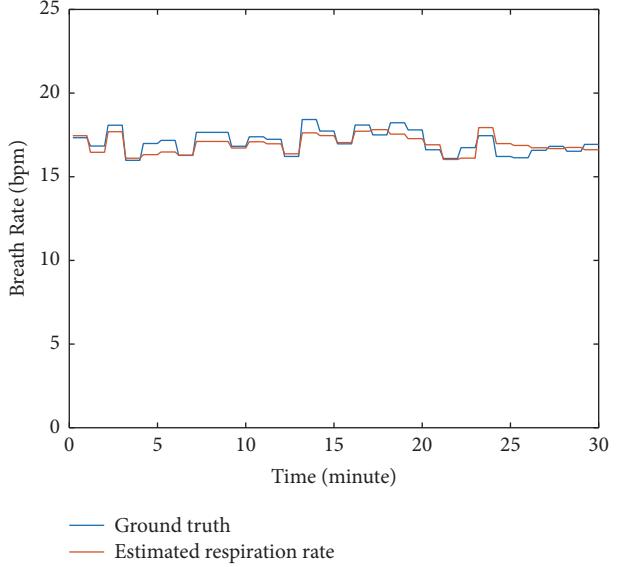


FIGURE 8: Respiration rate estimation result for 30-minute testing.

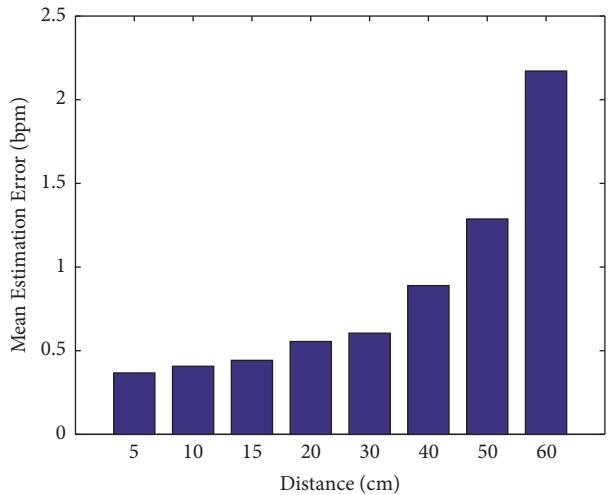


FIGURE 9: Estimation error versus the distance between the user and the smartphone.

They employed sonar phase data to get the breath rate, resulting in high complexity in algorithm. In their work, the estimation error at a distance of 10 cm is about 0.3 bpm, which is very close to our results, but we achieve similar results using a much simpler algorithm.

4.3. Impact of Different Factors

Impact of Distance. Figure 9 shows the impact of distance between the microphone and people's chest. When the distance gets longer, the mean error gets larger. With the increase of distance, the signal attenuation becomes severe and the amount of reflected ultrasound signal becomes less. Thus, the energy change caused by the reflected ultrasound signal becomes more inconspicuous; therefore, the estimation accuracy will be reduced. We can see that our algorithm

TABLE 1: Estimation error versus the orientation of the smartphone.

Orientation (degree)	Mean estimation error (bpm)
0	0.39
45	0.50
90	0.62

TABLE 2: Estimation errors on different persons.

Person number	Mean estimation error (bpm)
1	0.41
2	0.45
3	0.42

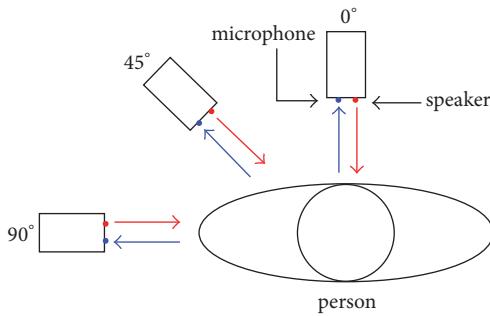


FIGURE 10: Orientation of the smartphone.

works well when the distance is under 40 cm. The mean estimation error is around 0.5 bpm under a distance of 30 cm and it is acceptable for daily use. When the distance is 40 cm, the mean estimation error is almost 1 bpm.

Impact of Orientation. Table 1 shows how the orientation influences the mean estimation error. In this experiment, we keep the distance at 10 cm and control the angle between the smartphone and the tester as Figure 10 shows. We consider three cases when the angle is equal to 0, 45, and 90 degrees, respectively. At 0 degrees, the person directly faces the smartphone, and we get the best performance at a mean estimation error of 0.39 bpm. At 45 degrees, the chest movements become less clear compared with 0 degrees. So, we get a worse mean estimation error at 0.50 bpm. At 90 degrees, the estimation error of 0.62 bpm is the worst due to the minimum chest movements in these three situations. Despite being much worse than the situation of facing the smartphone, the results at 90 degrees are still acceptable. The system can work in all orientations, because during breathing, the chest moves at both the front and side directions. However, the performance is optimal when the tester directly faces the smartphone.

Results on Different Persons. Table 2 shows whether the tester influences the estimation error a lot. In the experiment, we test three persons (two men and one woman) as shown in Table 3. The experiment is done at 10 cm distance in the office. We can see that there are small differences between

different persons. These small differences may be caused by environmental noise, different chest movement length, and experimental error. The mean estimation errors of these three persons are all between 0.4 and 0.5 bpm, which is an acceptable result. Thus, our algorithm works well on different persons.

Impact of Different Smartphones. In this experiment, we use two smartphones, Xiaomi MI5 and Samsung Galaxy S4. From Table 4, we can see that the MI5 has a smaller mean estimation error than the S4. This is because the structures of the two smartphones are slightly different. For MI5, the speaker and microphone are both at the bottom of the phone. This means that when you put it on the desk, the speaker and microphone are directly pointing to the chest of the person. In contrast, the speaker of the S4 is on its back and its microphone is at the bottom. Thus, when we put the S4 on the desk, the volume may be reduced due to the position of the speaker. So, the received signal of MI5 is stronger than that of S4 and MI5 achieves a better performance in the experiment. We now know that the position of the speaker and the microphone does matter. A smartphone with its speaker and microphone at the bottom usually yields a better result. However, even when the speaker is at a relatively bad position, the performance is still acceptable.

Impact of Various Testing Scenarios. Table 5 shows the impact of different scenarios. We test four scenarios including office, dormitory, library, and office with music playing. They achieved mean estimation errors of 0.48, 0.43, 0.31, and 0.78 bpm, respectively. We get the minimum mean estimation error at the library, because in the library, it is quiet and there are a few objects around, which results in the minimum impact of the environment, and thus it achieves the minimum mean estimation error. In the dormitory and office, the situation is similar: crowded room with computer and air conditioners running, even some people talking with others. The noise generated from machines and people does influence the performance of our experiment. Because our algorithm is based on the energy of the signal, a loud noise may cover the signal that we want, resulting in bad performance. We further run experiments in the office with music playing to verify the impact of noise. The mean estimation error with music is 0.78 bpm, which is much larger than the general situation. This shows that a loud noise does make the performance worse. Furthermore, the music has a more severe impact than people's talking. That is because the frequency domain of music is relatively higher than people's voice; therefore, music has a higher influence on the 20 kHz signal that we use to monitor the breath. We did not test our system under the scenario when the tester is running, because all existing works [17, 21] conduct their experiments in a stable scenario. The experiments validate that, even in noisy scenarios, we can get mean estimation errors less than 1 bpm. In general scenarios, the mean estimation error is about 0.5 bpm, which is accurate enough for daily use.

Impact of Different Parameters. Figures 11 and 12 show the reason why we choose the sample frequency of signal strength

TABLE 3: Subjects information.

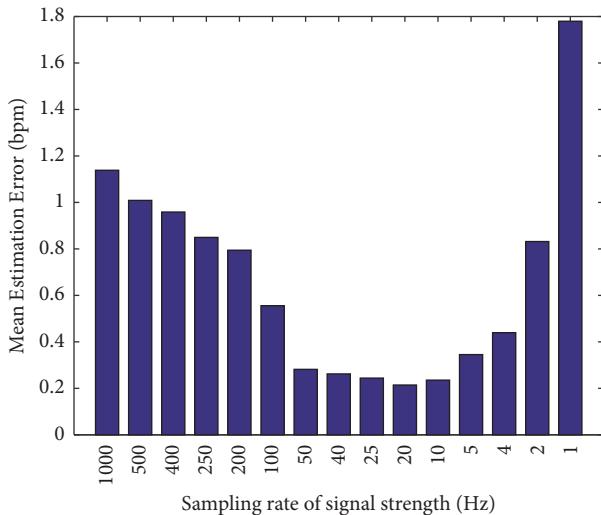
Person number	Gender	Age	Height (cm)	Weight (kg)
1	Male	23	172	65
2	Male	22	175	50
3	Female	22	160	50

TABLE 4: Estimation errors on different smartphones.

Phone	Mean estimation error (bpm)
Xiaomi MI5	0.37
Samsung Galaxy S4	0.41

TABLE 5: Estimation errors in different testing scenarios.

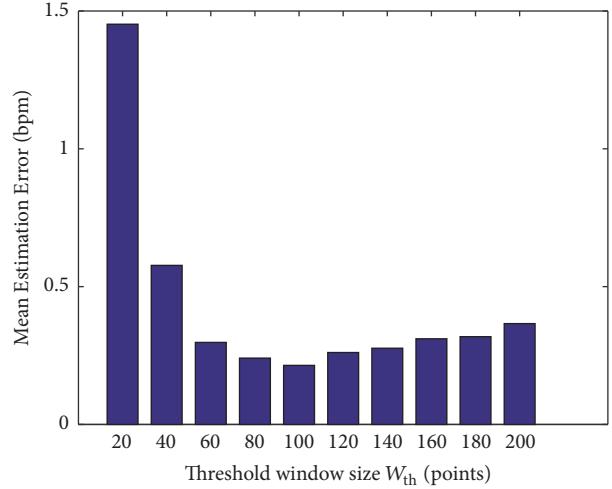
Scenarios	Mean estimation error (bpm)
Office	0.49
Dormitory	0.44
Library	0.31
Office (with music)	0.78

FIGURE 11: Mean estimation error versus sampling rate of signal strength F_p .

F_p equal to 20 Hz and the moving window size W_{th} equal to 100 points in our algorithm. We tried different parameters in our algorithm. For sampling rate F_p , we get the minimum mean estimation error at 20 Hz. Although it is acceptable from 5 Hz to 50 Hz, we choose 20 Hz because the medical device SCHUHFRIED Biofeedback Xpert [26], which we use as our ground truth, also works at a sampling rate of 20 Hz. And for window size W_{th} of dynamic threshold estimation, we get the minimum estimation error at $W_{th} = 100$ points.

5. Conclusion

In this paper, we proposed a smartphone-based respiration rate detection system based on single-frequency ultrasound signals. The proposed system can track the movement of

FIGURE 12: Mean estimation error versus threshold window size W_{th} .

the human chest by observing the signal strength of the recorded ultrasound data. We implemented our system on an Android smartphone and conducted extensive experiments to show the feasibility and accuracy of our system. The results show that this system can achieve accurate respiration rate estimation under various scenarios.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant no. 61701216, by Guangdong Innovative and Entrepreneurial Research Team Program under contract no. 2016ZT06G587, and by Shenzhen Science, Technology and Innovation Commission Basic Research Project under Grants nos. JCYJ 20160531190935987 and JCYJ 20160531191011045.

References

- [1] W. Website, "Who chronic respiratory diseases (crds)," <http://www.who.int/respiratory/en/>.
- [2] T. Young, M. Palta, J. Dempsey, J. Skatrud, S. Weber, and S. Badr, "The occurrence of sleep-disordered breathing among middle-aged adults," *The New England Journal of Medicine*, vol. 328, no. 17, pp. 1230–1235, 1993.

- [3] R. B. Berry, R. Brooks, C. E. Gamaldo, S. M. Harding, C. Marcus, and B. Vaughn, "The aasm manual for the scoring of sleep and associated events," in *Rules, Terminology and Technical Specifications*, American Academy of Sleep Medicine, Darien, Ill, USA.
- [4] V. respiration monitor belt, Vernier software and technology vernier," <http://www.vernier.com/>.
- [5] O. P. A. nasal mask, Af531 oro-nasal mask philips," <http://www.philips.com.hk/healthcare/product/HCNOCTN102/respironics-af531-patient-interface>.
- [6] H. Abdelnasser, K. A. Harras, and M. Youssef, "Ubibreathe: A ubiquitous non-invasive wifi-based breathing estimator," in *Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, pp. 277–286, ACM, Hangzhou, China, June 2015.
- [7] H. Wang, D. Zhang, J. Ma et al., "Human respiration detection with commodity wifi devices," in *Proceedings of the the 2016 ACM International Joint Conference*, pp. 25–36, Heidelberg, Germany, September 2016.
- [8] X. Wang, C. Yang, and S. Mao, "PhaseBeat: Exploiting CSI Phase Data for Vital Sign Monitoring with Commodity WiFi Devices," in *Proceedings of the 37th IEEE International Conference on Distributed Computing Systems, ICDCS 2017*, pp. 1230–1239, USA, June 2017.
- [9] J. Liu, Y. Wang, Y. Chen, J. Yang, X. Chen, and J. Cheng, "Tracking Vital Signs During Sleep Leveraging Off-the-shelf WiFi," in *Proceedings of the the 16th ACM International Symposium*, pp. 267–276, Hangzhou, China, June 2015.
- [10] X. Wang, C. Yang, and S. Mao, "ResBeat: Resilient Breathing Beats Monitoring with Realtime Bimodal CSI Data," in *Proceedings of the 2017 IEEE Global Communications Conference (GLOBECOM 2017)*, pp. 1–6, Singapore, December 2017.
- [11] A. Lazaro, D. Girbau, and R. Villarino, "Analysis of vital signs monitoring using an IR-UWB radar," *Progress in Electromagnetics Research*, vol. 100, pp. 265–284, 2010.
- [12] J. Salmi and A. F. Molisch, "Propagation parameter estimation, modeling and measurements for ultrawideband mimo radar," *Institute of Electrical and Electronics Engineers. Transactions on Antennas and Propagation*, vol. 59, no. 11, pp. 4257–4267, 2011.
- [13] T. Rahman, A. T. Adams, R. V. Ravichandran et al., "DoppleSleep," in *Proceedings of the the 2015 ACM International Joint Conference*, pp. 39–50, Osaka, Japan, September 2015.
- [14] S. Suzuki, T. Matsui, H. Imuta et al., "A novel autonomic activation measurement method for stress monitoring: non-contact measurement of heart rate variability using a compact microwave radar," *Medical & biological engineering & computing*, vol. 46, no. 7, pp. 709–714, 2008.
- [15] G. G. Senaratne, R. B. Keam, W. L. Sweatman, and G. C. Wake, "Solutions of inverse problems with potential application for breast tumour detection using microwave measurements," *Computational and Mathematical Methods in Medicine. An Interdisciplinary Journal of Mathematical, Theoretical and Clinical Aspects of Medicine*, vol. 8, no. 4, pp. 245–261, 2007.
- [16] H. Aly and M. Youssef, "Zephyr demo: Ubiquitous accurate multi-sensor fusion-based respiratory rate estimation using smartphones," in *Proceedings of the IEEE INFOCOM 2016 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 1-2, San Francisco, CA, USA, April 2016.
- [17] Y. Ren, C. Wang, J. Yang, and Y. Chen, "Fine-grained sleep monitoring: Hearing your breathing with smartphones," in *Proceedings of the 34th IEEE Annual Conference on Computer Communications and Networks, IEEE INFOCOM 2015*, pp. 1194–1202, Hong Kong, May 2015.
- [18] R.-C. Peng, X.-L. Zhou, W.-H. Lin, and Y.-T. Zhang, "Extraction of heart rate variability from smartphone photoplethysmograms," *Computational and Mathematical Methods in Medicine*, Article ID 516826, 11 pages, 2015.
- [19] L. J. Mena, V. G. Felix, R. Ostos et al., "Mobile personal health system for ambulatory blood pressure monitoring," *Computational and Mathematical Methods in Medicine*, vol. 2013, Article ID 598196, 2013.
- [20] R. Nandakumar, S. Gollakota, and N. Watson, "Contactless Sleep Apnea Detection on Smartphones," in *Proceedings of the the 13th Annual International Conference*, pp. 45–57, Florence, Italy, May 2015.
- [21] X. Wang, R. Huang, and S. Mao, "SonarBeat: Sonar phase for breathing beat monitoring with smartphones," in *Proceedings of the 26th International Conference on Computer Communications and Networks, ICCCN 2017*, Canada, August 2017.
- [22] W. Wang, A. X. Liu, and K. Sun, "Device-free gesture tracking using acoustic signals," in *Proceedings of the the 22nd Annual International Conference on Mobile Computing and Networking, MobiCom 2016*, pp. 82–49, ACM, New York City, New York, October 2016.
- [23] W. Ruan, Q. Z. Sheng, L. Yang, T. Gu, P. Xu, and L. Shangguan, "Audiogest: enabling fine-grained hand gesture detection by decoding echo signal," in *Proceedings of the the 2016 ACM International Joint Conference*, pp. 474–485, Heidelberg, Germany, September 2016.
- [24] W. Mao, J. He, and L. Qiu, "CAT: High-precision acoustic motion tracking," in *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking, MobiCom 2016*, pp. 69–81, New York, NY, USA, October 2016.
- [25] R. Nandakumar, V. Iyer, D. Tan, and S. Gollakota, "Fingerio: Using active sonar for fine-grained finger tracking," in *Proceedings of the the 2016 CHI Conference*, pp. 1515–1525, Santa Clara, California, USA, May 2016.
- [26] S. GmbH, Schuhfried - biofeedback xpert, <https://www.schuhfried.com/biofeedback/biofeedback-xpert/>.

Research Article

Does the Temporal Asymmetry of Short-Term Heart Rate Variability Change during Regular Walking? A Pilot Study of Healthy Young Subjects

Xinpei Wang,¹ Chang Yan,¹ Bo Shi,² Changchun Liu,¹ Chandan Karmakar,³ and Peng Li¹ 

¹School of Control Science and Engineering, Shandong University, Jinan, Shandong 250061, China

²Department of Medical Imaging, Bengbu Medical College, Bengbu, Anhui 233030, China

³School of Information Technology, Deakin University, Burwood, VIC 3125, Australia

Correspondence should be addressed to Peng Li; pli@sdu.edu.cn

Received 1 December 2017; Revised 21 February 2018; Accepted 21 March 2018; Published 30 April 2018

Academic Editor: Andrzej Kloczkowski

Copyright © 2018 Xinpei Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The acceleration and deceleration patterns in heartbeat fluctuations distribute asymmetrically, which is known as heart rate asymmetry (HRA). It is hypothesized that HRA reflects the balancing regulation of the sympathetic and parasympathetic nervous systems. This study was designed to examine whether altered autonomic balance during exercise can lead to HRA changes. Sixteen healthy college students were enrolled, and each student undertook two 5-min ECG measurements: one in a resting seated position and another while walking on a treadmill at a regular speed of 5 km/h. The two measurements were conducted in a randomized order, and a 30-min rest was required between them. RR interval time series were extracted from the 5-min ECG data, and HRA (short-term) was estimated using four established metrics, that is, Porta's index (PI), Guzik's index (GI), slope index (SI), and area index (AI), from both raw RR interval time series and the time series after wavelet detrending that removes the low-frequency component of ~ 0.03 Hz. Our pilot data showed a reduced PI but unchanged GI, SI, and AI during walking compared to resting seated position based on the raw data. Based on the wavelet-detrended data, reduced PI, SI, and AI were observed while GI still showed no significant changes. The reduced PI during walking based on both raw and detrended data which suggests less short-term HRA may underline the belief that vagal tone is withdrawn during low-intensity exercise. GI may not be sensitive to short-term HRA. The reduced SI and AI based on detrended data suggest that they may capture both short- and long-term HRA features and that the expected change in short-term HRA is amplified after removing the trend that is supposed to link to long-term component. Further studies with more subjects and longer measurements are warranted to validate our observations and to examine these additional hypotheses.

1. Introduction

Under healthy physiological conditions, the human heart does not beat at a constant frequency; instead, heart rate changes all the time. This phenomenon has been recognized as heart rate variability (HRV) [1, 2]. For a given observation scale, the acceleration and deceleration patterns in beat-to-beat heart rate fluctuations distribute asymmetrically rather than contribute equally to HRV [3–7]. This suggests that the underlying heart rate control mechanisms—the regulation of sympathetic and parasympathetic nervous systems—are physiologically disproportionate over fixed temporal scales [8–12]. This asymmetry of acceleration

and deceleration runs is defined as heart rate asymmetry (HRA).

In clinical settings, the electrocardiograms (ECGs) are commonly collected under well-controlled conditions such as resting supine or seated position and within a short time range (e.g., 5 min or shorter). Increasing attention nowadays has been drawn to the ambulatory ECG monitoring [13], which facilitates the tracking of heart rate and HRV with activities of free living, such as walking and exercise [14]. Long-term ambulatory measurement also assists to examine whether and how HRV properties respond to these daily activities [15]. Besides, daily activities may also evoke changes that may mask the effects of interest, for example, the changes

that are related to alterations of health status or different times of the day. Thus, the examination of the changes of different HRV measures with daily activities may help better understand the variation profile of these measures, providing opportunities to comprehend the knowledge of how these novel properties respond to the changing physiological conditions that eventually should be of great help to develop sensitive and specific makers for cardiovascular diseases. With such a motivation, this study focused on elucidating whether and how the daily activities alter HRA.

The high-frequency power of HRV is accepted to be related to the parasympathetic tone while HRA has shown to be positively correlated with the high-frequency power [16], offering the link between HRA and parasympathetic activity. This link has further been strengthened by the observations that parasympathetic block leads to less prevalence of HRA [16] and that the deceleration patterns have a larger contribution to short-term HRA than acceleration patterns [9, 12]. Based on these existing results, we expect to see a significantly reduced short-term HRA level during low-intensity daily exercises that are assumed to be accompanied with the withdrawal of parasympathetic modulation [17]. In the current study, we applied treadmill-based regular walking protocol to imitate daily exercises in laboratory. To examine the within-subject changes, each participant undertook a walking protocol and a rest protocol. During each protocol, ECG data were collected continuously for 5 minutes. The next section explains in detail the subjects, experimental protocols, and analysis methods. Experimental results are summarized in the Results, followed by discussions in the Discussions.

2. Methods

2.1. Subjects. Subjects include 16 college students (4 females, 12 males; age: 20.1 ± 0.6 years [mean \pm standard deviation]) with their physical and mental health status confirmed by questionnaire on the history of cardiovascular diseases, diabetes, depression, and neurological disorders. No subject has been taking any medications that have known effects on ANS within two weeks before participation. Adequate sleep during the night before coming to the laboratory, as well as avoidance of vigorous exercises during the test day and the day before, was requested. Written informed consent was obtained from all subjects. The study was approved by the Ethics Committee in Clinical Study of Bengbu Medical College.

2.2. Protocols. For each subject, ECG was recorded twice in random order with the subject seating on a chair or walking on a treadmill (ZRII, Reebok, Canton, MA, USA) at a speed of 5 km/h. Both ECGs last for 5 min and a 30-min rest was scheduled between the two measurements. Holter monitors (DiCare-mlCP, DimeTek Digital Medical Tech., Ltd., Shenzhen, China) were used to collect ECG data. The sampling frequency was 200 Hz, and standard unipolar chest lead V5 was applied. All the measurements were undertaken in a quiet, temperature-controlled (23 ± 1 degree Celsius) room.

2.3. Construction of HRV Time Series. ECGs were first subjected to a visual quality inspection assisted by a self-designed MATLAB program with user interface, which confirmed that all recordings were with high signal qualities. A template-matching process was then applied to extract the R peaks [18] followed by a second-round visual inspection for the correction of misidentified peaks and ectopic beats using the same MATLAB program. During this visual inspection, false positive detection was removed while false negatives were filled with the actual location of R peaks read manually from the program. We confirmed that no ectopic beats occurred in those data. HRV time series were finally constructed by the consecutive R-R intervals.

2.4. HRA Metrics. The following four well-established metrics derived from the Poincaré plot were calculated.

2.4.1. Porta's Index (PI). Conceptually, PI renders symmetry when the numbers of points in the two regions in Poincaré plot separated by the line of identity (LI) are the same and renders asymmetry if they differ [19]. Different levels of asymmetry can be estimated by how much the numbers differ. Thus, PI can be calculated by

$$\text{PI} = \frac{a}{m} \times 100, \quad (1)$$

wherein a is the number of points above LI and m the total number of points (points on LI excluded).

2.4.2. Guzik's Index (GI). GI uses the distances between points and LI as a measure to assess whether the contributions of points in the two different regions in Poincaré plot are equal or not [20]. Specifically,

$$\text{GI} = \frac{\sum_{i=1}^a D_i}{\sum_{i=1}^m D_i} \times 100, \quad (2)$$

wherein D_i is the Euclidian distance of point i to LI. For the RR interval time series, the Poincaré plot is actually to plot the current RR interval versus its subsequent interval. Thus, $D_i = |\text{RR}_{i+1} - \text{RR}_i|/\sqrt{2}$.

2.4.3. Slope Index (SI). The average phase angles of points in the two different regions in Poincaré plot are calculated and used to assess the asymmetry [21]. Specifically,

$$\text{SI} = \frac{\sum_{i=1}^a |R\theta_i|}{\sum_{i=1}^m |R\theta_i|} \times 100, \quad (3)$$

wherein $R\theta_i = \pi/4 - \theta_i$. $\theta_i = \text{atan}(\text{RR}_{i+1}/\text{RR}_i)$ is the phase angle of point i and $\pi/4$ is the point angle of LI, that is, $\text{atan}(1)$.

2.4.4. Area Index (AI). The average areas of sectors formed by the points and LI are calculated and used to assess the asymmetry [22]. Specifically,

$$\text{AI} = \frac{\sum_{i=1}^a S_i}{\sum_{i=1}^m S_i} \times 100, \quad (4)$$

wherein $S_i = 1/2 \times R\theta_i \times r^2$ is the area of the sector formed by point i and LI. r is the radius of the sector.

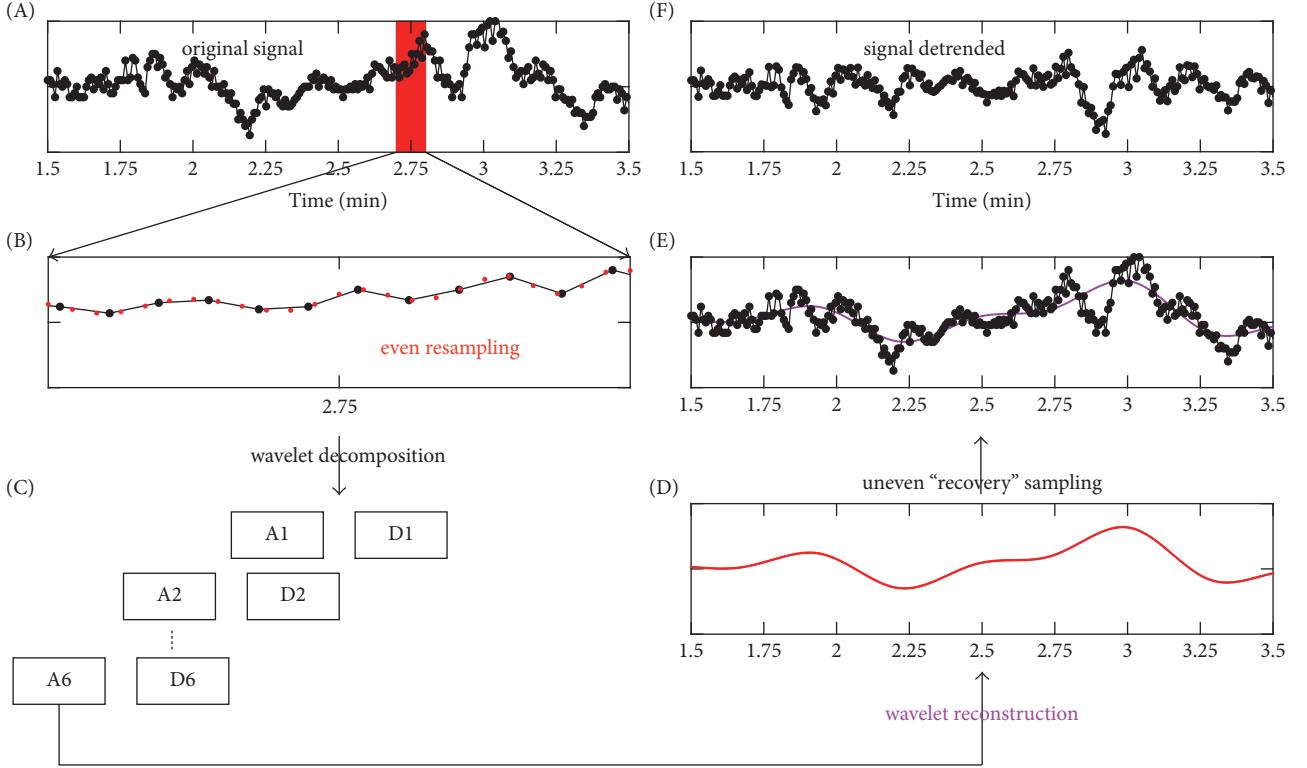


FIGURE 1: Wavelet-based nonstationary trend removal procedure. (A) The original RR interval time series. (B) The RR interval time series after the even resampling. In order to clearly demonstrate the even resampling time points (red dots) and the original time points (black dots), a segment of data from (A) is zoomed in and shown in this panel. (C) The 6-level wavelet decomposition. (D) Trend with even sampling points is obtained from wavelet reconstruction of the approximate coefficients on the 6th level. (E) The actual nonstationary trend (purple) is obtained from the uneven recovering sampling from the trend component in panel (D). (F) The detrended RR interval time series is obtained by subtracting the actual nonstationary trend from the original RR interval time series.

2.5. HRA Analysis of Short-Term HRV. The four HRA metrics were performed on HRV data collected under both conditions. The asymmetry level was further defined as the deviation of a specific HRA metric from its level for completely symmetrical data, that is, $|x - 50|$ (x denotes an HRA metric), and was denoted as ΔPI , ΔGI , ΔSI , and ΔAI , respectively. Besides, to explore the potential effect of nonstationary trend, wavelet detrending was performed and the above four asymmetrical indices were recalculated using the detrended data. To perform the wavelet detrending, raw HRV data were first evenly resampled to 4 Hz by spline interpolation. A 6-level wavelet decomposition using the coif5 wavelet was then conducted. The approximation coefficients on the 6th level were reconstructed to the original scale and were nonevenly “recovered” by spline interpolation which resulted in the trend that would be subtracted. The 6-level decomposition was used so that the frequency band of the trend would be less than ~ 0.03 Hz. Figure 1 intuitively demonstrates this wavelet detrending procedure.

2.6. Statistical Analysis. The Shapiro–Wilk W test suggested nonnormal distribution of all the HRA results. Therefore, the Wilcoxon signed-rank test of each pair was used to examine the within-subject differences under the two measurement conditions. In addition, Cohen’s d static was calculated for

statistically significant observations to examine the effect size of the corresponding metric. A medium effect size was considered if $d \geq 0.5$ and large if $d \geq 0.8$ [23]. As secondary analysis, we also performed the Wilcoxon signed-rank tests by restricting to male subjects ($N = 12$) only. We did not perform these tests separately on females as we only had 4 females. All the statistical analyses were performed using the JMP software (Pro 13, SAS Institute, Cary, NC, USA).

3. Results

A typical RR interval time series for resting seated position and the corresponding RR interval time series from the same subject during walking are shown in Figure 2. Overall, the RR intervals become shorter (i.e., heart beats faster) during walking, such that the points distribute more compactly on the Poincaré plot than those during rest if the same scale is used. The Poincaré plots also become more compact after nonstationary trend removal, which is expected because of the effect of detrending on long-term HRA.

3.1. Asymmetry Based on Raw HRV Time Series. HRV data collected under both conditions displayed asymmetry as assessed by the four HRA metrics (all four p ’s < 0.001 under

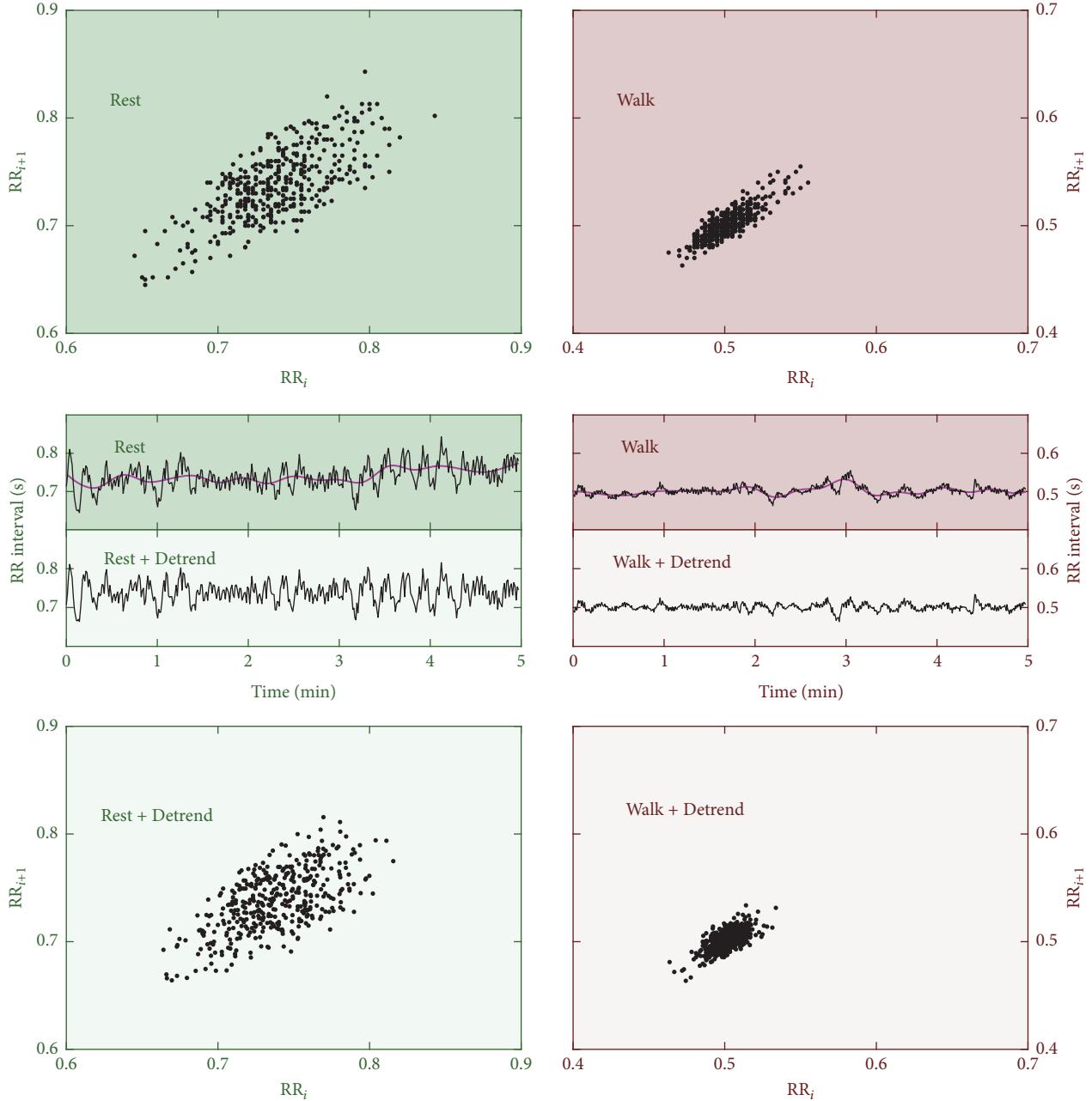


FIGURE 2: Exemplary RR interval time series (middle four panels) and the corresponding Poincaré plots (upper two and lower two panels). Left: data during seated position, right: data during walking. Data after nonstationary trend removal are shown with light-shaded background colors on lower four panels with “+Detrend” legend.

both conditions as revealed by Wilcoxon signed-rank test of each measure versus symmetrical level; i.e., index = 0). Compared to the resting seated position, a significant reduction of HRA during walking was observed by PI (Wilcoxon signed-rank test of each pair: $p = 0.001$; Cohen's $d = 1.0$; out of the 16 subjects, 14 including all the four females showed reduction; Figure 3(A1)). No significant HRA changes during walking were suggested by the remaining three metrics (all p 's > 0.1 ; Figures 3(B1)–3(D1)). The results persisted when restricting the Wilcoxon signed-rank tests to male subjects only (Figures 3(A2)–3(D2)).

3.2. Asymmetry Based on Detrended HRV Time Series. Wavelet detrending did not change the HRA levels significantly under resting seated position (all p 's > 0.05 versus results from raw HRV data as revealed by the Wilcoxon signed-rank test). Similarly, the HRA levels during walking did not show significant changes after wavelet detrending (all p 's > 0.1 for PI, GI, and SI) except that assessed by AI which indicated a significant reduction ($p = 0.04$; 11 out of 16 subjects showed reduced AI after wavelet detrending). As a consequence, AI indicated significantly lower HRA during walking than that under resting seated position ($p = 0.025$; $d = 0.7$; 12 out of 16

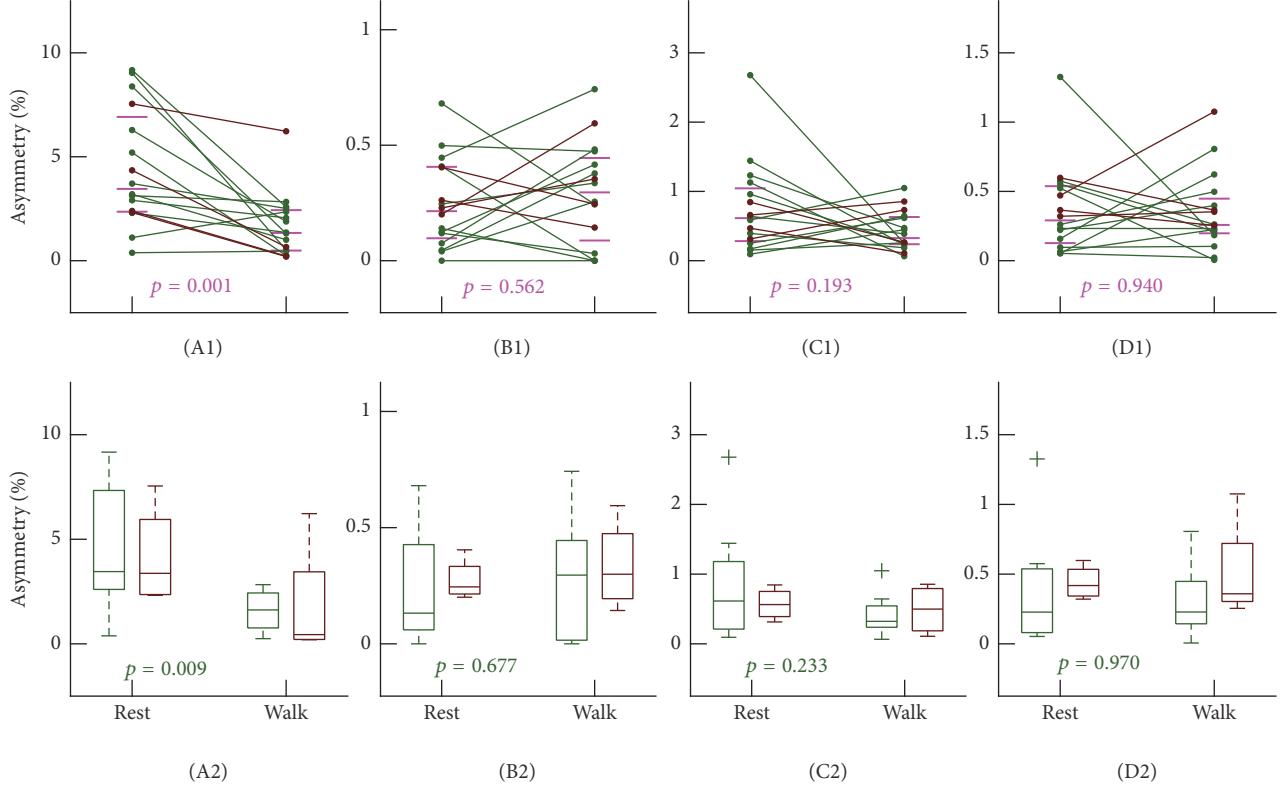


FIGURE 3: The asymmetries of short-term heart rate variability without detrending. In order to show the changes, results from the same individual were connected by lines. Horizontal bars indicate the median and [1st, 3rd] quartiles. p values were from Wilcoxon signed-rank test of each pair. ((A1) and (A2)) Δ PI; ((B1) and (B2)) Δ GI; ((C1) and (C2)) Δ SI; ((D1) and (D2)) Δ AI. Rest: results under resting seated position; Walk: results during regular walking. 14 individuals show reduction from Rest to Walk in (A1). Results from males and from females are marked in different colors (male: green; female: brown). Lower panels summarize the box plots (that show from top to bottom the max, 3rd quartile, median, 1st quartile, and min) for males and females, separately. Outliers, if there are any, are marked by “+.”

subjects showed reduction; Figure 4(D1)). SI also indicated significantly lower HRA during walking ($p = 0.044$; $d = 0.4$; 12 out of 16 subjects showed reduction; Figure 4(C1)). The remaining two metrics showed consistent results as compared with those based on raw HRV data; that is, PI reduced significantly ($p = 0.050$; $d = 0.6$; 13 out of 16 subjects showed reduction) while GI showed no significant changes ($p = 0.562$; Figures 4(A1) and 4(B1)). Within the three metrics that showed statistical significance (i.e., PI, SI, and AI), the four female subjects did not display consistent changing patterns (i.e., for each metric there are both decrease and increase during walking across the four female subjects). The between-condition changes remain when restricting data to male subjects only (Figures 4(A2)–4(D2)), except that the reduction during walking in PI becomes borderline significant ($p = 0.077$; Figure 4(A2)).

4. Discussions

Asymmetry is an accepted intrinsic property of HRV. It imparts the time irreversibility of HRV—an important marker of the nonlinearity in HRV dynamics that can be perturbed by many pathologies [19]. For example, perturbed HRA has been observed in diseases including arrhythmia [21], heart failure [24], obstructive sleep apnea [25],

myocardial infarction [26, 27], postoperative myocardial ischemia [28], and type 1 diabetes [29]. Most interestingly, HRA has suggested potential for postinfarction risk prediction [30]. The current pilot study explores whether and how HRA changes during regular walking. To answer the question, we used 5 min ECG data that applied a within-subject, randomized “crossover” design to examine changes of short-term HRV during exercise [15]. ECG data of each participant were monitored two times that correspond to a resting seated position and a regular walking protocol on the treadmill, respectively. We assessed the HRA using four established HRA metrics, that is, PI, GI, SI, and AI. With the 5 min ECG data, mainly the short-term HRA is expected to be captured [12, 19, 30] while the components related to long-term HRA may only have slight contributions to results, which limits the availability of long-term HRA to be examined fairly. Therefore, in this study we focused only on short-term HRA, and in order to further get rid of the potential weak contributions of long-term HRA, we repeated the calculations of the four HRA metrics on HRV recordings after a wavelet detrending process that removes the low-frequency components of ~ 0.03 Hz which contribute primarily to long-term HRA.

Our pilot data on 16 healthy college students showed a reduced PI while unchanged GI, SI, and AI during walking

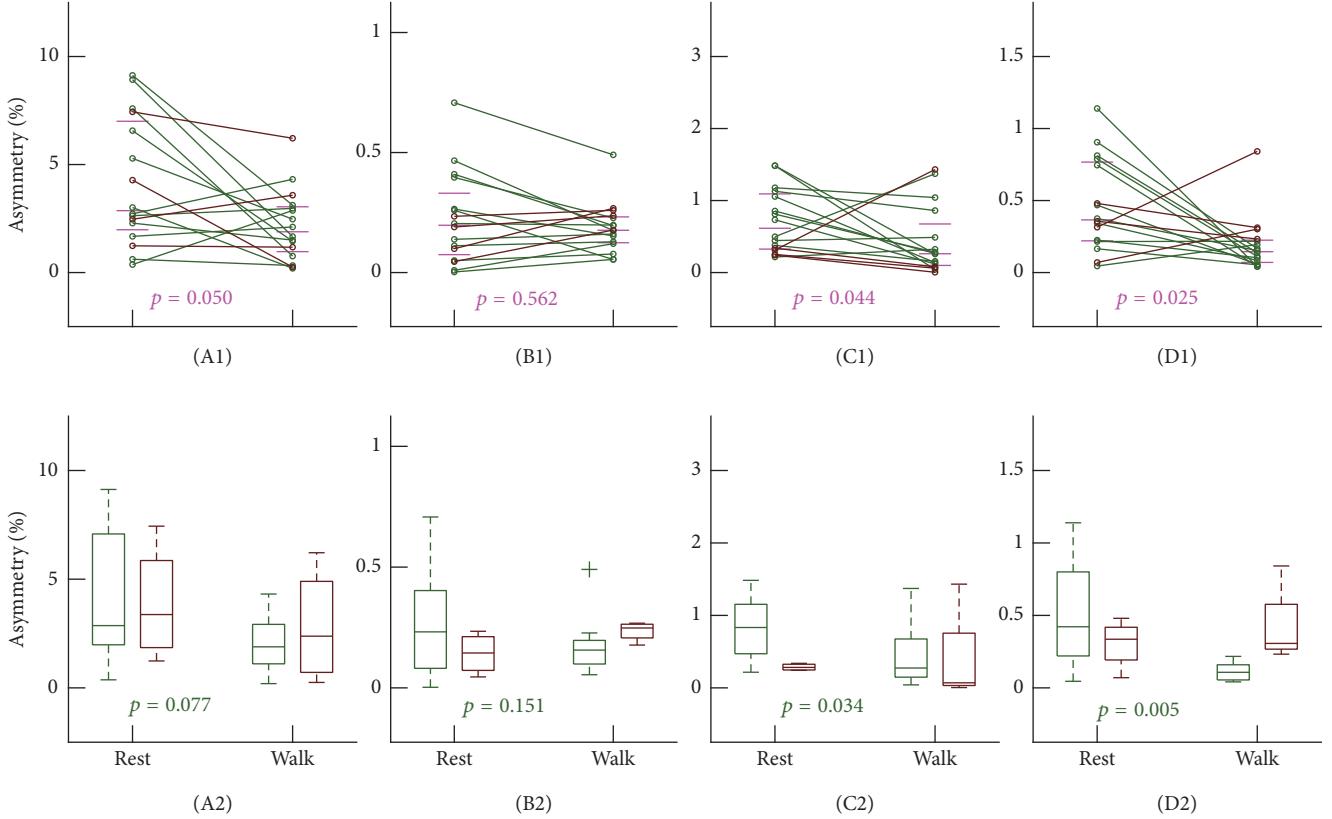


FIGURE 4: The asymmetries of short-term heart rate variability after wavelet detrending. In order to show the changes, results from the same individual were connected by lines. Horizontal bars indicate the median and [1st, 3rd] quartiles. p values were from Wilcoxon signed-rank test of each pair. ((A1) and (A2)) ΔPI ; ((B1) and (B2)) ΔGI ; ((C1) and (C2)) ΔSI ; ((D1) and (D2)) ΔAI . Rest: results under resting seated position; Walk: results during regular walking. 13 individuals show reduction from Rest to Walk in (A1), 12 individuals show reduction from Rest to Walk in (C1), and 12 individuals show reduction from Rest to Walk in (D1). Results from males and from females are marked in different colors (male: green; female: brown). Lower panels summarize the box plots (that show from top to bottom the max, 3rd quartile, median, 1st quartile, and min) for males and females, separately. Outliers, if there are any, are marked by “+.”

on treadmill based on raw HRV data. Based on wavelet-detrended data, reduced PI, SI, and AI were documented while GI still indicated no significant changes. It has been hypothesized that short-term HRA possesses a dominant contribution of vagal activity [9, 12]. Thus, the reduced PI observed from both raw HRV and detrended HRV data may underline the belief that vagal tone is withdrawn during low-intensity exercise [17, 31, 32]. However, none of the remaining three metrics, that is, GI, SI, and AI, showed significant changes based on raw HRV data, suggesting a possible lack of sensitivity to vagal withdrawal. Furthermore, SI and AI indicated significant decreases during walking using detrended HRV data, suggesting that, in addition to short-term HRA, SI and AI may also capture long-term HRA that confounds the changes of short-term HRA even though the contribution of long-term component in 5 min ECG data is low. GI was almost unchanged after detrending, implying that GI, a second-dimensional metric that relies on the distances, may capture mostly long-term HRA. We note that even with significant observations, the changing directions of these metrics with regular walking in several individuals are totally opposite (see Figures 3 and 4). Different changing directions may reflect different autonomic responses across

individuals to the walking stimuli. The difference may come from different exercise habits, different levels of college study stress, or even autonomic disorders [33]. This information will be collected in our future studies in order to uncover what leads to the differences.

Consistently, all our results still held when using data of male subjects. However, with only four females, we could not reliably perform any statistical analyses. Besides, the changing directions of HRA from resting to walking conditions seemed not consistent. Together, they limited our ability to conclude anything for female subjects. In a previous study, an interesting sex difference in HRA in particularly younger subjects has been reported [34]. Further studies are thus warranted to examine whether the effect of regular walking on HRA differs across sexes. In addition, participants in the current study were all quite young. How age influences the effect of regular walking is yet another concern that requires further elucidations.

Our results also show consistency with some published work. For example, there are studies that observed decreased HRA during acute mental stress (i.e., Stroop and arithmetic test) [35] and aerobic exercise [36], both corresponding to an autonomic balance shift towards sympathetic predominance

or vagal withdrawal. However, in the study that applied acute mental stress [35], GI was found to better reflect vagal withdrawal than PI did, which is different from what we observed. This difference may partially due to different data lengths used (i.e., 6 min in the mentioned study versus 5 min in ours). And another possible reason might be that we calculated the absolute difference of the actual HRA metrics and 50 (see Section 2.5, and more discussions regarding this can be found in the next paragraph). In a different study, the same group (i.e., the group of the mental stress study) also showed that HRA increased significantly during orthostasis and that GI was more sensitive to the stimulus [37]. HRV of ~15 min was used in that study. The increased HRA during orthostasis based on this relatively long data might reflect mainly the sympathetic activation, and the better performance of GI could thus be understandable as our results before and after detrending provide a hint that GI may be more sensitive to the sympathetic modulation and thus the long-term HRA. What is interesting is that the changing directions of short- and long-term HRA during vagal withdrawal or sympathetic activation are completely opposite which is worth further elucidations. An increase in HRA has also been observed with respiratory maneuver (e.g., inspiration/expiration = 2:1 or 1:1) [38]. Three 4.5 s metronome breathing patterns (1:1, 2:1, and normal pattern ~1:2) were administered for each participant while ECGs were recorded for 5 min at each breathing pattern. However, in that study no significant differences in traditional HRV parameters (such as power of higher frequency—the marker for vagal activity) were found. Further examinations to clearly figure out what led to the observed HRA changes are thus still required and this also limits the direct comparison between our study and the other three studies reviewed above, that is, [35–37], which attribute HRA changes mainly to autonomic responses.

It is worth noting that we used the absolute difference of an HRA metric to 50 as the index of asymmetric level (see Section 2.5). By calculating the absolute difference, we lost the power to differentiate the contributions of instantaneously accelerated and decelerated patterns. However, in the currently study, we focused mainly on the “asymmetry” phenomenon, which is believed to be existing especially during resting state as reported by many previous studies [3, 12], instead of the unbalanced sympathetic or vagal tones. The calculation of absolute changes provides the possibility of screening more asymmetric patterns out, as stated by a previous study [39]. In addition, the regulation of heart rate is not instantaneous. Instead, it takes a couple of seconds [40], which imparts the importance of measuring the symmetry of changes rather than the exact acceleration or deceleration patterns [39].

Our pilot study also touches a potential important point in short-term HRA analysis—the influence of nonstationary trend. To the best of our knowledge, this has not been considered seriously in previous work. We note that the very low-frequency component of ~ 0.03 Hz is usually considered nonstationary trend and its removal will hardly affect the beat-to-beat decelerating/accelerating patterns. However, it may affect how much the pattern deviates from symmetric.

Therefore, if an HRA algorithm takes the position of the patterns in the Poincaré plot (either above or below the line of identity) into consideration, the results after trend removal would rarely get affected (such as the case for the metric PI; see Figures 3 and 4). If an algorithm considers the distance or area characteristics, it is possible that a pattern will be considered deviated a little bit more from symmetric before trend removal than from afterwards. This effect will be important when considering long-term asymmetry. In this scenario, a decreased asymmetry (increased symmetry) for the same recording after trend removal would thus be expected when using metrics such as SI and AI, especially for data during walking (which are true when comparing Figures 3 and 4). This also provides a possible explanation that PI decreases significantly during exercise both before and after trend removal while significant changes in SI and AI are only observed after trend removal. Based on our pilot data, nonstationary trend removal is recommended for short-term HRA analysis, and to validate this, definitely further examinations with more participants and different stimulus are warranted.

Recent advances in smart wearables open a new avenue for the monitoring and management of individual’s health during daily routine. Perhaps one of the most common wearable devices is the ECG or heart rate monitor that can be used to assess the cardiovascular function and the underlying autonomic control status. A simplest idea in using such devices is to implement the algorithms that are previously developed based on data episodes collected during one time clinic or laboratory visit into continuous data, which provides the opportunity for sporadic, health-related alterations to be picked up, as well as the feasibility to look at the variations of these markers with time throughout a day.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the Natural Science Foundation of Shandong Province, China (no. BS2012DX019), the National Natural Science Foundation of China (nos. 61601263, 61471223), and the Key Program on Natural Scientific Research from the Department of Education of Anhui Province, China (no. KJ2016A470).

References

- [1] E. H. Hon and S. T. Lee, “Electronic evaluations of the fetal heart rate patterns preceding fetal death, further observations,” *American Journal of Obstetrics & Gynecology*, vol. 87, pp. 814–826, 1965.
- [2] Task Force of the European Society of Cardiology, “Heart rate variability, standards of measurement, physiological interpretation, and clinical use,” *Circulation*, vol. 93, pp. 1043–1065, 1996.
- [3] M. Costa, A. L. Goldberger, and C.-K. Peng, “Broken asymmetry of the human heartbeat: loss of time irreversibility in aging

- and disease," *Physical Review Letters*, vol. 95, no. 19, Article ID 198102, 2005.
- [4] M. D. Costa, C.-K. Peng, and A. L. Goldberger, "Multiscale analysis of heart rate dynamics: entropy and time irreversibility measures," *Cardiovascular Engineering*, vol. 8, no. 2, pp. 88–93, 2008.
 - [5] K. R. Casali, A. G. Casali, N. Montano et al., "Multiple testing strategy for the detection of temporal irreversibility in stationary time series," *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 77, no. 6, Article ID 066204, 2008.
 - [6] C. Cammarota and E. Rogora, "Time reversal, symbolic series and irreversibility of human heartbeat," *Chaos, Solitons & Fractals*, vol. 32, no. 5, pp. 1649–1654, 2007.
 - [7] F. Hou, J. Zhuang, C. Bian et al., "Analysis of heartbeat asymmetry based on multi-scale time irreversibility test," *Physica A: Statistical Mechanics and its Applications*, vol. 389, no. 4, pp. 754–760, 2010.
 - [8] U. R. Acharya, K. P. Joseph, N. Kannathal, C. M. Lim, and J. S. Suri, "Heart rate variability: a review," *Medical & Biological Engineering & Computing*, vol. 44, no. 12, pp. 1031–1051, 2006.
 - [9] J. Piskorski and P. Guzik, "Compensatory properties of heart rate asymmetry," *Journal of Electrocardiology*, vol. 45, no. 3, pp. 220–224, 2012.
 - [10] R. Hainsworth, "Physiology of the Cardiac Autonomic System," in *Clinical Guide to Cardiac Autonomic Tests*, M. Malik, Ed., pp. 10–1007, Springer, Dordrecht, Netherlands, 1998.
 - [11] P. Nicolini, M. M. Ciulla, C. D. Asmundis, F. Magrini, and P. Brugada, "The Prognostic Value of Heart Rate Variability in the Elderly, Changing the Perspective: From Sympathovagal Balance to Chaos Theory," *Pacing and Clinical Electrophysiology*, vol. 35, no. 5, pp. 621–637, 2012.
 - [12] J. Piskorski and P. Guzik, "Asymmetric properties of long-term and total heart rate variability," *Medical & Biological Engineering & Computing*, vol. 49, no. 11, pp. 1289–1297, 2011.
 - [13] A. A. Akintola, V. van de Pol, D. Bimmel, A. C. Maan, and D. van Heemst, "Comparative analysis of the equivital EQ02 lifemonitor with holter ambulatory ECG device for continuous measurement of ECG, heart rate, and heart rate variability: A validation study for precision and accuracy," *Frontiers in Physiology*, vol. 7, article no. 391, 2016.
 - [14] J. Kristiansen, M. Korshøj, J. H. Skotte et al., "Comparison of two systems for long-term heart rate variability monitoring in free-living conditions—a pilot study," *Biomedical Engineering Online*, vol. 10, article 27, 2011.
 - [15] B. Shi, Y. Zhang, C. Yuan, S. Wang, and P. Li, "Entropy Analysis of Short-Term Heartbeat Interval Time Series during Regular Walking," *Entropy*, vol. 19, no. 10, p. 568, 2017.
 - [16] C. Karmakar, A. Khandoker, and M. Palaniswami, "Investigating the changes in heart rate asymmetry (HRA) with perturbation of parasympathetic nervous system," *Australasian Physical & Engineering Sciences in Medicine*, vol. 35, no. 4, pp. 465–474, 2012.
 - [17] S. Boettger, C. Puta, V. K. Yeragani et al., "Heart rate variability, QT variability, and electrodermal activity during exercise," *Medicine & Science in Sports & Exercise*, vol. 42, no. 3, pp. 443–448, 2010.
 - [18] P. Li, C. Liu, M. Zhang, W. Che, and J. Li, "A Real-Time QRS Complex Detection Method," *Acta Biophysica Sinica*, vol. 27, no. 3, pp. 222–230, 2011.
 - [19] A. Porta, K. R. Casali, A. G. Casali et al., "Temporal asymmetries of short-term heart period variability are linked to autonomic regulation," *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, vol. 295, no. 2, pp. R550–R557, 2008.
 - [20] P. Guzik, J. Piskorski, T. Krauze, A. Wykretowicz, and H. Wysocki, "Heart rate asymmetry by Poincaré plots of RR intervals," *Biomedizinische Technik. Biomedical Engineering*, vol. 51, no. 4, pp. 272–275, 2006.
 - [21] C. K. Karmakar, A. H. Khandoker, and M. Palaniswami, "Phase asymmetry of heart rate variability signal," *Physiological Measurement*, vol. 36, no. 2, article no. 303, pp. 303–314, 2015.
 - [22] C. Yan, P. Li, L. Ji, L. Yao, C. Karmakar, and C. Liu, "Area asymmetry of heart rate variability signal," *Biomedical Engineering Online*, vol. 16, no. 1, 2017.
 - [23] S. S. Sawilowsky, "Very large and huge effect sizes," *Journal of Modern Applied Statistical Methods*, vol. 8, no. 2, pp. 597–599, 2009.
 - [24] M. A. Woo, W. G. Stevenson, D. K. Moser, R. B. Trelease, and R. M. Harper, "Patterns of beat-to-beat heart rate variability in advanced heart failure," *American Heart Journal*, vol. 123, no. 3, pp. 704–710, 1992.
 - [25] P. Guzik, J. Piskorski, K. Awan, T. Krauze, M. Fitzpatrick, and A. Baranchuk, "Obstructive sleep apnea and heart rate asymmetry microstructure during sleep," *Clinical Autonomic Research*, vol. 23, no. 2, pp. 91–100, 2013.
 - [26] H. V. Huikuri, T. Seppänen, M. J. Koistinen et al., "Abnormalities in beat-to-beat dynamics of heart rate before the spontaneous onset of life-threatening ventricular tachyarrhythmias in patients with prior myocardial infarction," *Circulation*, vol. 93, no. 10, pp. 1836–1844, 1996.
 - [27] P. K. Stein, P. P. Domitrovich, H. V. Huikuri, and R. E. Kleiger, "Traditional and nonlinear heart rate variability are each independently associated with mortality after myocardial infarction," *Journal of Cardiovascular Electrophysiology*, vol. 16, no. 1, pp. 13–20, 2005.
 - [28] T. T. Laitio, T. H. Mäkipallio, H. V. Huikuri et al., "Relation of heart rate dynamics to the occurrence of myocardial ischemia after coronary artery bypass grafting," *American Journal of Cardiology*, vol. 89, no. 10, pp. 1176–1181, 2002.
 - [29] P. Guzik, "Heart rate variability by Poincaré plot and spectral analysis in young healthy subjects and patients with type 1 diabetes," *Folia Cardiol*, vol. 12, no. suppl D, pp. 64–67, 2005.
 - [30] P. Guzik, J. Piskorski, P. Barthel et al., "Heart rate deceleration runs for postinfarction risk prediction," *Journal of Electrocardiology*, vol. 45, no. 1, pp. 70–76, 2012.
 - [31] D. W. White and P. B. Raven, "Autonomic neural control of heart rate during dynamic exercise: revisited," *The Journal of Physiology*, vol. 592, part 12, pp. 2491–2500, 2014.
 - [32] J. P. Fisher, "Autonomic control of the heart during exercise in humans: Role of skeletal muscle afferents," *Experimental Physiology*, vol. 99, no. 2, pp. 300–305, 2014.
 - [33] Q. Fu and B. D. Levine, "Exercise and the autonomic nervous system," in *Journal of the Autonomic Nervous System*, vol. 117 of *Handbook of Clinical Neurology*, pp. 147–160, Elsevier, 2013.
 - [34] A. Voss, R. Schroeder, A. Heitmann, A. Peters, and S. Perz, "Short-term heart rate variability - Influence of gender and age in healthy subjects," *PLoS ONE*, vol. 10, no. 3, Article ID e0118308, 2015.
 - [35] Z. Visnovcova, M. Mestanik, M. Javorka et al., "Complexity and time asymmetry of heart rate variability are altered in acute mental stress," *Physiological Measurement*, vol. 35, no. 7, pp. 1319–1334, 2014.

- [36] B. D. L. C. Torres and J. N. Orellana, "Multiscale time irreversibility of heartbeat at rest and during aerobic exercise," *Cardiovascular Engineering*, vol. 10, no. 1, pp. 1–4, 2010.
- [37] L. Chladekova, B. Czippelova, Z. Turianikova et al., "Multiscale time irreversibility of heart rate and blood pressure variability during orthostasis," *Physiological Measurement*, vol. 33, no. 10, article no. 1747, pp. 1747–1756, 2012.
- [38] A. Klintworth, Z. Ajtay, A. Paljunite, S. Szabados, and L. Hejjel, "Heart rate asymmetry follows the inspiration/expiration ratio in healthy volunteers," *Physiological Measurement*, vol. 33, no. 10, article 1717, 2012.
- [39] A. H. Khandoker, C. Karmakar, M. Brennan, M. Palaniswami, and A. Voss, *Poincaré Plot Methods for Heart Rate Variability Analysis*, Springer US, Boston, MA, USA, 2013.
- [40] D. L. Eckberg, "Nonlinearities of the human carotid baroreceptor-cardiac reflex," *Circulation Research*, vol. 47, no. 2, pp. 208–216, 1980.

Research Article

Kernel Principal Component Analysis of Coil Compression in Parallel Imaging

Yuchou Chang  ¹ and Haifeng Wang  ²

¹Computer Science and Engineering Technology Department, University of Houston-Downtown, Houston, TX 77002, USA

²Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, Guangdong 518055, China

Correspondence should be addressed to Haifeng Wang; hf.wang1@siat.ac.cn

Received 26 November 2017; Accepted 7 March 2018; Published 19 April 2018

Academic Editor: Sheng-Yu Peng

Copyright © 2018 Yuchou Chang and Haifeng Wang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A phased array with many coil elements has been widely used in parallel MRI for imaging acceleration. On the other hand, it results in increased memory usage and large computational costs for reconstructing the missing data from such a large number of channels. A number of techniques have been developed to linearly combine physical channels to produce fewer compressed virtual channels for reconstruction. A new channel compression technique via kernel principal component analysis (KPCA) is proposed. The proposed KPCA method uses a nonlinear combination of all physical channels to produce a set of compressed virtual channels. This method not only reduces the computational time but also improves the reconstruction quality of all channels when used. Taking the traditional GRAPPA algorithm as an example, it is shown that the proposed KPCA method can achieve better quality than both PCA and all channels, and at the same time the calculation time is almost the same as the existing PCA method.

1. Introduction

Parallel imaging methods [1, 2] have been widely used to accelerate MRI acquisitions. Due to the increased number of coils in parallel Magnetic Resonance Imaging, the numbers of coils (e.g., 128 channels) have been developed to improve the image quality of reconstruction and the sampling speed of acquisition [3–5]. On the other hand, the calculated cost increases as the number of coils increases, especially for coil-based reconstruction methods such as GRAPPA [2]. A number of coil compression methods have been proposed [6–19] to reduce computational time. They can be divided into two categories, one based on the hardware approach [5] and the other based on the software approach [6–19]. Those software-based coil compression methods provide a more flexible way to reduce computation workload. For example, principal component analysis (PCA) has been applied on compressing large array coils [10, 19]. The coil compression process produces a smaller set of virtual channels that can be represented as a linear combination of physical channels. The method has been successfully applied to most existing reconstruction methods such as SENSE [1],

GRAPPA [2], and SPIRiT [20]. All existing coil compression methods have demonstrated that the number of channels can be greatly reduced without significant loss of SNR or image degradation, thereby increasing computational efficiency. In addition to saving computing time, PCA-based channel reduction methods have been shown to have noise reduction effects [10, 19, 21]. However, this denoising effect has been discussed in [21] without significant improvement.

The purposed method is to study the noise reduction capability of software-based coil compression methods while achieving noise suppression and channel reduction simultaneously. And we present a PCA-based approach, which is a nonlinear extension of the conventional PCA method [10, 19]. In contrast to the linear combination used in the conventional PCA, the proposed channel reduction technique nonlinearly combines physical channels to generate a new reduced set of virtual channels. Actually, the conception of nonlinear reconstruction using kernel methods has been studied in nonlinear GRAPPA [22], and the advantages of nonlinear combination over linear techniques have been demonstrated. The proposed kernel PCA (KPCA) method can reduce the

usage of nonlinear combination on additional dimensions and more effectively enhance the quality of coil channels. In experiments, we used the GRAPPA method [2] as the reconstruction demos to achieve the final images from the data reduced channels. When generating the same small number of virtual channels, the proposed KPCA can reduce GRAPPA calculation time the same as the previous PCA-GRAPPA reconstruction [10] calculation time; however, the signal-to-noise ratio (SNR) is higher than the conventional GRAPPA [2] and PCA-GRAPPA [10].

2. Background

Generally, the GRAPPA reconstruction [2] can be represented as

$$S_j(k_y + r\Delta k_y, k_x) = \sum_{l=1}^L \sum_{b=B_1}^{B_2} \sum_{h=H_1}^{H_2} w_{j,r}(l, b, h) \times S_l(k_y + bR\Delta k_y, k_x + h\Delta k_x), \quad (1)$$

where the unacquired k -space signal S_j (the left side of (1)) is calculated by a linear combination of k -space signals (the right side of (1)). Here, w represents the coefficient set, R is the outer reduction factor (ORF), j is the target coil, l counts all coils, b and h are calculated by taking neighbored k -space data in k_y and k_x directions, respectively, and the variables k_x and k_y represent the coordinates encoded along the frequency and phase, respectively. The GRAPPA formulation can be simplified as a matrix equation:

$$\mathbf{b}_{M \times 1} = \mathbf{D}_{M \times K} \mathbf{x}_{K \times 1}, \quad (2)$$

where \mathbf{D} denotes the matrix consisting of the acquired data, \mathbf{b} represents the vector of the missing data, and \mathbf{x} represents the coefficients.

In general, the coefficients are dependent on the coil sensitivity, which are a priori unknown. In GRAPPA, auto-calibration data (ACS) are obtained and used as the vector \mathbf{b} to estimate the coefficient vector \mathbf{x} . The least-squares method is usually used to calculate the coefficients:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{b} - \mathbf{D}\mathbf{x}\|^2. \quad (3)$$

When the matrix \mathbf{D} changes with a higher reduction factors, the noise in the estimation coefficients can be greatly amplified.

As a dimension reduction technique, PCA has been successfully used to reduce the number of effective channels in GRAPPA reconstruction [8, 9]. The PCA finds an orthogonal linear transformation that converts the data to a new coordinate system so that the largest change in any projection of data comes from the first coordinate, the second largest change is on the second coordinate, and so on. When applied to channel reduction, the ACS data is used to obtain the transformation and then applied to all acquired data to obtain a new dataset in the new coordinate system. Mathematically, the linear transformation \mathbf{W} can be calculated by

the eigen-decomposition of the covariance matrix of the ACS data:

$$\mathbf{A}^H \mathbf{A} = \mathbf{W}^H \Sigma \mathbf{W}, \quad (4)$$

where $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_L]$ consisted of vector \mathbf{a}_l generated from the ACS data of the l th channel (a total of L channels) after removing the average; \mathbf{W} and Σ are, respectively, eigenvectors and eigenvalues of the matrices. The new coordinates based on eigenvectors are called principal components. Assuming that the direction of largest variance represents interesting information and the direction of the minimum variance indicates noise that is not of interest. For simplicity, only a few first eigenvectors corresponding to the largest eigenvalues are retained to form a linear transformation \mathbf{T} . The transformation matrix is then applied to the acquired k -space data to obtain an orthogonal projection of the eigenvectors, resulting in a new set of reduced virtual channels. Then, the undersampled data is reconstructed in the transform domain by conventional GRAPPA. Note that [9] the number of source channels (N_{sch}) and the number of target channels (N_{tch}) may differ after PCA reduction. One may be bigger than the other, with the same calculation time to get the best result. The final image is produced by combining the virtual channels with root sum-of-square (SOS). Obviously, the assumptions in the PCA are not necessarily kept because of the possibility of small variance in the direction of interesting signals, in which case the useful information is lost after reductions.

3. Proposed Method

3.1. Kernel PCA. The kernel method [23] is a widely used machine learning method. The main idea of the kernel method is that a set of points which cannot be linearly segmented in a low-dimensional space is likely to become linearly separable when transformed into a set of points in a high-dimensional space. For a given linear algorithm, the data is mapped from the input space A to the feature space H through a nonlinear mapping $\Phi(\cdot): A \rightarrow H$, and then the algorithm is applied on the vector representation $\Phi(\mathbf{a})$ of the data. When the PCA method is a nonlinear mapping algorithm, the approach becomes a kernel PCA (KPCA) method.

KPCA is a process of attribute dependency. The correlation here mainly refers to the linear correlation. So, for nonlinear situation, it involves kernel PCA called KPCA [24]. Intuitively, the kernel PCA is the PCA dimensionality reduction based on the kernel space after the original sample has passed the kernel mapping. KPCA formula derivation and PCA are very similar, but there are two innovations. In order to deal with nonlinear data better, a nonlinear mapping function $\Phi(\mathbf{A})$ is introduced to map the data in the original space into a high-dimensional feature space. For any vector in space, even if it is a basis vector, all samples lie in the linear representation. After kernel mapping, we make a linear PCA on the new data in the feature space constructed by the product of vector elements, thus taking into account higher-order statistics. We applied kernel PCA on parallel imaging reconstruction methods such as GRAPPA [2].

3.2. Nonlinear Mapping Function. In order to achieve a smooth relationship, a nonuniform polynomial kernel is selected for Φ mapping. It has the following form:

$$\kappa(\mathbf{a}, \mathbf{b}) = (\lambda \mathbf{a}^T \mathbf{b} + r)^d, \quad (5)$$

where λ and r are scalars; d represents the degree of the polynomial. Due to explicit representation of nonlinear mapping $\Phi(\mathbf{A})$ of $\kappa(\mathbf{a}, \mathbf{b}) = \langle \Phi(\mathbf{a}), \Phi(\mathbf{b}) \rangle$, polynomial kernel is also suitable for mapping MRI data. For instance, $\Phi(\mathbf{A})$ maps the original L -channel data \mathbf{A} to \mathbf{a}^2 , when d is 2,

$$\begin{aligned} \Phi(\mathbf{A}) \\ = [r^2, \sqrt{2\lambda r} \mathbf{a}_1, \dots, \sqrt{2\lambda r} \mathbf{a}_L, \lambda \mathbf{a}_1^{(2)}, \dots, \lambda \mathbf{a}_L^{(2)}, \sqrt{2\lambda} \mathbf{a}_1 \\ \times \mathbf{a}_2, \dots, \sqrt{2\lambda} \mathbf{a}_i \times \mathbf{a}_j, \dots, \sqrt{2\lambda} \mathbf{a}_{L-1} \times \mathbf{a}_L]^T, \end{aligned} \quad (6)$$

where $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_L$ are vectors representing different channels; superscript (2) means piecewise square; \times denotes piecewise multiplication. It can be seen that the vector includes the constant, linear, second-orders in the original data, and $\Phi(\mathbf{A})$ has $(L + 2)(L + 1)/2$ terms in total.

In order to avoid overfitting, some second-order terms are removed. In particular, the second-order terms are rearranged in the following order. The square terms are selected within each coil at first, and then the product terms between the nearest neighbors are chosen, and then the next-nearest neighbors are selected in k -space and so on. The vector $\Phi(\mathbf{a})$ is removed by using sorted terms based on the desired dimension of the feature space. If all second-order terms are truncated, the proposed method is the same as linear PCA-based channel compression algorithm.

The target channels are corresponding to data on the left side of (1) and source channels as those for the right side of (1). The original space for the target channel is selected to avoid the complexity of converting the data from the feature space back to the original space. The source channels are used for estimation only, so there is no need to convert it back to the original space. The number of second-order terms to be three times of that of the first-order terms is chosen for building the source channels. Since MRI noise is generated in a very complicated procedure, which can be considered as non-Gaussian distribution [25]. Noise also exists in sensitivities of acquired channel data. Noise and true signal can be considered as error-in-variable model [22]. The traditional linear space is mapped to nonlinear feature space to capture noise characteristics existing in coil sensitivities. Nonlinearity is added to modulate sensitivities in the channel compression procedure. The benefit of the proposed method is the simultaneous channel compression and noise suppression in reconstruction procedure.

To balance linearity and nonlinearity of the new coordinate system, the parameters r and λ are finely tuned. If the nonlinearity dominates the coordinate, the reconstructed image quality is distorted since the original channel

information is lost and overridden by nonlinearity information. By contrast, if the nonlinearity is too tiny, reconstruction is almost equivalent to original PCA-based channel reduction method, so that nonlinearity does not have effect on suppressing noise. $\sqrt{2\lambda r} = 1$ and adjustable λ are set to obtain the better performance. The maximum absolute value M_{2nd} of the second-order terms is identified for building the feature space. λ_s sets the value within the range within $(1/M_{2nd}, 10/M_{2nd})$ based on the experience that the reconstruction is insensitive to the values in the above range.

3.3. Proposed Algorithm. The proposed method is presented in the following steps.

Step 1. Extracted calibration data is the input data of KPCA for target channels and source channels, respectively. The calibration data in each channel is arranged into a vector; therefore, there are overall N vectors $\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_N$ corresponding to overall N channels of original k -space data.

Step 2. Nonlinear mapping Φ is applied on random variable \mathbf{V} here to construct the covariance matrix C_t and C_s of target channels and source channels, respectively. The new vectors \mathbf{U} are constructed as follows:

$$\begin{aligned} \mathbf{U} = [\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_N, \lambda \mathbf{V}_{11}, \lambda \mathbf{V}_{22}, \dots, \lambda \mathbf{V}_{NN}, \lambda \mathbf{V}_{NI}, \lambda \mathbf{V}_{12}, \\ \lambda \mathbf{V}_{23}, \dots, \lambda \mathbf{V}_{(N-1)N}, \lambda \mathbf{V}_{(N-1)1}, \lambda \mathbf{V}_{N2}, \lambda \mathbf{V}_{13}, \lambda \mathbf{V}_{24}, \dots, \\ \lambda \mathbf{V}_{(N-2)N}], \end{aligned} \quad (7)$$

where $\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_N$ denote vectors obtained from original k -space ACS data; second-order terms represent the vector from the point-wise multiplication by $\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_N$. For example, $\mathbf{V}_{11} = [v_1^2, v_2^2, \dots, v_n^2]^T$. Furthermore, the dimension of \mathbf{U} is $n \times 4N$, where n is the total number of k -space data obtained at the central strip, and Nyquist rate ($n = N_{nyq} \times N_x$, where N_{nyq} is the number of phase-encoding lines fully sampled with Nyquist rate, and N_x is the number of points along the frequency-encoding direction) is fully sampled. If ACS lines are defined, which are fully sampled k -space data at the central strip, $N_{acs} = N_{nyq}$ can be derived. Nonlinearity controlled by the parameter λ is added into the new coordinates. Since the target channels will be used for final reconstruction which can't be incorporated large nonlinearity, so both parameters λ_t and λ_s are tuned for constructing \mathbf{U}_t and \mathbf{U}_s , respectively. Generally, λ_t is much smaller than λ_s .

Step 3. For target and source vectors \mathbf{U}_t and \mathbf{U}_s produced in Step 2, mean of zero is calculated to make sure \mathbf{U}_t and \mathbf{U}_s will be the direction of maximal variance. The mean of zero can be calculated as follows:

$$\widehat{\mathbf{U}} = \mathbf{U} - \frac{\sum_{p=1}^n u^p}{n}. \quad (8)$$

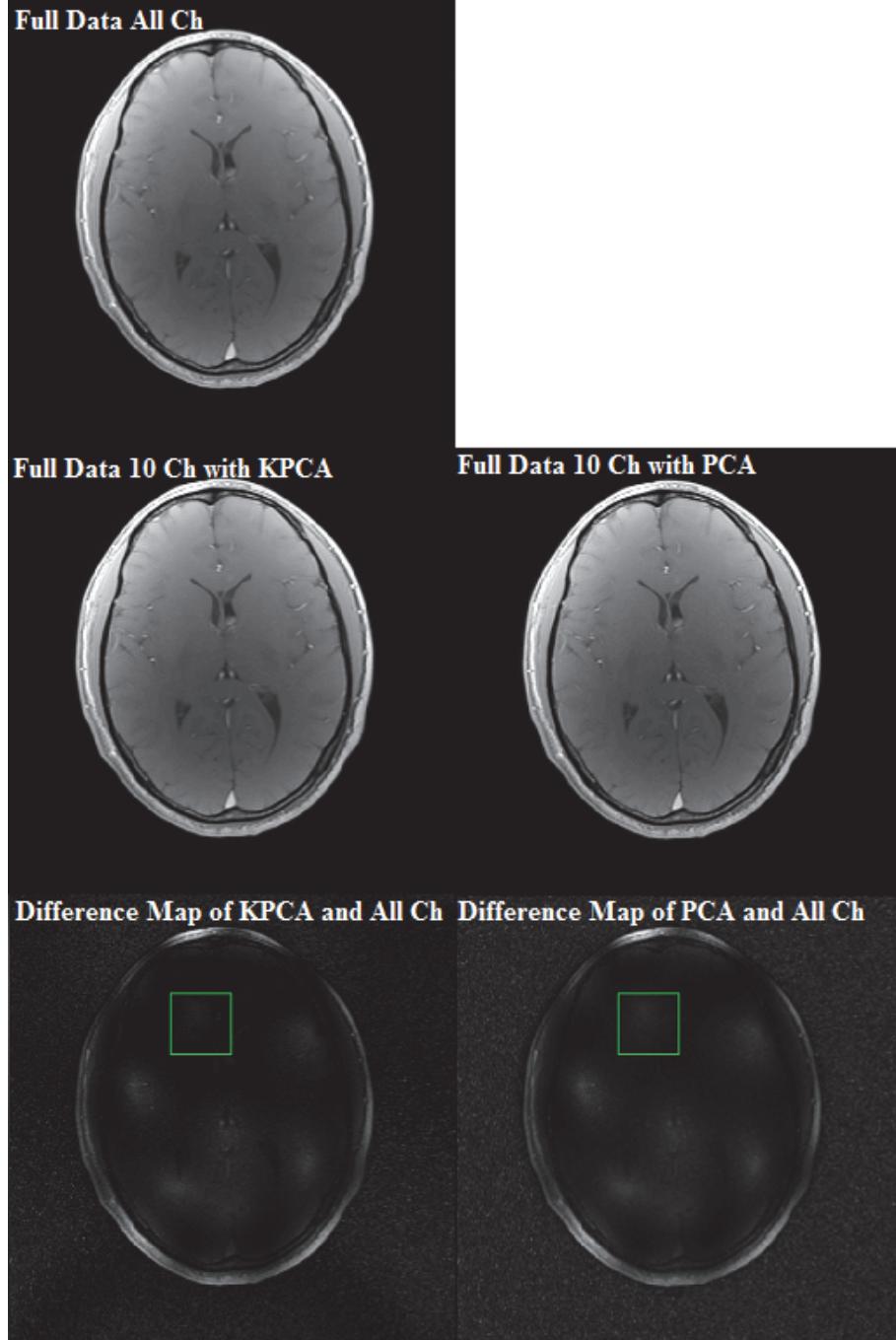


FIGURE 1: Reconstruction comparison of PCA and KPCA with compressed data. All channel (All Ch) data are used for the reference image. And the reconstructed images of KPCA and PCA with 10-channel (10 Ch) compressed data are, respectively, compared with the reference image.

Step 4. For target channels, covariance matrix C_t is generated as follows:

$$C_t = \text{cov} (\widehat{\mathbf{U}}_t(i), \widehat{\mathbf{U}}_t(j)), \quad 1 \leq i \leq 4N, \quad 1 \leq j \leq 4N, \quad (9)$$

where one component $C_t(i, j)$ represents covariance between random variables $\widehat{\mathbf{U}}_t(i)$ and $\widehat{\mathbf{U}}_t(j)$. The parameter λ_t can be set as zero to keep uniform with PCA-based channel reduction.

Step 5. Similarly to Step 4, covariance matrix C_s is constructed for source channels. The difference is that the parameter λ_s is chosen, which is generally larger than λ_t in Step 4.

Step 6. Calculate eigenvalues and eigenvectors using singular value decomposition (SVD) on covariance matrix C_t and C_s , respectively. Since nonlinear mapping is directly used here

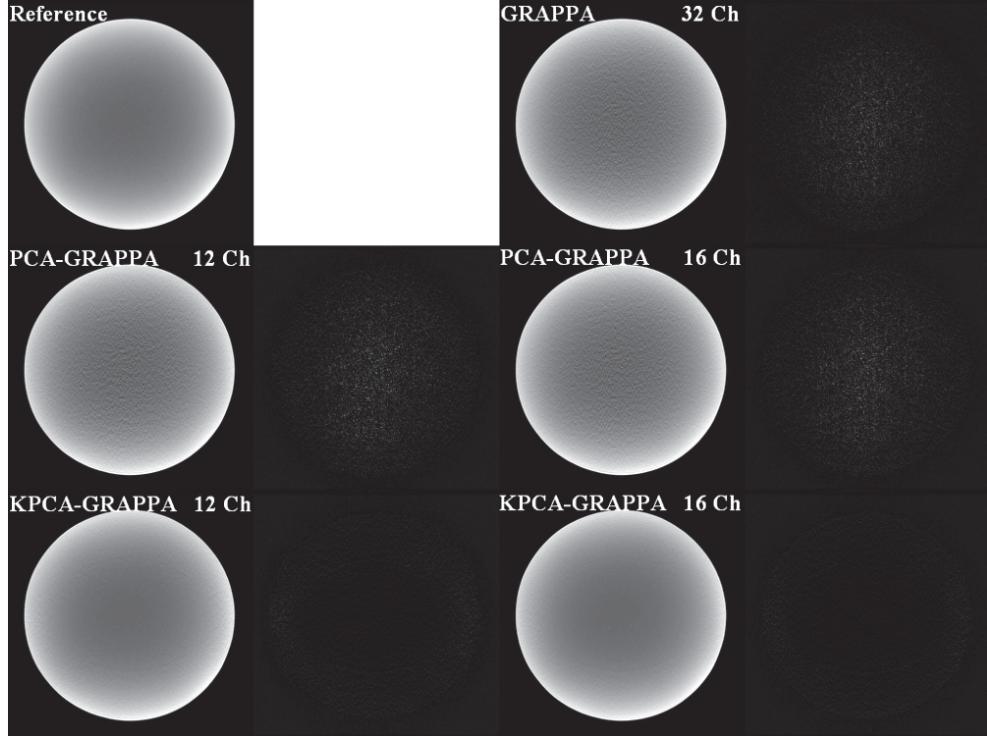


FIGURE 2: Reconstruction results of the 32-channel (32 Ch) experimental phantom images. Three methods were compared, the GRAPPA using full channel data, the PCA-GRAPPA, and KPCA-GRAPPA using, respectively, 12-channel (12 Ch) and 16-channel (16 Ch) compressed data.

and kernel trick matrix is not needed to be computed here, SVD can be directly used here to calculate eigenvalues and eigenvectors, like conventional linear PCA does [23]. The transformation matrix \mathbf{T} is composed of eigenvectors of the covariance matrix, which transforms data in \mathbf{U}_t and \mathbf{U}_s into new coordinates.

Step 7. Generate the transformed data in the new coordinates for N_{tch} target channels and N_{sch} source channels, respectively. Similarly to [26], N_{tch} and N_{sch} are not needed to be necessarily equal. In the calibration step, S_l in (1) are obtained from target channels and S_j in (1) are obtained from source channels to calculate weights. In the synthesis step, calculated weights and acquired data on source channels are combined to predict missing values on target channels, which are used for final image reconstruction.

4. Results

We validate the proposed algorithm performance by using three MRI datasets. At first, a uniform water phantom was scanned using a gradient echo (GRE) sequence (15.63 kHz bandwidth, FOV = 250 mm², matrix size = 256 × 256, TE/TR = 10/100 ms, and slice thickness = 3.0 mm). Then, a coronary brain image was acquired by using a 2D spin echo (SE) sequence (slice thickness = 3.0 mm, matrix size = 256 × 256, FOV = 240 mm², and TE/TR = 2.29/100 ms.). The third set of axial brain data was acquired on a 3T scanner (SIEMENS AG, Erlangen, German) with a 32-channel head coil using a

2D gradient echo sequence (TE/TR = 2.29/100 ms, flip angle = 25, matrix size = 256 × 256, slice thickness = 3 mm, and FOV = 24 cm²). The conventional GRAPPA [2] and PCA-based GRAPPA [10] were implemented for comparing with the proposed method on the Matlab platform (Mathworks, Natick, MA, USA). For reference image, fully sampled data with all channels are reconstructed via root sum of squares (SoS).

To measure signal loss in channel compression, KPCA and PCA channel reduction based reconstructions with fully sampled data are evaluated firstly. Both of KPCA and PCA are applied to reduce the total 32 channels to 10 channels without undersampling k -space data. The compressed channels are used to reconstruct the images with inverse Fourier transform, respectively. Both reconstructed images are compared to the reference image with fully sampled data of all 32 channels. As shown in Figure 1, KPCA channel compression based reconstruction can suppress more noise than PCA channel compression based reconstruction in the region of interest (ROI), as demonstrated in the difference images.

Figure 2 shows the reconstructions of the first dataset (phantom) reconstructed by the conventional GRAPPA with full channels, GRAPPA with PCA-based channel reduction, and the proposed method using KPCA-based channel reduction. The first dataset was undersampled with an outer reduction factor (ORF) of 5 and the ACS of 42 (net acceleration of 3.01). The configuration of the reconstruction coefficients was 15 columns and 2 blocks. The number of target channels

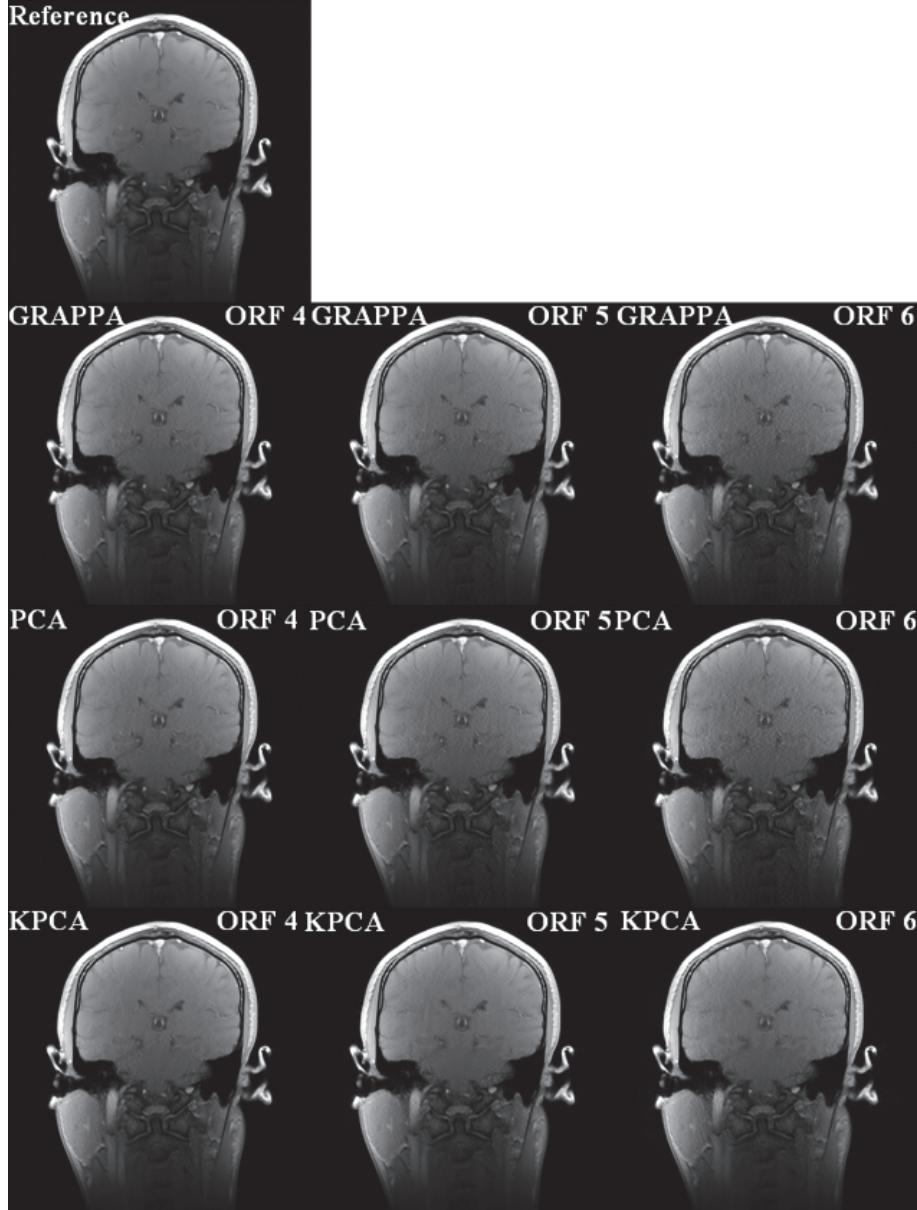


FIGURE 3: Reconstruction comparisons of the full 32-channel coronal human brain images. Three methods, the GRAPPA using full 32-channel data, the PCA-GRAPPA, and KPCA-GRAPPA using 20-channel compressed data, were compared, respectively, at the outer reduction factors (ORF) of 4, 5, and 6.

is set as 12 and the number of source channels is set as 16 for comparing performance. All sources of errors are displayed by subtracting reconstructed image by reference images. The error resources include blurring, aliasing, and noise. The proposed method is able to suppress more noise in comparison with traditional GRAPPA with full channels and GRAPPA with PCA-based channel reduction.

For the second dataset, reconstruction parameters are 48 ACS lines, 15 columns, and 2 blocks. Outer reduction factor (ORF) is set as 4, 5, and 6 for multiple comparisons. Correspondingly, the net reduction factors are, respectively, 2.56, 2.81, and 3.01. For PCA-GRAPPA and KPCA-GRAPPA, the number of reduced channels is set as 20 for both N_{tch}

target channels and N_{sch} source channels. The parameter λ_{sch} is set as 1.16×10^{-6} . From left column to right column, it is seen that reconstructions of GRAPPA, PCA-GRAPPA, and KPCA-GRAPPA are deteriorated when ORF is increasing. For the same ORF (from top to bottom), reconstruction of the proposed method always suppress more noise than that of traditional GRAPPA and GRAPPA with PCA-based channel reduction. For the column images of ORF 5, although trivial aliasing artifacts exist in all reconstructions of GRAPPA, PCA-GRAPPA, and KPCA-GRAPPA, the proposed method can reconstruct the image with cleaner contents as shown in Figure 3. Due to loss of information, reconstruction of PCA-GRAPPA is a very little worse than GRAPPA.

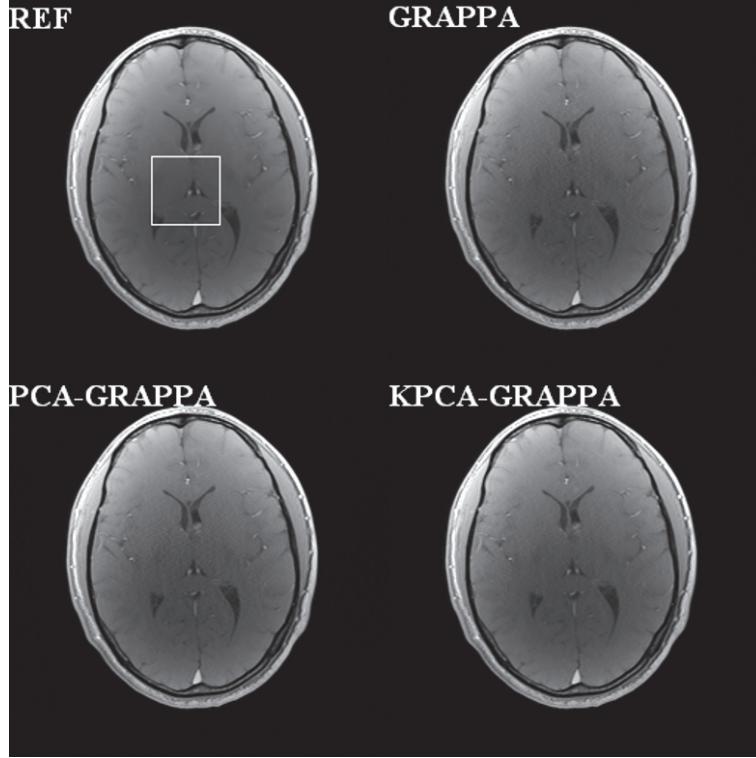


FIGURE 4: Reconstruction results of the 32-channel axial human brain images. We compared the GRAPPA reconstructions without channel reduction, with the PCA and KPCA-based channel reductions, when ORF = 4 and ACS = 48.

The proposed method is also evaluated and compared with the PCA-based channel reduction method shown in Figure 4 when using the traditional GRAPPA as the reconstruction method of the third dataset. It is also shown that there is no channel reduction in GRAPPA reconstruction and complete sampling from all channels as the square root of the reference image reconstructed for comparison. For visual comparison, the corners of each image contain an enlarged area of interest. Using 4 ORF and 48 ACS, the net acceleration was 2.56. The number of target and sources channels is 16, and λ_s is 1.23×10^{-9} . Spatially varying noise is suppressed in both conventional GRAPPA and PCA-reduced GRAPPA reconstructions. Furthermore, the proposed method also preserves the image details without obvious blurring. The computation speed of the proposed method is almost the same as the PCA-based channel reduction method (around 863 seconds), which takes only about 11% reconstruction time of the traditional GRAPPA (7771 seconds) to reconstruct the image but the quality is better. Selecting a small area of interest to calculate the SNR of each reconstruction is shown in Figure 5. The SNRs of reference, conventional GRAPPA, PCA, and KPCA reduced GRAPPA reconstructions are 16.69 dB, 14.86 dB, 14.47 dB, and 15.42 dB, respectively.

5. Conclusion

A new nonlinear kernel PCA-based channel compression method is proposed in parallel MRI. The method maps data to higher dimensional space by nonlinear transformation and performs PCA to generate a set of compressed coils. The experimental results show the proposed coil compression method based on KPCA can not only reduce the computation time cost but also suppress more noise in GRAPPA reconstructions than previous PCA method. In the future, we plan to do more experiments and investigate how to automatically choose the optimal nonlinear mapping parameters of the proposed algorithm for more parallel imaging reconstructions [27–29].

Conflicts of Interest

The authors do not have any conflicts of interest.

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China (61471350, 81729003), the Basic Research Program of Shenzhen (JCYJ20150831154213680),

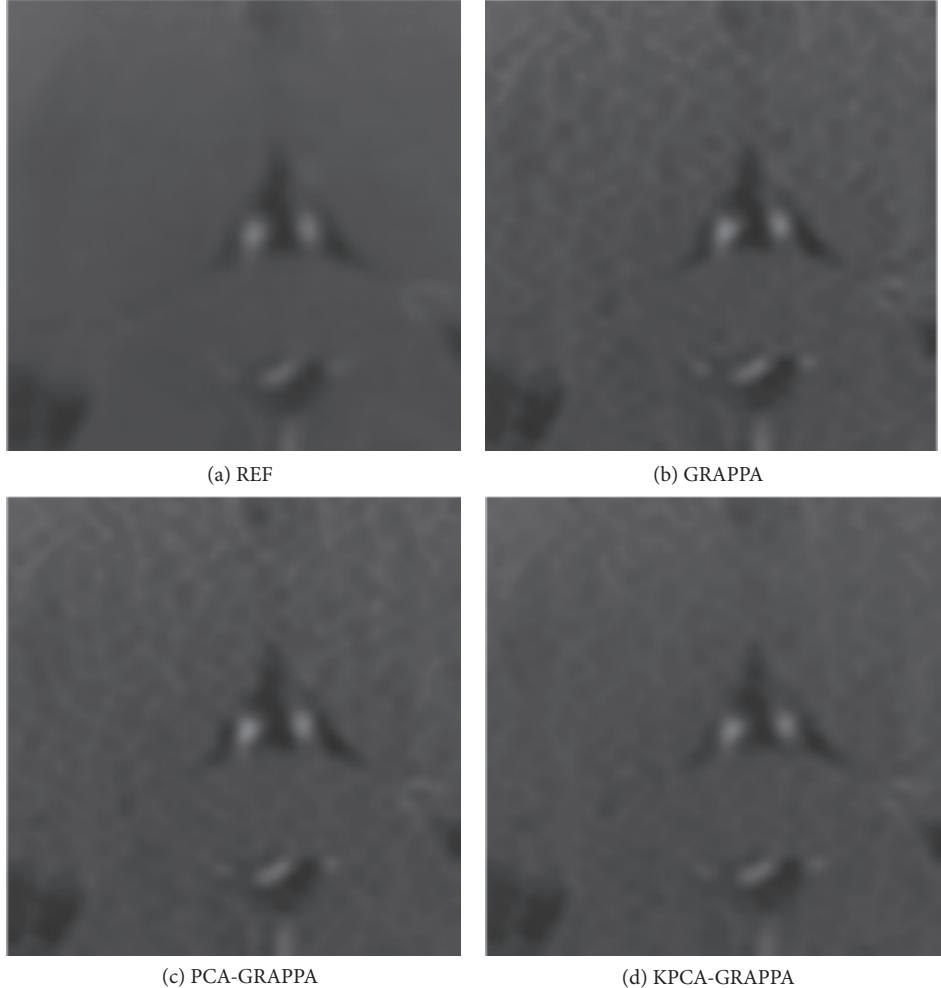


FIGURE 5: Zoom-in regions (white color box) of the reconstruction images in Figure 4. The proposed KPCA-GRAPPA method has higher SNR performance than GRAPPA and PCA-GRAPPA.

and the Key Laboratory for Magnetic Resonance and Multimodality Imaging of Guangdong Province (2014B030301013).

References

- [1] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "SENSE: sensitivity encoding for fast MRI," *Magnetic Resonance in Medicine*, vol. 42, no. 5, pp. 952–962, 1999.
 - [2] M. A. Griswold, P. M. Jakob, and R. M. Heidemann, "Generalized autocalibrating partially parallel acquisitions (GRAPPA)," *Magnetic Resonance in Medicine*, vol. 47, no. 6, pp. 1202–1210, 2002.
 - [3] J. F. Villena, A. G. Polimeridis, Y. Eryaman et al., "Fast Electromagnetic Analysis of MRI Transmit RF Coils Based on Accelerated Integral Equation Methods," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 11, pp. 2250–2261, 2016.
 - [4] D. Jiao and J.-M. Jin, "Fast frequency-sweep analysis of RF coils for MRI," *IEEE Transactions on Biomedical Engineering*, vol. 46, no. 11, pp. 1387–1390, 1999.
 - [5] V. Soumya, A. Varghese, and N. Neetha, "A review on noise estimation methods in multi-coil MRI," in *Proceedings of the*

International Conference on Computing and Network Communications, CoCoNet 2015, pp. 859–866, India, December 2015.

- [6] S. B. King, S. M. Varosi, and G. R. Duensing, "Optimum SNR data compression in hardware using an Eigencoil array," *Magnetic Resonance in Medicine*, vol. 63, no. 5, pp. 1346–1356, 2010.
 - [7] M. Buehrer, K. P. Pruessmann, P. Boesiger, and S. Kozerke, "Array compression for MRI with large coil arrays," *Magnetic Resonance in Medicine*, vol. 57, pp. 1131–1139, 2007.
 - [8] M. Doneva and P. Börnert, "Automatic coil selection for channel reduction in SENSE-based parallel imaging," *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 21, no. 3, pp. 187–196, 2008.
 - [9] F. Huang, S. Vijayakumar, Y. Li, S. Hertel, and G. R. Duensing, "A software channel compression technique for faster reconstruction with many channels," *Magnetic Resonance Imaging*, vol. 26, no. 1, pp. 133–141, 2008.
 - [10] F. Huang, W. Lin, G. R. Duensing, and A. Reykowski, "A hybrid method for more efficient channel-by-channel reconstruction with many channels," *Magnetic Resonance in Medicine*, vol. 67, no. 3, pp. 835–843, 2012.
 - [11] Z. Dong, F. Wang, X. Ma et al., "Interleaved EPI diffusion imaging using SPIRiT-based reconstruction with virtual coil compression," *Magnetic Resonance in Medicine*, 2017.

- [12] S. Feng and J. Ji, "Channel reduction in massive array parallel MRI," in *Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society: Engineering the Future of Biomedicine (EMBC '09)*, pp. 4045–4048, IEEE, Minneapolis, Minn , USA, September 2009.
- [13] J. Wang, Z. Chen, Y. Wang, L. Yuan, and L. Xia, "A Feasibility Study of Geometric-Decomposition Coil Compression in MRI Radial Acquisitions," *Computational and Mathematical Methods in Medicine*, vol. 2017, Article ID 7685208, 2017.
- [14] A. Chu and D. C. Noll, "Coil compression in simultaneous multislice functional MRI with concentric ring slice-GRAPPA and SENSE," *Magnetic Resonance in Medicine*, vol. 76, no. 4, pp. 1196–1209, 2016.
- [15] S. Bhavé, S. G. Lingala, and M. Jacob, "A variable splitting based algorithm for fast multi-coil blind compressed sensing MRI reconstruction," in *Proceedings of the 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014*, pp. 2400–2403, USA, August 2014.
- [16] J. Wang, Z. Chen, Y. Wang, L. Yuan, and L. Xia, "A Feasibility study of geometric-decomposition coil compression in MRI radial acquisitions," *Computational and Mathematical Methods in Medicine*, vol. 2017, Article ID 7685208, 9 pages, 2017.
- [17] T. Zhang, J. M. Pauly, S. S. Vasanawala, and M. Lustig, "Coil compression for accelerated imaging with Cartesian sampling," *Magnetic Resonance in Medicine*, vol. 69, no. 2, pp. 571–582, 2013.
- [18] M. Buehrer, K. P. Pruessmann, P. Boesiger, and S. Kozerke, "Array compression for MRI with large coil arrays," *Magnetic Resonance in Medicine*, vol. 57, no. 6, pp. 1131–1139, 2007.
- [19] F. Huang, S. Vijayakumar, and J. Akao, "Software Compression for Partially Parallel Imaging with Multi-channels," in *Proceedings of the 27th Annual International Conference of the Engineering in Medicine and Biology Society, (IEEE-EMBS '05)*, pp. 1348–13451, IEEE, Shanghai, China, 2005.
- [20] M. Lustig and J. M. Pauly, "SPIRiT: iterative self-consistent parallel imaging reconstruction from arbitrary k -space," *Magnetic Resonance in Medicine*, vol. 64, no. 2, pp. 457–471, 2010.
- [21] S. Vijayakumar, F. Huang, J. H. Akao, M. K. Limkeman, and G. R. Duensing, "Channel compression and denoising," in *Proceedings of the 15th ISMRM Scientific Meeting*, Berlin, Germany, May 2007.
- [22] Y. Chang, D. Liang, and L. Ying, "Nonlinear GRAPPA: A kernel approach to parallel MRI reconstruction," *Magnetic Resonance in Medicine*, vol. 68, no. 3, pp. 730–740, 2012.
- [23] B. Schölkopf, A. Smola, and K. Müller, "Kernel principal component analysis," in *International Conference on Artificial Neural Networks*, vol. 1327 of *Lecture Notes in Computer Science*, pp. 583–588, Springer Berlin Heidelberg, Berlin, Germany, 1997.
- [24] S. Mika, B. Schölkopf, A. Smola, K. R. Muller, M. Scholz, and G. Ratsch, "Kernel principal component analysis," in *Artificial Neural Networks—ICANN '97: 7th International Conference Lausanne, Switzerland, October 8–10, 1997*, Springer, Berlin, Germany, 1997.
- [25] H. Gudbjartsson and S. Patz, "The Rician distribution of noisy MRI data," *Magnetic Resonance in Medicine*, vol. 34, no. 6, pp. 910–914, 1995.
- [26] F. Huang and G. R. Duensing, "A theoretical analysis of errors in GRAPPA," in *Proceedings of the 14th ISMRM Scientific Meeting*, Seattle, Wash, USA, May 2006.
- [27] H. Wang, D. Liang, K. F. King, G. Nagarsekar, Y. Chang, and L. Ying, "Improving GRAPPA using cross-sampled autocalibration data," *Magnetic Resonance in Medicine*, vol. 67, no. 4, pp. 1042–1053, 2012.
- [28] H. Wang, L. Tam, E. Kopanoglu, D. C. Peters, R. T. Constable, and G. Galiana, "Experimental O-space turbo spin echo imaging," *Magnetic Resonance in Medicine*, vol. 75, no. 4, pp. 1654–1661, 2016.
- [29] Y. Chang, H. Wang, Y. Zheng, and H. Lin, "Instrument Variables for Reducing Noise in Parallel MRI Reconstruction," *BioMed Research International*, vol. 2017, Article ID 9016826, 2017.

Research Article

Scenario Screen: A Dynamic and Context Dependent P300 Stimulator Screen Aimed at Wheelchair Navigation Control

Omar Piña-Ramirez , Raquel Valdes-Cristerna , and Oscar Yanez-Suarez 

Neuroimaging Research Laboratory, Electrical Engineering Department, Universidad Autonoma Metropolitana Unidad Iztapalapa, Av. San Rafael Atlixco 186, Col. Vicentina, Iztapalapa, 09340 Ciudad de México, Mexico

Correspondence should be addressed to Omar Piña-Ramirez; delozath@gmail.com

Received 4 August 2017; Revised 21 December 2017; Accepted 4 January 2018; Published 14 February 2018

Academic Editor: Sheng-Yu Peng

Copyright © 2018 Omar Piña-Ramirez et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

P300 spellers have been widely modified to implement nonspelling tasks. In this work, we propose a “scenario” stimulation screen that is a P300 speller variation to command a wheelchair. Our approach utilized a stimulation screen with an image background (scenario snapshot for a wheelchair) and stimulation markers arranged asymmetrically over relevant landmarks, such as suitable paths, doors, windows, and wall signs. Other scenario stimulation screen features were green/blue stimulation marker color scheme, variable Interstimulus Interval, single marker stimulus mode, and optimized stimulus sequence generator. Eighteen able-bodied subjects participated in the experiment; 78% had no experience in BCI usage. A waveform feature analysis and a Mann–Whitney *U* test performed over the pairs of target and nontarget coherent averages confirmed that 94% of the subjects elicit P300 ($p < .005$) on this modified stimulator. Least Absolute Shrinkage and Selection Operator optimization and Linear Discriminant Analysis were utilized for the automatic detection of P300. For evaluation with unseen data, target detection was computed (median sensitivity = 1.00 (0.78–1.00)), together with nontarget discrimination (median specificity = 1.00 (0.98–1.00)). The scenario screen adequately elicits P300 and seems suitable for commanding a wheelchair even when users have no previous experience on the BCI spelling task.

1. Introduction

Noninvasive Brain-Computer Interfaces (BCI) are systems that translate the electrical brain activity measured through Electroencephalography (EEG) into executable commands for any enabled device [1]. BCI were developed to augment the communication and environment interaction possibilities for patients with severe motor disabilities such as amyotrophic lateral sclerosis, spinal cord injury, or locked-in syndrome [2–4].

The P300 speller is a computer based dictation device controlled through the P300 Event-Related Potential (ERP) which is a cognitive brain response elicited by the stimulation-dependent (synchronous) *oddball* paradigm [1, 2, 4]. The P300 speller is one of the most used BCI, whose conventional stimulation screen has solid-homogeneous black color background with a symmetric and homogeneous matrix arrangement of 6×6 stimulation markers: 26 English alphabet characters, nine decimal digits, and the

underscore for blank space. The stimulation consists of a random sequence of marker flashes on single or multimarker mode. Then, when the users perceive a flash stimulus on the symbol to which they are focused on, a P300 ERP is elicited [2, 4]. Finally, the spelling task process relates an automatic detection of the P300 to the letter that generates it [2, 4–6].

P300 speller variations have included matrix size, marker arrangement, marker types, stimulus sequence, and stimulus presentation and have been tested to increase the information transfer rate and the detection rate and even to perform nonspelling tasks. A summary of these variations is described next, to contextualize the stimulation screen presented in this work.

Sellers et al. [7] used a 3×3 P300 speller matrix and estimated the optimal Interstimulus Interval (ISI) and Stimulus Duration (SD) to achieve detection and transfer rates similar to the 6×6 conventional size speller. Colwell et al. [8], Jin et al. [9–11], and Shi et al. [12] utilized rectangular speller matrices of 9×8 , 7×12 , and 6×12 , respectively. Regarding the stimulus

presentation, some successful approaches were the submatrix stimulation sequences [13], disperse multimarker stimulation [9–11, 14], blue/green color scheme of stimulation markers for nonflashing/flashing states [15–17], and face paradigm [9, 18–20]. Furthermore, all these variants had a solid black color as background. In respect of P300 detection performances, high rates were reached for all the variations mentioned above.

The geometric variation named *Geospell* [21] displayed screen sequences instead of flash events. Each screen had six characters in a circular arrangement, whose center has a cross symbol where the subjects fix their attention. Each group of six characters presented corresponded to the conventional 6×6 matrix rows and columns. Another geometric variation was the lateral single character speller, proposed by Pires et al. [5], which rearranged the stimulation markers laterally on the screen preserving a regular geometry. Its stimulus was presented in single marker mode that pseudorandomly alternates between right and left. Both geometric variations demonstrated high target detection and transfer rates. For all these speller variants, backgrounds remain solid black color.

An approach of the Internet browsing task using a P300 speller-like variation was reported by Mugler et al. [22]. A pair of monitors were used, one for stimulus presentation and the other for web browser display. A homogeneous 8×8 stimulation matrix in row/column mode was used. On the other hand, Yu et al. [23] implemented an oddball paradigm to move the mouse cursor into the web browser. Additionally, Halder et al. [18] also used two monitors, one of them for face paradigm stimulus presentation of a P300 speller that was reshaped according to the number of available browsing commands, each represented with alphabet characters. On another monitor, the web browsing task was performed.

Carabalona et al. [24] and Aloise et al. [25] adapted the conventional speller to control a domotic system. Both cases used icons as stimulation markers instead of alphabet letters. However, they used the typical flashing and background color schemes. A different stimulation screen for domotics was proposed by Hoffmann et al. [26] where RGB images were placed over the solid white background as stimulation markers arranged on a rectangular 3×2 matrix. Meanwhile, the smart house control proposed by Corralejo et al. [27] used RGB images and their corresponding text as stimulation markers; the background was solid bright green; the text and its outline were, respectively, white and black. In addition, Ganin et al. [28, 29] demonstrated that moving and consequently asymmetrically arranged stimulation markers elicited P300 on a three-trial paradigm obtaining an average hit rate of 0.8. This paradigm was used in [29] to control a puzzle game, whose stimulation markers are circular puzzle pieces, each labeled with a Cyrillic character.

P300-based wheelchair navigation consists of two components: the P300-based selection of the displacement command and the robotic system that performs the navigation. In this context, there are, essentially, two schemes for selecting the navigation routes: selecting a direction in which a fixed displacement will occur (with stimulation sequences after each fixed displacement) or selecting a destination from some sort of local or global map (with stimulation sequences after attaining each selected destination). Rebsamen et al. [35, 36]

introduced a P300 3×3 stimulation matrix that had names of apartment rooms and appliances like TV, lights, bed, and so on, as stimulation markers. Its robotic navigation system managed to reach the localization with no additional stimulation sequences while the wheelchair was in transit. All the paths and destinations were predefined. That is to say, the approach implemented stimulation sequences favoring long and complex navigation routes on known scenarios. On the other hand, two approaches introduced a stimulation screen that utilized image-arrows as stimulation markers over a solid color background, emulating a joystick wheelchair. Lopes et al. [34] presented the stimulation in single marker mode for commanding an actual robotic navigation of the wheelchair. Differently, Gentiletti et al. [33] utilized row/column mode in a virtual navigation controller. Both works implemented stimulation sequences after fixed displacements, since each selection corresponds to a short piece of route. Though the joystick stimulators enable navigation on unknown scenarios, a high amount of selections is needed to perform long pathways; therefore, the participants fatigue and lower P300 detection rates were obtained [38, 39]. Notwithstanding, Gentiletti et al. [33] incorporated some kind of destination-oriented stimulation by including a control automata for *wall-following*. Other three approaches incorporated the current information of the navigation scenario into the stimulation screen. Notably, Iturrate et al. [32] integrated a real-time rendered virtual representation of the actual navigation scenario into the stimulation screen, where floor, walls, and obstacles were distinguishable. The stimulation markers for navigation were presented as dots in an arrangement that indicates distance and direction using a polar coordinate system grid centered at the wheelchair. Additional stimulation markers for auxiliary commands were shown on a menu bar; the stimulus presentation was in row/column stimulation mode with a fixed displacement sequence scheme.

Similarly, Escolano et al. [30] implemented a polar grid and menu bar commands, but using augmented reality. That is to say, the virtual representation of the scenario screen, as in [32], was set to low opacity and overlaid to its real counterpart. Thus, a mixed image was presented as stimulation screen whose markers performed row/column stimulation mode with stimulation sequences for fixed displacement according to which a teleoperated robot was controlled. In order to change the robot field of view, a stimulation screen with a gray scale image as background and a square grid of circles as stimulation markers was implemented. In another study, Escolano et al. [30] reported another P300-based approach for commanding a teleoperated robot. It used a gray scale image of 60×60 pixels resolution. The stimulation markers were six icons split in two menu bars, and six circles for navigation points which were arranged in a rectangular grid of 2×3 . All of them were in fixed positions over the screen performing row/column stimulation for fixed displacement.

This work introduces the *scenario screen* stimulator, which is a P300 speller variant. The scenario screen has an image background that is a snapshot of the wheelchair perspective and whose stimulation markers are located over relevant landmarks. To contrast the scenario screen features with other reported P300-based stimulation screens for

TABLE 1: Comparison among several works where P300 speller variants were used to generate commands for a wheelchair navigation control system. t - nt diff: target–nontarget difference. *Included due to the asymmetric markers arrangement.

	Stimulation <i>Color scheme</i> <i>Mode</i> <i>Presentation</i>	Background <i>Type</i> <i>Color depth</i>	Markers <i>Number</i> <i>Encoding</i> <i>Arrangement</i>	P300 analysis # subjects <i>Latency</i> <i>t</i> - <i>nt</i> diff.
Scenario screen (this work)	Blue/green Single marker Destination-based and fixed displacements	Real image 24-bit	12 Unicode Asymmetrical	18 Yes Yes
Escolano [30, 31]	White-green/blue Row/column Fixed displacements	Real image Gray scale and 24-bit	12 Icon-images Symmetrical	2 No, yes No
Iturrate [32]	Black/blue/white Row/column Fixed displacements	Virtual image 8-bit	20 Icon-images Symmetrical	5 No No
Gentiletti [33]	Gray/white Row/column Fixed displacements	Solid black 5-bit	12 Images + text Symmetrical	2 No No
Lopes [34]	Gray/green Single marker Fixed displacements	Solid black; 5-bit	7 Arrows + text Symmetrical	11 No No
Rebsamen [35, 36]	Blue/white Single marker Destination-based	Solid gray 5-bit	7 Box + text Symmetrical	5 No Yes
Ganin [28, 29]*	Image/white Single marker Fixed displacements	Solid gray 24-bit	9 Image + text Asymmetrical	14 Yes Yes

navigation tasks, a comparative chart is shown in Table 1. Except for the scenario screen, all the other approaches utilized symmetric stimulation markers arrangements and were focused on the evaluation of the wheelchair navigation task. Although some studies indirectly reflected that an image background does not impact the target detection, further analysis in terms of the P300 elicitation, stimulus presentation, and classifier adaptation is needed.

2. Materials and Methods

2.1. Overview. A scheme of the wheelchair navigation system within which the scenario screen is incorporated is shown in Figure 1. It is briefly described in the following, for the sake of establishing the role of the stimulator that is the subject of this document.

The Navigation Core component coordinates the workflow and the datapath throughout all the components of the system. The Scenario Analysis component takes photographs of the current frontal wheelchair perspective which is analyzed by computer vision methods searching for landmarks, in particular navigation paths, doors, windows, and wall signals. The P300 Path Selector component produces two different stimulation screens. The first is the proper scenario screen, which is the default stimulator screen. Its stimulation markers are located over the most relevant landmarks according to Scenario Analysis component results; the second,

for unexpected, finer-control situations, a P300 joystick-like variant is displayed instead. The robotic navigation system controls all the devices and sensors mounted on the wheelchair for driving it safely to its destination, either with the destination-based selections from the scenario screen or following fixed displacement commands from the joystick screen.

Figure 2 exemplifies two realizations for navigating from point 1 to point F, based on the scheme detailed above. The navigation based on the scenario screen (S NAV) only performs stimulation sequences at points 1, 2, and 3 (destination-based sequences), where stimulation marker localizations vary according to the image background. For H NAV navigation, the scenario screens and routes are the same as in S NAV. Nonetheless, an eventuality (obstacle) occurs at point 3A (between 2 and 3) and then the joystick-like screen is shown. Several fixed displacement stimulation sequences are run with this screen to sort the obstacle, and the previous navigation path is resumed to reach point 3 where scenario screen is displayed again.

The robotic navigation system is currently under active development in our research team.

2.2. Scenario Screen Development. Figure 1 shows the three components of the scenario screen: stimulation screen, stimulus generator, and pattern recognition system. In this section, the first two components are detailed.

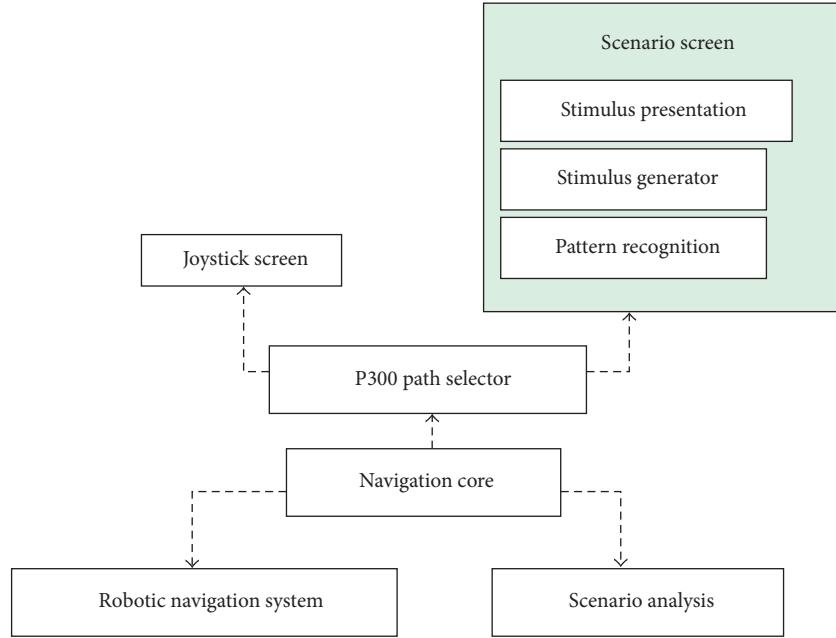


FIGURE 1: Schematic diagram of a robotic wheelchair navigation system where the scenario screen is utilized for the navigation command. See text for the description.

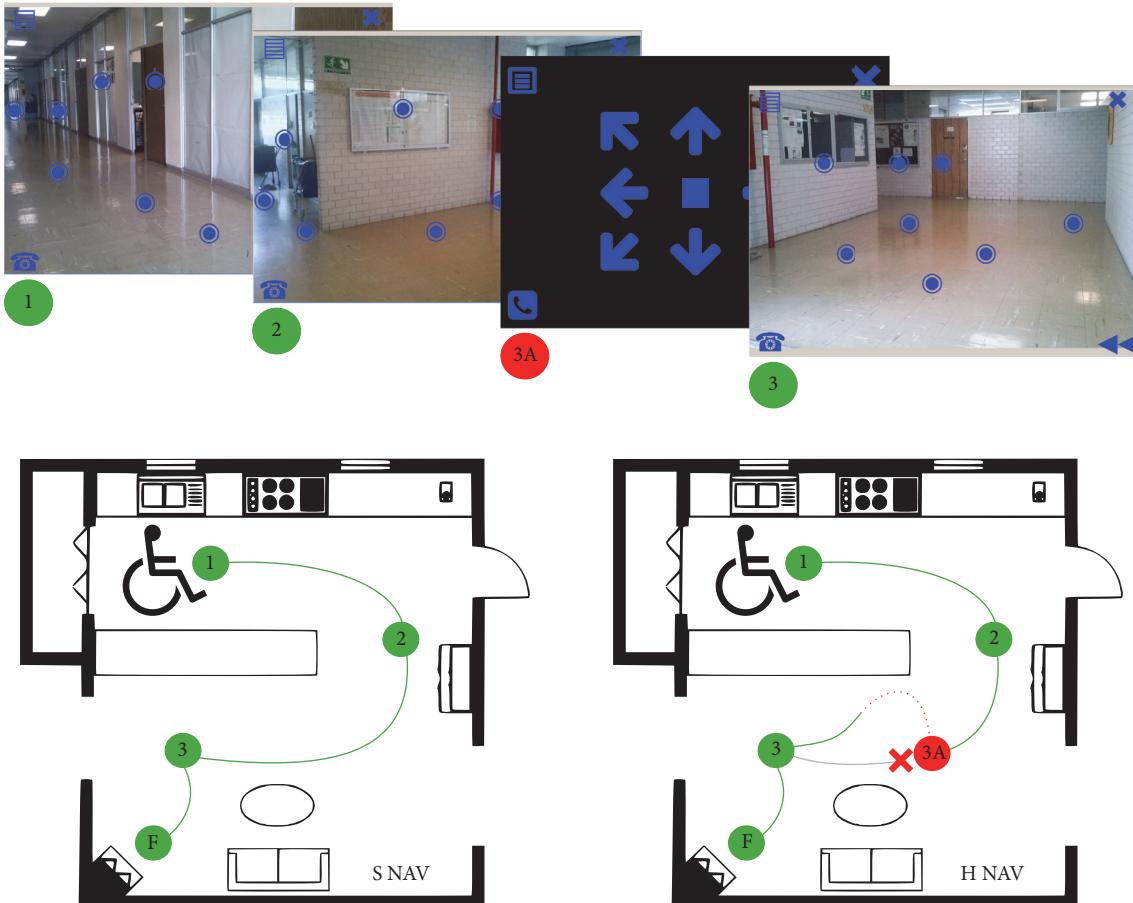


FIGURE 2: Scenario screen navigation scheme. Each green circle represents where the wheelchair stops to generate the scenario screen. The green lines are the paths followed by the wheelchair. The red point (3A) is where an obstacle interrupts the navigation. The dotted red line represents a fixed displacement commanded navigation.



(a) Stimulation screen with no stimulus presentation
(b) Stimulation screen showing a single marker stimulus

FIGURE 3: Scenario screen. In this instance, eight markers are arranged in landmarks as follows: five on the floor, two on wall signs, and one on a door. Four markers placed at corners were reserved for other tasks. The background is an unprocessed RGB image of a corridor in the research building.

2.2.1. Design and Implementation. The scenario screen is a P300 speller variant that enables users to select a destination of navigation. Figure 3 shows the scenario screen evaluated in this work. As it can be seen, two were the major variations implemented. First, it incorporates the current scene information into the stimulation screen through an RGB image background that depicts the wheelchair navigation perspective, instead of using the conventional solid color background or a virtual-real combination. In this case the image corresponded to an RGB 1024×825 pixel snapshot of the research building hall at the University, photographed with a 2.2 Mpx mobile phone camera, with no preprocessing. Second, only twelve stimulation markers, grouped into navigation and auxiliary, replaced the original 36 characters. Using the Unicode fisheye character, the eight navigation markers were located over the most relevant landmarks, in this case: three markers for the main corridor, two for the secondary corridor, one for the door, and two for the wall signals. The four auxiliary markers also used Unicode characters, each having a static placement on screen corners. Their placement, meaning, and encoding are as follows: top left: “open menu” task; bottom left: “call to a predefined contact”; top right: “close stimulation screen”; bottom right: “return to previous wheelchair location.” For this work’s purposes, all the stimulation markers were manually located. The scenario screen was implemented by modifying the Native OpenViBE P300 speller GUI [40]. The 6×6 stimulation matrix was enlarged to 9×9 . Except for the twelve stimulation markers, the 81 available labels of the enlarged matrix were not displayed and all the label backgrounds were set to zero opacity to make GUI canvas visible, with the image background set within [41]. Thus, users see the stimulation screen as shown in Figure 3.

2.2.2. Stimulus Presentation. Nonflashing and flashing color scheme were blue/green with RGB encoding $(0, 0, 255)$ and $(0, 255, 0)$, respectively, as used in [15, 16]. The font size of the twelve stimulation markers was fixed to 120 points for both states. These customizations were made with OpenViBE designer tool v0.14 [40].

2.2.3. Sequence Generator Development. A new stimulus/target sequence generator engine was developed to deal with the shortcomings of its OpenViBE P300 speller stimulus generator counterpart. The stimulator engine developed implemented the following features: pseudorandom ISI, single marker stimulus mode, and a strategy to avoid two consecutive stimulations of the same marker. In order to perform the stimulus presentation, this engine interfaced with OpenViBE designer core through proper native OpenViBE ID directives. Furthermore, the stimulus generation was implemented with two modules: the first of them was a Python script that computed and stored a large set of paired stimulus-timing sequences, while the second module was a lua script executed by OpenViBE designer during the acquisition process; it loads the stored stimulation sequences and implements the scenario screen stimulation paradigm described in the following sections. Figure 3(b) shows the single marker stimuli as seen by users.

2.3. Scenario Screen Paradigm. This section describes the paradigm to evaluate the ERP elicitation capabilities of scenario screen. EEG acquisition setup and subject selection are also detailed.

2.3.1. Scenario Screen Stimulation Paradigm. Following the OpenViBE nomenclature, a trial was a random flashing sequence of the 12 available scenario screen stimulation markers. All flash durations were 0.125 s. The ISI produced by a stimulus generator took values among 0.12, 0.14, 0.16, 0.18, 0.20, and 0.22 s. Intertrial Interval (ITI) was set to 0.5 s. In this form, a repetition (Figure 4: S4) was the sequence of 10 trials performed to select one marker. All repetitions run in copy mode indicate the target to subject by blinking it five times at 1.0 Hz (Figure 4: S3). Finally the Interrepetition Interval was set to 5.0 s.

The initial stimulation screen state is shown in Figure 4: S0. To indicate task starting time, the twelve available stimulation markers were turned to green during 5 s (Figure 4: S1) after which they all returned to blue. 20 s baseline EEG was acquired (Figure 4: S2); previously the subjects were

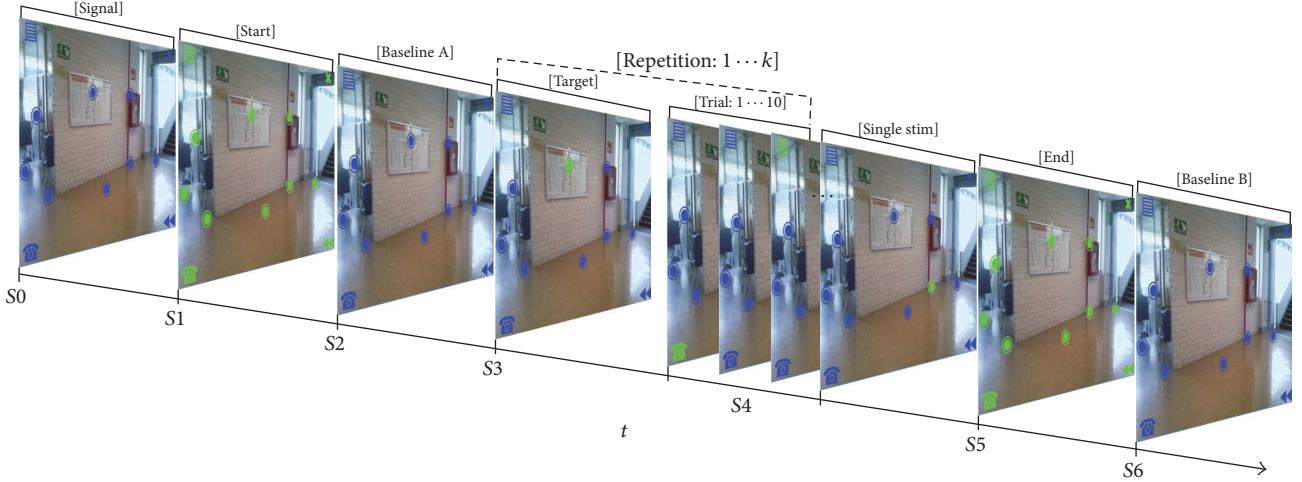


FIGURE 4: Stimulation sequence repetition. The sequences outline how each stimulation screen state is shown to the user. S_i represents the i th state and $k = \{3, 5, 6, 6\}$ the number of repetitions.

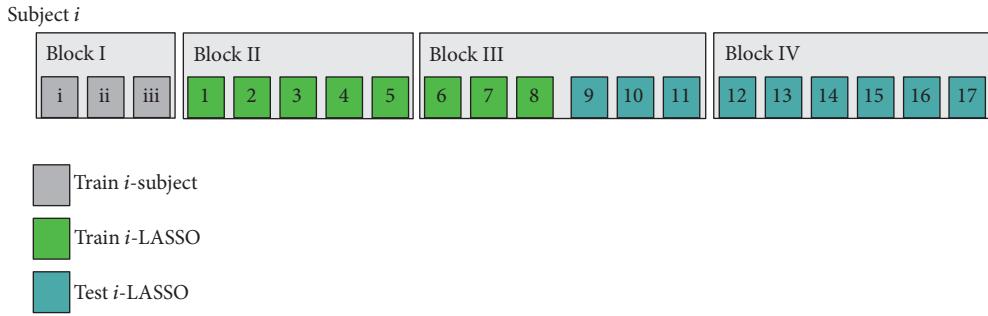


FIGURE 5: The four blocks which comprise the acquisition session for the i th subject. Each color filled square represents a repetition.

instructed to stay calm and to remain with open eyes but blink and breathe normally. Subjects performed four blocks of $k = \{3, 5, 6, 6\}$ repetitions as Figures 5 and 4: S3 and Figure 4: S4 show. Once the last repetition of the current block was accomplished, the block's ending indication was shown with 5 s of green markers (Figure 4: S5). Then, another 20 s of baseline EEG was recorded (Figure 4: S6).

Block order was preserved for all subjects. Block I was used to train subjects on the correct usage of the scenario screen; thus the data from this block was not considered for analysis purposes. Data from Blocks II, III, and IV provided a set of 17 repetitions for analysis; see Figure 5. Given that all these 17 repetitions were performed in copy mode, a pseudorandom strategy that only allowed each of the available markers to be selected as the target once or twice was implemented in the developed stimulus generator engine. Finally, resting time within blocks was from 30 to 60 s during which the subjects answered the questionnaire previously detailed.

2.3.2. Experimental Setup. EEG recordings utilized eight gold surface passive electrodes (Fz, C3, Cz, C4, P3, Pz, P4, and Oz) fixed to an extended 10–20 g.tec cap (Guger Technologies™). The reference was the joint A1-A2 connected through a pair of gold earlobe clip electrodes. A gold cup electrode at right mastoid was the ground. In all cases the

electrode-gel-skin impedance was lower than $10\text{ k}\Omega$ [33, 42]. A g.USBamp biosignal amplifier (Guger Technologies) was used for the recordings with the following configuration: 512.0 Hz sample rate, 8th-order 0.1–30.0 Hz passband, and 4th-order 58.0–62.0 Hz notch Butterworth hardware filters. For this configuration, the OpenViBE acquisition server v0.14 tool was used [40].

Nineteen able-bodied subjects (20–35 years) participated in the experiment; 15 of them had no previous experience in BCI usage. All subjects gave their informed consent by signing the corresponding form and were asked to have at least six hours of sleep the night before the study. Exclusion criteria were intake of stimulant or depressive substances within the six previous hours to the study, any pathological or psychological condition, and light flashing sensitivity. Subjects sat in front of a 22" LCD in a comfortable position. They were asked to always blink normally and move only if it was necessary. Eyes–LCD distance was approximately 1.2 m. They were instructed on how to use the scenario screen and also to count mentally how many times the target flashed [42].

2.4. Data Analysis. Data from each subject was separately analyzed in two stages:

- Global P300 elicitation: the 17 target coherent averages of each subject were compared with their

corresponding nontarget counterpart. A waveform analysis of target averages searching P300 features and statistical tests of difference between target and nontarget averages were performed.

- (ii) Evaluation of the automatic target/nontarget detection: a classifier was trained for each subject for automatic labeling of the target and nontarget stimulation markers. Trained classifiers labeled unseen data and their performances metrics were reported, namely, sensitivity, specificity, and accuracy. The behavior of these metrics was analyzed with respect to the number of trials used to make the decisions; group and individual analysis are reported.

2.4.1. Preprocessing Data. The same offline preprocessing was applied for the two analysis stages. The EEG signals were filtered with a 40th-order 1.5–10.0 Hz FIR passband filter designed with a Chebyshev window. Event-related epochs of 307 samples (600 ms) were extracted from signals. Each epoch was detrended and normalized to zero mean and unit variance.

2.4.2. Global P300 Elicitation. Each subject's target and non-target coherent grand averages were computed using the full data set (Blocks II, III, and IV) combining the eight electrode epochs. A waveform analysis was performed on the target averages observing for the occurrence of characteristic P300 features, such as the peak within 150–450 ms, in addition to N100, N200, and P200 ERP. Two tailed corrected Mann–Whitney *U* tests (MWUT) were computed for each pair in order to test whether there were differences between target average and its nontarget counterpart. Those statistical comparisons were computed using a window spanning 150 to 450 ms after stimulus onset, averages where typically the P300 appears [5, 43–45]; two tailed corrected *p* values are reported.

2.4.3. Automatic Target Detection: Classifier Training. Feature vectors \mathbf{x} were obtained directly from the preprocessed epochs; each of them were decimated by a factor of two and concatenated giving vectors of 1228 features (*d*). These vectors were arranged by temporal repetition occurrence, stimulation marker ID, and temporal trial occurrence.

Let $\{(\mathbf{x}_i, y_i) : \mathbf{x}_i \in \mathbb{R}^d, y_i \in \{-1, 1\}\}$ be the training data set, where \mathbf{x}_i is the *i*th feature vector and y_i is its corresponding target or nontarget label. The *Least Absolute Shrinkage and Selection Operator* (LASSO) is a Linear Discriminant Analysis (LDA) variation that uses the l_1 -norm as regularizer:

$$\begin{aligned} \beta^{\text{LASSO}} &= \arg \min_{\beta} \left\{ \sum_{i=1}^N \left(y_i - \beta_0 - \sum_{j=1}^d \beta_j x_{ij} \right)^2 + \sqrt{\lambda} \sum_{j=1}^d |\beta_j| \right\}. \quad (1) \end{aligned}$$

The sparsity parameter λ needs to be adjusted; however, the optimal λ can be estimated through a cross-validation process. LASSO approach is preferred since some β components become equal to zero according to the sparsity

parameter $\lambda \in (0, 1]$. Thus, it is also considered a feature selection method [46–48].

Let the reduced feature vector \mathbf{x} correspond to the nonzero features of $\mathbf{w} = \beta^{\text{LASSO}}$. The score denoted by $s(\mathbf{x})$ is computed with the discrimination function (see (2)) [49–51] that measures the feature vector membership to the target or nontarget class, as a distance to the regression hyperplane:

$$s(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b. \quad (2)$$

A LASSO-LDA was trained for each subject using the data from their corresponding Blocks II and III as Figure 5 shows. No class balancing method was utilized (Python machine learning library: scikit-learn LASSO through LARS method and sparsity parameter λ was estimated by cross-validation [52]). Target labeling was performed in two stages: (1) to score the feature vectors \mathbf{x} (see (2)) of a given repetition and (2) to select the target through a voting scheme based on the score accumulation (see (3)) [49, 51, 53]:

$$\text{target} = \arg \max \left\{ \sum_{l=1}^L \mathbf{w}^T \mathbf{x}_{(m,l)} : m \in \{1, \dots, 12\} \right\}, \quad (3)$$

where $\mathbf{x}_{(m,l)}$ is the feature vector associated with the *m*th stimulation marker on the *l*th trial over a given repetition. Note that the bias b of (2) was not considered for calculation since it is constant across all feature vectors [50, 51].

2.4.4. Automatic Target Detection: Classifier Performance Evaluation. Unseen epochs from the last nine repetitions of Blocks III and IV were labeled as target and nontarget via trained LASSO-LDA and voting scheme described above. Three were the metrics for classifier evaluation: *sensitivity*, also known as true positive rate, hit rate, or recall, *specificity* or true negative rate, and *accuracy* or the fraction of data correctly classified [53]. These metrics were reported for user sample and individually with respect to the number of trials scored (*l*).

3. Results and Discussion

3.1. Stimulation Screen Design and Development. A seven-item questionnaire was answered by users after the stimulation session. Closed-form yes/no questions asked about stimulus perception, discomforts, losing attention, and stimulus sequence patterns. The responses reflected that none of the users perceived tearing, pain, or any discomfort related to stimulation, even in large stimulus sequences (19.7 ± 0.3 min for the four blocks). Also, all stimulation markers with the blue/green color scheme were perceptible in any circumstance. Five of the novice BCI subjects reported failing to perceive some target flashes or getting somnolent, both circumstances at the end of the last block, which probably affected the classification performance, but not the global coherent averages. This somnolence phenomenon was, probably, caused by the copy mode static scenario screen that provides no feedback to users. Consequently, the P300 elicitation and its detection rate were affected. Decision to maintain the same background image and marker

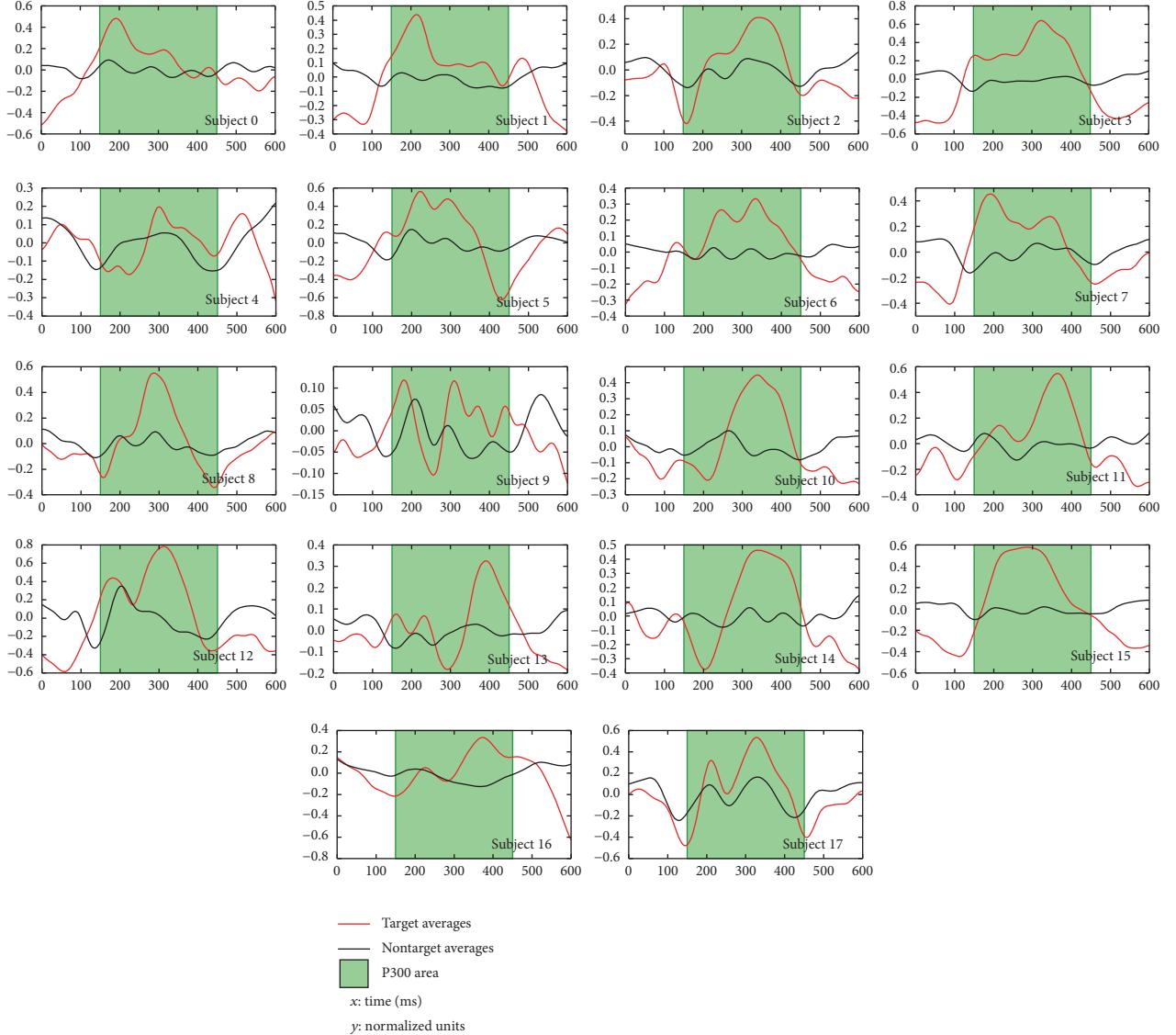


FIGURE 6: Each subject target (red) and nontarget (black) coherent grand average plots. X -axis is the 0–600 ms span, and Y -axis is the average amplitude in μV ; these averages were computed with z -scored epochs (mean: 0 and SD: 1). The green filled areas are the 150–450 ms span.

distribution in performing the evaluation for all subjects was taken to avoid biases due to these factors. With the evidence that scenario screens are useful, there is confidence for implementing scenario and marker changes as described in Section 2.1. Regarding the stimulus generator developed, no subject reported having perceived target or stimulus patterns. Furthermore, they also reported perceiving the ISI variability that likely caused more expectation. Therefore, our developed target and stimulus sequences generators demonstrated being usable since they were comfortable for users. In addition, while OpenViBE was sufficient for scenario screen evaluation, it currently presents limitations for implementing a free-mode scenario screen session (the actual implementation for navigation is developed over a different architecture).

These results overcame the red/white color scheme used in [41] where four of eight subjects reported some type of discomfort. This red/white scheme was said to present no

clearly perceptible stimulus when the markers were placed over bright image areas. A further analysis of the native OpenViBE stimulus generator showed that its sequences always alternated row and column stimulus for square arranged stimulation markers [41]. Furthermore, the scenario screen used a high resolution real image background with the twelve available stimulus perfectly perceivable; furthermore, the stimulation markers' meaning depends on where they were located over the image background.

3.2. Global P300 Elicitation. Paired target and nontarget great coherent averages plots from each of the 18 subjects are shown in Figure 6. One subject was discarded as he slept during two blocks of the experiment. The target averages, in all cases, were different from their respective nontarget averages. 72% of the target averages (subjects 1, 2, 4–8, 10–15, and 17) presented positive peaks within 150–450 ms (Figure 6 green filled

area) that are consistent with a P300 elicitation. Furthermore, these peaks are preceded by a negative peak around 200 ms that may be associated with the N200 elicitation. Four subject target averages (0, 3, 6, and 16) had other features not directly associated with expected P300 elicitation, but their respective target and nontarget averages are still different, as is the case for subjects (2, 3, 5, 7, 8, and 17) who had a negative peak within 400–550 ms that might be related to the long latency N400. However, further analysis should be done to identify the meaning of these negative deflections. Those results were consistent with the findings reported by [18, 54, 55]. Finally, subject 9 target average waveform is noisy and does not represent a typical P300 elicitation, yet it is different with respect to its nontarget average counterpart.

The stimulus generator engine developed, particularly, the pseudorandom ISI feature, showed effectiveness on visual steady-state periodic components reduction, as is shown in Figure 6. Except for subject 9, all the coherent grand averages do not have steady-state artifacts taking into account the fact that the preprocessing used a passband 40th FIR filter within 1.5–10.0 Hz; in contrast, when the native OpenViBE stimulus generator was used [41] those filter parameters were 100th and 0.1–2.0 Hz due to the ISI-related components.

Selecting P300 of each subject as the highest peak within 150–450 ms on the target coherent averages (see Figure 6 green areas), all these magnitudes were contrasted with their nontarget analogues by a signed rank Wilcoxon test obtaining a statistical difference ($W = 0.0$; $p < 0.001$) which suggests that attending the target marker generates a different brain response. Furthermore, the median latency of those P300 peaks was 310 (217–337) ms which corresponds to P300 span and is consistent with the results reported in [54–56].

In regard to the statistical analysis of the paired waveforms, it was found that target coherent averages are statistically different from their nontarget pairs (MWUT; $p < 0.005$) considering the 150–450 ms spans, except for the fourth subject. Summarizing, the target perception elicits a response within (150–450) ms differentiable from the nontarget stimulation. To put it differently, the stimulation screen with an image in the background and whose stimulation markers are asymmetrically arranged elicits an adequate response when the subjects perceive the target stimulation in 94% of the cases.

3.3. Target Detection in Scenario Screen

3.3.1. Target Detection Performance Analysis over the Subject Sample. Sample median sensitivity, specificity, and accuracy [57, 58] with respect to number of trials (l) accumulated for the score calculation are summarized in Table 2. These performance metrics were calculated using the unseen nine repetition data of each subject. More precisely, the sensitivity or true positive rate was the metric related to the proportion of target markers correctly classified. From the 6th scored trial ($l = 6$) the sample median was above 0.75 sensitivity which is the minimum accepted on P300 spelling tasks [2, 4, 6, 8]. Moreover, from the 7th trial ($l = 7$), 0.7 sensitivity was found for 75% of the subjects regardless of the fact that 78% of them had no experience on the task. This is graphically

shown in the boxplots of Figure 7(a) where it is seen that the target detection is improved and also the dispersion becomes lower as more trials were scored. This was consistent with the conventional P300 speller knowledge [5, 16, 59], but using a stimulation screen with a snapshot background and no orthographic or semantic stimulation markers.

The performance of five commonly used classifiers for conventional P300 spelling task (LASSO-LDA, shrinkage-LDA, Linear-SVM, Radial Basis Function SVM, and SWLDA) were evaluated for the scenario screen [60]. For this study all the hyperparameters of each classifier were appropriately optimized. No statistical differences were found among the five classifier performances but LASSO-LDA presented a lower variance nevertheless. Thus, only LASSO-LDA was utilized in this work; a reduction from 4 to 2 in the decimation factor with respect to [60] was implemented reflected in lower variances in the subject sample sensitivity as Figure 7 shows. Additionally, partial coherent averages as feature vectors and linear-SVM were tested for P300 detection [41], although the sensitivity was at least 0.7 in most of the cases using 10 trials; the target marker detection rates shown in Table 2 are higher even from five scored trials ($l = 5$).

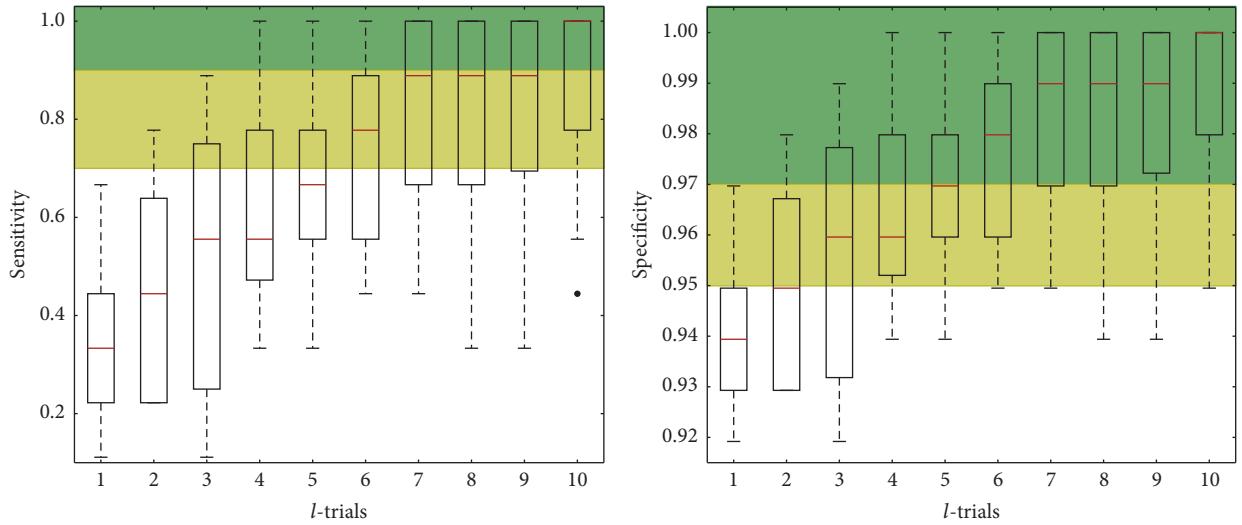
In terms of specificity or true negative rate which was the metric related to the proportion of well labeled nontargets shown in Table 2, the median specificity was higher than 0.97 from $l = 5$ which means that 3 false targets were detected. Figure 7(b) shows the corresponding specificity boxplots with respect to the number of scored trials. From $l = 7$, 75% of the subjects got at most 3 misclassified nontargets. Table 2 also shows the accuracies which might be considered good given that they are always higher than 0.85; however, this is a biased measure given the unbalanced number of target and nontarget feature vectors. For this reason we prefer measuring the performance in terms of sensitivity and specificity. Additionally, when ten scores are accumulated ($l = 10$) the average time to perform a detection is 40 s ($12 \times 10 \times [0.125 + 0.17] + 9 \times 0.5$); these values correspond to the number of available target markers, number of trials, IS, average ISI, number of ITI, and ITI.

On the other hand, the summarized performances of the screens with image backgrounds reported in [30, 31] suggested that a row/column stimulation mode was feasible for symmetrically arranged stimulation markers. However, when this stimulation mode was used in a scenario screen [41], a minimum distance constraint among stimulation markers was used to avoid peripheral stimulus perception of nontargets. In this work, that constraint was no longer necessary because of the single marker mode usage and the high performances obtained.

In sum, the scenario screen was capable of eliciting detectable P300 through LASSO-LDA with performances higher than 0.75 from $l = 6$. The highest median performances were obtained accumulating all the trials ($l = 10$), with 1.0 (0.78–1.00) and 1.0 (0.98–1.00) being the median sensitivity and specificity, respectively. Thus, the scenario screen resulted in a very suitable P300 speller variation for selecting the path for a wheelchair, even when more than 78% of the subjects had no experience in P300 spelling tasks. These results were consistent with [5, 18, 28, 29], even

TABLE 2: Subject sample median sensitivities, specificities, and accuracies with respect to the number of trials scored $l = \{1, \dots, 10\}$. IQR: interquartile range.

l trials	Sensitivity (IQR)	Specificity (IQR)	Accuracy (IQR)
(1)	0.33 (0.22–0.44)	0.94 (0.93–0.95)	0.89 (0.87–0.91)
(2)	0.44 (0.22–0.64)	0.95 (0.93–0.97)	0.91 (0.87–0.94)
(3)	0.56 (0.25–0.75)	0.96 (0.93–0.98)	0.93 (0.88–0.96)
(4)	0.56 (0.47–0.78)	0.96 (0.95–0.98)	0.93 (0.91–0.96)
(5)	0.67 (0.56–0.78)	0.97 (0.96–0.98)	0.94 (0.93–0.96)
(6)	0.78 (0.56–0.89)	0.98 (0.96–0.99)	0.96 (0.93–0.98)
(7)	0.89 (0.67–1.00)	0.99 (0.97–1.00)	0.98 (0.94–1.00)
(8)	0.89 (0.67–1.00)	0.99 (0.97–1.00)	0.98 (0.94–1.00)
(9)	0.89 (0.69–1.00)	0.99 (0.97–1.00)	0.98 (0.95–1.00)
(10)	1.00 (0.78–1.00)	1.00 (0.98–1.00)	1.00 (0.96–1.00)



(a) Median sensitivities with respect to the number of scored trials

(b) Median specificities with respect to the number of scored trials

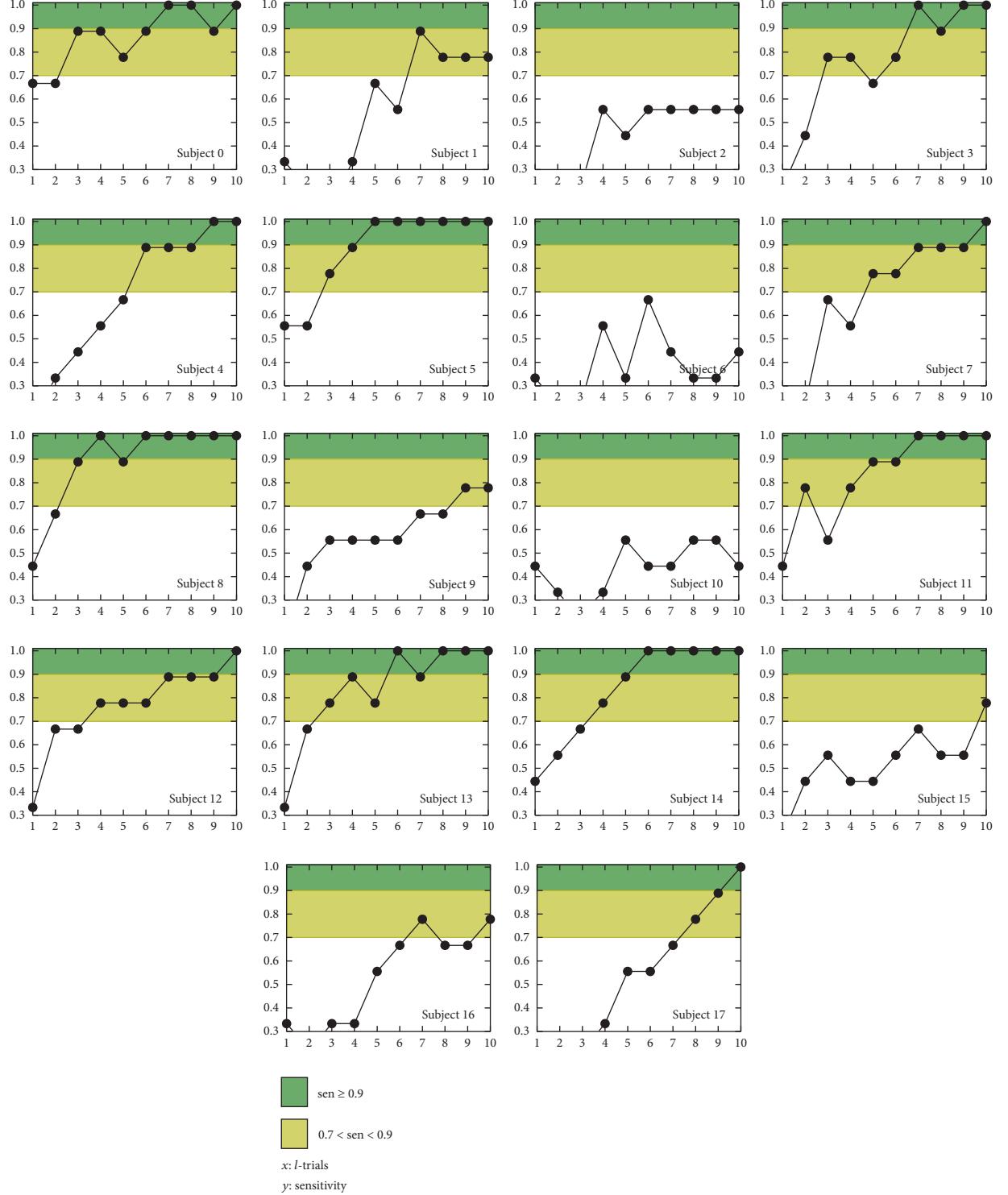
FIGURE 7: Sensitivity and specificity performances with respect to the number of trials.

when the scenario screen stimulation markers did not use an orthographic or semantic scheme; thus the implicit findings reported in [30, 31] in terms of stimulation screen evaluation and target marker detection were also complemented.

3.3.2. Individual Analysis of the Target Detection Performances. Taking into account the 44% of the subjects (0, 3, 4, 5, 8, 11, 13, 14) who reached a plateau with at least two 1.0 sensitivity points (see Figure 8), the median sensitivities for l equal to five, six, and seven are, respectively, 0.78 (0.64–0.81), 0.94 (0.89–1.00), and 1.00 (0.97–1.00) which are high performances even with $l = 5$. For all these subjects the number of trials might be reduced to seven with no performance cost; this repetition shortening would imply a 27.5% time reduction. Moreover 28% of subjects reached high performances with seven scored trials, despite their inexperience on P300 spelling tasks. On the other hand, the 39% of the subjects (1, 7, 9, 12, 15, 16, 17) that had not reached a plateau, got a median sensitivity of 0.78 (0.67–0.89) with seven trials scored; furthermore, the lower quartile

included the minimum performance for spelling tasks that was consistent with the findings reported in [5, 18, 28, 29].

In regard to the three lowest sensitivity performances (subjects 2, 6, and 10) whose average was 0.48 (SD = 0.1), all these subjects presented a performance consistently low despite the number of scored trials. Although this behavior might be associated with subjects' wrongly executed task, when LASSO-LDA was substituted by a Linear Support Vector Machine (Linear-SVM) for subject 2, the detection rate had improved up to 57% in terms of sensitivity. That behavior suggested that a customized classifier selection is likely to improve the detection rates [54, 55, 60]. A similar case was subject 15 whose performance had improved when a Radial Basis Function SVM was utilized. A whole behavior comparison between LASSO-LDA and the best classifier of subjects 2 and 15 with respect to the number of scored trials is shown in Figure 9. Notwithstanding other four classifiers (Linear-SVM, Radial Basis Function SVM, Shrinkage-LDA, and SWLDA) tested in subjects 6 and 10, no sensitivity improvement was reached. However as Figure 8 shows



subjects 6 and 10 had sensitivities of 0.67 and 0.56 with six and eight scored trials, respectively. These results reinforced the necessity of an exhaustive customization of the pattern recognition method.

Performing a similar analysis for the specificity, the results obtained for all subjects were high, always greater than 0.92, which implied that only a small amount of nontargets were labeled as targets (Figure 10 and Table 2). Median

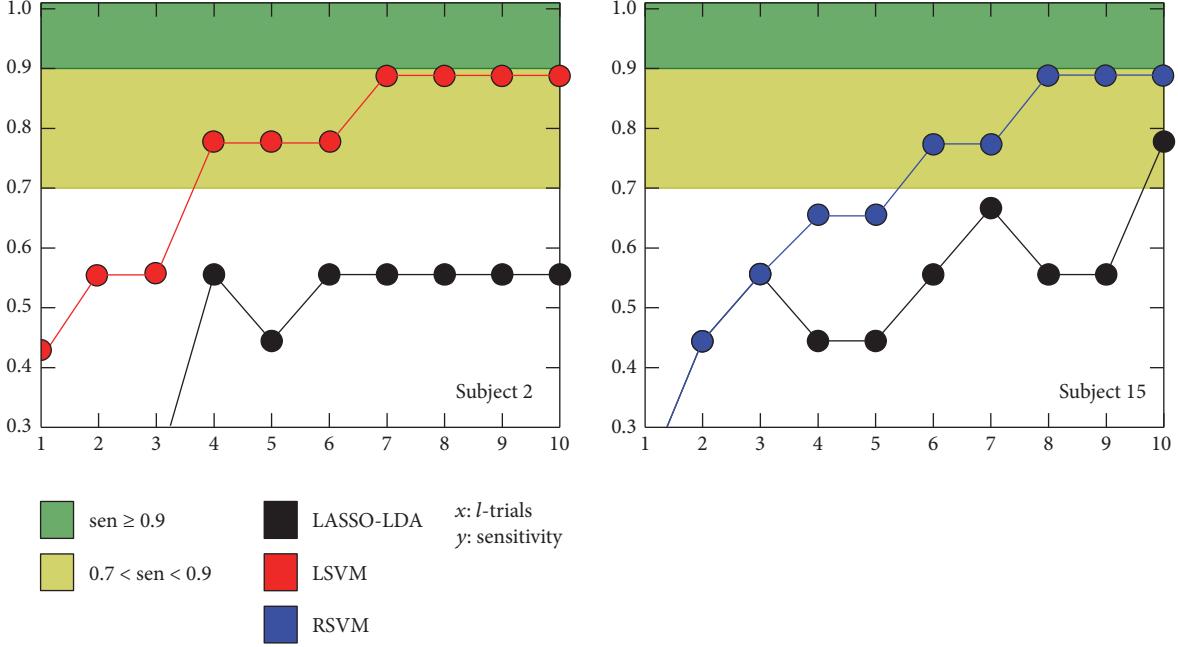


FIGURE 9: Sensitivity with respect the number of scored trials. The black line is LASSO-LDA, while red and blue ones correspond to LSVM and RSVM, respectively.

TABLE 3: Individual sensitivity, specificity, and accuracy achieved when ten trials are scored ($l = 10$). j and k are, respectively, the number of correctly classified targets and nontargets. Meanwhile, the correctly labeled target and nontarget count is m . AUROC: area under the receiver operating characteristic curve [37]. SD: standard deviation.

Subject	Sensitivity (j -correct)	Specificity (k -correct)	Mean AUROC (SD)	Accuracy (m -correct)
17	1.00 (9)	1.00 (99)	0.98 (0.02)	1.00 (108)
11	1.00 (9)	1.00 (99)	0.99 (0.01)	1.00 (108)
3	1.00 (9)	1.00 (99)	1.00 (0.00)	1.00 (108)
4	1.00 (9)	1.00 (99)	1.00 (0.01)	1.00 (108)
5	1.00 (9)	1.00 (99)	1.00 (0.00)	1.00 (108)
7	1.00 (9)	1.00 (99)	0.98 (0.02)	1.00 (108)
8	1.00 (9)	1.00 (99)	1.00 (0.01)	1.00 (108)
12	1.00 (9)	1.00 (99)	1.00 (0.00)	1.00 (108)
13	1.00 (9)	1.00 (99)	0.85 (0.09)	1.00 (108)
14	1.00 (9)	1.00 (99)	0.99 (0.01)	1.00 (108)
0	1.00 (9)	1.00 (99)	0.96 (0.05)	1.00 (108)
9	0.78 (7)	0.98 (97)	0.95 (0.06)	0.96 (104)
16	0.78 (7)	0.98 (97)	0.91 (0.05)	0.96 (104)
15	0.78 (7)	0.98 (97)	0.95 (0.03)	0.96 (104)
1	0.78 (7)	0.98 (97)	0.86 (0.06)	0.96 (104)
2	0.56 (5)	0.96 (95)	0.92 (0.04)	0.93 (100)
10	0.44 (4)	0.95 (94)	0.87 (0.06)	0.91 (98)
6	0.44 (4)	0.95 (94)	0.87 (0.07)	0.91 (98)

misclassified number of targets for $l = 10$ was 1.5 (1.0–3.75); this result, on the scenario screen, meant that the destination was not correct.

When ten scored trials were considered, 11 of 18 subjects (61%) accomplished the correct selection of 100% of the targets, despite eight of them (44%) having no previous experience on P300 spelling task. This is summarized in

Table 3. Additionally, four subjects (22%) detected correctly seven of nine targets while three subjects (17%) detected between four and five. However, when LSVM and RSVM are, respectively, used for subjects 2 and 15, there was a global correct target improvement due to 72% of the subjects having a sensitivity higher than 0.87. In contrast, only subjects 6 and 10 (11%) had four of nine correctly detected targets. Although

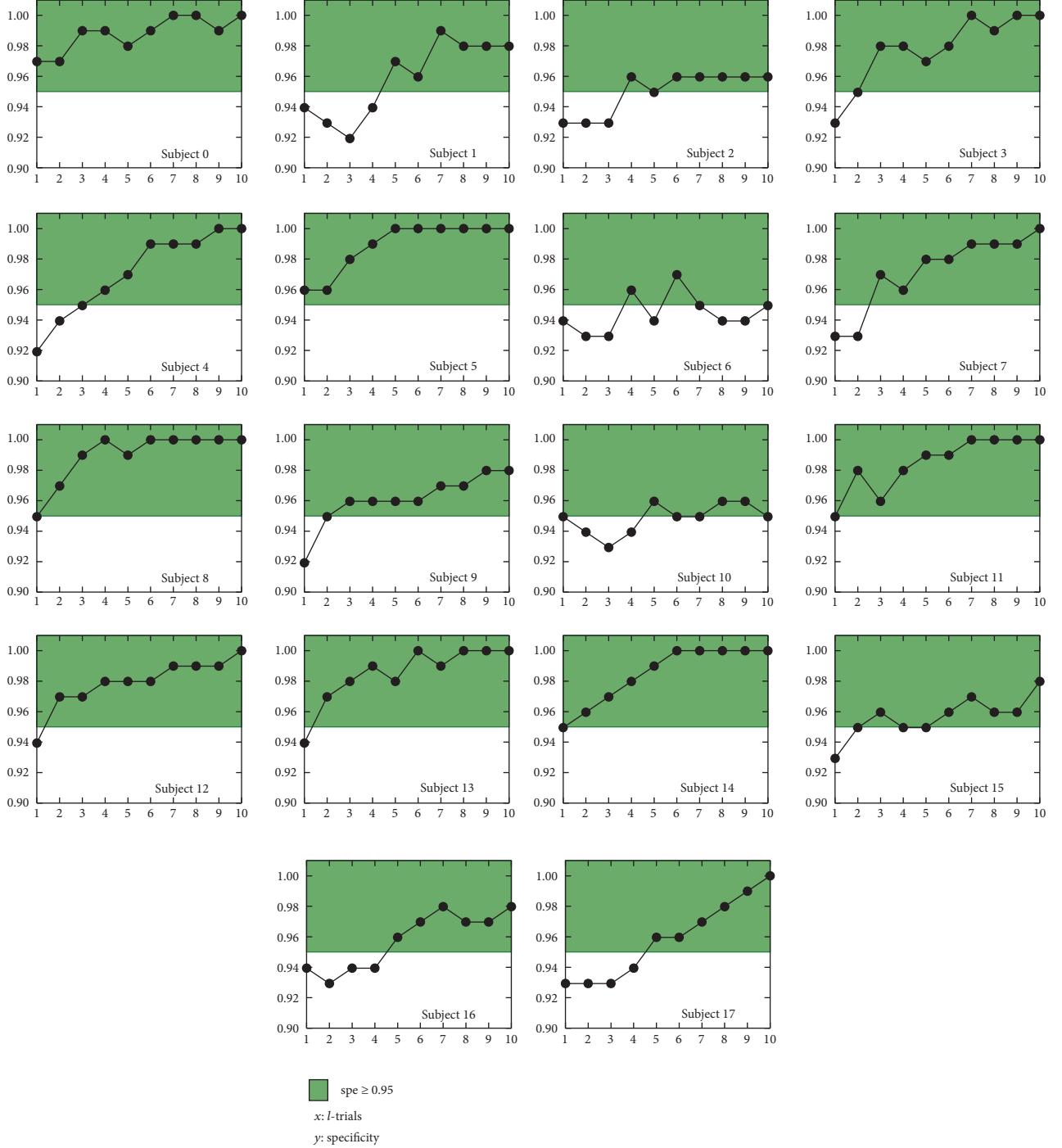


FIGURE 10: Individual specificities with respect to the number of trials scored $l = \{1, \dots, 10\}$. Green filled area represents specificities greater than 0.95.

those two subjects presented a target coherent grand average with P300 features and their target and nontarget averages were statistically different, none reached the minimum performance. Yet, subject 4 had no statistical differences between his coherent averages; he reached 1.0 in both sensitivity and specificity. Furthermore, subject 9 (see Figure 6) had a target

coherent average with no P300 features but he was capable of correctly selecting seven of nine trials nevertheless, therefore reaching more than the minimum performances for the tasks. In conclusion, those four cases showed that a global P300 elicitation was not necessarily related to the target detection performance. In regard to accuracy, metric defined in [53],

Tables 2 and 3 show accuracies higher than 0.87 from single trial ($l = 1$); additionally, the AUROC with ten trials ($l = 10$) is higher than 0.87 ± 0.07 for all subjects.

All recordings were performed in realistic conditions with no sound isolation; therefore, subjects were prone to listen to distracting noises, for instance, voices, fan noise, alarms, or phones. Yet, 89% of the subjects accomplished the correct selection of more than seven targets (sensitivity > 0.75) and discriminated at least 97 nontargets (specificity > 0.97). Furthermore, 12 of the 16 highest performances (75%) were obtained by unexperienced subjects. Hence the P300 stimulation screen variation with an image background of asymmetrically arranged stimulation markers seems reliable for a contextual and real-time generated stimulation screen for commanding a robotic navigation system of a wheelchair using destination-based stimulus sequences.

4. Conclusions

A stimulation screen with an image background whose stimulation markers were asymmetrically arranged and a stimulation sequence generator were developed and evaluated. Using this stimulator, target coherent grand averages were statistically different from their corresponding nontarget averages, except for one subject, who nonetheless reached high sensitivity and specificity. It was also corroborated that target and nontarget averages within 150–450 ms are statistically different.

The scenario screen and its stimulus sequencer showed that they are feasible for use on the wheelchair navigation task for destination-based sequences given that 89% of subjects were able to select the target with high performances. Regardless of the fact that 78% of the subjects had no previous experience in using BCI systems, 72% of the sample reached high sensitivity and specificity performances. In addition, only two unexperienced subjects had performances lower than 0.7. LASSO-LDA altogether with the voting scheme was able to detect the target markers with high performances in almost all subjects. Further analysis of LASSO feature selection is needed to explain why in some cases a direct relationship between the P300 global elicitation and the correct target detection rates cannot be established. Notwithstanding the fact that LASSO-LDA approach had the highest performances on 89% of the subjects, an individualized classifier method might still be considered.

The occasional occurrence of clusters of stimulation markers on a particular scenario screen might contribute to false target detection due to the target and nontargets nearness. An error analysis of those clusters is required in order to verify spatial correlations between the misclassified targets and the true targets. More importantly, the misclassified targets might not have a high impact on the wheelchair navigation task when they are near the true targets; the navigation route and final wheelchair position will be close to the desired one. That is to say, misclassified targets might have lower impact on the navigation accomplishment than their misclassified counterparts on the conventional P300 spelling task where a mistake distorts the message orthographically or semantically [5, 11, 14, 18, 25, 29, 33, 54].

Although all the recordings were performed under realistic conditions, high target and nontarget detection performances were achieved on most subjects. Causes that might impact the performances were task misunderstanding and confusion with the scenario screen usage due to its eight markers with the same character or due to loss of attention. Only two subjects had performances lower than 0.7 sensitivity. These two subjects were able to select correctly 4 of 9 targets which are not directly associated with a random selection given the amount of data (number of epochs and stimulation markers) to be processed to get a target.

Despite the fact that almost all subjects had no discomfort in using the scenario screen, some of them reported somnolence and failing to perceive some targets on the fourth block. This somnolence might be linked to no provision of feedback to the user. This is in accordance with results of other works that relate the ERP elicitation to user motivation [39, 61]. In regard to the stimulation screen features, a blue/green stimuli color scheme seems the correct election since no subject reports discomfort related to stimulus attention, in contrast to the previous reported red/white scheme [41].

Although Escolano et al. [30, 31] reported two stimulation screens with image background, those were implemented differently from scenario screen as described in Introduction. Nonetheless, both reports had an entirely distinct aim with respect to this paper; that is, they were focused on the evaluation of navigation not on the stimulation screen evaluation nor the stimulus presentation features that improve the target detection.

Finally, the stimulus sequence generator developed substitutes the native OpenViBE sequencer overcoming its limitations at implementing the single marker stimulus mode. Other OpenViBE inconveniences, which did not compromise the realization of the evaluations reported herein, indicate that other software architectures should be used for a release version of the navigation control.

Conflicts of Interest

All authors expressly declare that they have no conflicts of interest.

Acknowledgments

M.S. Piña-Ramírez gratefully acknowledges the CONACYT (Mexican Council for Science and Technology) for Ph.D. Scholarship 175754. The authors also acknowledge SECITI (Mexico City Agency for Science, Technology and Innovation) for funding this research project through PICSA12-216.

References

- [1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.
- [2] E. Donchin and M. G. H. Coles, "Is the P300 component a manifestation of cognitive updating?" *Behavioral and Brain Sciences*, vol. 11, no. 3, pp. 357–427, 1988.

- [3] M. Marchetti and K. Priftis, "Effectiveness of the P3-speller in brain-computer interfaces for amyotrophic lateral sclerosis patients: A systematic review and meta-analysis," *Frontiers in Neuroengineering*, 2014.
- [4] J. R. Wolpaw, "Brain-computer interfaces," in *Neurological Rehabilitation*, vol. 110 of *Handbook of Clinical Neurology*, pp. 67–74, Elsevier, 2013.
- [5] G. Pires, U. Nunes, and M. Castelo-Branco, "Comparison of a row-column speller vs. a novel lateral single-character speller: Assessment of BCI for severe motor disabled patients," *Clinical Neurophysiology*, vol. 123, no. 6, pp. 1168–1181, 2012.
- [6] C. Guger, S. Daban, E. Sellers et al., "How many people are able to control a P300-based brain-computer interface (BCI)?" *Neuroscience Letters*, vol. 462, no. 1, pp. 94–98, 2009.
- [7] E. W. Sellers, D. J. Krusinski, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "A P300 event-related potential brain-computer interface (BCI): the effects of matrix size and inter stimulus interval on performance," *Biological Psychology*, vol. 73, no. 3, pp. 242–252, 2006.
- [8] K. Colwell, D. Ryan, C. Throckmorton, E. Sellers, and L. Collins, "Channel selection methods for the P300 Speller," *Journal of Neuroscience Methods*, vol. 232, pp. 6–15, 2014.
- [9] J. Jin, B. Z. Allison, T. Kaufmann et al., "The changing face of P300 BCIs: a comparison of stimulus changes in a P300 BCI involving faces, emotion, and movement," *PLoS ONE*, vol. 7, no. 11, Article ID e49688, 2012.
- [10] J. Jin, B. Z. Allison, X. Wang, and C. Neuper, "A combined brain-computer interface based on P300 potentials and motion-onset visual evoked potentials," *Journal of Neuroscience Methods*, vol. 205, no. 2, pp. 265–276, 2012.
- [11] J. Jin, E. W. Sellers, and X. Wang, "Targeting an efficient target-to-target interval for P300 speller brain-computer interfaces," *Medical & Biological Engineering & Computing*, vol. 50, no. 3, pp. 289–296, 2012.
- [12] J. Shi, J. Shen, Y. Ji, and F. Du, "A submatrix-based P300 brain-computer interface stimulus presentation paradigm," *Journal of Zhejiang University SCIENCE C*, vol. 13, no. 6, pp. 452–459, 2012.
- [13] J. Shen, J. Liang, J. Shi, and Y. Wang, "A dynamic submatrix-based P300 online brain-computer interface," *Biomedical Signal Processing and Control*, vol. 15, pp. 27–32, 2015.
- [14] C. Polprasert, P. Kukieattikool, T. Demechai, J. A. Ritcey, and S. Siwamogsatham, "New stimulation pattern design to improve P300-based matrix speller performance at high flash rate," *Journal of Neural Engineering*, vol. 10, no. 3, Article ID 036012, 2013.
- [15] S. Ikegami, K. Takano, M. Wada, N. Saeki, and K. Kansaku, "Effect of the green/blue flicker matrix for P300-based brain-computer interface: An EEG-fMRI study," *Frontiers in Neurology*, 2012.
- [16] K. Takano, T. Komatsu, N. Hata, Y. Nakajima, and K. Kansaku, "Visual stimuli for the P300 brain-computer interface: A comparison of white/gray and green/blue flicker matrices," *Clinical Neurophysiology*, vol. 120, no. 8, pp. 1562–1566, 2009.
- [17] K. Takano, H. Ora, K. Sekihara, S. Iwaki, and K. Kansaku, "Coherent activity in bilateral parieto-occipital cortices during P300-BCI operation," *Frontiers in Neurology*, vol. 5, Article ID Article 74, 2014.
- [18] S. Halder, A. Pinegger, I. Käthner et al., "Brain-controlled applications using dynamic P300 speller matrices," *Artificial Intelligence in Medicine*, vol. 63, no. 1, pp. 7–17, 2015.
- [19] T. Kaufmann, S. M. Schulz, C. Grünzinger, and A. Kübler, "Flashing characters with famous faces improves ERP-based brain-computer interface performance," *Journal of Neural Engineering*, vol. 8, no. 5, pp. 173–178, 2011.
- [20] J. Jin, E. W. Sellers, Y. Zhang, I. Daly, X. Wang, and A. Cichocki, "Whether generic model works for rapid ERP-based BCI calibration," *Journal of Neuroscience Methods*, vol. 212, no. 1, pp. 94–99, 2013.
- [21] F. Aloise, P. Aricò, F. Schettini, S. Salinari, D. Mattia, and F. Cincotti, "Asynchronous gaze-independent event-related potential-based brain-computer interface," *Artificial Intelligence in Medicine*, vol. 59, no. 2, pp. 61–69, 2013.
- [22] E. M. Mugler, C. A. Ruf, S. Halder, M. Bensch, and A. Kübler, "Design and implementation of a P300-based brain-computer interface for controlling an internet browser," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 6, pp. 599–609, 2010.
- [23] T. Yu, Y. Li, J. Long, and Z. Gu, "Surfing the internet with a BCI mouse," *Journal of Neural Engineering*, vol. 9, no. 3, p. 036012, 2012.
- [24] R. Carabalona, F. Grossi, A. Tessadri, P. Castiglioni, A. Caraciolo, and I. de Munari, "Light on! Real world evaluation of a P300-based brain-computer interface (BCI) for environment control in a smart home," *Ergonomics*, vol. 55, no. 5, pp. 552–563, 2012.
- [25] F. Aloise, F. Schettini, P. Aricò et al., "P300-based brain-computer interface for environmental control: an asynchronous approach," *Journal of Neural Engineering*, vol. 8, no. 2, p. 025025, 2011.
- [26] U. Hoffmann, J.-M. Vesin, T. Ebrahimi, and K. Diserens, "An efficient P300-based brain-computer interface for disabled subjects," *Journal of Neuroscience Methods*, vol. 167, no. 1, pp. 115–125, 2008.
- [27] R. Corralejo, L. F. Nicolás-Alonso, D. Álvarez, and R. Hornero, "A P300-based brain-computer interface aimed at operating electronic devices at home for severely disabled people," *Medical & Biological Engineering & Computing*, vol. 52, no. 10, pp. 861–872, 2014.
- [28] I. P. Ganin, S. L. Shishkin, A. G. Kochetova, and A. Y. Kaplan, "P300-based brain-computer interface: The effect of the stimulus position in a stimulus train," *Human Physiology*, vol. 38, no. 2, pp. 121–128, 2012.
- [29] I. P. Ganin, S. L. Shishkin, and A. Y. Kaplan, "A P300-based brain-computer interface with stimuli on moving objects: four-session single-trial and triple-trial tests with a game-like task design," *PLoS ONE*, vol. 8, no. 10, Article ID e77755, 2013.
- [30] C. Escolano, J. Antelis, and J. Minguez, "Human brain-teleoperated robot between remote places," in *Proceedings of the 2009 IEEE International Conference on Robotics and Automation, ICRA '09*, pp. 4430–4437, Japan, May 2009.
- [31] C. Escolano and J. Minguez, "Sistema de Teleoperación Multi-Robot basado en Interfaz Cerebro-Computador," *Revista Iberoamericana de Automática e Informática Industrial RIAI*, vol. 8, no. 2, pp. 16–23, 2011.
- [32] I. Iturrate, J. M. Antelis, A. Kübler, and J. Minguez, "A noninvasive brain-actuated wheelchair based on a P300 neurophysiological protocol and automated navigation," *IEEE Transactions on Robotics*, vol. 25, no. 3, pp. 614–627, 2009.
- [33] G. G. Gentiletti, J. G. Gebhart, R. C. Acevedo, O. Yáñez-Suárez, and V. Medina-Bañuelos, "Command of a simulated wheelchair on a virtual environment using a brain-computer interface," *IRBM*, vol. 30, no. 5–6, pp. 218–225, 2009.

- [34] A. C. Lopes, G. Pires, and U. Nunes, "Assisted navigation for a brain-actuated intelligent wheelchair," *Robotics and Autonomous Systems*, vol. 61, no. 3, pp. 245–258, 2013.
- [35] B. Rebsamen, E. Burdet, C. Guan et al., "A brain-controlled wheelchair based on P300 and path guidance," in *Proceedings of the 1st IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics, 2006, BioRob 2006*, pp. 1101–1106, Italy, February 2006.
- [36] B. Rebsamen, C. Guan, H. Zhang et al., "A brain controlled wheelchair to navigate in familiar environments," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 6, pp. 590–598, 2010.
- [37] M. H. Zweig and G. Campbell, "Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine," *Clinical Chemistry*, vol. 39, no. 4, pp. 561–577, 1993.
- [38] I. Käthner, S. C. Wriessnegger, G. R. Müller-Putz, A. Kübler, and S. Halder, "Effects of mental workload and fatigue on the P300, alpha and theta band power during operation of an ERP (P300) brain-computer interface," *Biological Psychology*, vol. 102, no. 1, pp. 118–129, 2014.
- [39] L. da Silva-Sauer, L. Valero-Aguayo, A. de la Torre-Luque, R. Ron-Angevin, and S. Varona-Moya, "Concentration on performance with P300-based BCI systems: A matter of interface features," *Applied Ergonomics*, vol. 52, article no. 2090, pp. 325–332, 2016.
- [40] Y. Renard, F. Lotte, G. Gibert et al., "OpenViBE: an open-source software platform to design, test, and use brain-computer interfaces in real and virtual environments," *Presence: Teleoperators and Virtual Environments*, vol. 19, no. 1, pp. 35–53, 2010.
- [41] O. Piña-Ramírez, O. Yañez-Suárez, and R. Valés-Cristerna, "Visual P300 stimulator with non uniform background and non symmetric stimulation markers," in *Proceedings of the VI Latin American Congress on Biomedical Engineering CLAIB 2014*, pp. 576–579, Paraná, Argentina, 2014.
- [42] D. J. Krusienski, E. W. Sellers, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "Toward enhanced P300 speller performance," *Journal of Neuroscience Methods*, vol. 167, no. 1, pp. 15–21, 2008.
- [43] P. Brunner, S. Joshi, S. Briskin, J. R. Wolpaw, H. Bischof, and G. Schalk, "Does the 'P300' speller depend on eye gaze?" *Journal of Neural Engineering*, vol. 7, no. 5, Article ID 56013, 2010.
- [44] P. Aricò, F. Aloise, F. Schettini, S. Salinari, D. Mattia, and F. Cincotti, "Influence of P300 latency jitter on event related potential-based brain-computer interface performance," *Journal of Neural Engineering*, vol. 11, no. 3, Article ID 035008, 2014.
- [45] J. Polich, "Updating P300: an integrative theory of P3a and P3b," *Clinical Neurophysiology*, vol. 118, no. 10, pp. 2128–2148, 2007.
- [46] R. Tibshirani, "Regression shrinkage and selection via the lasso: a retrospective," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 73, no. 3, pp. 273–282, 2011.
- [47] R. Tibshirani, "Regression shrinkage and selection via the lasso: A retrospective," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 58, pp. 267–288, 1994.
- [48] R. Tomioka and K.-R. Müller, "A regularized discriminative framework for EEG analysis with application to brain-computer interface," *NeuroImage*, vol. 49, no. 1, pp. 415–432, 2010.
- [49] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, New York, NY, USA, 2006.
- [50] D. J. Krusienski, E. W. Sellers, F. Cabestany et al., "A comparison of classification techniques for the P300 Speller," *Journal of Neural Engineering*, vol. 3, no. 4, pp. 299–305, 2006.
- [51] D. J. Krusienski and J. J. Shih, "Control of a visual keyboard using an electrocorticographic brain-computer interface," *Neurorehabilitation and Neural Repair*, vol. 25, no. 4, pp. 323–331, 2011.
- [52] F. Pedregosa, G. Varoquaux, A. Gramfort et al., "Scikit-learn: machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [53] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006.
- [54] F. Aloise, F. Schettini, P. Aricò, S. Salinari, F. Babiloni, and F. Cincotti, "A comparison of classification techniques for a gaze-independent P300-based brain-computer interface," *Journal of Neural Engineering*, vol. 9, no. 4, p. 045012, 2012.
- [55] N. V. Manyakov, N. Chumerin, A. Combaz, and M. M. Van Hulle, "Comparison of classification methods for P300 brain-computer interface on disabled subjects," *Computational Intelligence and Neuroscience*, vol. 2011, Article ID 519868, 12 pages, 2011.
- [56] D. J. McFarland, W. A. Sarnacki, G. Townsend, T. Vaughan, and J. R. Wolpaw, "The P300-based brain-computer interface (BCI): Effects of stimulus rate," *Clinical Neurophysiology*, vol. 122, no. 4, pp. 731–737, 2011.
- [57] J. Demšar, "Statistical comparisons of classifiers over multiple data sets," *Journal of Machine Learning Research*, vol. 7, pp. 1–30, 2006.
- [58] S. García and F. Herrera, "An extension on 'statistical comparisons of classifiers over multiple data sets' for all pairwise comparisons," *Journal of Machine Learning Research*, vol. 9, pp. 2677–2694, 2008.
- [59] F. Akram, S. M. Han, and T. Kim, "An efficient word typing P300-BCI system using a modified T9 interface and random forest classifier," *Computers in Biology and Medicine*, vol. 56, pp. 30–36, 2015.
- [60] O. Pina-Ramrez, R. Valdes-Cristerna, and O. Yanez-Suarez, "Classifiers' comparison for P300 detection in a modified speller screen," in *Proceedings of the 12th International Symposium on Medical Information Processing and Analysis, SIPAIM 2016*, E. Romero, N. Lepore, J. Brieva et al., Eds., Argentina, December 2017.
- [61] S. C. Kleih, F. Nijboer, S. Halder, and A. Kübler, "Motivation modulates the P300 amplitude during brain-computer interface use," *Clinical Neurophysiology*, vol. 121, no. 7, pp. 1023–1031, 2010.

Research Article

Novel Signal Noise Reduction Method through Cluster Analysis, Applied to Photoplethysmography

William Waugh,¹ John Allen ,^{1,2} James Wightman,²
Andrew J. Sims,^{1,2} and Thomas A. W. Beale ^{1,2}

¹*Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne NE2 4HH, UK*

²*Northern Medical Physics and Clinical Engineering, Newcastle upon Tyne NHS Foundation Trust, Newcastle upon Tyne NE7 7DN, UK*

Correspondence should be addressed to John Allen; john.allen@nuth.nhs.uk

Received 4 August 2017; Accepted 31 December 2017; Published 29 January 2018

Academic Editor: Chung-Min Liao

Copyright © 2018 William Waugh et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Physiological signals can often become contaminated by noise from a variety of origins. In this paper, an algorithm is described for the reduction of sporadic noise from a continuous periodic signal. The design can be used where a sample of a periodic signal is required, for example, when an average pulse is needed for pulse wave analysis and characterization. The algorithm is based on cluster analysis for selecting similar repetitions or pulses from a periodic signal. This method selects individual pulses without noise, returns a clean pulse signal, and terminates when a sufficiently clean and representative signal is received. The algorithm is designed to be sufficiently compact to be implemented on a microcontroller embedded within a medical device. It has been validated through the removal of noise from an exemplar photoplethysmography (PPG) signal, showing increasing benefit as the noise contamination of the signal increases. The algorithm design is generalised to be applicable for a wide range of physiological (physical) signals.

1. Introduction

Signal quality or signal-to-noise ratio requires consideration in almost all signal measurements. This is especially true in physiological measurements where the signals tend to be small and prone to measurement artefacts and the noise is often difficult to control. In this paper, a novel cluster analysis method is described to reduce the influence of noise on photoplethysmography (PPG) signals. PPG is an optical measurement technique that can be used to detect blood volume changes in the microvascular bed of tissue [1]. The peripheral pulse, as measured by PPG, is often used in the assessment of health and disease and can provide important valuable information about the cardiovascular system [2–5]. Our research group is evaluating PPG for the diagnosis of peripheral arterial disease in a primary care situation using a fully automated diagnostic device [6]. The clinical utility of such a device relies on its ability to identify and eliminate noise from PPG signals.

Noise minimisation starts with removing the source of the noise; this can be through electrical isolation or, for

example, by keeping the subject relaxed and still during measurements to eliminate muscle and movement artefact. There is also inherent noise produced through the amplification of small signals; however modern physiological amplifiers and analogue-to-digital converters tend to minimise this for all but the smallest input signals. When the sources of the noise have been reduced as far as possible, various active noise reduction techniques can be used. The most common kind of noise minimisation is filtering [7, 8] that can be used to reduce any noise frequencies that do not overlap the signal frequencies. More sophisticated methods such as wavelet denoising [9] can be employed where filtering is insufficient. Physiological signals, in particular ECG and PPG, have been the focus of noise reduction using a signal quality index, whereby each pulse has attributed a signal quality, which is then used to assess the validity of that pulse [10–13].

Cluster analysis is a method of arranging features into groups such that those with similar characteristics lie within a single group. Cluster analysis is common in data analysis and there are many algorithms [14]. In this paper, we have applied a simple cluster analysis to remove noise from a physiological

PPG signal. This signal is periodic and the disease diagnosis is performed from a representative sample pulse. Therefore, characteristics of individual pulses are not needed but rather characteristics from a representative single pulse (be this a selected good pulse or an average of pulses). In order to output a representative pulse, a trace with many pulses is recorded and the average pulse from these calculated. This method performs well when noise contamination is low, for example, when recorded in established physiological measurement settings and by trained researchers studying PPG. However, moving diagnostic devices into real-world clinical settings to provide a robust and automated assessment can be challenging. For example, patients may not stay still during the recording and the device must be designed to return a valid clinical result with any reasonable expected clinical setting and level of staff training. When significant measurement noise is present, this noise can dominate the average pulse such that this is no longer a true representation of the subject's PPG pulse. This paper describes an algorithm using cluster analysis to select a subset of pulses to return a representative pulse returned for subsequent diagnosis.

2. Method

The algorithm was developed using Matlab® version 2016b; the photoplethysmography and electrocardiogram signals were measured using a multichannel PPG and ECG recorder as used in a clinical study [15]. Prerecorded signals from normal subjects were used as an input to the algorithm. A variety of finger and/or toe pulse signals with noise implemented through on-purpose patient movement were used to train the algorithm. The algorithm was developed to run in real time such that the signal capture could be terminated when sufficient signal has been received. For a high-quality signal, this can result in a shorter recording time. For signals with a small signal-to-noise ratio, this allows the device to collect sufficient data such that a diagnosis is possible, up to a time-out limit.

2.1. Algorithm Development. The design requirements for the algorithm were as follows:

- (i) Compute on a continuous digital data stream, with a minimal signal delay (pseudo-real-time).
- (ii) Remove low-frequency noise (DC drift).
- (iii) Remove high-frequency noise.
- (iv) Remove sporadic mixed-frequency noise.
- (v) Terminate when sufficient “good” pulses are recorded.

The algorithm was developed to reduce noise from a PPG signal. This signal has a periodic frequency equal to the heart rate of the subject. The signal structure mainly exists in the low-frequency domain, with the desirable frequencies for analysis lying between 0.15 Hz and 20 Hz.

The algorithm's steps are shown in Figure 2. The algorithm can be divided into three sections: initial filtering and slicing of the data, pulse clustering, and termination.

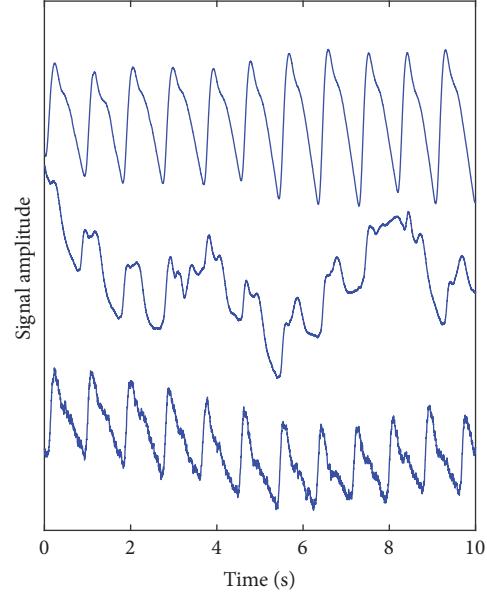


FIGURE 1: Three recorded photoplethysmography (PPG) traces measured from the great toe site. Upper: a clean PPG signal; middle: a signal dominated by low-frequency noise, usually caused by movement of the subject's limb; lower: a PPG trace contaminated with high-frequency noise, typical of electrical interference. Typically, these noise features can appear intermittently within a recording made over a measurement period of 1-2 minutes.

2.2. Initial Filtering and Slicing Stage. The incoming PPG signal is subject to a digital bandpass filter to remove unwanted noise and signal drift. This is implemented through a low-pass filter and a high-pass filter, designed to minimise both the signal distortion and the signal phase delay. A minimal delay is imperative for any device where a live trace is shown, especially where operator feedback is a possibility (e.g., adjustment of the sensor at the measurement site). Any substantial delay can render such operator feedback confusing and nonintuitive. The options for digital filters fall into two main categories: FIR (Finite Impulse Response) and IIR (Infinite Impulse Response) [17]. Although symmetric FIR filters have the advantage that they have linear phase and are always stable, they have substantial delays when designed with low cut-off frequencies. IIR filters generally have a non-linear phase response and therefore cause a frequency-related signal delay; however, they can be faster than an FIR filter.

The information in the PPG signal lies in the low-frequency range (below ~20 Hz); however, the signal is often contaminated by high-frequency noise (Figure 1). This is often due to measurements in an electrically noisy environment or optical pick-up from external lighting sources.

The high frequencies are removed by a low-pass filter with a cut-off frequency close to 22 Hz. This is achieved using a moving average filter, which is a simple implementation of an FIR filter. This filter has a linear phase so as not to distort the waveforms and a low roll-off rate. With careful implementation of the filter taps, this can also be designed to minimise multiples of 50 Hz noise (Figure 3) [18].

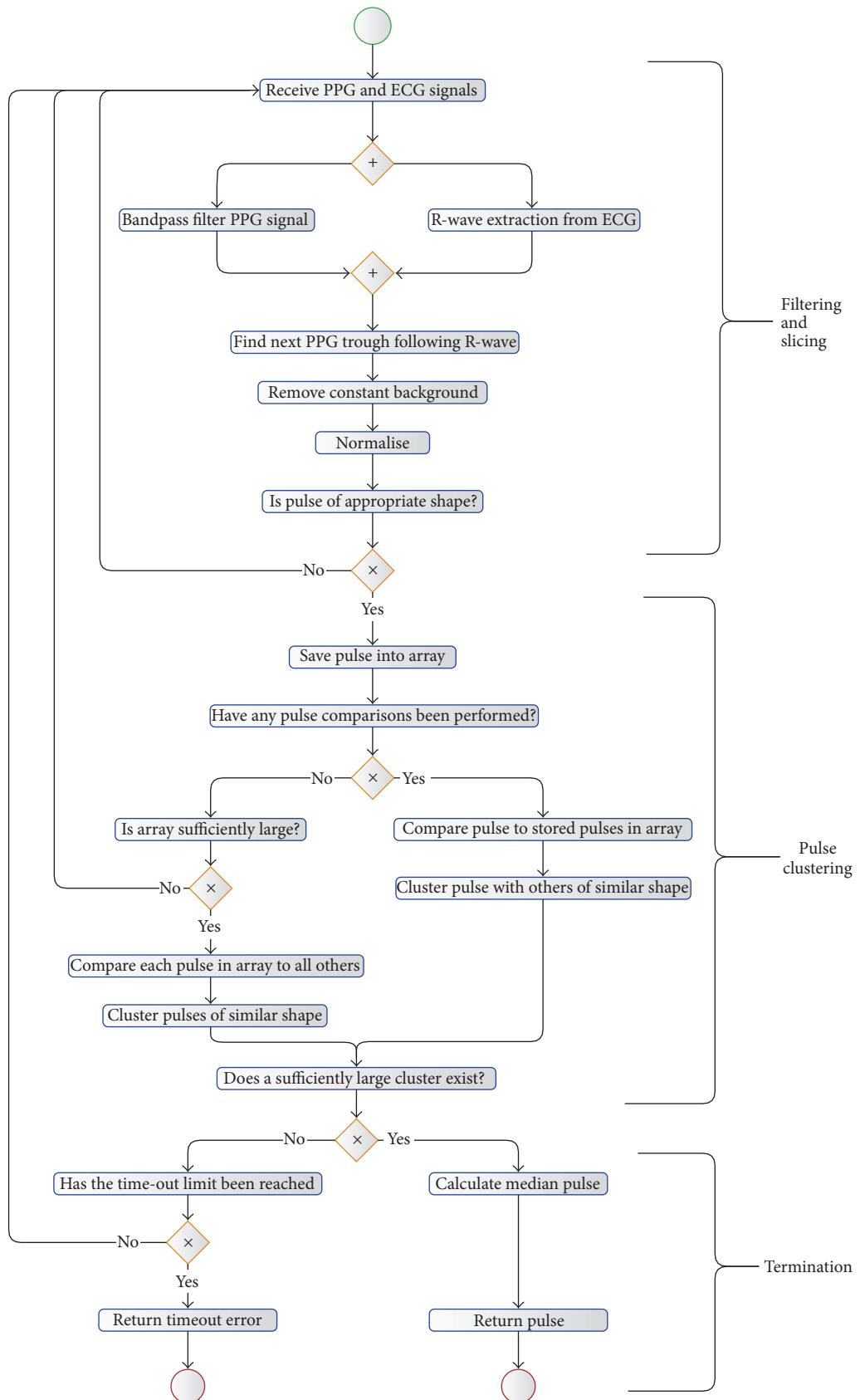


FIGURE 2: Algorithm's steps represented as a flow diagram, utilising BPMN Notation [16].

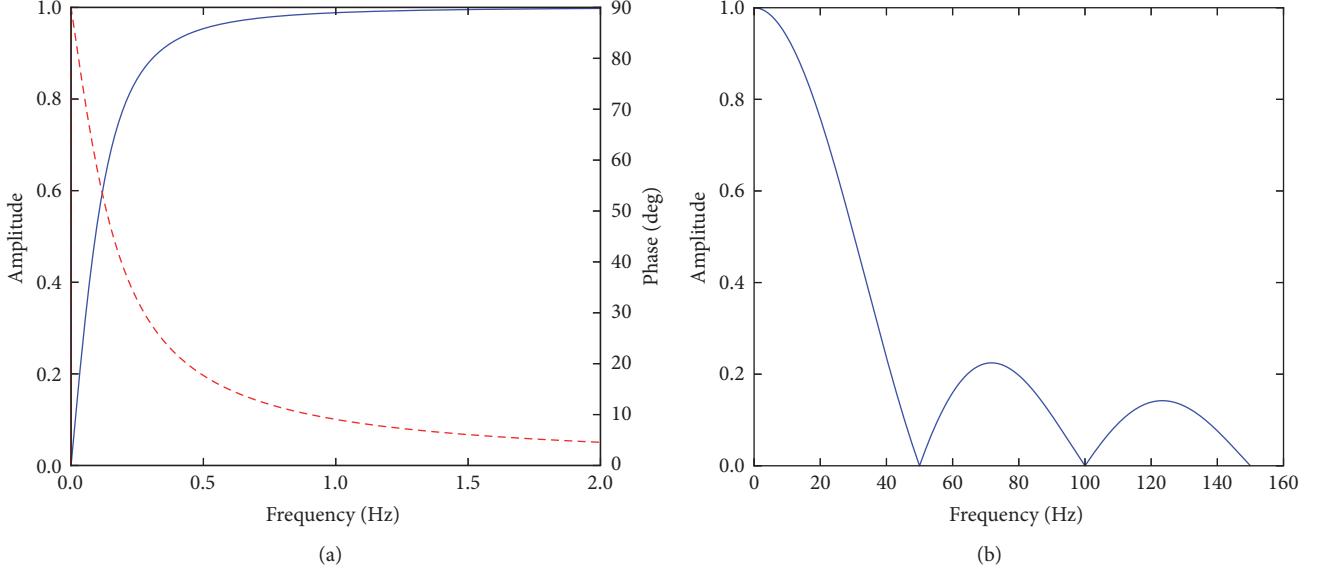


FIGURE 3: The (a) high-pass and (b) low-pass filters. The phase response for the high-pass filter is shown as a dashed red line, showing an increasing nonlinear effect at very low frequencies. The phase for the low pass is linear, such that all frequencies are delayed by the same time.

The high-pass filter is more complicated to design due to the very low (0.15 Hz) cut-off frequency. The primary purpose of this filter is to block the dominant DC background on which the PPG signal is superimposed. For speed of processing and response, we have adopted a first-order digital high-pass IIR filter with a cut-off of 0.15 Hz, and the transfer function is as follows:

$$H(z) = \frac{1 - z^{-1}}{1 - \alpha \cdot z^{-1}} \quad \text{where } \alpha < 1. \quad (1)$$

Although this filter is an IIR filter and has a nonlinear phase response, this nonlinearity is concentrated at very low frequencies below the filter cut-off frequency. The filter responses are shown in Figure 3 and combined give a bandpass with the required attributes.

The filtered signal is then sliced into individual pulses. This could be done with the PPG traces, finding the troughs between the pulses; however, this can be problematic either with a weak signal or when there is substantial noise. A more reliable method is using the R-wave gating from an ECG signal. In this study, the R-waves from the ECG signal have been extracted using the method developed by Pan and Tompkins [19] and the troughs between the pulses found from the subsequent minima following each R-wave. Each resulting pulse then has a constant background removed and is normalised in amplitude and duration. As the clustering method is processor-intensive, there is a “sanity check” on the pulse to check that it is pulse-like in form. This is designed to be computationally fast and is used to discard obvious nonpulses. This check averages the amplitude of the samples in the first 5%, middle 90%, and last 5% of the pulse. The average of the middle section must be 1.5x larger than the biggest of the average of the first and the average of the last sections. This ensures that the pulse amplitude starts low, goes

up, and then returns low, giving confidence that a periodic pulse-like feature is present for subsequent analysis.

2.3. Pulse Clustering. The algorithm saves the pulse into an array. This pulse is then compared to all previous pulses by comparing the amplitude of each sample within the pulse. In order to compare the pulses, distance metrics were tested, including calculating the Pearson correlation coefficient, the Kendall rank correlation coefficient, the Spearman rank correlation coefficient, and the root mean square error (RMSE). Each of these distance metrics is optimised differently; by using a subset of data and visual comparisons of the clusters, RMSE produced the most appropriate clustering. RMSE also has the advantage of being computationally simple and therefore fast.

Each pulse forms a new cluster and is the centre of that cluster. In addition, each pulse is placed into any other cluster, where the RMSE between this pulse and the pulse at the centre of that cluster is below a threshold value. In this way, N pulses create N clusters, each populated with pulses with an RMSE from the centre pulse less than a preset threshold.

2.4. Termination. After each pulse has been clustered, the number of pulses in each cluster is calculated. If any cluster has sufficient pulses for the algorithm requirements, then the loop is terminated, and an averaged (normalised) pulse is returned. As the pulses are normalised in time, the median pulse is calculated by finding the median of each point on the pulse. If there is no cluster with sufficient pulses, then the algorithm accepts more data, or if a predefined time-limit has been reached, then the algorithm terminates with a time-out error. This protects the algorithm from running continuously with no output.

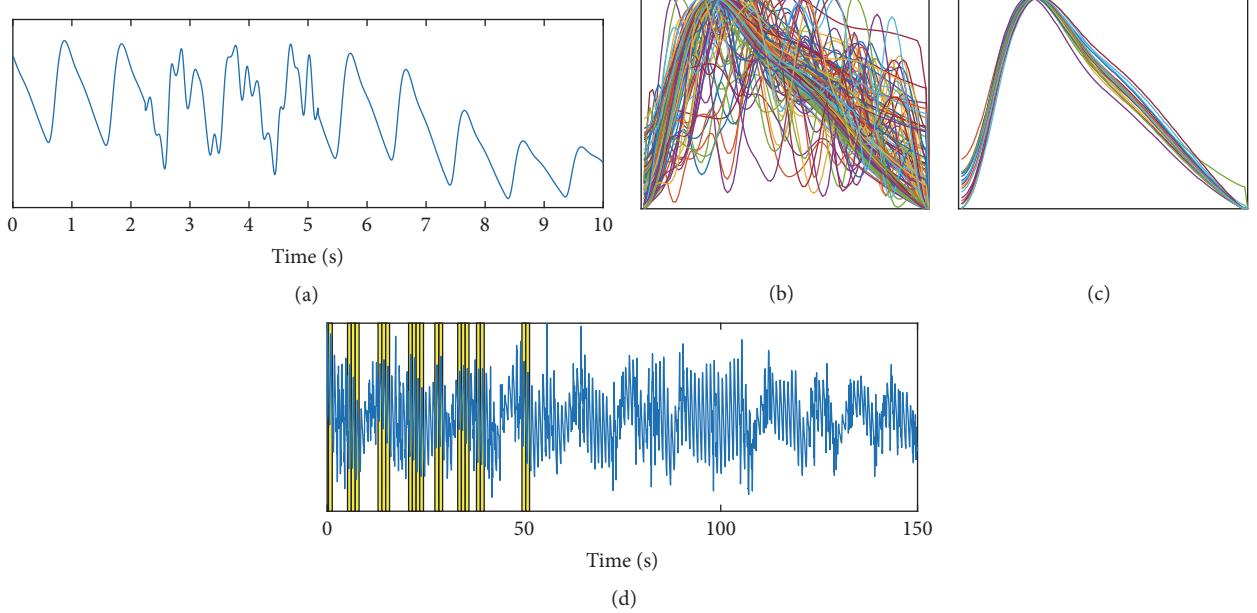


FIGURE 4: (a) PPG trace with artificial “movement” noise added. (b) and (c) show full pulse set and clustered pulses. (d) shows full trace with cluster pulses highlighted.

2.5. Algorithm Validation. To validate the algorithm, a clean PPG signal was analysed using the filtering described above, however with the clustering turned off. This returned an averaged representative pulse shape. The signal was then digitally contaminated with noise, and the analysis was repeated with and without clustering. The output of these two methods of analysis was compared to the representative pulse from the clean signal.

3. Results and Discussion

Figure 4(a) shows an extract from a PPG trace with artificial noise added to approximately 30% of the signal. The noise has been designed to replicate movement noise as seen in the middle trace in Figure 1. Figure 4(b) shows all of the pulses (without clustering), and Figure 4(c) shows just those pulses within the largest cluster. The duration of the input PPG signal is 150 seconds; however, the algorithm with clustering self-terminates when any cluster contains 20 pulses. Figure 4(d) shows the full trace of the PPG with the 20 pulses contributing to the largest cluster highlighted in yellow. Note that there are no pulses selected from the end of the trace as the algorithm is analysing the trace as if in real time and therefore terminates when there are sufficient (in this case 20) pulses.

It is clear from Figure 4 that the clustering successfully selects pulses of a similar shape and these visually appear to be a physiologically representative set. Figure 5 shows the median of all the pulses and the median from the cluster pulses together with the median of all the pulses from the original clean data. It can be seen in Figure 5 that the cluster set produces a median pulse much closer to the clean data median pulse than the median from all pulses.

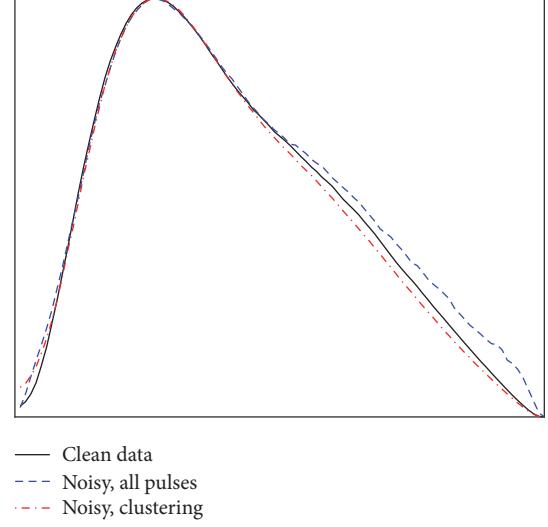


FIGURE 5: Median peaks from the clean data pulses, all pulses after noise is added, and the median from using clustering on the noisy data.

In order to validate the algorithm, we have quantified the difference between the clean, cluster, and noncluster median pulses using the same RMSE comparison. We have simulated movement noise, constant electrical noise (similar to the lower trace in Figure 1), a combination of movement and electrical noise, no noise, and white noise with no PPG signal, all shown in Figure 6. The right-hand panels show a comparison between the median of the clean data, the median of the data with noise added, and the median of the cluster. In Figure 6(c), a combination of movement and electrical noise prevented a cluster forming within the defined

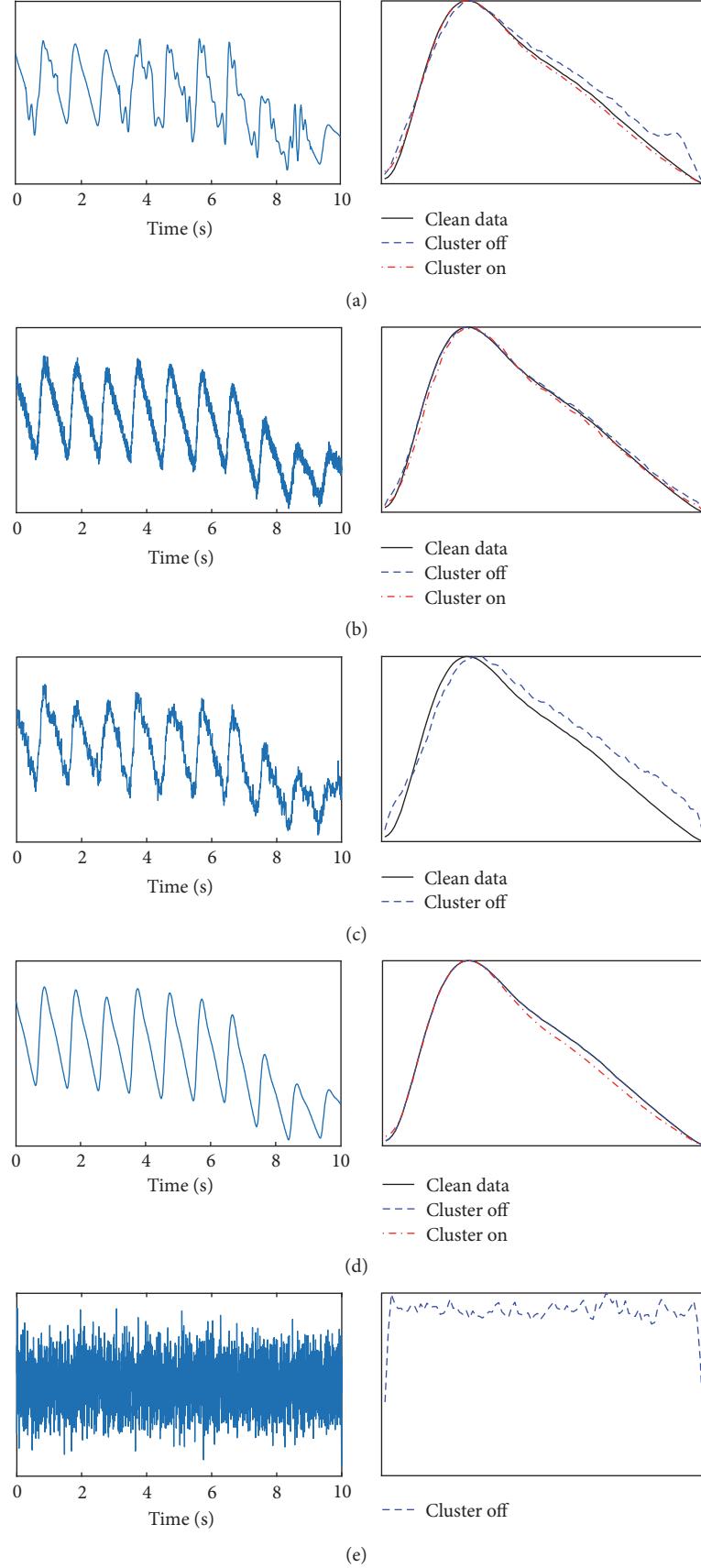


FIGURE 6: Comparison between clustering and nonclustering for a variety of inputs. (a) Simulated movement artefacts, (b) simulated electrical noise, (c) simulated electrical artefact and movement noise, (d) no noise, and (e) white noise as an input. Note that in (c) and (e) a cluster of sufficient size was not formed; therefore the algorithm reported no result.

TABLE 1: RMSE comparisons between noncluster and cluster algorithms for different noise situations. Note that where there is significant noise the cluster method returns a time-out error.

Panel in Figure 6	Simulated noise	RMSE between clean median pulse and median of all noisy pulses (153 pulses)	RMSE between clean median pulse and median of cluster (20 pulses)
(a)	Movement artefact (10% noise)	0.03	0.14
	Movement artefact (20% noise)	0.12	0.14
	Movement artefact (30% noise)	0.23	0.11
	Movement artefact (40% noise)	0.27	0.07
	Movement artefact (50% noise)	0.43	0.12
	Movement artefact (60% noise)	0.64	—
(b)	Electrical noise	0.12	0.09
(c)	Movement (30%) and electrical noise	0.66	—

thresholds, and the median of all of the noisy pulses is deviating significantly from the clean signal. In Figure 6(e), we used white noise as the input. Again, there was no cluster formed; however, the data was still gated by the ECG R-waves and therefore a median of pulses of noise is produced. The ability of the cluster algorithm to not produce a median pulse is extremely important as it prevents returning a false signal for disease diagnosis.

Table 1 shows the RMSE values for each of these cases, showing in all situations where a cluster was formed that this produced a median pulse closer to the median pulse from the original clean data. No RMSE values can be calculated for a white noise input as seen in Figure 6(e), as there is no original “clean” data signal. The biggest improvement of the cluster algorithm is seen with a high proportion of movement noise on the signal. The movement noise is sporadic and only affects individual pulses; removing these pulses can result in a dramatic improvement. By contrast, electric noise is applied to all pulses, and therefore it is more difficult for a cluster based algorithm to select individual pulses without noise. Despite this, with simulated electrical noise, the cluster analysis produced an RMSE almost half the size of the RMSE from a median of all the pulses.

The performance of the cluster algorithm increased as the proportion of the signal contaminated by movement noise was set from 10% to 50%. The comparison between the RMSE of the cluster algorithm and a median of all pulses is shown in Figure 7. This shows that the cluster result returns broadly consistent results independent of the noise added to the signal, until it is no longer able to produce a result. By comparison, the median of all pulses becomes increasingly poor at representing the original signal. At very low noise levels, a better result can be achieved through averaging over more pulses; therefore the cluster algorithm limiting the number of pulses averaged to 20 performs worse. However, as the noise level increases, the clustering algorithm is superior. Note also that the time for the algorithm to return a result increases with the cluster method as the noise content increases, and indeed noise contamination greater than 50% could be achieved by increasing the time-out limit. By comparison, a median of all pulses will take a fixed duration of time independent of the signal quality

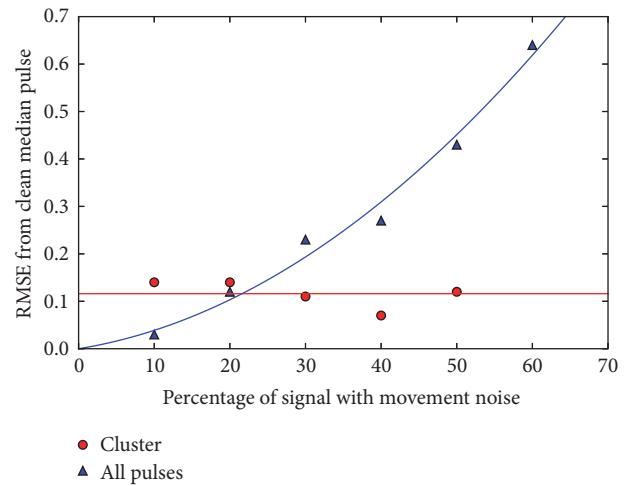


FIGURE 7: Comparison between the cluster and noncluster algorithms with increasing noise.

and will therefore take significantly longer than the cluster method where there is low noise contamination. For real-world clinical application, the algorithm has the scope to indicate when a probe has become unattached from the measurement site or has failed.

Further work linked to photoplethysmography can include assessments of our method approach across a wide range of recordings from healthy subjects and vascular patients and also for different peripheral measurement sites such as the ear lobe and finger pads.

4. Conclusions

We have shown that clustering can be used within a real-time algorithm to minimise the effects of noise on a periodic physiological signal, with an algorithm that can be tailored to individual signal type. For this paper, we have explored its value for photoplethysmography waveforms as the input signal, where a dramatic reduction in the effect of noise on the output result has been demonstrated. Furthermore, if there is insufficient quality of data, the algorithm returns a null result

rather than an incorrect median pulse. The algorithm returns a consistent result as the noise on the signal is increased and at low noise levels can produce a result quickly and efficiently. This algorithm was developed to be computationally fast, such that it could be run in real time on an embedded microcontroller within a portable medical device.

Disclosure

The views expressed in this publication are those of the authors and not necessarily those of the NHS, the National Institute for Health Research, or the Department of Health.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors would like to thank Eden Winlow, Charalambos Theodorou, James Brockhurst, Jonathan Gilmour, and Christopher Saunter for stimulating discussions. This report is independent research funded by the National Institute for Health Research (Invention for Innovation, “Innovative photoplethysmography technology for rapid non-invasive assessment of peripheral arterial disease in primary care,” II-C1-0412-20003).

References

- [1] A. V. J. Challoner, “Photoelectric plethysmography for estimating cutaneous blood flow,” in *Non-Invasive Physiological Measurements*, P. Rolfe, Ed., pp. 1125–1151, Academic, London, UK, 1979.
- [2] W. M. Nichols and M. F. O’Rourke, *McDonald’s Blood Flow in Arteries*, Arnold, London, UK, 3rd edition, 1990.
- [3] I. B. Wilkinson, J. R. Cockcroft, and D. J. Webb, “Pulse wave analysis and arterial stiffness,” *Journal of Cardiovascular Pharmacology*, vol. 3, pp. S33–S37, 1998.
- [4] A. A. R. Kamal, J. B. Harness, G. Irving, and A. J. Mearns, “Skin photoplethysmography - a review,” *Computer Methods and Programs in Biomedicine*, vol. 28, no. 4, pp. 257–269, 1989.
- [5] J. Allen, “Photoplethysmography and its application in clinical physiological measurement,” *Physiological Measurement*, vol. 28, no. 3, pp. R1–R39, 2007.
- [6] S. Wilkes, G. Stansby, A. Sims, S. Haining, and J. Allen, “Peripheral arterial disease: diagnostic challenges and how photoplethysmography may help,” *British Journal of General Practice*, vol. 65, no. 635, pp. 323–324, 2015.
- [7] K. T. Sweeney, T. E. Ward, and S. F. McLoone, “Artifact removal in physiological signals-practices and possibilities,” *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, no. 3, pp. 488–500, 2012.
- [8] T. R. H. Cutmore and D. A. James, “Identifying and reducing noise in psychophysiological recordings,” *International Journal of Psychophysiology*, vol. 32, no. 2, pp. 129–150, 1999.
- [9] M. Mamun, M. Al-Kadi, and M. Marufuzzaman, “Effectiveness of wavelet denoising on electroencephalogram signals,” *Journal of Applied Research and Technology*, vol. 11, no. 1, pp. 156–160, 2013.
- [10] C. Orphanidou, T. Bonnici, P. Charlton, D. Clifton, D. Vallance, and L. Tarassenko, “Signal-quality indices for the electrocardiogram and photoplethysmogram: Derivation and applications to wireless monitoring,” *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 3, pp. 832–838, 2015.
- [11] J. A. Sukor, S. J. Redmond, and N. H. Lovell, “Signal quality measures for pulse oximetry through waveform morphology analysis,” *Physiological Measurement*, vol. 32, no. 3, pp. 369–384, 2011.
- [12] I. Silva, J. Lee, and R. G. Mark, “Signal quality estimation with multichannel adaptive filtering in intensive care settings,” *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 9, pp. 2476–2485, 2012.
- [13] M. Elgendi, “Optimal Signal Quality Index for Photoplethysmogram Signals,” *Bioengineering*, vol. 3, no. 4, p. 21, 2016.
- [14] V. Estivill-Castro, “Why so many clustering algorithms: a position paper,” *ACM SIGKDD Explorations Newsletter*, vol. 4, no. 1, pp. 65–75, 2002.
- [15] J. Allen, K. Overbeck, A. F. Nath, A. Murray, and G. Stansby, “A prospective comparison of bilateral photoplethysmography versus the ankle-brachial pressure index for detecting and quantifying lower limb peripheral arterial disease,” *Journal of Vascular Surgery*, vol. 47, no. 4, pp. 794–802, 2008.
- [16] <http://www.omg.org/spec/BPMN/2.0>.
- [17] R. G. Lyons, *Understanding Digital Signal Processing*, Pearson, 2004.
- [18] S. W. Smith, *Digital Signal Processing: A Practical Guide for Engineers and Scientists*, Newnes, 2002.
- [19] J. Pan and W. J. Tompkins, “A real-time QRS detection algorithm,” *IEEE Transactions on Biomedical Engineering*, vol. 32, no. 3, pp. 230–236, 1985.

Research Article

Linearized and Kernelized Sparse Multitask Learning for Predicting Cognitive Outcomes in Alzheimer's Disease

Xiaoli Liu,^{1,2} Peng Cao , Jinzhu Yang,^{1,2} and Dazhe Zhao^{1,2}

¹Computer Science and Engineering, Northeastern University, Shenyang, China

²Key Laboratory of Medical Image Computing of Ministry of Education, Northeastern University, Shenyang, China

Correspondence should be addressed to Peng Cao; caopeng@cse.neu.edu.cn

Received 4 August 2017; Revised 18 December 2017; Accepted 26 December 2017; Published 24 January 2018

Academic Editor: Peng Li

Copyright © 2018 Xiaoli Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Alzheimer's disease (AD) has been not only the substantial financial burden to the health care system but also the emotional burden to patients and their families. Predicting cognitive performance of subjects from their magnetic resonance imaging (MRI) measures and identifying relevant imaging biomarkers are important research topics in the study of Alzheimer's disease. Recently, the multitask learning (MTL) methods with sparsity-inducing norm (e.g., $\ell_{2,1}$ -norm) have been widely studied to select the discriminative feature subset from MRI features by incorporating inherent correlations among multiple clinical cognitive measures. However, these previous works formulate the prediction tasks as a linear regression problem. The major limitation is that they assumed a linear relationship between the MRI features and the cognitive outcomes. Some multikernel-based MTL methods have been proposed and shown better generalization ability due to the nonlinear advantage. We quantify the power of existing linear and nonlinear MTL methods by evaluating their performance on cognitive score prediction of Alzheimer's disease. Moreover, we extend the traditional $\ell_{2,1}$ -norm to a more general $\ell_q\ell_1$ -norm ($q \geq 1$). Experiments on the Alzheimer's Disease Neuroimaging Initiative database showed that the nonlinear $\ell_{2,1}\ell_q$ -MKMTL method not only achieved better prediction performance than the state-of-the-art competitive methods but also effectively fused the multimodality data.

1. Introduction

Alzheimer's disease (AD) is a severe neurodegenerative disorder that results in a loss of mental function due to the deterioration of brain tissue, leading directly to death [1]. It accounts for 60–70% of age related dementia, affecting an estimated 30 million individuals in 2011 and the number is projected to be over 114 million by 2050 [2]. The cause of AD is poorly understood and currently there is no cure for AD. AD has a long preclinical phase, lasting a decade or more. There is increasing research emphasis on detecting AD in the preclinical phase, before the onset of the irreversible neuron loss that characterizes the dementia phase of the disease, since therapies/treatment are most likely to be effective in this early phase. The Alzheimer's Disease Neuroimaging Initiative (ADNI, <http://adni.loni.usc.edu/>) has been facilitating the scientific evaluation of neuroimaging data including magnetic resonance imaging (MRI) and positron emission tomography (PET), along with other biomarkers and clinical

and neuropsychological assessments for predicting the onset and progression of MCI (mild cognitive impairment) and AD. Early diagnosis of AD is key to the development, assessment, and monitoring of new treatments for AD.

Recently, rather than predicting categorical variables in the classification, various studies started to estimate continuous clinical variables from brain images. Therefore, instead of classifying a subject into binary or multiple predetermined categories or stages of the disease, regression focus is on estimating continuous values which may help to assess patient's disease progression. The most commonly used cognitive measures are Alzheimer's Disease Assessment Scale (ADAS) cognitive total score, Mini Mental State Exam (MMSE) score, and Rey Auditory Verbal Learning Test (RAVLT). Regression analyses were commonly used to predict cognitive scores from imaging measures. The relationship between commonly used cognitive measures and structural changes with MRI has been previously studied by regression models and the results demonstrated that there exists a relationship between

baseline MRI features and cognitive measures [3, 4]. For example, Wan et al. proposed an elegant regression model called CORNLIN that employs a sparse Bayesian learning algorithm to predict multiple cognitive scores based on 98 structural MRI regions of interests (ROIs) for Alzheimer's disease patients. The polynomial model used in CORNLIN can detect either a nonlinear or a linear relationship between brain structure and cognitive decline [3]. Stonnington et al. adopted relevance vector regression, a sparse kernel method formulated in a Bayesian framework, to predict four sets of cognitive scores using MRI voxel based morphometry measures [4]. One of the biggest challenges in the prediction of inferring cognitive outcomes with MRI is the high dimensionality, which affects the computational performance and leads to a wrong estimation and identification of the relevant predictors. To reduce the high dimensionality and identify the relevant biomarkers, the sparse methods have attracted a great amount of research efforts in the neuroimaging field due to its sparsity-inducing property. Ye et al. applied sparse logistic regression with stability selection to ADNI data for robust feature selection [5] and successfully predicted the conversion from MCI into probable AD and identified a small subset of biosignatures.

It is known that there exist inherent correlations among multiple clinical cognitive variables of a subject. However, many works do not model dependence relation between multiple tasks and neglect the correlation between clinical tasks which is potentially useful. When the tasks are believed to be related, learning multiple related tasks jointly can improve the performance relative to learning each task separately. Multitask learning (MTL) is a statistical learning framework which aims at learning several models in a joint manner. It has been commonly used to obtain better generalization performance than learning each task individually [6, 7]. The critical issues in MTL are to identify how the tasks are related and build learning models to capture such task relatedness. The most recent studies [6, 8, 9] employed multitask learning with $\ell_{2,1}$ -norm [7] regularization and aimed to select features that could predict all or most clinical scores. The $\ell_{2,1}$ -norm is chosen to be the regularization. Thus, the $\ell_{2,1}$ -norm regularized regression model is able to select some common features across all the tasks. However, in these learning methods, each task is traditionally performed by formulating a linear regression problem, in which the cognitive score is a linear function of the neuroimaging measures.

Kernel methods have been studied to model the cognitive scores as nonlinear functions of neuroimaging measures. Recently, many kernel-based classification or regression methods with faster optimization speed or stronger generalization performance have been proposed and investigated by theoretically analyzing and experimentally evaluating [10, 11]. Multiple kernel learning (MKL) [12], which learns the optimal kernel for a given task by a weighted, linear combination of predefined candidate kernels, has been introduced to handle the problem of kernel selection. The multiple kernel learning method not only learns an optimal combination of given base kernels but also provides a flexible framework to exploit the nonlinear relationship between MRI measures and cognitive scores.

In building the predictive model for classification or regression in AD, kernel has been widely used; therefore, it is important to extend the existing kernel-based learning methods to the case of multitask learning. In this paper, we propose two nonlinear multikernel-based multiple learning methods in [13] for building regression models, to exploit and investigate the nonlinear relationship between MRI measures and cognitive scores. Moreover, an $\ell_q\ell_1$ -norm is used to extend the traditional $\ell_2\ell_1$ -norm. The goal of our work is to (1) predict subjects' cognitive scores in a number of neuropsychological assessments using their MRI measures across the entire brain, (2) identify what the performance of the nonlinear method is compared with the linear $\ell_q\ell_1$ -norm MTL and other MTL methods with different assumption. No previous studies have systematically and extensively examined the prediction performance by linear MTL and nonlinear MTL methods, and (3) identify what the learning capacity of the multikernel framework on fusing multimodality data is.

The rest of the paper is organized as follows. In Section 2, we provide a description of the multitask learning formulation. A linearized MTL and two multikernel-based MTL methods with $\ell_q\ell_1$ -norm are provided in Section 3. In Section 4, we present the experimental results and compare the performance of linearized and kernelized MTL methods from the ADNI-1 dataset. The conclusion is drawn in Section 5.

2. Multitask Learning

Consider a multitask learning (MTL) setting with T tasks. Let p be the number of covariates, shared across all the tasks, and m be the number of samples. Let $X \in \mathbb{R}^{m \times p}$ denote the matrix of covariates, $Y \in \mathbb{R}^{m \times T}$ be the matrix of responses with each row corresponding to a sample, and $\Theta \in \mathbb{R}^{p \times T}$ denote the parameter matrix, with column $\theta_{t \cdot} \in \mathbb{R}^p$ corresponding to task t , $t = 1, \dots, T$, and row $\theta_{h \cdot} \in \mathbb{R}^T$ corresponding to feature h , $h = 1, \dots, p$.

The MTL formulation focuses on the following regularized loss function:

$$\min_{\Theta \in \mathbb{R}^{p \times T}} F(Y, X, \Theta) + \lambda R(\Theta), \quad (1)$$

where $F(\cdot)$ denotes the loss function and $R(\cdot)$ is the regularizer. In the current context, we assume the loss to be square loss; that is,

$$F(Y, X, \Theta) = \|Y - X\Theta\|_F^2 = \sum_{i=1}^m \|\mathbf{y}_i - \mathbf{x}_i\Theta\|_2^2, \quad (2)$$

where $\mathbf{y}_i \in \mathbb{R}^{1 \times T}$ and $\mathbf{x}_i \in \mathbb{R}^{1 \times p}$ are the i th rows of Y and X , respectively, corresponding to the multitask response and covariates for the i th sample. We note that the MTL framework can be easily extended to other loss functions. Base on some prior knowledge, we then add penalty $R(\Theta)$ to encode the relatedness among tasks.

3. $\ell_q\ell_1$ -Norm Regularized Linearized Multitask Learning, $\ell_q\ell_1$ -MTL

The $\ell_2\ell_1$ -norm was popularly used in multitask feature learning [14]. All the existing algorithms for multitask feature learning assume a linear relationship between MRI features and cognitive scores and aim to learn a common subset of features for all tasks. Since the $\ell_2\ell_1$ -norm regularizer imposes the sparsity between all features and nonsparsity between tasks, the features that are discriminative for all tasks will get large weights. However, the $\ell_2\ell_1$ -norm is a fixed nonadaptive penalty. To obtain an adaptive regularization and better suit different data structures, we extend the $\ell_{2,1}$ -norm to a larger class of mixed norm $\ell_q\ell_1$ that can be adapted to the data. The objective function of linear $\ell_q\ell_1$ -MTL is formulated:

$$\min_{\Theta} \frac{1}{2} \|Y - X\Theta\|_F^2 + \lambda \|\Theta\|_{q,1}. \quad (3)$$

When $q = 1$, problem (3) reduces to the ℓ_1 -regularized problem; when $q = 2$, problem (3) reduces to the $\ell_{2,1}$ -regularized problem.

An efficient algorithm is based on the accelerated gradient method for solving the $\ell_q\ell_1$ -regularized problem, which is applicable for all values of q larger than 1.

First, construct the following model for approximating the composite function $\mathcal{M}(\cdot)$ at the point $\Theta^{(l)}$:

$$\begin{aligned} \mathcal{M}_{L,\Theta^{(l)}}(\Theta) := & F(\Theta^{(l)}) + \langle \Theta - \Theta^{(l)}, \nabla F(\Theta^{(l)}) \rangle \\ & + \frac{L}{2} \|\Theta - \Theta^{(l)}\|_F^2 + R(\Theta), \end{aligned} \quad (4)$$

where $L > 0$. In the model $\mathcal{M}_{L,\Theta^{(l)}}(\Theta)$, apply the first-order Taylor expansion at the point Θ (including all terms in the square bracket) for the smooth loss function $F(\cdot)$, and directly put the nonsmooth penalty $R(\cdot)$ into the model. The regularization term $(L/2)\|\Theta - \Theta^{(l)}\|_F^2$ prevents Θ from walking far away from $\Theta^{(l)}$, and thus the model can be a good approximation to $\Phi(\Theta)$ in the neighborhood of $\Theta^{(l)}$, where $\Phi(\Theta) \equiv F(\Theta) + R(\Theta)$.

The accelerated gradient method is based on two sequences $\{\Theta^{(l)}\}$ and $\{\Gamma^{(l)}\}$ in which $\{\Theta^{(l)}\}$ is the sequence of approximate solutions and $\{\Gamma^{(l)}\}$ is the sequence of search points. The search point $\Gamma^{(l)}$ is the affine combination of $\Theta^{(l-1)}$ and $\Theta^{(l)}$ as

$$\Gamma^{(l)} = \Theta^{(l)} + \beta^{(l)} (\Theta^{(l)} - \Theta^{(l-1)}), \quad (5)$$

where $\beta^{(l)}$ is a properly chosen coefficient. The approximate solution $\Theta^{(l+1)}$ is computed as the minimizer of $\mathcal{M}_{L^{(l)},\Gamma^{(l)}}(\Theta)$:

$$\Theta^{(l+1)} = \arg \min_{\Theta} \mathcal{M}_{L^{(l)},\Gamma^{(l)}}(\Theta), \quad (6)$$

where $L^{(l)}$ is determined by line search, for example, the Armijo-Goldstein rule, so that $L^{(l)}$ should be appropriate for $\Gamma^{(l)}$.

The key subroutine is (6), which can be computed as $\Theta^{(l+1)} = \pi_{1q}(\Gamma^{(l)} - \nabla F(\Gamma^{(l)})/L^{(l)}, \lambda/L^{(l)})$, where $\pi_{1q}(\cdot)$ is the $\ell_q\ell_1$ -regularized Euclidean projection (EP_{1q}) problem:

$$\pi_{1q}(V, \lambda) = \arg \min_{\Theta \in \mathbb{R}^{p \times T}} \frac{1}{2} \|\Theta - V\|_F^2 + \lambda \sum_{h=1}^p \|\theta_h\|_q. \quad (7)$$

Note that the h features in (7) are independent. In [15], the method can be used for ease of different independent groups; that is, $\pi_{1q}(V, \lambda) = \arg \min_{W \in \mathbb{R}^n} (1/2) \|W - V\|_2^2 + \lambda \sum_{i=1}^G \|w_i\|_q$, where \mathcal{G} is the independent groups. In our paper, we focus on how the method deals with multitask learning problem in (7), where \mathcal{G} is equal to p , and each group denotes the corresponding feature shared across the multiple tasks. Thus, the optimization in (7) decouples into a set of p independent ℓ_q -regularized Euclidean projection problems:

$$\pi_q(v_h) = \arg \min_{\theta_h \in \mathbb{R}^T} \frac{1}{2} \|\theta_h - v_h\|_2^2 + \lambda \|\theta_h\|_q. \quad (8)$$

Then, the optimal solution θ_h^* of (8) can be gotten as follows:

$$\begin{aligned} & \text{if } \|v_h\|_{\bar{q}} \leq \lambda, \\ & \quad \theta_h^* = \mathbf{0}; \\ & \text{else if } \|v_h\|_{\bar{q}} \geq \lambda, \quad q = 1, \\ & \quad \theta_h^* = \text{sgn}(v_h) \odot \max(|v_h| - \lambda, 0); \\ & \text{else if } \|v_h\|_{\bar{q}} \geq \lambda, \quad q = 2, \\ & \quad \theta_h^* = \frac{\|v_h\|_2 - \lambda}{\|v_h\|_2} v_h; \\ & \text{else if } \|v_h\|_{\bar{q}} \geq \lambda, \quad q = \infty, \\ & \quad \theta_h^* = \text{sgn}(v_h) \odot \min(|v_h|, u^*); \\ & \text{else } \|v_h\|_{\bar{q}} \geq \lambda, \quad 1 < q < \infty, \quad q \neq 2, \\ & \quad \theta_h^* \text{ is the unique root of } \varphi_{c^*}^{v_h}, \end{aligned} \quad (9)$$

where $\bar{q} = q/(q-1)$, and thus q and \bar{q} satisfy the following relationship: $1/\bar{q} + 1/q = 1$, u^* is the unique root of $\zeta(u) = \sum_{h=1}^p \max(|v_h| - u, 0) - \lambda$, and $\zeta(\cdot)$ is an auxiliary function, defined as $\zeta_c^v(\theta) = \theta + c\theta^{q-1} - v$ with $0 \leq \theta \leq v$; And $\varphi_c^v(\theta) = \theta + c\theta^{(q-1)} - v$, $0 < c < v$ and $c^* = \lambda \|\theta_h^*\|_q^{1-q}$. Note that $\mathbf{z} = \mathbf{x} \odot \mathbf{y}$ denotes $z_i = x_i y_i$.

The algorithm $\ell_q\ell_1$ -MTL is summarized in Algorithm 1.

4. Kernelized Multitask Learning

4.1. Multikernel Learning. The limitation in this traditional $\ell_{2,1}$ -norm MTL model is that subjects cognitive score under a task is modeled as a linear function of his/her MRI measures. The kernel methods, for example, SVM or SVR, can model the nonlinear distribution of the data by mapping the input

```

Input:  $\lambda > 0, L^{(0)} > 0, X, Y$ 
Output:  $\Theta$ .
(1) Initialize  $\Theta^{(1)} = \Theta^{(0)}, \alpha^{(-1)} = 0, \alpha^{(0)} = 1$  and  $L = L^{(0)}$ .
(2)  $l = 1$ 
(3) repeat
(4) Set  $\beta^{(l)} = (\alpha^{(l-2)} - 1)/\alpha^{(l-1)}, \Gamma^{(l)} = \Theta^{(l)} + \beta^{(l)}(\Theta^{(l)} - \Theta^{(l-1)})$ 
(5) Find the smalles  $L = L^{(l-1)}, 2L^{(l-1)}, \dots$  such that

$$\Phi(\Theta^{(l+1)}) \leq \mathcal{M}_{L,\Gamma^{(l)}}(\Theta^{(l+1)}),$$

where  $\Theta^{(l+1)} = \arg \min_{\Theta} \mathcal{M}_{L,\Gamma^{(l)}}(\Theta)$ 
(6)  $L^{(l)} = L$  and  $\alpha^{(l+1)} = (1 + \sqrt{1 + 4\alpha^{(l)^2}})/2$ 
(7)  $l = l + 1$ 
(8) until convergence criterion is satisfied

```

ALGORITHM 1: $\ell_q\ell_1$ -MTL.

data into a nonlinear feature space by kernel embedding. In this section, we consider the case that $\ell_{2,1}$ -norm regularized MTL is extended to kernel method. Let us define the kernel function $\phi_j(\mathbf{x}) : \mathbb{R}^p \rightarrow \mathbb{R}^{\hat{p}}$, which maps the data samples from an input space to a feature space (a high-dimensional Hilbert space \mathcal{H}), where \hat{p} denotes the dimensionality of the feature space and \mathbf{x} is a sample from the input space. A kernel function k is capable of attaining the inner product of two mapped datasets in \mathcal{H} : $k(\mathbf{x}, \mathbf{x}') = \phi(\mathbf{x}) \cdot \phi(\mathbf{x}')$ in the original space without explicitly computing the mapped data. The associated Gram matrix has entries $K(i, j) = k(\mathbf{x}_i, \mathbf{x}_j)$.

The most suitable types and parameters of the kernels for a particular task are often unknown, and the selection of the optimal kernel by exhaustive search on a predefined pool of kernels is usually time-consuming and sometimes causes overfitting. Multiple kernel learning (MKL) attempts to achieve better results by combining several base kernels instead of using only one specific kernel. MKL assumes that \mathbf{x}_i can be mapped to k different Hilbert spaces, $\mathbf{x}_i \rightarrow \phi_j(\mathbf{x}_i)$, $j = 1, \dots, k$, implicitly with k nonlinear mapping functions, and the objective of MKL is to seek the optimal kernel combination $\widehat{k}(\mathbf{x}, \mathbf{x}') = \sum_{j=1}^k d_j k_j(\mathbf{x}, \mathbf{x}')$, $d_j \geq 0$, $\sum_{j=1}^k d_j = 1$, where \mathbf{d} is the kernel weight vector. The primal objective function of multiple kernel regression model is written as follows:

$$\begin{aligned}
\min_{\tilde{\theta}, \xi} \quad & \frac{1}{2} \sum_{j=1}^k \frac{\|\tilde{\theta}_j\|_2^2}{d_j} + \frac{\lambda}{2} \sum_{i=1}^m \xi_i^2, \\
\text{s.t.} \quad & \sum_{j=1}^k \tilde{\theta}_j^T \phi_j(x_i) - y_i = \xi_i, \\
& \sum_{j=1}^k d_j = 1, \quad d_j \geq 0.
\end{aligned} \tag{10}$$

MKL learns both the weights of the kernel combination \mathbf{d} and the parameters of the regression $\tilde{\theta}$ by solving a single joint optimization problem.

Using α to denote the Lagrange multipliers, the objective value of the dual problem of (10) can be written as follows:

$$\begin{aligned}
J(\mathbf{d}) = \max_{\alpha} \quad & -\alpha^T \mathbf{y}_t - \frac{1}{2} \alpha^T \widehat{\mathbf{K}} \alpha - \frac{1}{2C} \alpha^{*T} \alpha, \\
\text{s.t.} \quad & \sum_{j=1}^k d_j = 1, \quad d_j \geq 0,
\end{aligned} \tag{11}$$

where $\widehat{\mathbf{K}} = \sum_{j=1}^k d_j \mathbf{K}_j$ is the combined Gram matrix and \mathbf{K}_j , $j = 1, \dots, k$, is the given set of base kernels.

4.2. $\ell_q\ell_1$ -Norm Regularized Multikernel Multitask Learning, $\ell_q\ell_1$ -MKMTL. We follow the multiple kernel learning scheme and use the $\ell_{q,1}$ -norm to model the relationship between the tasks to learn a common kernel representation by imposing sparsity constraint on the kernel weight. The method, called $\ell_q\ell_1$ -MKMTL, assumes that few base kernels are important for the tasks and encourages a linear combination of only few kernels and assumes few selected kernels are similar across the tasks. The formulation of $\ell_q\ell_1$ -MKMTL can be expressed as follows:

$$\begin{aligned}
\min_{\tilde{\theta}, \xi} \quad & \frac{1}{2} \left(\sum_{j=1}^k \left(\sum_{t=1}^T \|\tilde{\theta}_{tj}\|_2^q \right)^{1/q} \right)^2 + \frac{\lambda}{2} \sum_{t=1}^T \sum_{i=1}^{m_t} \xi_{ti}^2, \\
\text{s.t.} \quad & \sum_{j=1}^k \tilde{\theta}_{tj}^T \phi_j(x_{ti}) - y_{ti} = \xi_{ti}.
\end{aligned} \tag{12}$$

We now rewrite this formulation in a convenient form which can be efficiently solved using mirror-descent based algorithms. We introduce some more notations: let $\Delta_{d,r} = \{\mathbf{z} \equiv [z_1, \dots, z_d]^T \mid \sum_{i=1}^d z_i^r \leq 1, z_i \geq 0, i = 1, \dots, d\}$ and with slight abuse of notation let $\Delta_{d,1} = \Delta_d$. Next, we note the following [16].

Lemma 1. Let $a_i \geq 0$, $i = 1, \dots, d$ and $1 < r < \infty$. Then, for $\Delta_{d,r}$ defined as before,

$$\min_{\eta \in \Delta_{d,r}} \sum_i \frac{a_i}{\eta_i} = \left(\sum_{i=1}^d a_i^{r/(r+1)} \right)^{(r+1)/r}, \quad (13)$$

and the minimum is attained at

$$\eta_i = \frac{a_i^{1/(r+1)}}{\left(\sum_{i=1}^d a_i^{r/(r+1)} \right)^{1/r}}, \quad (14)$$

with the convention that $a/0$ is 0 if $a = 0$ and is ∞ if $a \neq 0$.

Using the result of the lemma (with $r = 1$) and introducing variables $\mu = [\mu_1, \dots, \mu_k]^T$, we have

$$\begin{aligned} & \left(\sum_{j=1}^k \left(\sum_{t=1}^T (\|\tilde{\theta}_{tj}\|_2)^q \right)^{1/q} \right)^2 \\ &= \min_{\mu \in \Delta_k} \sum_{j=1}^k \frac{\left(\sum_{t=1}^T (\|\tilde{\theta}_{tj}\|_2)^q \right)^{2/q}}{\mu_j}. \end{aligned} \quad (15)$$

Now introducing dual variables $\gamma_j = [\gamma_{j1}, \dots, \gamma_{jT}]^T$, $j = 1, \dots, k$, and using the notion of dual norm [17], we obtain

$$\left(\sum_{t=1}^T (\|\tilde{\theta}_{tj}\|_2^2)^{q/2} \right)^{2/q} = \max_{\gamma_j \in \Delta_{T,\bar{q}}} \sum_{t=1}^T \gamma_{jt} \|\tilde{\theta}_{tj}\|_2^2, \quad (16)$$

where $\bar{q} = q/(q-2)$. With this, the objective in the $\ell_q \ell_1$ -MKMTL formulation can now be written as

$$\min_{\mu \in \Delta_k} \max_{\tilde{\theta}, \xi} \max_{\gamma_j \in \Delta_{T,\bar{q}}} \frac{1}{2} \sum_{j=1}^k \frac{\sum_{t=1}^T \gamma_{jt} \|\tilde{\theta}_{tj}\|_2^2}{\mu_j} + \frac{\lambda}{2} \sum_{t=1}^T \sum_{i=1}^{m_t} \xi_{ti}^2. \quad (17)$$

Using α to denote the Lagrange multipliers, this has the Lagrangian

$$\begin{aligned} \mathcal{L} = & \frac{1}{2} \sum_{j=1}^k \frac{\sum_{t=1}^T \gamma_{jt} \|\tilde{\theta}_{tj}\|_2^2}{\mu_j} + \frac{\lambda}{2} \sum_{t=1}^T \sum_{i=1}^{m_t} \xi_{ti}^2 \\ & + \sum_{t=1}^T \sum_{i=1}^{m_t} \alpha_{ti} \left(\sum_{j=1}^k \tilde{\theta}_{tj}^T \phi_j(x_{ti}) - y_{ti} - \xi_{ti} \right). \end{aligned} \quad (18)$$

Recall our foray into Lagrange duality. We can solve the original problem by doing

$$\max_{\alpha} \min_{\tilde{\theta}, \xi} \mathcal{L}(\tilde{\theta}, \xi, \alpha). \quad (19)$$

To begin, we attack the inner minimization: For fixed α , we would like to solve for the minimizing $\tilde{\theta}$ and ξ . We can do this by setting the derivatives of \mathcal{L} with respect to ξ_{ti} and $\tilde{\theta}$ to be zero. Doing this, we can find

$$\tilde{\theta}_{tj}^* = -\alpha_t^T \left[\sum_{j=1}^k \frac{\mu_j \Phi_{tj}}{\gamma_{jt}} \right], \quad (20a)$$

$$\xi_{ti}^* = \frac{\alpha_{ti}}{\lambda}, \quad (20b)$$

where α_t is a vector corresponding to the t th task in the $\ell_q \ell_1$ -MKMTL formulation and Φ_{tj} is the data matrix with columns as $\phi_j(x_{ti})$, $i = 1, \dots, m_t$. So, we can solve the problem by maximizing the Lagrangian (with respect to α), where we substitute the above expressions for ξ and $\tilde{\theta}$. Thus, we have an unconstrained maximization.

$$\max_{\alpha} \sum_{t=1}^T \left\{ -\alpha_t^T y_t - \frac{1}{2} \alpha_t^T \left[\sum_{j=1}^k \frac{\mu_j K_{tj}}{\gamma_{jt}} \right] \alpha_t - \frac{1}{2\lambda} \alpha_t^T \alpha_t \right\}. \quad (21)$$

Here, y_t is vector of scores of the t th task training data points and K_{tj} represents the Gram matrix of the t th task training data points with respect to the j th kernel. Equation (21) is just a quadratic in α . As such, we can find the optimum as the solution of a linear system.

Then, (17) can be written as follows:

$$\min_{\mu \in \Delta_k} \max_{\gamma_j \in \Delta_{T,\bar{q}}} \max_{\alpha} \sum_{t=1}^T \left\{ -\alpha_t^T y_t - \frac{1}{2} \alpha_t^T \left[\sum_{j=1}^k \frac{\mu_j K_{tj}}{\gamma_{jt}} \right] \alpha_t - \frac{1}{2\lambda} \alpha_t^T \alpha_t \right\}. \quad (22)$$

The formulation can be transformed as follows:

$$\min_{\mu \in \Delta_k} \max_{\gamma_j \in \Delta_{T,\bar{q}}} \max_{\alpha} \sum_{t=1}^T \left\{ -\alpha_t^T y_t - \frac{1}{2} \alpha_t^T \left[\sum_{j=1}^k \frac{\mu_j K_{tj}}{\gamma_{jt}} \right] \alpha_t - \frac{1}{2\lambda} \alpha_t^T \alpha_t \right\}. \quad (23)$$

The algorithm $\ell_q \ell_1$ -MKMTL is summarized in Algorithm 2.

4.3. $\ell_{2,1}-\ell_q$ -Norm Regularized Multikernel Multitask Learning, $\ell_{2,1}\ell_q$ -MKMTL. The linearized $\ell_q \ell_1$ -MTL assumed linear

```

Input:  $\lambda > 0, X, Y$ 
Output:  $\alpha, \nu, \mu$ 
(1)  $n = 0$ 
(2) repeat
(3)   initiate  $\mu$  and  $\nu$ 
(4)   for  $t = 1$  to  $T$  do
(5)     With fixed  $\mu$  and  $\nu$ , compute  $\alpha_t^*$  by using an SVR solver
(6)   end for
(7)   optimize  $\mu$  with mirror-descent algorithm
(8)   optimize  $\nu: -\sum_{j=1}^k \min_{\nu_j} \in \Delta_{T,\bar{q}} \sum_{t=1}^T (D_{jt}/\nu_{jt})$  where  $D_{jt} = (1/2)\mu_j \alpha_t^T K_{jt} \alpha_t$ .
(9)    $n = n + 1$ 
(10)  until convergence criterion is satisfied

```

ALGORITHM 2: $\ell_q \ell_1$ -MKMTL.

relationship between the MRI features and the cognitive outcomes. Such a model is the lack of capability to capture nonlinear predictive information from the features. Although the $\ell_q \ell_1$ -MKMTL builds the nonlinear relationship for the features and task by mapping to high-dimensional space, it considers that tasks to be learned share a common subset of kernel representations without capturing the interrelationships between different cognitive measures over the feature space.

To overcome the weaknesses of the previous two methods, we project the original feature vectors to a high-dimensional space using multiple nonlinear mapping functions for performing regression task in a nonlinear manner and utilize multitask learning in the multiple kernel spaces for modeling the disease's cognitive scores with a joint $\ell_{2,1}$ - ℓ_q sparsity-inducing regularizers. Moreover, we construct new features as orthogonal transforms of the given features, that is, $L_j \phi_j(x)$, where L_j is an orthogonal matrix which is to be learned. Again, low empirical risk over each task would imply minimizing the following quadratic loss: $\sum_{t=1}^T \sum_{i=1}^{m_t} \min(\sum_{j=1}^k \tilde{\theta}_{tj} L_j^T \phi_j(x_{ti}) - y_{ti})^2$. Before describing the regularization term, we introduce some more notations: Let the entries of $\tilde{\theta}_{tj}$ be $\tilde{\theta}_{tjl}$, $l = 1, \dots, p_j$, where p_j is the dimensionality of the feature space induced by the j th kernel. By $\tilde{\theta}_{.jl}$ we denote the vector with entries $\tilde{\theta}_{tjl}$, $t = 1, \dots, T$. The regularization term we employ is $(\sum_{j=1}^k (\sum_{l=1}^{p_j} \|\tilde{\theta}_{.jl}\|_2)^q)^{2/q}$, where $q \in [1, 2]$. Different from $\ell_q \ell_1$ -MKMTL, the ℓ_q -norm in $\ell_{2,1} \ell_q$ -MKMTL is employed over the kernels rather than the tasks.

Mathematically, the $\ell_{2,1} \ell_q$ -MKMTL formulation can be expressed as follows:

$$\begin{aligned} \min_{\tilde{\theta}, \xi, L} \quad & \frac{1}{2} \left(\sum_{j=1}^k \left(\sum_{l=1}^{p_j} \|\tilde{\theta}_{.jl}\|_2 \right)^q \right)^{2/q} + \frac{\lambda}{2} \sum_{t=1}^T \sum_{i=1}^{m_t} \xi_{ti}^2, \\ \text{s.t.} \quad & \sum_{j=1}^k \tilde{\theta}_{tj}^T L_j^T \phi_j(x_{ti}) - y_{ti} = \xi_{ti}, \quad L_j \in O^{p_j}, \end{aligned} \quad (24)$$

where O^{p_j} represents the set of all orthogonal matrices of dimensionality p_j . In the following text, we rewrite this formulation in a form which is convenient to solve using an MD based algorithm.

Using the result of Lemma 1 and introducing new variables $\nu = [\nu_1, \dots, \nu_k]^T$, we have

$$\left(\sum_{j=1}^k \left(\sum_{l=1}^{p_j} \|\tilde{\theta}_{.jl}\|_2 \right)^q \right)^{2/q} = \min_{\nu \in \Delta_{k,\bar{q}}} \sum_{j=1}^k \frac{\left(\sum_{l=1}^{p_j} \|\tilde{\theta}_{.jl}\|_2 \right)^2}{\nu_j}, \quad (25)$$

where $\bar{q} = q/(2-q)$. Again using the lemma and introducing new variables $\mu_j = [\mu_{j1}, \dots, \mu_{jp_j}]^T$, $j = 1, \dots, k$, the regularizer can be written as

$$\min_{\nu \in \Delta_{k,\bar{q}}} \min_{\mu_j \in \Delta_{p_j}} \sum_{t=1}^T \sum_{j=1}^k \sum_{l=1}^{p_j} \frac{\tilde{\theta}_{tjl}^2}{\mu_{jk} \nu_j}. \quad (26)$$

Now, we perform a change of variables: $\tilde{\theta}_{tjl}/\sqrt{\mu_{jk} \nu_j} = \bar{\theta}_{tjl}$, $l = 1, \dots, p_j$. Using this, one can rewrite the $\ell_{2,1} \ell_q$ -MKMTL formulation as

$$\begin{aligned} \min_{\nu, \mu_j, L} \quad & \sum_{t=1}^T \min_{\bar{\theta}_t, \xi_t} \frac{1}{2} \sum_{j=1}^k \bar{\theta}_{tj}^T \bar{\theta}_{tj} + \frac{\lambda}{2} \sum_{t=1}^T \sum_{i=1}^{m_t} \xi_{ti}^2, \\ \text{s.t.} \quad & \sum_{j=1}^k \bar{\theta}_{tj}^T \Lambda_j^{1/2} L_j^T \phi_j(x_{ti}) - y_{ti} = \xi_{ti}, \end{aligned} \quad (27)$$

$$\nu \in \Delta_{k,\bar{q}}, \mu_j \in \Delta_{p_j}, L_j \in O^{p_j},$$

where Λ_j is a diagonal matrix with entries as $\nu_j \mu_{jl}$, $l = 1, \dots, p_j$.

Now, using α to denote the Lagrange multipliers, this has the Lagrangian of

$$\begin{aligned} \mathcal{L} = \sum_{t=1}^T & \left(\frac{1}{2} \sum_{j=1}^k \bar{\theta}_{tj}^T \bar{\theta}_{tj} + \frac{\lambda}{2} \sum_{i=1}^{m_t} \xi_{ti}^2 \right. \\ & \left. + \sum_{i=1}^{m_t} \alpha_{ti} \left(\sum_{j=1}^k \bar{\theta}_{tj}^T \Lambda_j^{1/2} L_j^T \phi_j(x_{ti}) - y_{ti} - \xi_{ti} \right) \right). \end{aligned} \quad (28)$$

This can be solved like $\ell_q \ell_1$ -MKMTL:

$$\tilde{\theta}_{tj}^* = -\alpha_{ti}^T \Lambda_j^{1/2} L_j^T \Phi_{tj}, \quad (29a)$$

$$\xi_{ti}^* = \frac{\alpha_{ti}}{\lambda}. \quad (29b)$$

```

Input:  $X, Y, \lambda > 0$ 
Output:  $\alpha^*, \bar{\mathbf{Q}}$ 
(1) repeat
(2)   optimize  $\bar{\mathbf{Q}}$  with mirror-descent algorithm
(3)   for  $t = 1$  to  $T$  do
(4)     with fixed  $\bar{\mathbf{Q}}$ , compute  $\alpha_t^*$  by using an SVR solver
(5)   end for
(6)    $n = n + 1$ 
(7) until convergence criterion is satisfied

```

ALGORITHM 3: $\ell_{2,1}$ - ℓ_q -MKMTL.

Again, we substitute the above expressions for ξ and $\tilde{\theta}$. Thus, we have the following form:

$$\begin{aligned} \min_{\nu, \mu_j, \mathbf{L}_j} \quad & \sum_{t=1}^T \max_{\alpha_t} -\alpha_t^T y_t - \frac{1}{2} \alpha_t^T \left(\sum_{j=1}^k \Phi_{tj}^T \mathbf{L}_j^T \Lambda_j \mathbf{L}_j \Phi_{tj} \right) \alpha_t \\ \text{s.t.} \quad & \nu \in \Delta_{k, \bar{q}}, \mu_j \in \Delta_{p_j}, \mathbf{L}_j \in O^{p_j}. \end{aligned} \quad (30)$$

Denoting $\mathbf{L}_j^T \Lambda_j \mathbf{L}_j$ by $\bar{\mathbf{Q}}_j$ and eliminating variables ν, μ , and \mathbf{L} 's lead to

$$\begin{aligned} \min_{\bar{\mathbf{Q}}} \quad & \sum_{t=1}^T \max_{\alpha_t} -\alpha_t^T y_t - \frac{1}{2} \alpha_t^T \left(\sum_{j=1}^k \Phi_{tj}^T \bar{\mathbf{Q}}_j \Phi_{tj} \right) \alpha_t \\ \text{s.t.} \quad & \bar{\mathbf{Q}}_j \succeq 0, \sum_{j=1}^k (\text{tr}(\bar{\mathbf{Q}}_j))^{\bar{q}} \leq 1. \end{aligned} \quad (31)$$

The difficulty in working with this formulation is that the explicit mappings ϕ_j 's are required. We now describe a way of overcoming this problem and efficiently kernelizing the formulation (refer to [1] also). Let $\Phi_j \equiv [\Phi_{1j}, \dots, \Phi_{Tj}]$ and the compact SVD of Φ_j be $\mathbf{U}_j \Sigma_j \mathbf{V}_j^T$. Then, we introduce a symmetric positive semidefinite \mathbf{Q}_j with the same rank as that of Φ_j such that $\bar{\mathbf{Q}}_j = \mathbf{U}_j \mathbf{Q}_j \mathbf{U}_j^T$. By eliminating $\bar{\mathbf{Q}}_j$, we can rewrite the above problem using \mathbf{Q}_j as

$$\begin{aligned} \min_{\mathbf{Q}} \quad & \sum_{t=1}^T \max_{\alpha_t} -\alpha_t^T y_t - \frac{1}{2} \alpha_t^T \left(\sum_{j=1}^k \mathbf{M}_{tj}^T \mathbf{Q}_j \mathbf{M}_{tj} \right) \alpha_t \\ \text{s.t.} \quad & \mathbf{Q}_j \succeq 0, \sum_{j=1}^k (\text{tr}(\mathbf{Q}_j))^{\bar{q}} \leq 1, \end{aligned} \quad (32)$$

where $\mathbf{M}_{tj} = \Sigma_j^{-1} \mathbf{V}_j^T \Phi_j^T \Phi_{tj}$. Note that calculation of \mathbf{M}_{tj} does not require the kernel-induced features explicitly and

hence the formulation is kernelized. It can be transformed as follows:

$$\min_{\mathbf{Q}} \quad f(\mathbf{Q}) = \sum_{t=1}^T -\alpha_t^T y_t - \frac{1}{2} \text{tr}(\mathbf{Q}\mathbf{B}) - \frac{1}{2\lambda} \alpha_t^T \alpha_t, \quad (33)$$

where \mathbf{B} is a block diagonal matrix with entries as $\mathbf{B}_j = \sum_{t=1}^T \mathbf{M}_{tj} \alpha_t \alpha_t^T \mathbf{M}_{tj}^T$.

\mathbf{Q} can be solved by mirror-descent. The gradient of ∇f with respect to \mathbf{Q} is calculated as follows:

$$\nabla f(\mathbf{Q}^{(l)}) = -\frac{1}{2} \mathbf{B}^{(l)}, \quad (34)$$

where $\mathbf{B}^{(l)}$ is the value obtained using optimal α_t obtained while evaluating $f(\mathbf{Q}^{(l)})$.

The algorithm $\ell_{2,1}$ - ℓ_q MKMTL is summarized in Algorithm 3.

5. Experimental Results and Discussions

5.1. Experimental Setup. We use 10-fold cross validation to evaluate our model and conduct the comparison. In each of ten trials, a 5-fold nested cross validation procedure is employed to tune the regularization parameters. Data was z -scored before applying regression methods. The range of each parameter varied from 10^{-1} to 10^3 . The candidate kernels are as follows: six different kernel bandwidths ($2^{-2}, 2^{-1}, \dots, 2^3$), polynomial kernels of degrees 1 to 3, and a linear kernel, which totally yields 10 kernels. The kernel matrices were precomputed and normalized to have unit trace. The reported results were the best results of each method with the optimal parameter. For the quantitative performance evaluation, we employed the metrics of Correlation Coefficient (CC) and Root Mean Squared Error (rMSE) between the predicted clinical scores and the target clinical scores for each regression task. Moreover, to evaluate the overall performance on all the tasks, the normalized mean squared error (nMSE) [7, 18] and weighted R-value (wR) [4] are used. The nMSE and wR are defined as follows:

$$\text{nMSE}(Y, \hat{Y}) = \frac{\sum_{t=1}^T (\|Y_t - \hat{Y}_t\|_2^2 / \sigma(Y_t))}{\sum_{t=1}^T m_t}, \quad (35)$$

$$\text{wR}(Y, \hat{Y}) = \frac{\sum_{t=1}^T \text{Corr}(Y_t, \hat{Y}_t) m_t}{\sum_{t=1}^T m_t}, \quad (36)$$

where Y and \hat{Y} are the ground truth cognitive scores and the predicted cognitive scores, respectively.

A smaller (higher) value of nMSE and rMSE (CC and wR) represents better regression performance. We report the mean and standard deviation based on 10 iterations of experiments on different splits of data for all comparable experiments.

In ADNI, all participants received 1.5-Tesla (T) structural MRI. The MRI features used in our experiments are based on the imaging data from the ADNI database processed by a team from UCSF (University of California at San Francisco), who performed cortical reconstruction and volumetric segmentations with the FreeSurfer image analysis suite (<http://surfer.nmr.mgh.harvard.edu/>) according to the atlas generated in [19]. Totally, 48 cortical regions and 44 subcortical regions are generated. For each cortical region, the cortical thickness average (TA), standard deviation of thickness (TS), surface area (SA), and cortical volume (CV) were calculated as features. For each subcortical region, subcortical volume was calculated as features. The SA of left and right hemisphere and total intracranial volume (ICV) were also included. This yielded a total of $p = 319$ MRI features extracted from cortical/subcortical ROIs in each hemisphere (including 275 cortical and 44 subcortical features). Details of the analysis procedure are available at <http://adni.loni.usc.edu/methods/mri-analysis/>.

Ten widely used clinical/cognitive assessment scores [3, 20, 21] were employed in this study, including Alzheimer’s Disease Assessment Scale (ADAS) cognitive total score, Mini Mental State Exam (MMSE) score, Rey Auditory Verbal Learning Test (RAVLT) involving total score of the first 5 learning trials (TOTAL), Trial 6 total number of words recalled (TOT6), 30-minute delay score (T30), and 30-minute delay recognition score (RECOG), FLU involving animal total score (ANIM) and vegetable total score (VEG), and TRAILS including Trail Making test A score and B score.

5.2. Comparison with the State-of-the-Art MTL Methods. To compare the kernelized MTL with the other linearized one and illustrate how well the two multikernel-based MTL methods work by means of modeling the correlation among the tasks, we comprehensively compare our proposed methods with several popular state-of-the-art related methods. Representative comparable algorithms include

- (1) Ridge [22]: $\min_{\Theta} L(X, Y, \Theta) + \lambda \|\Theta\|_F^2$
- (2) Lasso [23]: $\min_{\Theta} L(X, Y, \Theta) + \lambda \|\Theta\|_1$
- (3) MKL [24]: $\min_{\tilde{\theta}, \xi} (1/2) \|f\|_{\mathcal{H}}^2 + \lambda \sum_i \xi_i$, such that $y_i(f(x_i) + b) \geq 1 - \xi_i$ and $\xi_i \geq 0$, $\forall i$
- (4) Robust Multitask Feature Learning (RMTL) [25]: RMTL ($\min_{\Theta} L(X, Y, \Theta) + \lambda_1 \|P\|_* + \lambda_2 \|S\|_{2,1}$, subject to $\Theta = P + S$), which assumes that the model Θ can be decomposed into two components: a shared feature structure P capturing task relatedness and a group-sparse structure S detecting outliers
- (5) Clustered Multitask Learning (CMTL) [16]: CMTL ($\min_{\Theta, M: M^T M = I_c} L(X, Y, \Theta) + \lambda_1 (\text{tr}(\Theta^T \Theta) - \text{tr}(M^T \Theta^T \Theta M)) + \lambda_2 \text{tr}(\Theta^T \Theta)$, where $M \in \mathbb{R}^{c \times k}$ is

an orthogonal cluster indicator matrix and the tasks are clustered into $c < k$ clusters) incorporating a regularization term to induce clustering between tasks and then sharing information only to tasks belonging to the same cluster. In the CMTL, the number of clusters is set to 11 since the 20 tasks belong to 11 sets of cognitive functions

- (6) Trace-norm regularized multitask learning (Trace) [17]: assuming that all models share a common low-dimensional subspace ($\min_{\Theta} L(X, Y, \Theta) + \lambda \|\Theta\|_*$)
- (7) Sparse regularized multitask learning formulation (SRMTL) [26]: SRMTL ($\min_{\Theta} L(X, Y, \Theta) + \lambda_1 \|\Theta\|_F^2 + \lambda_2 \|\Theta\|_1$, where $\mathcal{Z} \in \mathbb{R}^{T \times T}$) containing two regularization processes: (1) all tasks are regularized by their mean value, and therefore knowledge from one task can be utilized by other tasks via the mean value; (2) sparsity is enforced in the learning with ℓ_1 -norm.

Experimental results are reported in Tables 1 and 2 where the best results are boldfaced. A first glance at the results shows that $\ell_{2,1}\ell_q$ -MKMTL generally outperforms all the other compared methods on both metrics and across all the cognitive tasks. Additionally, a statistical analysis is performed on the results. As can be seen, our proposed method achieves statistically significant results compared to all the other methods on most of the results. These results reveal several interesting points:

- (1) All the compared multitask learning methods ($\ell_q\ell_1$ -MTL, $\ell_q\ell_1$ -MKMTL, and $\ell_{2,1}\ell_q$ -MKMTL) improve the predictive performance over the independent regression algorithms (Ridge, Lasso, and MKL). This justifies the motivation of learning multiple tasks simultaneously.
- (2) The two multikernel-based MTL methods outperform the linearized $\ell_q\ell_1$ -MTL in terms of nMSE, and $\ell_{2,1}\ell_q$ -MKMTL outperforms the linearized $\ell_q\ell_1$ -MTL in terms of wR. It indicates that the nonlinear MTL models via kernel functions can capture complex patterns between brain images and the corresponding cognitive measures.
- (3) By the appropriate $\ell_{2,1}\ell_q$ regularization, the $\ell_{2,1}\ell_q$ -MKMTL model enables us (1) to obtain capture nonlinear associations between MRI and cognitive outcomes, (2) to obtain the intrinsic relationships between multiple related tasks in \mathcal{H} , and (3) to promote the sparse kernel combinations to support the interpretability and scalability. The outcomes demonstrate that $\ell_{2,1}\ell_q$ -MKMTL outperforms $\ell_q\ell_1$ -MTL and $\ell_q\ell_1$ -MKMTL, both of which neglect the inherently nonlinear relationship between MRI and cognitive outcomes, and the correlation among multiple related tasks in the feature space.
- (4) Compared with the other multitask learning methods with different assumptions, our proposed methods belong to the multitask feature learning methods with

TABLE 1: Performance comparison of various methods in terms of rMSE and nMSE on 10 cross validation cognitive prediction tasks.

Method	ADAS	MMSE	TOTAL	RAVLT		
				TOT6	T30	RECOG
Ridge	7.556 ± 0.294	2.656 ± 0.134	11.41 ± 0.498	3.907 ± 0.236	4.052 ± 0.224	4.331 ± 0.294
Lasso	6.846 ± 0.361	2.216 ± 0.098	10.02 ± 0.548	3.320 ± 0.195	3.443 ± 0.177	3.639 ± 0.213
MKL	6.893 ± 0.528	2.214 ± 0.106	9.911 ± 0.695	3.424 ± 0.296	3.570 ± 0.340	3.745 ± 0.237
Robust MTL	7.651 ± 0.442	3.326 ± 0.266	11.02 ± 0.590	3.574 ± 0.235	3.704 ± 0.171	3.858 ± 0.310
CMTL	7.642 ± 0.373	3.083 ± 0.461	11.56 ± 0.510	3.907 ± 0.260	4.038 ± 0.244	4.381 ± 0.226
Trace	8.180 ± 0.605	6.113 ± 2.038	13.09 ± 3.128	3.782 ± 0.491	3.906 ± 0.431	4.520 ± 0.859
SRMTL	6.882 ± 0.325	2.331 ± 0.271	9.961 ± 0.561	3.320 ± 0.152	3.445 ± 0.116	3.639 ± 0.261
$\ell_q\ell_1$ -MTL	6.772 ± 0.312	2.206 ± 0.081	9.606 ± 0.448	3.344 ± 0.154	3.440 ± 0.151	3.644 ± 0.247
$\ell_q\ell_1$ -MKMTL	6.825 ± 0.455	2.417 ± 0.197	9.699 ± 0.505	3.396 ± 0.188	3.495 ± 0.144	3.653 ± 0.243
$\ell_{2,1}\ell_q$ -MKMTL	6.806 ± 0.447	2.185 ± 0.106	9.628 ± 0.510	3.331 ± 0.196	3.467 ± 0.172	3.627 ± 0.199
Method	FLU		TRAILS		nMSE	
	ANIM	VEG	A	B		
Ridge	6.521 ± 0.418	4.322 ± 0.178	27.18 ± 1.702	83.72 ± 5.713	16.44 ± 1.725	
Lasso	5.352 ± 0.447	3.701 ± 0.093	23.75 ± 1.398	71.23 ± 2.812	12.05 ± 0.758	
MKL	5.342 ± 0.510	3.761 ± 0.137	24.71 ± 1.781	78.09 ± 6.916	13.56 ± 1.133	
Robust MTL	5.946 ± 0.398	3.988 ± 0.083	27.78 ± 1.922	90.12 ± 7.098	17.68 ± 2.303	
CMTL	6.608 ± 0.561	4.398 ± 0.284	27.46 ± 1.980	83.66 ± 5.418	16.67 ± 1.912	
Trace	6.743 ± 1.425	4.672 ± 0.778	28.82 ± 3.278	89.68 ± 7.838	20.23 ± 5.215	
SRMTL	5.327 ± 0.334	3.713 ± 0.088	25.09 ± 1.421	80.00 ± 4.637	14.01 ± 1.169	
$\ell_q\ell_1$ -MTL	5.298 ± 0.439	3.704 ± 0.096	23.42 ± 1.110	71.32 ± 2.945	11.92 ± 0.969	
$\ell_q\ell_1$ -MKMTL	5.304 ± 0.350	3.676 ± 0.094	23.09 ± 1.438	70.28 ± 0.898	11.72 ± 0.222	
$\ell_{2,1}\ell_q$ -MKMTL	5.232 ± 0.434	3.675 ± 0.157	23.13 ± 1.473	69.82 ± 1.236	11.56 ± 0.602	

TABLE 2: Performance comparison of various methods in terms of CC and wR on 10 cross validation cognitive prediction tasks.

Method	ADAS	MMSE	TOTAL	RAVLT		
				TOT6	T30	RECOG
Ridge	0.603 ± 0.031	0.407 ± 0.040	0.401 ± 0.084	0.361 ± 0.092	0.377 ± 0.096	0.261 ± 0.080
Lasso	0.655 ± 0.036	0.540 ± 0.046	0.493 ± 0.084	0.507 ± 0.100	0.523 ± 0.106	0.416 ± 0.087
MKL	0.658 ± 0.030	0.544 ± 0.052	0.502 ± 0.066	0.476 ± 0.095	0.506 ± 0.105	0.391 ± 0.072
Robust MTL	0.587 ± 0.022	0.338 ± 0.084	0.423 ± 0.090	0.432 ± 0.096	0.444 ± 0.094	0.354 ± 0.105
CMTL	0.603 ± 0.025	0.381 ± 0.042	0.397 ± 0.072	0.362 ± 0.090	0.381 ± 0.099	0.260 ± 0.068
Trace	0.548 ± 0.039	0.144 ± 0.091	0.342 ± 0.172	0.395 ± 0.159	0.402 ± 0.142	0.253 ± 0.130
SRMTL	0.655 ± 0.034	0.525 ± 0.058	0.492 ± 0.079	0.505 ± 0.097	0.523 ± 0.103	0.413 ± 0.092
$\ell_q\ell_1$ -MTL	0.662 ± 0.043	0.532 ± 0.056	0.532 ± 0.082	0.492 ± 0.109	0.522 ± 0.105	0.404 ± 0.091
$\ell_q\ell_1$ -MKMTL	0.661 ± 0.034	0.460 ± 0.099	0.519 ± 0.072	0.470 ± 0.089	0.494 ± 0.094	0.412 ± 0.090
$\ell_{2,1}\ell_q$ -MKMTL	0.660 ± 0.035	0.547 ± 0.045	0.529 ± 0.079	0.500 ± 0.095	0.508 ± 0.094	0.421 ± 0.075
Method	FLU		TRAILS		wR	
	ANIM	VEG	A	B		
Ridge	0.185 ± 0.090	0.396 ± 0.073	0.291 ± 0.097	0.330 ± 0.110	0.361 ± 0.041	
Lasso	0.365 ± 0.096	0.506 ± 0.059	0.363 ± 0.041	0.467 ± 0.096	0.484 ± 0.049	
MKL	0.375 ± 0.071	0.496 ± 0.067	0.374 ± 0.056	0.457 ± 0.060	0.478 ± 0.046	
Robust MTL	0.253 ± 0.096	0.443 ± 0.057	0.282 ± 0.113	0.292 ± 0.123	0.385 ± 0.038	
CMTL	0.180 ± 0.089	0.390 ± 0.071	0.287 ± 0.116	0.335 ± 0.112	0.358 ± 0.036	
Trace	0.212 ± 0.143	0.331 ± 0.112	0.270 ± 0.112	0.290 ± 0.122	0.319 ± 0.083	
SRMTL	0.362 ± 0.093	0.503 ± 0.064	0.340 ± 0.063	0.361 ± 0.095	0.468 ± 0.045	
$\ell_q\ell_1$ -MTL	0.379 ± 0.076	0.501 ± 0.063	0.399 ± 0.060	0.467 ± 0.098	0.489 ± 0.050	
$\ell_q\ell_1$ -MKMTL	0.381 ± 0.080	0.521 ± 0.067	0.421 ± 0.064	0.481 ± 0.076	0.482 ± 0.047	
$\ell_{2,1}\ell_q$ -MKMTL	0.409 ± 0.073	0.516 ± 0.065	0.417 ± 0.067	0.490 ± 0.087	0.500 ± 0.043	

TABLE 3: Performance comparison of various methods with fusing multiple modalities data in terms of rMSE and nMSE on 10 cross validation cognitive prediction tasks.

Method	ADAS	MMSE	FLU ANIM	TRAILS	
				A	B
$\ell_q\ell_1$ -MTL-MRI	6.494 ± 1.029	1.964 ± 0.306	4.911 ± 0.256	16.39 ± 2.906	55.82 ± 7.689
$\ell_q\ell_1$ -MTL-PET	6.941 ± 1.244	2.118 ± 0.298	5.192 ± 0.145	16.56 ± 3.533	56.88 ± 9.447
$\ell_q\ell_1$ -MTL-MP	6.219 ± 1.037	2.067 ± 0.293	4.928 ± 0.260	16.09 ± 2.768	53.70 ± 7.144
$\ell_q\ell_1$ -MTL-ALL	6.174 ± 0.978	2.062 ± 0.272	4.789 ± 0.206	15.97 ± 2.785	53.37 ± 7.243
$\ell_q\ell_1$ -MKMTL-MRI	6.369 ± 0.941	2.074 ± 0.291	4.993 ± 0.235	16.18 ± 3.089	55.95 ± 9.479
$\ell_q\ell_1$ -MKMTL-PET	6.812 ± 1.155	2.060 ± 0.364	5.151 ± 0.227	16.61 ± 3.588	57.85 ± 11.24
$\ell_q\ell_1$ -MKMTL-MP	6.112 ± 0.886	2.005 ± 0.258	4.966 ± 0.269	16.13 ± 2.988	54.13 ± 9.450
$\ell_q\ell_1$ -MKMTL-ALL	5.960 ± 0.834	1.959 ± 0.256	4.821 ± 0.224	16.00 ± 3.062	53.48 ± 9.592
$\ell_{2,1}\ell_q$ -MKMTL-MRI	6.425 ± 0.951	1.951 ± 0.308	4.886 ± 0.264	16.11 ± 2.939	54.96 ± 7.499
$\ell_{2,1}\ell_q$ -MKMTL-PET	6.783 ± 1.059	2.058 ± 0.323	5.107 ± 0.258	16.52 ± 3.515	55.51 ± 9.568
$\ell_{2,1}\ell_q$ -MKMTL-MP	6.086 ± 0.987	1.917 ± 0.299	4.855 ± 0.249	15.95 ± 2.996	52.44 ± 8.074
$\ell_{2,1}\ell_q$ -MKMTL-ALL	6.034 ± 0.978	1.905 ± 0.294	4.809 ± 0.244	15.88 ± 3.028	52.20 ± 8.120
RAVLT					
Method	TOTAL	TOT6	T30	RECOG	nMSE
$\ell_q\ell_1$ -MTL-MRI	10.18 ± 0.640	3.538 ± 0.147	3.735 ± 0.199	3.169 ± 0.306	10.24 ± 0.735
$\ell_q\ell_1$ -MTL-PET	10.41 ± 0.441	3.627 ± 0.140	3.796 ± 0.176	3.258 ± 0.360	10.72 ± 1.163
$\ell_q\ell_1$ -MTL-MP	10.01 ± 0.556	3.501 ± 0.149	3.693 ± 0.196	3.164 ± 0.314	9.710 ± 0.627
$\ell_q\ell_1$ -MTL-ALL	9.755 ± 0.575	3.450 ± 0.151	3.643 ± 0.200	3.172 ± 0.313	9.525 ± 0.608
$\ell_q\ell_1$ -MKMTL-MRI	10.09 ± 0.605	3.532 ± 0.081	3.731 ± 0.253	3.203 ± 0.304	10.21 ± 1.019
$\ell_q\ell_1$ -MKMTL-PET	10.30 ± 0.436	3.592 ± 0.145	3.754 ± 0.231	3.200 ± 0.357	10.82 ± 1.455
$\ell_q\ell_1$ -MKMTL-MP	9.787 ± 0.375	3.471 ± 0.089	3.664 ± 0.199	3.159 ± 0.302	9.713 ± 0.968
$\ell_q\ell_1$ -MKMTL-ALL	9.350 ± 0.460	3.402 ± 0.030	3.604 ± 0.221	3.196 ± 0.291	9.410 ± 0.985
$\ell_{2,1}\ell_q$ -MKMTL-MRI	9.984 ± 0.525	3.477 ± 0.130	3.678 ± 0.204	3.143 ± 0.314	9.937 ± 0.753
$\ell_{2,1}\ell_q$ -MKMTL-PET	10.19 ± 0.410	3.565 ± 0.146	3.745 ± 0.212	3.191 ± 0.351	10.31 ± 1.105
$\ell_{2,1}\ell_q$ -MKMTL-MP	9.727 ± 0.467	3.397 ± 0.136	3.593 ± 0.162	3.112 ± 0.323	9.282 ± 0.869
$\ell_{2,1}\ell_q$ -MKMTL-ALL	9.561 ± 0.442	3.361 ± 0.124	3.556 ± 0.170	3.104 ± 0.327	9.160 ± 0.860

sparsity-inducing norms, having an advantage over the other comparative multitask learning methods. Since not all the brain regions are associated with AD, many of the features are irrelevant and redundant. Sparse based MTL methods are appropriate for the task of predicting cognitive measures and better than the non-sparse-based MTL methods.

We also show the scatter plots of actual values versus predicted values for the score of ADAS, MMSE, TOTAL, and ANIM on testing data in Figure 1.

5.3. Multimodalities Fusion. To estimate the effect of combining multimodality image data with the linearized and kernelized MTL methods and provide a more comprehensive comparison of the results from the comparable MTL models, we further perform some experiments, and they are (1) using only MRI modality, (2) using only PET modality, (3) combining two modalities: PET and MRI (MP), and (4) combining three modalities: PET, MRI, and demographic information including age, gender, years of education, and

ApoE genotyping (MPD). Different from the above experiments, the samples from ADNI-2 are used instead of ADNI-1, since the amount of the patients with PET is sufficient. From the ADNI-2, we obtained all the patients with both MRI and PET, totally 756 samples. The PET imaging data are from the ADNI database processed by the UC Berkeley team, who use a native-space MRI scan for each subject that is segmented and parcellated with FreeSurfer to generate a summary cortical and subcortical ROI, and they coregister each florbetapir scan to the corresponding MRI and calculate the mean florbetapir uptake within the cortical and reference regions. The procedure of image processing is described in <http://adni.loni.usc.edu/updated-florbetapirav-45-pet-analysis-results/>. In the $\ell_q\ell_1$ -MKMTL and $\ell_{2,1}\ell_q$ -MKMTL, ten different kernel functions described in the first experiment are used for each modality. To show the advantage of the kernel-based methods, we compare them with linear $\ell_q\ell_1$ -MTL method, which concatenated the multiple modalities features into a long vector features.

The prediction performance results are shown in Tables 3 and 4. From the results, it is clear that the methods with

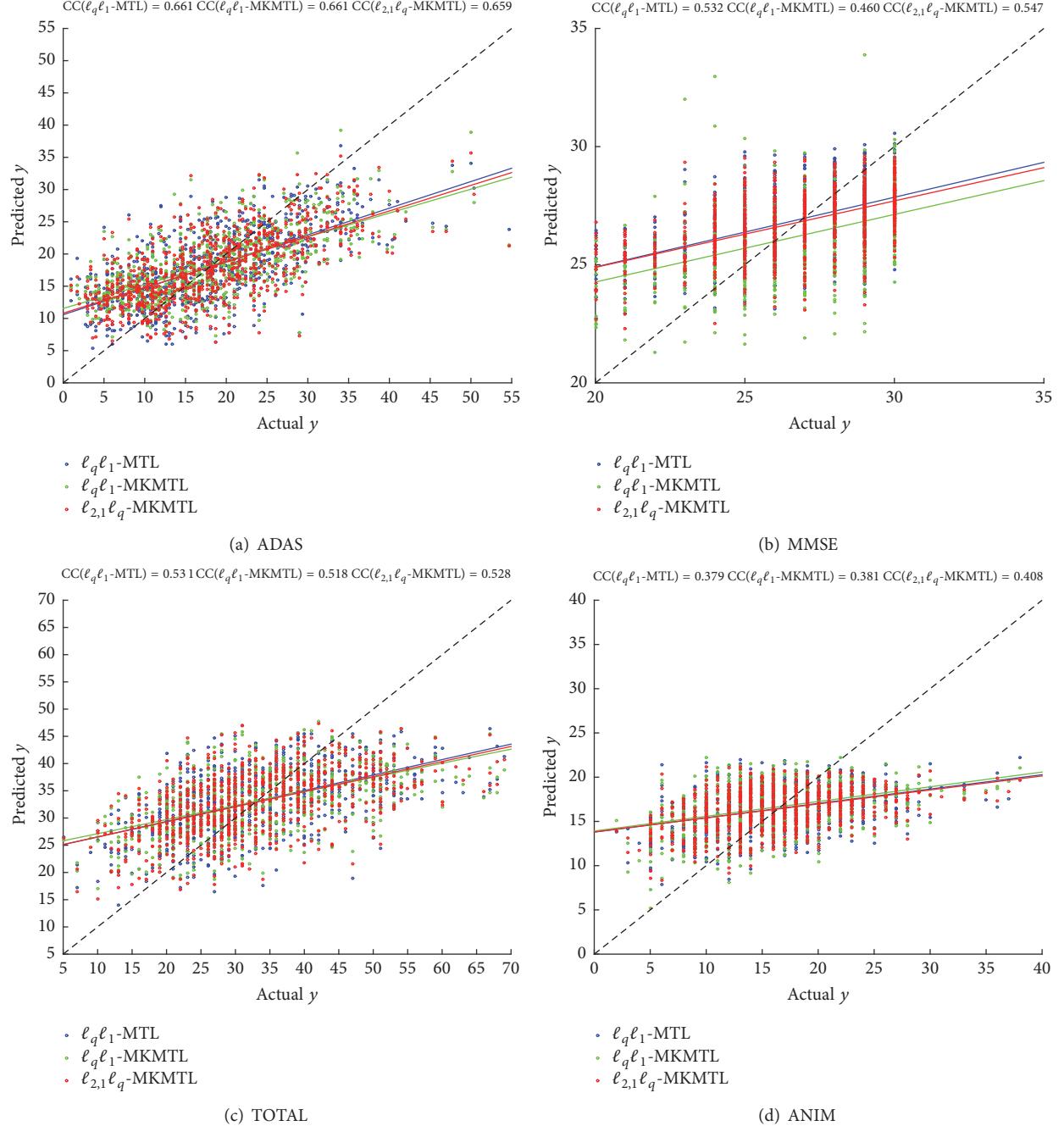


FIGURE 1: Scatter plots of actual versus predicted values of cognitive scores on each fold testing data using three comparable MTL methods based on MRI features.

multimodality outperform the methods using one single modality of data. This validates our assumption that the complementary information among different modalities is helpful for cognitive function prediction. Regardless of two or three modalities, $\ell_{2,1}\ell_q$ -MKMTL achieved better performances than the linear based multitask learning for the most cases, the same as for the single modality learning above.

6. Conclusion

Many multitask learning methods with sparsity-inducing regularization for modeling AD cognitive outcomes have

been proposed in the past decades. However, the current formulations remain restricted to the linear models and cannot capture the relationship between the MRI features and cognitive outcomes. To address these shortcomings, we applied two multikernel multitask learning methods with a joint sparsity-inducing regularization to model the more complicated but more flexible relationship between MRI features and cognitive outcomes and demonstrated their effectiveness compared with linearized multitask learning methods by applying them to the ADNI data for predicting cognitive outcomes from MRI scans. Extensive experiments

TABLE 4: Performance comparison of various methods with fusing multiple modalities data in terms of CC and wR on 10 cross validation cognitive prediction tasks.

Method	ADAS	MMSE	FLU	TRAILS	
			ANIM	A	B
$\ell_q\ell_1$ -MTL-MRI	0.670 ± 0.091	0.539 ± 0.117	0.481 ± 0.112	0.417 ± 0.115	0.525 ± 0.073
$\ell_q\ell_1$ -MTL-PET	0.619 ± 0.058	0.482 ± 0.087	0.395 ± 0.105	0.385 ± 0.120	0.501 ± 0.060
$\ell_q\ell_1$ -MTL-MP	0.700 ± 0.070	0.549 ± 0.108	0.486 ± 0.119	0.437 ± 0.119	0.567 ± 0.070
$\ell_q\ell_1$ -MTL-ALL	0.705 ± 0.067	0.560 ± 0.096	0.527 ± 0.102	0.450 ± 0.115	0.575 ± 0.064
$\ell_q\ell_1$ -MKMTL-MRI	0.677 ± 0.093	0.512 ± 0.113	0.464 ± 0.095	0.411 ± 0.113	0.529 ± 0.094
$\ell_q\ell_1$ -MKMTL-PET	0.634 ± 0.056	0.493 ± 0.100	0.410 ± 0.133	0.375 ± 0.090	0.478 ± 0.061
$\ell_q\ell_1$ -MKMTL-MP	0.710 ± 0.060	0.537 ± 0.106	0.472 ± 0.111	0.426 ± 0.105	0.566 ± 0.081
$\ell_q\ell_1$ -MKMTL-ALL	0.727 ± 0.062	0.551 ± 0.112	0.512 ± 0.097	0.444 ± 0.099	0.582 ± 0.065
$\ell_{2,1}\ell_q$ -MKMTL-MRI	0.673 ± 0.096	0.548 ± 0.124	0.491 ± 0.095	0.422 ± 0.135	0.528 ± 0.102
$\ell_{2,1}\ell_q$ -MKMTL-PET	0.631 ± 0.057	0.488 ± 0.108	0.418 ± 0.119	0.386 ± 0.095	0.524 ± 0.065
$\ell_{2,1}\ell_q$ -MKMTL-MP	0.714 ± 0.067	0.566 ± 0.107	0.499 ± 0.094	0.437 ± 0.122	0.583 ± 0.077
$\ell_{2,1}\ell_q$ -MKMTL-ALL	0.721 ± 0.064	0.574 ± 0.105	0.512 ± 0.094	0.445 ± 0.120	0.589 ± 0.073

Method	TOTAL	RAVLT			wR
		TOT6	T30	RECOG	
$\ell_q\ell_1$ -MTL-MRI	0.576 ± 0.077	0.536 ± 0.085	0.516 ± 0.041	0.444 ± 0.079	0.523 ± 0.082
$\ell_q\ell_1$ -MTL-PET	0.548 ± 0.103	0.497 ± 0.124	0.490 ± 0.092	0.409 ± 0.098	0.481 ± 0.081
$\ell_q\ell_1$ -MTL-MP	0.593 ± 0.079	0.547 ± 0.086	0.529 ± 0.038	0.450 ± 0.075	0.540 ± 0.077
$\ell_q\ell_1$ -MTL-ALL	0.618 ± 0.072	0.563 ± 0.077	0.546 ± 0.027	0.446 ± 0.085	0.554 ± 0.069
$\ell_q\ell_1$ -MKMTL-MRI	0.585 ± 0.069	0.533 ± 0.093	0.511 ± 0.044	0.434 ± 0.077	0.517 ± 0.079
$\ell_q\ell_1$ -MKMTL-PET	0.559 ± 0.110	0.508 ± 0.111	0.503 ± 0.085	0.432 ± 0.081	0.488 ± 0.075
$\ell_q\ell_1$ -MKMTL-MP	0.617 ± 0.080	0.561 ± 0.100	0.541 ± 0.057	0.462 ± 0.079	0.543 ± 0.075
$\ell_q\ell_1$ -MKMTL-ALL	0.654 ± 0.071	0.577 ± 0.082	0.560 ± 0.038	0.444 ± 0.087	0.561 ± 0.068
$\ell_{2,1}\ell_q$ -MKMTL-MRI	0.594 ± 0.070	0.554 ± 0.080	0.536 ± 0.033	0.459 ± 0.071	0.534 ± 0.082
$\ell_{2,1}\ell_q$ -MKMTL-PET	0.563 ± 0.104	0.510 ± 0.111	0.501 ± 0.081	0.436 ± 0.095	0.495 ± 0.072
$\ell_{2,1}\ell_q$ -MKMTL-MP	0.621 ± 0.075	0.582 ± 0.083	0.564 ± 0.046	0.475 ± 0.073	0.560 ± 0.071
$\ell_{2,1}\ell_q$ -MKMTL-ALL	0.637 ± 0.068	0.593 ± 0.077	0.575 ± 0.041	0.479 ± 0.081	0.570 ± 0.067

on ADNI dataset illustrate that the multikernel multitask learning method not only yields superior performance on regression performance but also is a powerful tool for fusing multimodalities data.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (no. 61502091), the Fundamental Research Funds for the Central Universities (no. N161604001 and no. N150408001), and the National Science Foundation for Distinguished Young Scholars of China under Grants no. 71325002 and no. 61225012.

References

- [1] Z. S. Khachaturian, “Diagnosis of Alzheimer’s disease,” *JAMA Neurology*, vol. 42, no. 11, pp. 1097–1105, 1985.

- [2] A. Wimo, B. Winblad, H. Aguero-Torres, and E. von Strauss, “The magnitude of dementia occurrence in the world,” *Alzheimer Disease & Associated Disorders*, vol. 17, no. 2, pp. 63–67, 2003.
- [3] J. Wan, Z. Zhang, B. D. Rao et al., “Identifying the neuroanatomical basis of cognitive impairment in Alzheimer’s disease by correlation-and nonlinearity-aware sparse bayesian learning,” *IEEE Transactions on Medical Imaging*, vol. 33, no. 7, pp. 1475–1487, 2014.
- [4] C. M. Stonnington, C. Chu, S. Klöppel, C. R. Jack, J. Ashburner, and R. S. J. Frackowiak, “Predicting clinical scores from magnetic resonance scans in Alzheimer’s disease,” *NeuroImage*, vol. 51, no. 4, pp. 1405–1413, 2010.
- [5] J. Ye, M. Farnum, E. Yang et al., “Sparse learning and stability selection for predicting MCI to AD conversion using baseline ADNI data,” *BMC Neurology*, vol. 12, article no. 46, 2012.
- [6] Y. Zhang and D.-Y. Yeung, “A convex formulation for learning task relationships in multi-task learning,” in *Proceedings of the 26th Conference on Uncertainty in Artificial Intelligence, UAI 2010*, pp. 733–742, Catalina Island, California, USA, July 2010.
- [7] A. Argyriou, T. Evgeniou, and M. Pontil, “Convex multi-task feature learning,” *Machine Learning*, vol. 73, no. 3, pp. 243–272, 2008.

- [8] H. Wang, F. Nie, H. Huang et al., "Sparse multi-task regression and feature selection to identify brain imaging predictors for memory performance," in *Proceedings of the 2011 IEEE International Conference on Computer Vision, ICCV 2011*, pp. 557–562, Barcelona, Spain, November 2011.
- [9] H. Wang, F. Nie, H. Huang et al., "High-Order Multi-Task Feature Learning to Identify Longitudinal Phenotypic Markers for Alzheimers Disease Progression Prediction," in *Advances in Neural Information Processing Systems (NIPS)*, 2012.
- [10] B. Gu and V. S. Sheng, "A robust regularization path algorithm for ν -support vector classification," *IEEE Transactions on Neural Networks and Learning Systems*, 2016.
- [11] B. Gu, V. S. Sheng, K. Y. Tay, W. Romano, and S. Li, "Incremental support vector learning for ordinal regression," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 7, pp. 1403–1416, 2015.
- [12] M. Gönen and E. Alpaydin, "Multiple kernel learning algorithms," *Journal of Machine Learning Research*, vol. 12, pp. 2211–2268, 2011.
- [13] P. Jawanpuria and J. S. Nath, "Multi-task multiple kernel learning," in *Proceedings of the 11th SIAM International Conference on Data Mining, SDM 2011*, pp. 828–838, Mesa, Arizona, USA, April 2011.
- [14] A. Argyriou, T. Evgeniou, and M. Pontil, "Multi-task feature learning," in *in: Advances in neural information processing systems*, pp. 41–48, 2007.
- [15] J. Liu and J. Ye, "Efficient l_1/l_q norm regularization," Arxiv preprint arXiv 1009.
- [16] J. Zhou, J. Chen, and J. Ye, "Clustered multi-task learning via alternating structure optimization in," *Advances in Neural Information Processing Systems*, pp. 702–710, 2011.
- [17] S. Ji and J. Ye, "An accelerated gradient method for trace norm minimization," in *Proceedings of the 26th International Conference On Machine Learning (ICML '09)*, pp. 457–464, ACM, June 2009.
- [18] J. Zhou, J. Liu, V. A. Narayan, and J. Ye, "Modeling disease progression via multi-task learning," *NeuroImage*, vol. 78, pp. 233–248, 2013.
- [19] R. S. Desikan, F. Ségonne, B. Fischl et al., "An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest," *NeuroImage*, vol. 31, no. 3, pp. 968–980, 2006.
- [20] X. Liu, P. Cao, D. Zhao, and A. Banerjee, "Multi-task Sparse Group Lasso for Characterizing Alzheimer's Disease in," *5th Workshop on Data Mining for Medicine and Healthcare*, p. 49, 2016.
- [21] J. Wan, Z. Zhang, J. Yan et al., "Sparse Bayesian multi-task learning for predicting cognitive outcomes from neuroimaging measures in Alzheimer's disease," in *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2012*, pp. 940–947, USA, June 2012.
- [22] N. R. Draper and H. Smith, *Applied Regression Analysis*, John Wiley & Sons, New York, NY, USA, 3rd edition, 1981.
- [23] J. Liu and J. Ye, "Efficient Euclidean projections in linear time," in *Proceedings of the 26th International Conference On Machine Learning, ICML 2009*, pp. 657–664, can, June 2009.
- [24] A. Rakotomamonjy, F. R. Bach, S. Canu, and Y. Grandvalet, "Simple MKL," *Journal of Machine Learning Research*, vol. 9, pp. 2491–2521, 2008.
- [25] J. Chen, J. Zhou, and J. Ye, "Integrating low-rank and group-sparse structures for robust multi-task learning," in *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD'11*, pp. 42–50, San Diego, Calif, USA, August 2011.
- [26] J. Zhou, "Multi-task learning in crisis event classification," Tech. Rep., <http://www.public.asu.edu/jzhou29>.

Research Article

Sequential Probability Ratio Testing with Power Projective Base Method Improves Decision-Making for BCI

Rong Liu,¹ Yongxuan Wang,² Xinyu Wu,³ and Jun Cheng³

¹Biomedical Engineering Department, Dalian University of Technology, Dalian, Liaoning 116024, China

²Affiliated Zhongshan Hospital of Dalian University, Dalian, Liaoning 116001, China

³Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China

Correspondence should be addressed to Yongxuan Wang; wyx8904@mail.dlut.edu.cn

Received 30 June 2017; Accepted 11 September 2017; Published 14 November 2017

Academic Editor: Fei Chen

Copyright © 2017 Rong Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Obtaining a fast and reliable decision is an important issue in brain-computer interfaces (BCI), particularly in practical real-time applications such as wheelchair or neuroprosthetic control. In this study, the EEG signals were firstly analyzed with a power projective base method. Then we were applied a decision-making model, the sequential probability ratio testing (SPRT), for single-trial classification of motor imagery movement events. The unique strength of this proposed classification method lies in its accumulative process, which increases the discriminative power as more and more evidence is observed over time. The properties of the method were illustrated on thirteen subjects' recordings from three datasets. Results showed that our proposed power projective method outperformed two benchmark methods for every subject. Moreover, with sequential classifier, the accuracies across subjects were significantly higher than that with nonsequential ones. The average maximum accuracy of the SPRT method was 84.1%, as compared with 82.3% accuracy for the sequential Bayesian (SB) method. The proposed SPRT method provides an explicit relationship between stopping time, thresholds, and error, which is important for balancing the time-accuracy trade-off. These results suggest SPRT would be useful in speeding up decision-making while trading off errors in BCI.

1. Introduction

Noninvasive brain-computer interface (BCI) based on the electroencephalogram (EEG) offers a new means of communication to locked-in or paralyzed patients [1, 2] and controlling a prosthesis [3, 4] without reliance on the usual neuromuscular pathways. The critical challenge of BCI technology is to classify the brain signals and mental tasks accurately and fast. However, the EEG recorded from the scalp has the characteristics of low strength, low SNR (signal noise ratio), and the EEG difference under different mental tasks is not significant. Therefore, various pattern recognition algorithms were used in BCI system to extract and classify EEG features.

Event-related desynchronization/synchronization (ERD/ERS) patterns of motor imagery are effective features for EEG-based BCI systems. The experiments show that the phenomenon of ERD/ERS varies among individuals. Therefore, a pattern recognition algorithm should be used to

facilitate decoding "motor intent," both to find subject-specific EEG features that maximize the separation between the patterns generated by executing the mental tasks and to train classifiers that minimize the classification error rates of these specific patterns. Currently, feature extraction for discrimination of left- and right-hand motor imagery EEG is usually based on EEG band power (BP). For example, autoregression (AR) model [5], discrete Fourier transformation (DFT) [6], and wavelet transforms (WT) [7] have been used to extract EEG features for classification. The wavelet method is one of the most effective algorithms. However, the success of wavelet application greatly depends on the proper selection of subject-specific parameters. Actually, the wavelet transform can be considered as projecting the EEG onto a wavelet basis and the band power as the modulus values of projective coefficients. Inspired by the wavelet method, we introduce a new feature extraction method based on power projective bases to classify EEGs without constrain of wavelet forms.

Moreover, the ability to make rapid decisions based on transient stimuli is a unique aspect of our brains' capacity to process information. Broadly speaking, signal detection theory (SDT) and sequential analysis (SA) are two branches of mathematical models that provide a theoretical framework for understanding how decisions are made [8]. SDT converts a single observation into a categorical choice. According to different decision rules, there are different testing approaches to this problem [9]. For example, Bayesian decision theory is derived by minimizing the posterior expected loss, while Neyman-Pearson (NP) criterion seeks to find the best error probability (α) level test. Like most statistical classification methods, for example, linear discriminant analysis (LDA) and support vector machines (SVM), the classification error is the only characteristic of the SDT decision strategies. The necessary number of observation samples determined by the criteria could be very large, which is especially impractical for BCI applications. To control brain-actuated devices, such as robotics and neuroprostheses, both fast decision-making and a stable control signal with a minimal error rate are important [10, 11]. Therefore, recent attentions have been paid to the variable-length sequential sampling model.

A systematic theory of optimal stopping emerged with the work by Wald on the optimality of the sequential probability ratio test (SPRT) [12]. The SPRT achieves a desired error rate with the smallest number of samples, on average. Therefore, in this paper, we introduce a new feature extraction method based on power projective base to classify the EEGs by combining the sequential probability ratio test (SPRT) approach to obtain a continuous dynamic estimate of brain state with accuracy and decision speed balance.

2. Methodology

2.1. Data Description. The EEG data used in this work were obtained from thirteen subjects from BCI Competitions II, III, and IV. The task was performed based on left- and right-hand motor imagination.

2.1.1. Dataset III from BCI Competition II. This dataset contains EEG data from one subject (S1) [13]. The data were recorded from three channels (C3, Cz, and C4) and sampled at 128 Hz. The data consist of 140 labelled and 140 unlabelled trials with an equal number of left- and right-hand trials. Each trial has a duration of 9 s, where a visual cue (arrow) is presented pointing to the left or the right after 3 s preparation period followed by a 6 s motor imagery (MI) task.

2.1.2. Dataset IIIb from BCI Competition III. The second dataset contains EEG data recorded over the channels C3 and C4 from three subjects (S2, S3, S4) with some corrections [14]. The data were sampled at 125 Hz. Training and testing sets were available for each subject. Except for the subject O3 that has only just 320 trials for each set, the subjects S4 and XI1 contain 540 labelled and 540 unlabelled trials. Each trial has duration of 7 s which consists of 3 s for preparation period and 1 s for a visual cue presentation, followed by another 3 s for the imagination task.

2.1.3. Dataset IIb from BCI Competition IV. This dataset contains EEG data from nine subjects (S5–S13) [15]. The data were recorded from three bipolar channels (C3, Cz, and C4) and 3 EOG channels. The sample frequency was 250 Hz. Training and testing set was available for each subject. Each subject participated in two screening sessions without feedback and three online feedback sessions with smiley feedback. The trials without feedback had duration of 7 s, and a visual cue was presented for 1.25 s followed by another 4 s for the imagination task. The trials with feedback had duration of 7.5 s, and a visual cue was presented for 4.5 s until the end of motor imagination.

2.2. Feature Extraction Method Based on Power Projective Base. Motor imagery can be regarded as mental rehearsal of a motor act without any obvious motor output. Recent studies show that when performing motor imagination, μ (8–13 Hz) and β (18–30 Hz) rhythms are found to reveal event-related desynchronization and synchronization (ERD/ERS) over sensorimotor cortex just as when one performs motor tasks. Due to nonstationary effects having often been observed in brain signals, we proposed a power projective base method to extract classification features from C3 and C4 channels. This method improves the classification accuracy by maximizing the difference of the average projective power between two-class signals. Specifically, the solution of the projective bases can be achieved by generalized eigenvalue decomposition for each subject.

Let $\mathbf{X}_k \in R^{M \times N_k}$ be the training dataset from one channel, where $k \in \{L, R\}$ denotes the left- or right-hand motor imagery tasks, M denotes the sampling points, and N_k is the number of trials. Moreover, let $\mathbf{u} \in R^M$ be the projective basis and $\|\mathbf{u}\| = 1$. The projective power of signal \mathbf{x}_{kj} , $j = 1, 2, \dots, N_k$ on the projective basis \mathbf{u} is

$$z_{kj}(\mathbf{u}) = (\mathbf{u}^T \mathbf{x}_{kj}) \cdot (\mathbf{x}_{kj}^T \mathbf{u}). \quad (1)$$

So the mean projective power $\bar{z}_k(\mathbf{u})$ can be calculated as

$$\begin{aligned} \bar{z}_k(\mathbf{u}) &= \frac{1}{N_k} \sum_{j=1}^{N_k} (\mathbf{u}^T \mathbf{x}_{kj}) \cdot (\mathbf{x}_{kj}^T \mathbf{u}) = \mathbf{u}^T \cdot \left(\frac{\mathbf{X}_k \mathbf{X}_k^T}{N_k} \right) \cdot \mathbf{u} \\ &= \mathbf{u}^T \mathbf{R}_k \mathbf{u}, \end{aligned} \quad (2)$$

where $\mathbf{R}_k = \mathbf{X}_k \mathbf{X}_k^T / N_k$ is the autocorrelation matrix and it is usually positively definite.

To formulate the objective function to be the ratio of the two-class average projective powers,

$$F(\mathbf{u}) = \frac{\bar{z}_L}{\bar{z}_R} = \frac{\mathbf{u}^T \mathbf{R}_L \mathbf{u}}{\mathbf{u}^T \mathbf{R}_R \mathbf{u}}. \quad (3)$$

By maximizing or minimizing $F(\mathbf{u})$ to be F_{\max} or F_{\min} , the corresponding eigenvector \mathbf{u}_{\max} or \mathbf{u}_{\min} is the optimal projective base to be solved. The optimization of (3) could be solved by taking a generalized eigenvalue decomposition

method. First of all, we can get the following decomposition as

$$\mathbf{T}^T \mathbf{R}_L \mathbf{T} = \boldsymbol{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_M),$$

$$\lambda_i \geq \lambda_j > 0 \quad (\forall i < j), \quad (4)$$

$$\mathbf{T}^T \mathbf{R}_R \mathbf{T} = \mathbf{I},$$

where $\mathbf{T} = [\mathbf{t}_1, \dots, \mathbf{t}_M]$ is the generalized eigenvector matrix and λ_i is the generalized eigenvalue. Therefore, the ratio of mean projection power $F(\mathbf{u})$ turns to

$$F(\mathbf{u}) = \frac{\mathbf{u}^T \mathbf{T}^{-T} \boldsymbol{\Lambda} \mathbf{T}^{-1} \mathbf{u}}{\mathbf{u}^T \mathbf{T}^{-T} \mathbf{T}^{-1} \mathbf{u}} = \frac{(\mathbf{u}^T \mathbf{T}^{-T}) \cdot \boldsymbol{\Lambda} \cdot (\mathbf{T}^{-1} \mathbf{u})}{(\mathbf{u}^T \mathbf{T}^{-T}) \cdot (\mathbf{T}^{-1} \mathbf{u})} \quad (5)$$

$$\triangleq \frac{\mathbf{v}^T \boldsymbol{\Lambda} \mathbf{v}}{\mathbf{v}^T \mathbf{v}} = \frac{\lambda_1 v_1^2 + \dots + \lambda_M v_M^2}{v_1^2 + \dots + v_M^2},$$

where $\mathbf{v} = \mathbf{T}^{-1} \mathbf{u}$. Since

$$\frac{\lambda_1 v_1^2 + \dots + \lambda_M v_M^2}{v_1^2 + \dots + v_M^2} \leq \frac{\lambda_1 v_1^2 + \dots + \lambda_1 v_M^2}{v_1^2 + \dots + v_M^2} = \lambda_1, \quad (6)$$

$$\frac{\lambda_1 v_1^2 + \dots + \lambda_M v_M^2}{v_1^2 + \dots + v_M^2} \geq \frac{\lambda_M v_1^2 + \dots + \lambda_M v_M^2}{v_1^2 + \dots + v_M^2} = \lambda_M$$

$F(\mathbf{u})$ has the maximum value $F_{\max} = \lambda_1$ and minimum value $F_{\min} = \lambda_M$. Obviously, the corresponding vectors \mathbf{v} can be obtained by

$$\mathbf{v}_{\max} = [1, 0, \dots, 0]^T, \quad (7)$$

$$\mathbf{v}_{\min} = [0, \dots, 0, 1]^T.$$

Then, we have

$$\mathbf{u}_{\max} = \mathbf{T} \cdot [1, 0, \dots, 0]^T, \quad (8)$$

$$\mathbf{u}_{\min} = \mathbf{T} \cdot [0, \dots, 0, 1]^T$$

which means that \mathbf{u}_{\max} and \mathbf{u}_{\min} are the first column \mathbf{u}_1 and last column \mathbf{u}_M of \mathbf{T} , respectively. Choosing \mathbf{u}_{\max} or \mathbf{u}_{\min} as the projective base depends on which is larger between λ_1 and $1/\lambda_M$. The projective power for the signals of the channels C3 and C4 onto their own projective bases is then stacked together into a 2-dimensional feature vector $\mathbf{z} = [e_{C3}, e_{C4}]^T$.

2.3. SPRT Classification Method. Sequential analysis is a statistical decision model that assumes decisions are formed by continuously sampling information until the response criterion is satisfied. Once a boundary has been reached, the decision process is concluded and a response is elicited. The number of observations needed for a decision is not determined in advance of the experiment, but by the observations obtained during the test. The data should be fed to the SPRT algorithm sequentially, so we divide each trial into segments with overlap and each one has the same length as that of the projective base used in feature extraction.

Taking into account the nonstationarity of the EEG sampling information, we assume that the probability distribution of the i th segment feature \mathbf{z}_i for class k is $g_k(\mathbf{z}_i)$, $k \in \{L, R\}$, and then the probability ratio for the i th segment is

$$l_i(\mathbf{z}_i) = \frac{g_{Li}(\mathbf{z}_i)}{g_{Ri}(\mathbf{z}_i)} \quad (9)$$

and the evidence accumulation turns to

$$l_q(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_q) = \frac{\prod_{i=1}^q g_{Li}(\mathbf{z}_i)}{\prod_{i=1}^q g_{Ri}(\mathbf{z}_i)}, \quad (10)$$

where q is the number of accumulated segments. Assuming the segments are independent for computation convenience [12], we have

$$l_q(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_q) = \frac{g_L(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_q)}{g_R(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_q)}, \quad (11)$$

where $g_k(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_q)$, $k \in \{L, R\}$ is the joint probability distribution of q dimensional vector $(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_q)$. This assumption is violated by our data in practice.

The decision rule with two thresholds η_L and η_R is

$$\begin{aligned} &\text{if } l_q(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_q) \geq \eta_L, \quad \hat{k} = L \\ &\text{if } l_q(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_q) \leq \eta_R, \quad \hat{k} = R \\ &\text{if } \eta_R < l_q(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_q) < \eta_L, \quad \text{compute } l_{q+1}. \end{aligned} \quad (12)$$

With two thresholds, we have the option to increase η_L or decrease η_R which will increase the probability to make a correct decision (by waiting to accumulate more data or evidence) but decrease the probability of making a wrong decision (by delaying the decision). The error probabilities are defined as

$$\begin{aligned} p_{L|R} &= P(\hat{k} = L \mid k = R), \\ p_{R|L} &= P(\hat{k} = R \mid k = L). \end{aligned} \quad (13)$$

If $l_q(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_q) \geq \eta_L$ is satisfied, we define the corresponding space of vector $(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_q)$ to be Z_L . With (11), we have

$$g_L(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_q) \geq \eta_L \cdot g_R(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_q). \quad (14)$$

Equation (14) is then integrated in Z_L yielding

$$\int_{Z_L} g_L(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_q) \geq \eta_L \cdot \int_{Z_L} g_R(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_q). \quad (15)$$

That is,

$$1 - p_{R|L} \geq \eta_L \cdot p_{L|R}. \quad (16)$$

Analogous reasoning for $l_q(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_q) \leq \eta_R$ yields

$$p_{R|L} \leq \eta_R \cdot (1 - p_{L|R}). \quad (17)$$

Thus, the two detection thresholds η_L and η_R are related to the error probabilities by

$$\begin{aligned} p_{L|R} &\leq \frac{1 - p_{R|L}}{\eta_L} \leq \frac{1}{\eta_L}, \\ p_{R|L} &\leq (1 - p_{L|R}) \cdot \eta_R \leq \eta_R. \end{aligned} \quad (18)$$

The two kinds of error probabilities can be lowered by either increasing η_L or decreasing η_R . However, due to the limited number of segments, the indecision ratio will increase as the error probability $p_{L|R}$ or $p_{R|L}$ is decreased. Therefore, we may not obtain the optimal result by simply increasing η_L or decreasing η_R . The suitable η_L and η_R could be achieved by the following optimization criteria.

Under the assumption that the features follow a Gaussian distribution, we can take logarithm on both sides of (10) to obtain the log probability ratio (log PR), which leads a sequential probability ratio test (SPRT) as

$$\begin{aligned} L_q &= \sum_{i=1}^q \frac{1}{2} \left(\log \frac{|\mathbf{S}_{Ri}|}{|\mathbf{S}_{Li}|} + D_M^2(\mathbf{z}_i, \mathbf{m}_{Ri}) - D_M^2(\mathbf{z}_i, \mathbf{m}_{Li}) \right) \\ &\triangleq \sum_{i=1}^q J_i, \end{aligned} \quad (19)$$

where $D_M(\mathbf{z}_i, \mathbf{m}_{ki}) = \sqrt{(\mathbf{z}_i - \mathbf{m}_{ki})^T \mathbf{S}_{ki}^{-1} (\mathbf{z}_i - \mathbf{m}_{ki})}$ is the Mahalanobis distance and J_i is the log PR of the i th segment.

We can derive the average J_i for each class:

$$\begin{aligned} E(J_i | k = L) &= \frac{1}{2} \left(\log \frac{|\mathbf{S}_{Ri}|}{|\mathbf{S}_{Li}|} \right. \\ &\quad \left. + \text{tr} \left((\mathbf{S}_{Li} + (\mathbf{m}_{Li} - \mathbf{m}_{Ri}) \cdot (\mathbf{m}_{Li} - \mathbf{m}_{Ri})^T) \cdot \mathbf{S}_{Ri}^{-1} \right. \right. \\ &\quad \left. \left. - \mathbf{I} \right) \right), \\ E(J_i | k = R) &= \frac{1}{2} \left(\log \frac{|\mathbf{S}_{Ri}|}{|\mathbf{S}_{Li}|} \right. \\ &\quad \left. - \text{tr} \left((\mathbf{S}_{Ri} + (\mathbf{m}_{Ri} - \mathbf{m}_{Li}) \cdot (\mathbf{m}_{Ri} - \mathbf{m}_{Li})^T) \cdot \mathbf{S}_{Li}^{-1} \right. \right. \\ &\quad \left. \left. - \mathbf{I} \right) \right). \end{aligned} \quad (20)$$

For any given threshold pair η_L and η_R , the number of accumulated segments to make a correct decision, that is, stopping time, for class k , is $E(q | k)$ which satisfies

$$\begin{aligned} E(q | k = L) &= \inf_n \left(\sum_{i=1}^n E(J_i | k = L) \geq (1 - \eta_R) \right. \\ &\quad \left. \cdot \log \eta_L + \eta_R \cdot \log \eta_R \right) \end{aligned}$$

$$\begin{aligned} E(q | k = R) &= \inf_n \left(\sum_{i=1}^n E(J_i | k = R) \leq \left(1 - \frac{1}{\eta_L} \right) \right. \\ &\quad \left. \cdot \log \eta_R + \frac{1}{\eta_L} \cdot \log \eta_L \right), \end{aligned} \quad (21)$$

where $\inf_n(\cdot)$ is the minimum element of set of n . Generally, $E(q | k = L)$ and $E(q | k = R)$ may be different. Since the stopping time is a key point in the sequential analysis, we constrain the two thresholds by unifying $E(q | k)$ of two classes to be equal, that is, $E(q | k = L) = E(q | k = R) \triangleq q_E$. Then the thresholds are given by

$$\begin{aligned} \xi_L &= \sum_{i=1}^{q_E} E(J_i | k = L), \\ \xi_R &= \sum_{i=1}^{q_E} E(J_i | k = R). \end{aligned} \quad (22)$$

For any given stopping time q_E , there is a corresponding threshold pair ξ_L and ξ_R . The decision rule with two thresholds ξ_L and η_R is

$$\begin{aligned} &\text{if } L_q \geq \xi_L, \quad \hat{k} = L, \\ &\text{if } L_q \leq \xi_R, \quad \hat{k} = R, \\ &\text{if } \xi_R < L_q < \xi_L, \quad \text{compute } L_{q+1}. \end{aligned} \quad (23)$$

From this decision policy, we can see that other than assigning one of the two classes L and R , the decision functions may still be undecided and continue testing to the next observation. The “undecided” response keeps the number of errors (false positives or false negatives) low, which is useful for avoiding making excessive mistakes to speed up decisions, for example, a BCI control wheelchair running into an obstacle [16]. In addition, when it is still undecided when reaching to the stopping time q_E , we specify that when $q = q_E$, the decision rule is

$$\begin{aligned} &\text{if } L_q > 0, \quad \hat{k} = L, \\ &\text{if } L_q < 0, \quad \hat{k} = R. \end{aligned} \quad (24)$$

Till now, with the above decision rules, the consequent results, such as accuracy, mutual information (MI) [17], the steepness of MI [18], and average decision time, will only rely on the stopping time q_E and the data to be analyzed. Depending on the actual specific needs, we can set the accuracy, MI, the steepness of MI, and average decision time as the optimization target, respectively, to determine the optimal stopping time q_{opt} . At the same time, the two thresholds are determined.

3. Results

3.1. Feature Extraction. To evaluate the performance of our method, we tested it using BCI Competition Datasets II and

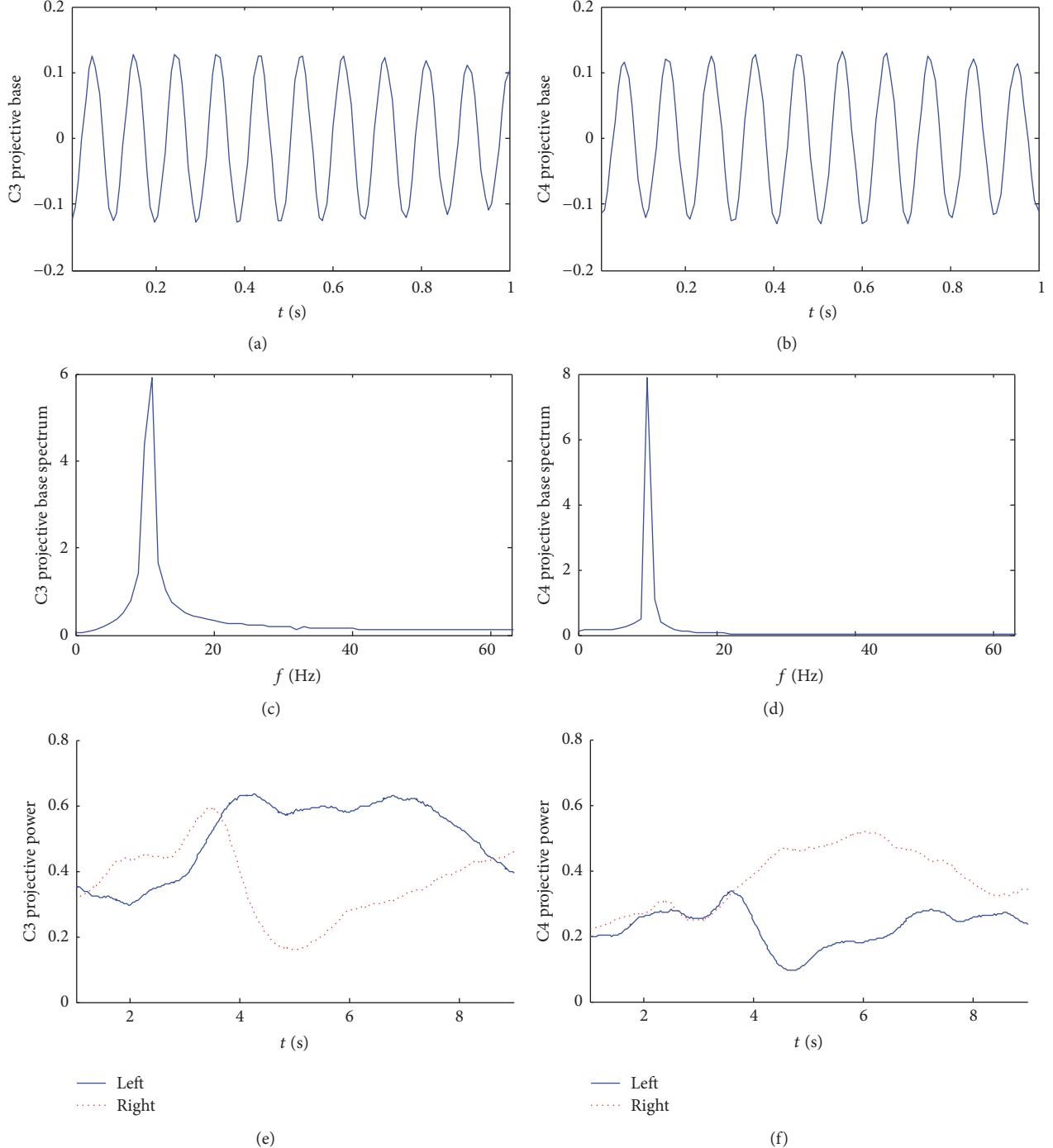


FIGURE 1: The power projective base of the subject S1. (a) C3 projective base; (b) C4 projective base; (c) C3 projective base spectrum; (d) C4 projective base spectrum; (e) the time-varied projective power of C3; (f) the time-varied projective power of C4.

Dataset IIIb from BCI Competition III. They were obtained from four subjects, denoted as S1–S4. The task performed was based on left and right-hand motor imagination.

The dimension of the projective base, that is, the length of the sliding window, is set to be 1 s. The time-domain waveforms of the optimal projective base of the two channels for subject S1 are shown in Figures 1(a) and 1(b). The corresponding frequency spectra are shown in Figures 1(c)

and 1(d). The average projective power time courses during the right-hand (dash line) and left-hand (solid line) imagined movement for the C3 and C4 are displayed in Figures 1(e) and 1(f). From this figure, we can see that the projective bases are similar to modulated sine signals and the spectra have band-pass characteristics which are similar to that of wavelet base. For this subject, the projective power dominates in the μ rhythm. During the first 3.5 s (0.5 s after cue presentation),

TABLE 1: A comparison of ACC and MI for three methods.

Method	Criterion	S1	S2	S3	S4	Avg.
DFT	Acc (%)	88.6	86.3	76.0	77.9	82.2
	MI (B)	0.488	0.424	0.205	0.238	0.339
WT	Acc (%)	89.3	88.4	79.2	81.1	84.5
	MI (B)	0.509	0.482	0.262	0.301	0.389
PP	Acc (%)	92.1	90.6	83.1	84.4	87.6
	MI (B)	0.601	0.550	0.345	0.375	0.468

the projective power curves under two conditions are close; after 3.5 s, distinct difference in the projective power can be observed which provides a good classification feature. The projective bases for subjects S2 and S3 are similar to that of subject S1.

The power projective base of subject S4 is shown in Figure 2. In contrast with subject S1, the reactivity patterns of the projective bases of this subject are quite different. The waveforms of the projective bases are oscillating faster. Obviously, the frequency of projective power is higher than that of subject S1. As seen in Figures 2(c) and 2(d), the projective bases display a peak in β rhythms. Moreover, the results of the average projective power time courses demonstrate that the patterns of ERD/ERS subject S4 are quite different. That means this projective power method is subject-adaptive as well as avoiding the parameter setup in advance.

3.2. Classification Results. Two kinds of experiments were performed to evaluate the performance of the proposed machine learning method. One is to evaluate the projective power feature extraction method and the other is to evaluate the SPRT classification performance. In the first one, for the purpose of benchmarking, we compared the classification accuracy (ACC), the mutual information (MI) with two benchmark feature extraction algorithms, and DFT and WT based on the sequential Bayesian classifier [17, 18]. These methods were also applied to the data of subjects S1–S4. The classification accuracy (ACC) and mutual information (MI) of the three methods in consideration are listed in Table 1, where Avg. denotes the averaged indexes over all four subjects. The results of WT are derived from Lemm’s method which won the BCI competitions in 2003 and 2005 for motor imagery datasets. From Table 1, we can see that the proposed power projective method outperforms two benchmark methods for every subject. Compared with wavelet method, the average ACC of our method increased from 84.5% to 87.6% and MI increased to 0.468 bits. Furthermore, a paired t -test analysis was used to compare the classification accuracies of the three methods. The paired t -test result confirms that with projective base method the ACC and MI across subjects are significantly higher than that with WT ($t = 15.1$, $p < 0.01$) and with DFT ($t = 8.432$, $p < 0.01$).

Then we compared the ACC and MI between several state-of-the-art nonsequential classifiers used in the BCI community and their sequential ones with the projective base feature. Those methods were LDA, SVM, and Bayesian,

and the corresponding sequential methods were denoted as SLDA, SSVM, and SB. The tenfold cross-validation was carried out for classification tests of EEG data in this study; that is, the datasets for each subject were divided into ten subsets, and the following procedure was repeated ten times. Each time, one of the ten subsets was used as the test set and the other nine were used as the training set. The average recognition rate was evaluated across all ten folds. These methods were applied to the three BCI competition datasets described in Section 2.1 from thirteen subjects S1–S13.

Table 2 shows the classification accuracy of all the methods on the competition test data, where “Avg.” denotes the averaged results over all the subjects. By applying the obtained effective features with the power projective method on the nonsequential methods, a classification accuracy of 76.0%, 71.6%, and 75% was achieved by LDA, SVM, and Bayesian, respectively. The accuracies achieved by all of the corresponding sequential classifiers were greater than the nonsequential one. The paired t -test result confirms that with sequential classifier the accuracies across subjects are significantly higher than that with nonsequential ones ($t = 6.15$, $p < 0.01$). Overall, the classification accuracy is higher when the sequential scheme is adopted.

Additionally, it is also important to consider the stopping time for speedy BCI applications without sacrificing accuracy. Therefore, we further analyze the classification accuracy, the MI, and stopping time of the SPRT and compared the time-accuracy trade-off between SPRT and SB. The classification results of the SPRT are provided in Table 3. By applying the obtained effective features with the power projective method on the SPRT, an average classification accuracy of 84.1% was achieved. The proposed SPRT classifier with the projective base features outperforms the SB classifier with the same feature extraction method in terms of classification accuracy across all the subjects except for subject S2. In terms of average accuracy, SPRT outperforms SB classifier. Analyzing the overall results, it can be concluded that SPRT with the projective base feature extraction method outperforms the state-of-the-art methods on the standard datasets.

The resulting time courses of the accumulative classification information, the so-called accumulative evidence, of our SPRT method and SB method are shown in Figure 3. Figure 3(a) shows that the SB classification method gains information from around 4 s (3 s preparation period and 1 s window length). The cumulative Bayesian posterior probabilities reach the extremum at around 5.5 s, indicating a peak decision confidence at this time. However, the accumulative information falls down at the end of trial. The result shows that the effective control takes place during the middle of a trial. More evidence could not increase the classification performance any further.

Compared with the SB method, the cumulative process of SPRT has monotonicity (shown in Figure 3) which makes it possible to improve the accuracy with more evidence available. During the SPRT classification process, once the accumulative evidence exceeds one of the thresholds, an immediate decision will be given. Moreover, the thresholds can be adjusted to change the decision time. That is to say, with two broader threshold choices, a larger number

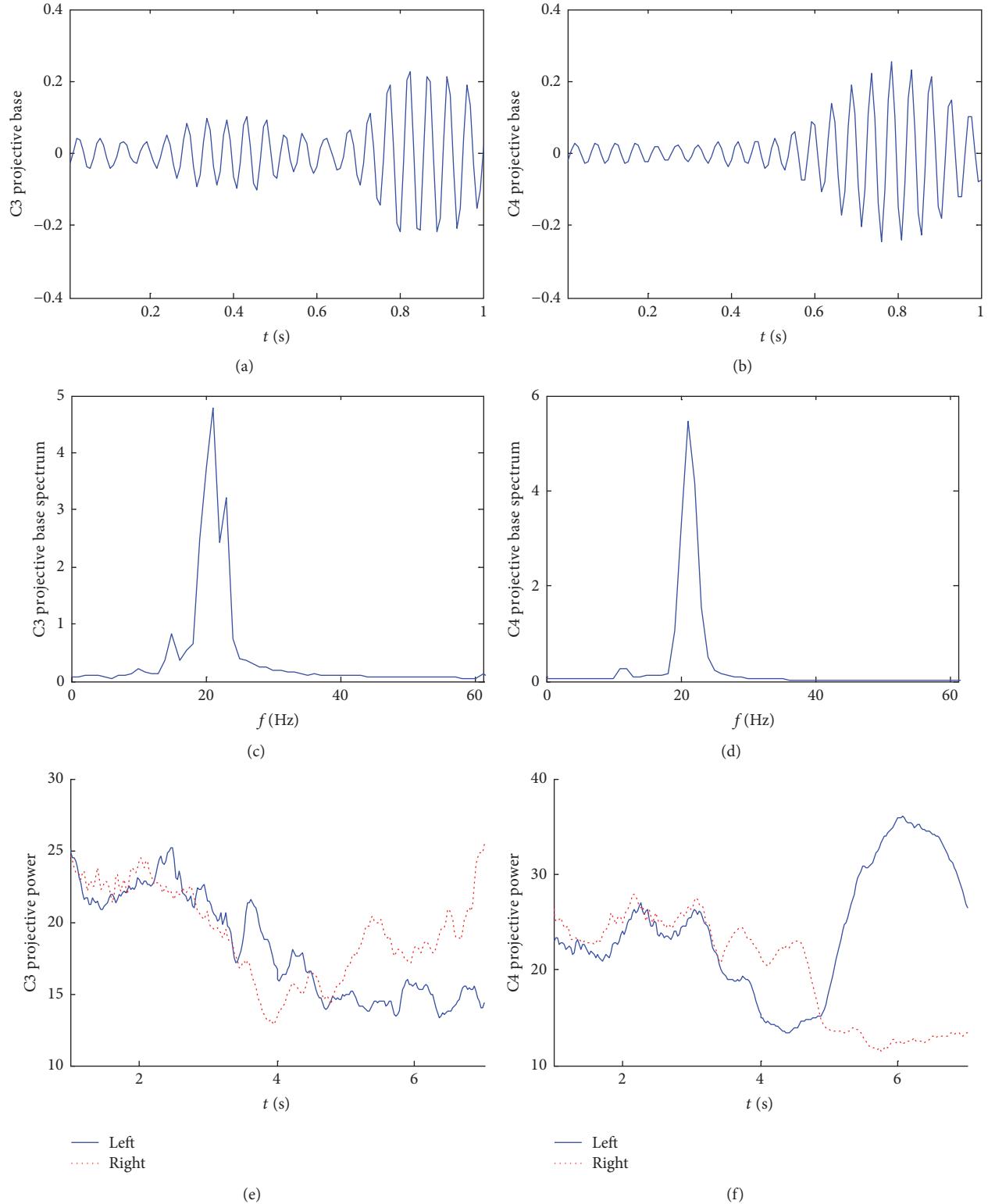


FIGURE 2: The power projective base of the subject S4. (a) C3 projective base; (b) C4 projective base; (c) C3 projective base spectrum; (d) C4 projective base spectrum; (e) the time-varied projective power of C3; (f) the time-varied projective power of C4.

TABLE 2: Accuracy (%) of different classifiers on the BCI competition datasets.

Subject	LDA	SLDA	SVM	SSVM	Bayesian	SB
S1	87.8	90.0	87.1	92.1	86.4	92.1
S2	74.6	88.3	74.6	85.7	76.1	90.6
S3	64.4	68.2	64.4	65.4	62.2	83.1
S4	71.8	77.1	66.4	66.4	71.2	84.4
S5	64.0	64.0	60.9	63.1	61.4	66.6
S6	61.2	61.2	59.1	57.1	60.4	64.1
S7	61.1	55.0	61.1	57.6	59.8	65.0
S8	95.7	98.0	87.3	96.0	97.1	98.1
S9	78.2	83.6	73.7	84.2	77.9	85.9
S10	77.5	86.4	66.1	73.5	72.1	80.5
S11	78.8	80.1	76.7	83.1	78.2	79.7
S12	90.4	92.4	80.5	88.4	90.8	92.8
S13	82.8	89.8	72.6	87.7	80.8	86.5
Avg.	76.0	79.5	71.6	76.9	75.0	82.3

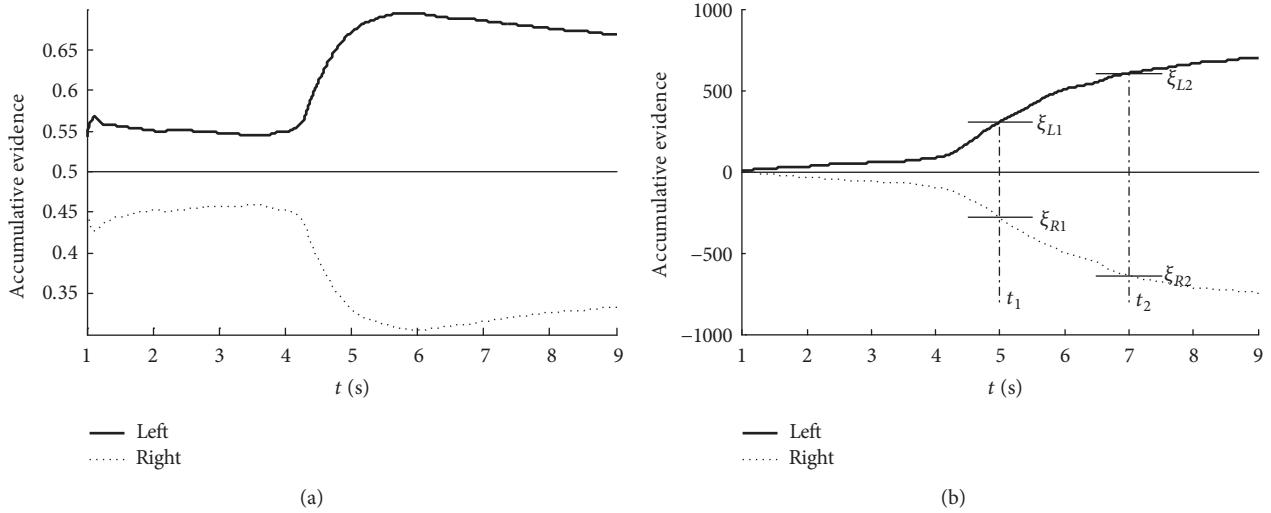


FIGURE 3: The average accumulative process of classification information for subject S1: (a) SB method; (b) SPRT method.

TABLE 3: The classification results of the SPRT method.

Subject	ACC (%)	MI (bit)	T (s)
S1	92.8	0.627	1.23
S2	88.7	0.491	1.95
S3	84.6	0.380	2.80
S4	85.0	0.390	2.35
S5	73.6	0.167	2.33
S6	68.2	0.098	1.23
S7	70.9	0.130	0.87
S8	97.9	0.853	1.33
S9	86.2	0.421	2.48
S10	83.8	0.361	2.54
S11	84.9	0.388	1.76
S12	93.6	0.657	1.69
S13	83.6	0.356	2.23
Avg.	84.1	0.409	1.91

of observations may be required to improve the accuracy, and vice versa. This can be seen from Figure 3(b), given the expected stopping time t_1 , the thresholds are ξ_{L1}, ξ_{R1} . When the expected stopping time is set to be t_2 , the thresholds change to be ξ_{L2}, ξ_{R2} . Obviously, it will achieve a higher accuracy with more decision time. This depicts the inherent trade-off between decision time (costs) and accuracy (benefits) of the SPRT method.

The resulting time courses for the classification accuracy, the MI, and SMI (calculated as $MI(t)/(t - 3\text{ s})$ for $t > 3.5\text{ s}$) for subject S1 are presented in Figure 4. The SMI quantifies the response time. During the first 4 s, the classification performs at a rate no better than chance. Afterwards a steep ascent in the classification accuracy can be observed, meanwhile reflecting a raising MI. The increase of MI indicates an increase in separation ability between left- and right-hand motor imagery. The maximum classification accuracy and the maximum MI are achieved at about 6.4 s and 7.1 s, respectively. The observation of nonmonotonic relationship

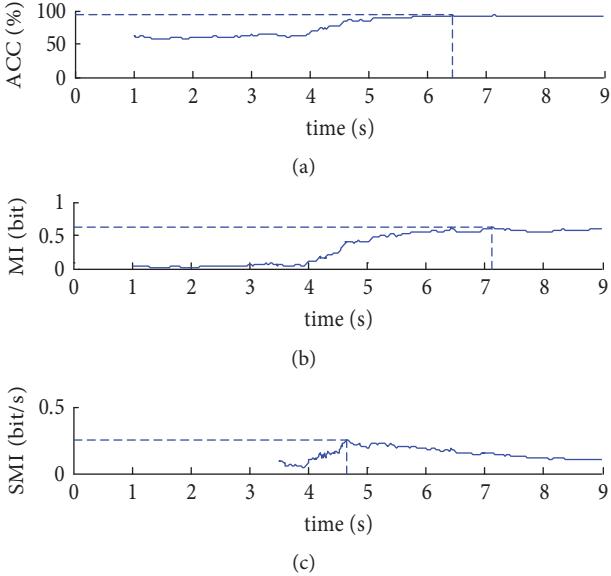


FIGURE 4: The time course of the classification accuracy, the MI, and the steepness of MI for subject S1: (a) ACC; (b) MI; (c) SMI.

between accuracy and the time is due to the limited data available which leads to the final decision. With consideration about time, the maximum steepness of MI is obtained at around 4.6 s.

The general sequential framework of the present approach is customizable to suit different task objectives, such as improving accuracy, MI, or steepness of MI. However, the optimization of one particular objective will come at the expense of the others.

4. Conclusion

In this paper, we present a SPRT method in conjunction with power projective base method to recognize mental states. The power projective method was first developed to determine features by maximizing the average projection energy difference of the two types of signals. With the accumulative evidence curve, the proposed SPRT sets the two-constrained thresholds based on a desired expected stopping time. The SPRT method adds the benefit of a customizable trade-off between accuracy and decision speed. Specifically, the thresholds in this method were determined without predefined error probabilities. Using standardized datasets, improved performances were demonstrated.

Although this study suggests that SPRT would be useful to balance the speed and accuracy for different BCI applications, we realize that further work requires investigation. Future work will attempt to validate the proposed method with a larger dataset and implement it in our BCI-actuated robotic system. Moreover, we will investigate the multiway SPRT theory and its application to the multiclass BCI systems.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported in part by the grants from the Natural Science Foundation of China (NSFC) (61005088 and 61573079), the Open Research Foundation of State Key Lab. of Digital Manufacturing Equipment & Technology in Huazhong University (DMETKF2015007), the Fundamental Research Funds for the Central Universities to Dr. Liu, and the NSFC-Shenzhen Robotics Research Center Project (U1613219) to Dr. Wu.

References

- C. E. Bouton, A. Shaikhouni, N. V. Annetta et al., "Restoring cortical control of functional movement in a human with quadriplegia," *Nature*, vol. 533, pp. 247–250, 2016.
- M. J. Vansteensel, E. G. M. Pels, M. G. Bleichner et al., "Fully implanted brain-computer interface in a locked-in patient with ALS," *The New England Journal of Medicine*, vol. 375, no. 21, pp. 2060–2066, 2016.
- G. R. Muller-Putz and G. Pfurtscheller, "Control of an electrical prosthesis with an SSVEP-based BCI," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 1, pp. 361–364, 2008.
- B. Rebsamen, C. L. Teo, Q. Zeng et al., "Controlling a wheelchair indoors using thought," *IEEE Intelligent Systems*, vol. 22, no. 2, pp. 18–24, 2007.
- V. Lawhern, W. D. Hairston, K. McDowell, M. Westerfield, and K. Robbins, "Detection and classification of subject-generated artifacts in EEG signals using autoregressive models," *Journal of Neuroscience Methods*, vol. 208, no. 2, pp. 181–189, 2012.
- K. Samiee, P. Kovács, and M. Gabbouj, "Epileptic seizure classification of EEG time-series using rational discrete short-time fourier transform," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 2, pp. 541–552, 2015.
- R. Liu, Y.-X. Wang, and L. Zhang, "An FDES-Based shared control method for asynchronous brain-actuated robot," *IEEE Transactions on Cybernetics*, vol. 46, no. 6, pp. 1452–1462, 2016.
- J. I. Gold and M. N. Shadlen, "The neural basis of decision making," *Annual Review of Neuroscience*, vol. 30, pp. 535–574, 2007.
- R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, Wiley-Interscience, New York, NY, USA, 2nd edition, 2001.
- G. Krausz, R. Scherer, G. Korisek, and G. Pfurtscheller, "Critical decision-speed and information transfer in the "Graz Brain-Computer Interface"," *Applied Psychophysiology and Biofeedback*, vol. 28, no. 3, pp. 233–240, 2003.
- B. Blankertz, G. Dornhege, M. Krauledat, K.-R. Müller, and G. Curio, "The non-invasive Berlin brain-computer interface: fast acquisition of effective performance in untrained subjects," *NeuroImage*, vol. 37, no. 2, pp. 539–550, 2007.
- A. Wald, *Sequential Analysis: Design Methods and Applications*, John Wiley & Sons, Inc., New York; Chapman & Hall, Ltd., London, 1947.
- B. Blankertz, K. R. Müller, G. Curio, T. M. Vaughan et al., "The BCI competition 2003 set III: probabilistic modeling of sensorimotor μ rhythms for classification of imaginary hand movements," *IEEE Transactions on Biomedical Engineering*, vol. 51, pp. 1044–1051, 2004.
- B. Blankertz, K.-R. Müller, D. J. Krusinski et al., "The BCI competition III: validating alternative approaches to actual BCI

- problems,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 153–159, 2006.
- [15] N. Brodu, F. Lotte, and A. Lécuyer, “Exploring two novel features for EEG-based brain-computer interfaces: Multifractal cumulants and predictive complexity,” *Neurocomputing*, vol. 79, pp. 87–94, 2012.
 - [16] J. D. R. Millan, F. Renkens, J. Mourino, and W. Gerstner, “Non-invasive brain-actuated control of a mobile robot by human EEG,” *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1026–1033, 2004.
 - [17] S. Lemm, C. Schäfer, and G. Curio, “BCI Competition 2003—data set III: modeling of sensorimotor μ rhythms for classification of imaginary hand movements,” *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1077–1080, 2004.
 - [18] S. Lemm, C. Schäfer, and G. Curio, “Aggregating classification accuracy across time: application to single trial EEG,” in *Advances in Neural Information Processing Systems*, vol. 19, pp. 825–832, Aggregating classification accuracy across time, application to single trial EEG, 2007.