Review of Sparse Representation-Based Classification Methods on EEG Signal Processing for Epilepsy Detection, Brain-Computer Interface and Cognitive Impairment

Dong Wen [1,2], Peilei Jia [1,2], Qiusheng Lian [1,2], Yanhong Zhou [3]* and Chengbiao Lu [4]*

[1] School of Information Science and Engineering, Yanshan University, Qinhuangdao, China, [2] The Key Laboratory for Computer Virtual Technology and System Integration of Hebei Province, Yanshan University, Qinhuangdao, China, [3] School of Mathematics and Information Science and Technology, Hebei Normal University of Science and Technology, Qinhuangdao, China, [4] School of Basic Medicine, Xinxiang Medical University, Xinxiang, China

At present, the sparse representation-based classification (SRC) has become an important approach in electroencephalograph (EEG) signal analysis, by which the data is sparsely represented on the basis of a fixed dictionary or learned dictionary and classified based on the reconstruction criteria. SRC methods have been used to analyze the EEG signals of epilepsy, cognitive impairment and brain computer interface (BCI), which made rapid progress including the improvement in computational accuracy, efficiency and robustness. However, these methods have deficiencies in real-time performance, generalization ability and the dependence of labeled sample in the analysis of the EEG signals. This mini review described the advantages and disadvantages of the SRC methods in the EEG signal analysis with the expectation that these methods can provide the better tools for analyzing EEG signals.

Keywords: sparse representation-based classification, sparse representation, EEG signal, preclinical mild cognitive impairment, mild cognitive impairment, Alzheimer's disease, epilepsy, brain computer interface

INTRODUCTION

Sparse representation (SR) is used to represent data with as few atoms as possible in a given overcomplete dictionary. By using the SR, we can concisely represent the data and easily extract the valuable information from the data. The sparse representation-based classification (SRC) methods have become a research hotspot for the data processing in many fields (Vialatte et al., 2009, 2012; Liu et al., 2012; Kaleem et al., 2013; Shin et al., 2015; Yuan et al., 2015), and can greatly simplify the processing of the multi-dimensional electroencephalograph (EEG) signals from epilepsy, mild cognitive impairment (MCI), Alzheimer's disease (AD) and brain computer interface (BCI).

Currently, studies on SRC methods used in the brain disorders and BCI involve mainly the preprocessing, SR and feature extraction, and have achieved accomplishments in computational accuracy, efficiency and robustness. Preclinical mild cognitive impairment (Pre-MCI) is a cognitive impairment status between normal aging and MCI, and also an earliest status of cognitive impairment which is more difficult to be diagnosed relative to MCI and AD (Sperling et al., 2011; Zhou et al., 2016). With the improvement of computational accuracy and efficiency, SRC methods may have potential to aid the diagnosis of Pre-MCI.
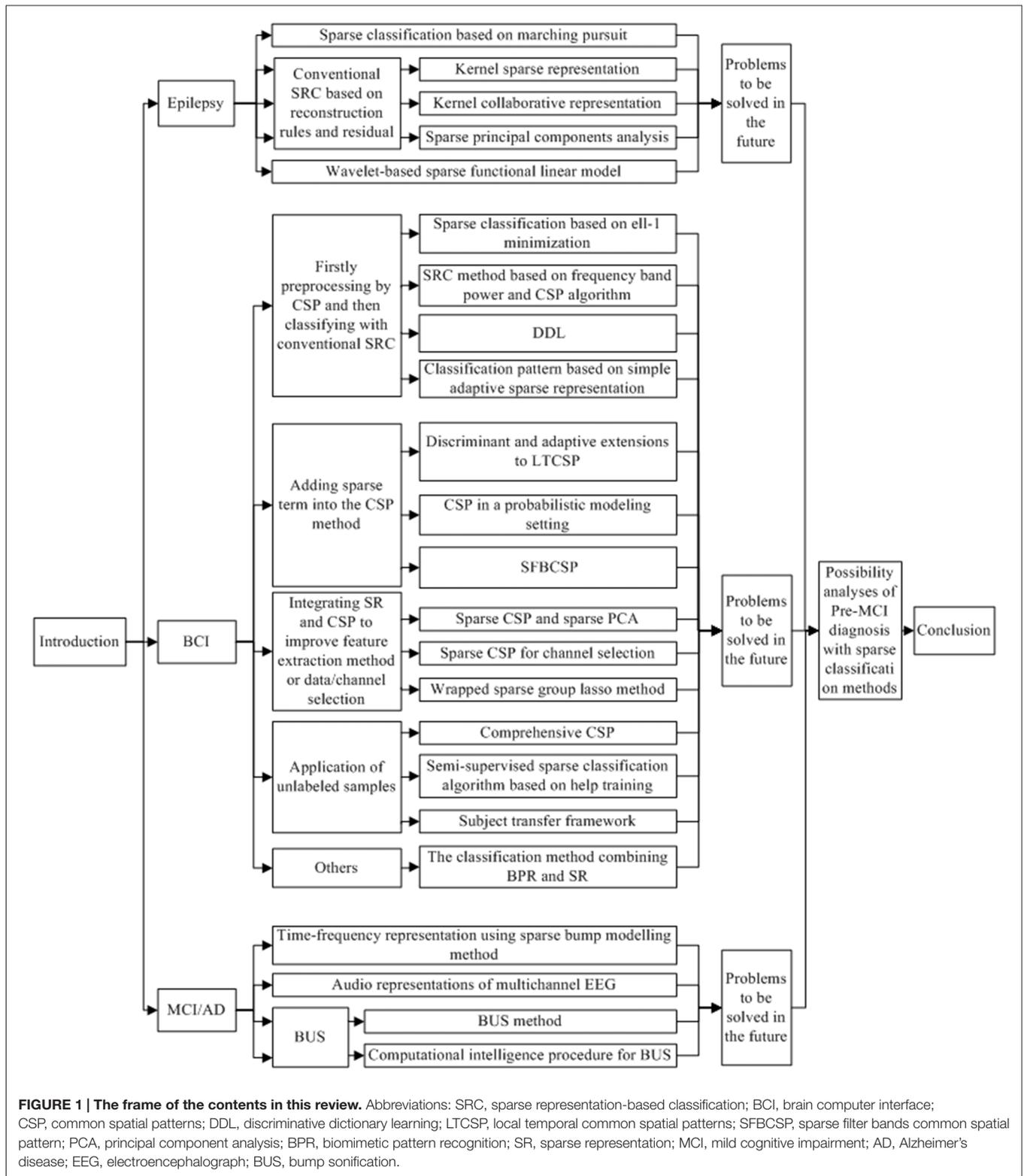
FIGURE 1 | The frame of the contents in this review. Abbreviations: SRC, sparse representation-based classification; BCI, brain computer interface; CSP, common spatial patterns; DDL, discriminative dictionary learning; LTCSP, local temporal common spatial patterns; SFBCSP, sparse filter bands common spatial pattern; PCA, principal component analysis; BPR, biomimetic pattern recognition; SR, sparse representation; MCI, mild cognitive impairment; AD, Alzheimer's disease; EEG, electroencephalograph; BUS, bump sonification.

However, there still exist some deficiencies needed to be solved.

This article reviewed the SRC methods in the analysis of EEG signals of epilepsy, MCI, AD and BCI, and discussed the possibility for the application of SRC methods in the diagnosis of Pre-MCI patients. The frame of this article was presented in **Figure 1**, and the main findings were listed in **Table 1**.

TABLE 1 | Summary of studies for epilepsy, brain computer interface (BCI), mild cognitive impairment (MCI) and Alzheimer's disease (AD) detection using sparse representation-based classification (SRC) methods.

(A) Summary of studies for epilepsy detection using SRC

Fields	Methods	Literatures	Datasets	Accuracy (%)	Sensitivity (%)	Specificity (%)	Main findings
Epilepsy		Wang and Guo (2011)		100	100	100	Obtaining the highest accuracy and robust to noise
	Sparse classification based on marching pursuit	Guo et al. (2012)	Datasets Z (ictal) and S (healthy)	100	–	–	Reducing running time and feature dimension
		Wang et al. (2013)		100	–	–	Enhancing the classification accuracy and the efficiency simultaneously
	Kernel sparse representation	Yuan et al. (2014)	Set A (healthy), D (interictal) and E (ictal)	98.63 ± 2.80	98.25 ± 4.37	99.00 ± 4.00	
	Kernel collaborative representation	Yuan et al. (2015)	Interictal and ictal IEEG database	99.99	100	100	Avoiding the choice and calculation of EEG features
	Sparse principal components analysis	Xie et al. (2012)	Set A,B (healthy) and C,D (interictal)	99.999 ± 0.0002	–	–	Focusing on the extraction of signal features with high discrimination power
	Wavelet-based sparse functional linear model	Xie and Krishnan (2013)	Set A,B (healthy), C,D (interictal) and E (ictal)	100			Effective combination of feature extraction and classification methods
			EEG data from University of Freiburg	99	–	–	

(B) Summary of studies for BCI using SRC

Fields	Methods	Literatures	Datasets	Accuracy (%)	Sensitivity (%)	Specificity (%)	Main findings
BCI	Sparse classification based on ell-1 minimization	Shin et al. (2011)	Motor imagery (left and right hand) EEG from four healthy subjects	91.67	–	–	SRC based on ell-1 minimization
	SRC method based on frequency band power and CSP algorithm	Shin et al. (2012)	INFONET dataset from experiment and Dataset IVa from BCI competition III	75.75/96.85	–	–	SRC based on L1 minimization.
		Shin et al. (2013)	Dataset Iva from BCI competition III	96.85	–	–	Dictionary based on the CSP filtering and the band power
	Discriminative dictionary learning (DDL)	Zhou et al. (2012)	Dataset IVa from BCI competition III	70.5/94.9	–	–	Lower computational complexity and higher accuracy than SRC

(Continued)

TABLE 1 | (Continued).

Fields	Methods	Literatures	Datasets	Accuracy (%)	Sensitivity (%)	Specificity (%)	Main findings
	Classification pattern based on simple adaptive sparse representation	Shin et al. (2015)	Motor imagery (left and right hand, foot) EEG from 10 healthy subjects and dataset IVc from BCI Competition III	98.0/96.07	–	–	Adaptive classification techniques based on sparse representation
	Discriminant and adaptive extensions to local temporal common spatial patterns	Wang (2013)	Dataset IVa from BCI competition III and Dataset IIa of BCI competition IV	98.21/93.75	–	–	Discriminant extension: combining the between-class and the within-class scatter information. Adaptive extension: defining the weights by utilizing the sparse representation
	CSP in a probabilistic modeling setting	Wu et al. (2015)	Dataset IIIa, IVa from BCI competition III	90.68 ± 9.93	–	–	Proposing probabilistic CSP (P-CSP) model
	Sparse filter bands common spatial pattern	Zhang et al. (2015)	Dataset IVa from BCI competition III and Dataset IIb from BCI competition IV	92.05 ± 2.45 81.17 ± 3.55	– –	– –	Automatically selecting the significant filter bands to improve classification performance
	Sparse CSP and sparse PC A	Shi et al. (2011)	Dataset IIIa from BCI competition III	90	–	–	Sparse subspace learning technique
	Sparse CSP for channel selection	Arvaneh et al. (2011)	Dataset IIa from BCI competition IV and Dataset IVa from BCI competition III	82.55 ± 12.8 73.5 ± 15.1	– –	– –	Improving performance in the case of noise interference and limited data
		Goksu et al. (2011, 2013)	ECoG dataset of BCI competition 2005	90	–	–	Extension of the greedy search based solution to multiple sparse filters
		Tomida et al. (2015)	Dataset IVa from BCI competition III and Dataset I from BCI competition IV	87.64/81.25	–	–	Introducing weighted averaging with weight coefficients rejecting low quality trials
	Wrapped sparse group lasso method	Wang et al. (2015)	Dataset I from BCI competition IV	84.72	–	–	Simultaneously achieving channel and feature selection with a lower error rate

(Continued)

TABLE 1 | (Continued).

Fields	Methods	Literatures	Datasets	Accuracy (%)	Sensitivity (%)	Specificity (%)	Main findings
	Comprehensive CSP	Wang and Xu (2012)	Dataset IVa from BCI competition III and EEG motor movement/ Imagery dataset	98.2/89.5	–	–	Contributing a comprehensive CSP (cCSP) that learns on both labeled and unlabeled trials
	Semi-supervised sparse classification Jiaetal., algorithm based on 2014 help training	Jia et al. (2014)	Dataset I from BCI competition I and Dataset II–IV from BCI competitin II	97/82	–	–	Selecting samples with high confidence according to sparse representation classifier
	Subject transfer framework	Tu and Sun (2012)	EEG data from the NIPS 2001 BCI workshop	72.49	–	–	Reducing the training sessions of the target subject by utilizing samples from other subjects
	The classification method combining BPR and SR	Ge and Wu (2012)	Dataset 1: Iva in 2005 BCIC III	94	–	–	Utilizing SR to solve the overlapping coverage problem of BPR
		Ren et al. (2014)	Dataset 1: Iva in 2005 BCIC III	97	–	–	
			Dataset 2: from Tongji University	82.6			
			Dataset 3: from BCIC III	88.03			

(C) Summary of studies for MCI/AD detection using SRC

Fields	Methods	Literatures	Datasets	Accuracy (%)	Sensitivity (%)	Specificity (%)	Main findings
MCI/AD	Time-frequency representation using sparse bump modeling method	Vialatte et al. (2005a,b)	Healthy control, MCI patients and AD patients	93	86.4	97.4	Compressing nformation contained in EEG time-frequency maps
	Bump sonification (BUS) method	Vialatte and Cichocki (2006)		–	–	–	Perceiving simultaneously every channel, and analyzing more tractably the time dynamics of the signals
	Computational intelligence procedure for BUS	Vialatte et al. (2009)	Mildly impaired patients progression towards AD group and Control group	89	–	–	Applicating BUS to online sonification
	Audio representations of multichannel EEG	Vialatte et al. (2012)	MCI group and Control group	89	–	–	Presenting a physiologically inspired method for generating music scores from multi-channel EEG

THE EEG SIGNAL ANALYSIS METHODS BASED ON SRC

Application and Performance Evaluation of SRC in Epilepsy Detection

Method Description and Evaluation

Currently, there are three perspectives of SRC methods used in epileptic detection, including reconstruction rules and residual error classifications on the whole classification stage, overcomplete dictionary on the preprocessing stage, and wavelet-based sparse functional linear model on the feature extraction stage.

For the first perspective, as the reconstruction rule classifications do not need to extract features or to design a classifier, the applied range of the methods is therefore greatly improved, and is superior than the traditional epilepsy detection methods. Using the classification method based on kernel SR and kernel collaborative representation, the classification accuracy in analyzing the epilepsy EEG signals reached up to 98.63% and 99.99% respectively, and the fast speed in computation can help to monitor epilepsy in real-time (Yuan et al., 2014, 2015).

Using above methods, good performance in classifications were demonstrated between epileptic patients with ictal EEG normal control group, or between epileptic patients with interictal EEG and ictal EEG However, for the classification between epileptic patients with interictal EEG and normal control group, whether these methods can achieve the similar performance remains to be further verified. Recently, using sparse principal components analysis method with reconstruction rules, the performance of classification between epilepsy patients with interictal EEG and normal control group was demonstrated to be excellent (Xie et al., 2012; Xie and Krishnan, 2013).

For the second perspective, Wang and Guo (2011) initially proposed SR based on matching pursuit and selected decomposition coefficients and atom parameters as features. However, the computation complexity was relatively high. To reduce the computation complexity, they then proposed Harmony Search method to find the optimal atom parameters, and selected the decomposition coefficients, FR parameters and restructured error to constitute a feature vector (Guo et al., 2012). By constituting the feature vector using decomposition coefficients, atom parameters and FR parameters, the computing time was further reduced (Wang et al., 2013).

For the third perspective, using wavelet-based sparse functional linear model, the accuracy in classifying epilepsy patients with interictal EEG from normal control group was up to 100% (Xie et al., 2012; Xie and Krishnan, 2013). However, the computation efficiency of feature extraction needs to be improved by using the methods, such as signal decomposition algorithms (Kaleem et al., 2013).

Problems to be Solved in the Future

A first problem is how to automatically determine the appropriate dictionary size and feature number of overcomplete dictionary. Secondly, the computation speed of SR needs to be improved. The aspects of the improvement may include dictionary learning algorithm and sparse coefficient solution algorithm. Thirdly, the difference between epileptic patients with interictal EEG and normal control group need to be analyzed in depth. It is the main reason why actual performance of different methods can be distinguished only when the difference between two kinds of signals is very small.

Application and Performance Evaluation of SRC in BCI

Method Description and Evaluation

Five perspectives of the SRC methods applied in BCI system were presented in this review. The main stream idea of the first three perspectives is to improve the classification performance, the feature extraction and data selection by combining SR with common spatial patterns (CSP). For the fourth perspective, researchers used unlabeled samples to improve the classification performance. As for the fifth perspective, some scholars proposed integrating SR with other traditional classification methods.

For the first perspective, some researchers used CSP and conventional SRC methods for signal preprocessing and classifying, respectively. SRC method based on ell-1 minimization has a classification accuracy of 91.67% (Shin et al., 2011), and the classification accuracy in constructing dictionary reached 96.85% when using the band power feature of signal filtered by CSP (Shin et al., 2012, 2013). However, it is difficult to select the appropriate number of CSP filters, and the computation complexity still needs to be reduced. In view of this, recently proposed discriminative dictionary learning (DDL) improved the classification accuracy and computational efficiency (Zhou et al., 2012). A new classification method based on simple adaptive SR also showed a relatively high classification accuracy (Shin et al., 2015).

For the second perspective, sparse term is often used to improve the performance of the CSP method. Wang (2013) integrated discriminant and adaptive extensions to local temporal CSP, which had better classification accuracy. CSP algorithm was cast in a probabilistic modeling setting to overcome overfitting problem of CSP by using of sparse Bayesian learning (Wu et al., 2015). Sparse filter bands common spatial pattern (SFBCSP) recently proposed by Zhang et al. (2015). showed an improved classification accuracy. However, the determination of the regularization parameter λ in SFBCSP is time consuming, and SFBCSP is not suitable for the analysis of the data set with small samples.

For the third perspective, SR and CSP are often integrated to improve the effectiveness of feature extraction or data/channel selection. In respect of feature extraction, the sparse component analysis (SCA) and CSP were utilized to construct a combined feature vector (Li et al., 2005). The sparse CSP and sparse principal component analysis (PCA) were applied to select relevant EEG components and extract EEG features in BCI system, respectively (Shi et al., 2011). However, there exists a

vast improvement space in the classification accuracy of these methods.

The classification performance can be improved according to the selection of different data/channels. A sparsity-aware method was proposed in order to select and remove low-quality trial data (Tomida et al., 2015). When applying L1 regularization term to CSP, Yong et al. (2008) showed that the average number of electrodes was reduced to 11% with a slight decrease of classification accuracy. To ensure the lowest reduction degree of classification performance, the minimal subset of EEG channels was selected for the classification. When L1/L2 norm was combined with CSP, the performance of channel selection algorithm was improved in the case of noise interference and limited data (Arvaneh et al., 2011). A sparse CSP (sCSP) method proposed by Goksu et al. (2011, 2013) showed a low computation complexity. However, the performance may be decreased when the different samples were used or the number of training samples is low.

A wrapped sparse group lasso method to select mixed EEG channel feature is suitable for high dimensional feature fusion. Stability and computing speed in this method were high, but the classification accuracy needs to be improved (Wang et al., 2015). The channel selection methods with CSP likely were trapped in a local minimum due to the non convexity of the optimization problem in CSP, which resulted in a decline in classification accuracy (Goksu et al., 2013).

For the fourth perspective, the less training samples will lead to the generalization performance deterioration caused by over-fitting, and it is easy to obtain unlabeled samples. Therefore, some researchers studied comprehensive learning mode to combine the labeled with unlabeled data, and showed that the classification performance was largely improved compared to the traditional CSP. The comprehensive learning mode includes the comprehensive CSP and semi-supervised SRC algorithm (Wang and Xu, 2012; Jia et al., 2014). A subject transfer framework reduced the training sessions of the target subjects by utilizing samples from other subjects and improved the classification accuracy (Tu and Sun, 2012). However, the computation complexity of this method was high, and the number of samples must be equal, which limited its application in reality.

For the fifth perspective, biomimetic pattern recognition (BPR) and SR were combined to accomplish the task of classification (Ge and Wu, 2012). A new classification method which combined BPR and SR under the semi-supervised co-training framework was recently proposed (Ren et al., 2014). These methods utilized SR to solve the overlapping coverage problem of BPR, and the classification accuracy was greatly increased compared to traditional classification methods. Mixed alternating least squares based on nonnegative matrix factorization were proposed to analyze event-related potential and event related spectral perturbation features. As a consequence, the performance of the algorithm was increased (Sburlea et al., 2015).

Problems to be Solved in the Future

Some problems remain to be solved in the field of BCI application. On account of channel selection in SRC, it is necessary not only to reduce channels, but also to maintain a high classification rate at the same time. Nevertheless, how to balance both is a challenge. It is still a research focus to determine the appropriate number of spatial filters in order to avoid over-fitting and meet the requirements of sparse coefficient solution. In addition to the principle based on the minimization of the reconstruction error, it is necessary to select new perspectives in the dictionary construction methods.

Application and Performance Evaluation of SRC in Detection of MCI and AD

Method Description and Evaluation

There are a few studies about SRC methods for the detection of MCI and AD. Most studies focused on the angle of sparse bump modeling. The classification accuracy was 93% when using the sparse bump modeling method in the analysis of the EEG signal (Vialatte et al., 2005a,b). However, it still needs validation with more datasets. A BUS method (Vialatte and Cichocki, 2006) and a computational intelligence procedure for online sonification were proposed by Vialatte et al. (2009, 2012). The results showed high identification accuracy and also confirmed the potential of these methods to be used in real-time diagnosis. In Vialatte et al. (2011) improved the classification specificity of clinical EEG by means of wavelet transform and sparse bump modeling. However, the application of sparse bump modeling method is limited to the analysis of the events at low frequency bands. And for the reason of using a low pass filter, gamma band activity did not suitable for the analysis in using this method.

Problems to be Solved in the Future

When utilizing SRC for the analysis of EEG signals from MCI and AD patients, the classification performance of SRC can be improved by using sparse Bayesian learning method to extract coupling and synchronization feature. For the MCI classification, the space sparsity of the brain areas and time sparsity of channel samples need to be considered. Reducing the amount of data participating in the classification by selecting channels will promote the classification performance.

Application of SRC Methods for the Analysis of EEG Signal of Pre-MCI Patients

Application of SRC method in the analysis of epilepsy, BCI, MCI and AD has achieved considerable achievements, however no relevant research literatures about Pre-MCI diagnosis using SRC methods can be found. We thus proposed to use SRC method for Pre-MCI diagnosis. The small difference in EEG signals between the Pre-MCI patients and normal control group makes the diagnosis of Pre-MCI difficult. However, if the accuracy, sensitivity, specificity and computing speed of SRC methods can be further improved, it is possible that these methods can be used for the diagnosis of the Pre-MCI. As the data dimension of Pre-MCI is high, we need to consider the space sparsity of brain areas and time sparsity of EEG signals of every channel, reduce the amount

of data used in the classification by selecting channels in order to improve classification performance and enhance the effectiveness in the dictionary design, feature extraction and SR.

CONCLUSION

We evaluated the SRC methods in the analysis of EEG signals from epilepsy, BCI, MCI and AD and illustrated the characteristics, advantages and disadvantages of various methods. The SRC methods have become an effective tool in aiding the diagnosis of brain disorder. Further improving the current SRC methods by such as combining SR with CSP will largely increase the classification accuracy and efficiency as well as sensitivity, making it potential for the application in diagnosis of Pre-MCI.

AUTHOR CONTRIBUTIONS

DW and YZ designed the study and wrote this article. PJ and QL wrote this article. CL designed the study and revised this article.

REFERENCES

Arvaneh, M., Guan, C., Ang, K. K., and Quek, H. C. (2011). "Spatially sparsed common spatial pattern to improve BCI performance," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Prague 2412–2415.

Ge, Y., and Wu, Y. (2012). A new hybrid method with biomimetic pattern recognition and sparse representation for EEG classification. *Emerg. Intell. Comput. Technol. Appl.* 304, 212–217. doi: 10.1007/978-3-642-31837-5_31

Goksu, F., Ince, N. F., and Tewfik, A. H. (2011). "Sparse common spatial patterns in brain computer interface applications," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Prague, 533–536.

Goksu, F., Ince, N. F., and Tewfik, A. H. (2013). Greedy solutions for the construction of sparse spatial and spatio-spectral filters in brain computer interface applications. *Neurocomputing* 108, 69–78. doi: 10.1016/j.neucom.2012.12.003

Guo, P., Wang, J., Gao, X. Z., and Tanskanen, J. (2012). "Epileptic EEG signal classification with marching pursuit based on harmony search method," in *IEEE International Conference on Systems, Man and Cybernetics*, Seoul, 283–288.

Jia, M., Wang, J., Li, J., and Hong, W. (2014). Application of semi-supervised sparse representation classifier based on help training in EEG classification. *Sheng Wu Yi Xue Gong Cheng Xue Za Zhi* 31, 1–6.

Kaleem, M., Guergachi, A., and Krishnan, S. (2013). EEG seizure detection and epilepsy diagnosis using a novel variation of empirical mode decomposition. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2013, 4314–4317. doi: 10.1109/EMBC.2013.6610500

Li, Y., Guan, C., and Qin, J. (2005). Enhancing feature extraction with sparse component analysis for brain-computer interface. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 5, 5335–5338. doi: 10.1109/iembs.2005.1615686

Liu, M., Zhang, D., and Shen, D. (2012). Ensemble sparse classification of Alzheimer's disease. *Neuroimage* 60, 1106–1116. doi: 10.1016/j.neuroimage.2012.01.055

Ren, Y., Wu, Y., and Ge, Y. (2014). A co-training algorithm for EEG classification with biomimetic pattern recognition and sparse representation. *Neurocomputing* 137, 212–222. doi: 10.1016/j.neucom.2013.05.045

Sburlea, A. I., Montesano, L., and Minguez, J. (2015). Continuous detection of the self-initiated walking pre-movement state from EEG correlates without session-to-session recalibration. *J. Neural Eng.* 12:036007. doi: 10.1088/1741-2560/12/3/036007

Shi, L., Li, Y., Sun, R., and Lu, L. (2011). A sparse common spatial pattern algorithm for brain-computer interface. *Neural Inf. Process.* 7062, 725–733. doi: 10.1007/978-3-642-24955-6_86

Shin, Y., Lee, S., Ahn, M., Cho, H., Jun, S. C., and Lee, H. N. (2015). Simple adaptive sparse representation based classification schemes for EEG based

brain-computer interface applications. *Comput. Biol. Med.* 66, 29–38. doi: 10.1016/j.compbiomed.2015.08.017

Shin, Y., Lee, S., Ahn, M., Jun, S. C., and Lee, H.-N. (2011). "Motor imagery based BCI classification via sparse representation of EEG signals," in *8th International Symposium on Noninvasive Functional Source Imaging of the Brain and Heart and 2011 8th International Conference on Bioelectromagnetism*, Banff, AB, 93–97.

Shin, Y., Lee, S., Lee, J., and Lee, H. N. (2012). Sparse representation-based classification scheme for motor imagery-based brain-computer interface systems. *J. Neural Eng.* 9:056002. doi: 10.1088/1741-2560/9/5/056002

Shin, Y., Lee, S., Woo, S., and Lee, H.-N. (2013). "Performance increase by using a EEG sparse representation based classification method," in *2013 IEEE International Conference on Consumer Electronics*, Las Vegas, NV, 201–203.

Sperling, R. A., Aisen, P. S., Beckett, L. A., Bennett, D. A., Craft, S., Fagan, A. M., et al. (2011). Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 7, 280–292. doi: 10.1016/j.jalz.2011.03.003

Tomida, N., Tanaka, T., Ono, S., Yamagishi, M., and Higashi, H. (2015). Active data selection for motor imagery EEG classification. *IEEE Trans. Biomed. Eng.* 62, 458–467. doi: 10.1109/TBME.2014.2358536

Tu, W., and Sun, S. (2012). A subject transfer framework for EEG classification. *Neurocomputing* 82, 109–116. doi: 10.1016/j.neucom.2011.10.024

Vialatte, F. B., and Cichocki, A. (2006). "Sparse bump sonification: a new tool for multichannel EEG diagnosis of mental disorders; application to the detection of the early stage of Alzheimer's disease," in *Neural Information Processing, Lecture Notes in Computer Science*, ed. I. King (Berlin, Heidelberg: Springer), 92–101.

Vialatte, F., Cichocki, A., Dreyfus, G., Musha, T., Rutkowski, T. M., and Gervais, R. (2005a). "Blind source separation and sparse bump modelling of time frequency representation of EEG signals: new tools for early detection of Alzheimer's disease," in *Proceedings of the IEEE Workshop on Machine Learning for Signal Processing* Mystic, CT, 27–32.

Vialatte, F., Cichocki, A., Dreyfus, G., Musha, T., Shishkin, S. L., and Gervais, R. (2005b). Early detection of Alzheimer's disease by blind source separation, time frequency representation and bump modeling of EEG signals. *Biol. Inspirations Artif. Neural Netw.* 3696, 683–692. doi: 10.1007/11550822_106

Vialatte, F. B., Dauwels, J., Musha, T., and Cichocki, A. (2012). Audio representations of multi-channel EEG: a new tool for diagnosis of brain disorders. *Am. J. Neurodegener. Dis.* 1, 292–304.

Vialatte, F.-B., Dauwels, J., Maurice, M., Musha, T., and Cichocki, A. (2011). Improving the specificity of EEG for diagnosing Alzheimer's disease. *Int. J. Alzheimers Dis.* 2011:259069. doi: 10.4061/2011/259069

Vialatte, F. B., Musha, T., and Cichocki, A. (2009). Sparse bump sonification: a new tool for multichannel EEG diagnosis of brain disorders. *Artif. Intell. Med.* 138–139.

Wang, H. (2013). Discriminant and adaptive extensions to local temporal common spatial patterns. *Pattern Recognit. Lett.* 34, 1125–1129. doi: 10.1016/j.patrec.2013.03.014

Wang, J., and Guo, P. (2011). "Epileptic electroencephalogram signal classification based on sparse representation," in *International Conference on Neural Computation Theory and Applications*, Seol 15–23.

Wang, J., Gao, X., and Guo, P. (2013). Feature extraction based on sparse representation with application to epileptic EEG classification. *Int. J. Imaging Syst. Technol.* 23, 104–113. doi: 10.1002/ima.22045

Wang, H., and Xu, D. (2012). Comprehensive common spatial patterns with temporal structure information of EEG data: minimizing nontask related EEG component. *IEEE Trans. Biomed. Eng.* 59, 2496–2505. doi: 10.1109/TBME.2012.2205383

Wang, J. J., Xue, F., and Li, H. (2015). Simultaneous channel and feature selection of fused EEG features based on sparse group lasso. *Biomed Res. Int.* 2015:703768. doi: 10.1155/2015/703768

Wu, W., Chen, Z., Gao, X., Li, Y., Brown, E. N., and Gao, S. (2015). Probabilistic common spatial patterns for multichannel EEG analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* 37, 639–653. doi: 10.1109/TPAMI.2014.2330598

Xie, S., and Krishnan, S. (2013). Wavelet-based sparse functional linear model with applications to EEGs seizure detection and epilepsy diagnosis. *Med. Biol. Eng. Comput.* 51, 49–60. doi: 10.1007/s11517-012-0967-8

Xie, S., Krishnan, S., and Lawniczak, A. T. (2012). "Sparse principal component extraction and classification of long-term biomedical signals," in *25th International Symposium on Computer-Based Medical Systems*, Rome, 1–6.

Yong, X., Ward, R. K., and Birch, G. E. (2008). "Sparse spatial filter optimization for EEG channel reduction in brain-computer interface," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Las Vegas, NV, 417–420.

Yuan, Q., Zhou, W., Yuan, S., Li, X., Wang, J., and Jia, G. (2014). Epileptic EEG classification based on kernel sparse representation. *Int. J. Neural Syst.* 24:1450015. doi: 10.1142/S0129065714500154

Yuan, S., Zhou, W., Yuan, Q., Li, X., Wu, Q., Zhao, X., et al. (2015). Kernel collaborative representation-based automatic seizure detection in intracranial EEG. *Int. J. Neural Syst.* 25:1550003. doi: 10.1142/S0129065715500033

Zhang, Y., Zhou, G., Jin, J., Wang, X., and Cichocki, A. (2015). Optimizing spatial patterns with sparse filter bands for motor-imagery based brain-computer interface. *J. Neurosci. Methods* 255, 85–91. doi: 10.1016/j.jneumeth.2015.08.004

Zhou, Y., Tan, C., Wen, D., Sun, H., Han, W., and Xu, Y. (2016). The biomarkers for identifying preclinical Alzheimer's disease via structural and functional magnetic resonance imaging. *Front. Aging Neurosci.* 8:92. doi: 10.3389/fnagi.2016.00092

Zhou, W., Yang, Y., and Yu, Z. (2012). "Discriminative dictionary learning for EEG signal classification in Brain-computer interface", *12th International Conference on Control Automation Robotics and Vision*, Guangzhou, 1582–1585.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.