An Improved Multiple LASSO Model for Steady-State Visual Evoked Potential Detection

Ruimin Wang, Keiji Iramina, and Sheng Ge

Abstract

Improving the classification accuracy in brain–computer interface (BCI) with a short data length is important to increase the BCI system's information transfer rate. Least absolute shrinkage and selection operator (LASSO) has been examined to be an effective way to detect the steady-state visual evoked potential (SSVEP) signals with a short time window. In this paper, an improved multiple LASSO model for SSVEP detection is proposed, which can process multichannel electroencephalogram (EEG) signals without electrode selection. EEG data from twelve healthy volunteers were used to test the improved multiple LASSO model. Compared with the traditional LASSO model, the improved multiple LASSO model gives a significantly better performance with multichannel EEG data.

Keywords

Brain–computer interface (BCI) • Steady-state visual evoked potential (SSVEP) • Electroencephalogram (EEG) • Least absolute shrinkage and selection operator (LASSO)

1 Introduction

Brain–computer interfaces (BCIs) support direct communication and control between a human brain and external devices [1]. This approach has shown a promising way to re-establish basic communication and autonomy abilities for individuals suffering from severe motor disabilities.

Electroencephalography (EEG) is a commonly used brain activity monitoring method. Over the past decades, many laboratories have built various paradigms using noninvasive scalp EEG that enable users to acquire conscious control over their brain activities [2]. Steady-state visual evoked potential (SSVEP) is an effective noninvasive electrophysiological response over the occipital region to a visual stimulus oscillating at a constant frequency higher than 4 Hz [3, 4]. It can be regarded as a typical sinusoidal oscillator with the same frequency as the stimulus, and often includes some higher harmonics. SSVEP-based BCIs have high information transfer rate and practical applications.

EEG signals acquired in an SSVEP-based BCI include quite much environmental noise, and many task-unrelated activities, i.e., spontaneous activity. Least absolute shrinkage and selection operator (LASSO) is a sparse regression model and gives high interpretable regression coefficients. It has shown a useful and robust way for signal detection [5, 6], and has been used for EEG channel selection and feature extraction [7]. A LASSO model for SSVEP feature extraction was proposed by Zhang et al. [8]. Compare with other traditional methods, power spectral density analysis, and canonical correlation analysis, LASSO achieves better performance within a short time window (TW) in a few-channel BCI system [8, 9].

R. Wang · S. Ge (✉)
Key Laboratory of Child Development and Learning Science, Ministry of Education, Research Center for Learning Science, Southeast University, 2 Sipailou, Nanjing, China
e-mail: shengge@seu.edu.cn

R. Wang · K. Iramina
Graduate School of Systems Life Sciences, Kyushu University, Fukuoka, Japan

K. Iramina
Graduate School of Information Science and Electrical Engineering, Kyushu University, Fukuoka, Japan

T. Vo Van et al. (eds.), *6th International Conference on the Development of Biomedical Engineering in Vietnam (BME6)*, IFMBE Proceedings 63, https://doi.org/10.1007/978-981-10-4361-1_72

Because LASSO can only compute the regression of a single one-dimension dependent variable on multidimensional independent variables. For a multichannel BCI system, each channel has to be processed separately. The classification results will be affected if some channels' data are interfered by noise [9]. Our earlier research also confirmed that, for a multichannel BCI, LASSO achieved best results under a channel selection process; and its performance dropped badly without channel selection [10].

In this study, we propose an improved multiple LASSO model for SSVEP detection, which can extract useful SSVEP features from multichannel EEG data directly without channel selection.

2 Methods

2.1 LASSO Estimate

As for an independent variable matrix $X = (x_1, x_2, \cdots, x_M) \in R^{N \times M}$ and a corresponding dependent variable vector $y \in R^{N \times 1}$, the linear regression of y on X is general defined as

$$y = X\beta + \varepsilon = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_M x_M + \varepsilon, \qquad (1)$$

where β is the regression coefficients and ε is the noise vector.

The LASSO estimate can be found by solving the following optimization problem [8]:

$$\hat{\beta} = \mathrm{argmin}_\beta \left(\|y - X\beta\|_2^2 + \lambda \|\beta\|_1 \right), \qquad (2)$$

where $\|\cdot\|_2$ and $\|\cdot\|_1$ denote the l_2-norm and l_1-norm, respectively. λ is a penalty parameter which encourages a sparse solution $\hat{\beta}$.

2.2 Traditional LASSO for SSVEP Frequency Detection

Sine-cosine waves, which are oscillating at frequencies or harmonics of visual stimuli, are commonly used as reference signals for SSVEP detection [8, 11]. The constructed reference signals $S_{i(i=1,2,\cdots,K)}$ can be expressed as follows:

$$S_i = \begin{pmatrix} \sin(2\pi f_i t) \\ \cos(2\pi f_i t) \\ \vdots \\ \sin(2\pi H f_i t) \\ \cos(2\pi H f_i t) \end{pmatrix}^T, \; t = \frac{1}{f_s}, \frac{2}{f_s}, \cdots, \frac{N}{f_s}, \qquad (3)$$

where K is the number of candidate stimulus frequencies, f_i is the stimulus frequency, f_s is the sampling frequency, N is the number of sampling points, and H is the number of harmonics. The superscript T denotes a matrix transpose operation.

For the traditional LASSO model, a single channel EEG data $y \in R^{N \times 1}$ is assumed to be a standard linear combination of the constructed reference signals $S = [S_1, S_2, \cdots, S_K]$:

$$y = S\beta + \varepsilon, \qquad (4)$$

where $\beta = [\beta_{1,1}, \beta_{1,2}, \cdots, \beta_{1,2H}, \cdots, \beta_{K,1}, \beta_{K,2}, \cdots, \beta_{K,2H}]$ is the regression coefficients estimated by LASSO, and ε is the noise vector [8].

The entries $[\beta_{i,1}, \beta_{i,2}, \cdots, \beta_{i,2H}]_{i=1,2,\cdots,K}$ imply the contribution degrees (CDs) of the ith stimulus frequency f_i and its harmonics to the one-channel EEG data y:

$$CD_i = \sum_{j=1}^{2H} |\beta_{i,j}|. \qquad (5)$$

For an M-channel BCI system, each channel's data $y_1, y_2, \cdots, y_M$ is processed respectively and equally, the CD values are computed as follows:

$$CD_i = \sum_{m=1}^{M} \sum_{j=1}^{2H} \left|\beta_{i,j}^m\right|. \qquad (6)$$

Target frequency contributes the most to EEG data. So, the frequency that corresponds to the maximum CD value is recognized as the target frequency.

2.3 Improved Multiple LASSO for SSVEP Detection

Instead of computing the LASSO regression of a single channel EEG data on sine-cosine waves (see (4)), in this study, we propose an improved multiple LASSO model that computes the LASSO regressions of sine-cosine waves on multichannel EEG data.

For a set of sine-cosine waves ($S_{i(i=1,2,\cdots,K)}$, see (3)) and M-channel EEG data $Y = [y_1, y_2, \cdots, y_M]$, the LASSO regression of each sine or cosine wave on Y is computed respectively:

$$\begin{aligned} \sin(2\pi f_i t)^T &= Y\beta_1^i + \varepsilon_1^i \\ \cos(2\pi f_i t)^T &= Y\beta_2^i + \varepsilon_2^i \\ &\vdots \\ \sin(2\pi H f_i t)^T &= Y\beta_{2H-1}^i + \varepsilon_{2H-1}^i \\ \cos(2\pi H f_i t)^T &= Y\beta_{2H}^i + \varepsilon_{2H}^i \end{aligned}, \qquad (7)$$

where $\beta^i_{j(j=1,2,\cdots,2H)} = [\beta^i_{j,1}, \beta^i_{j,2}, \cdots, \beta^i_{j,M}]^T \in R^{M\times 1}$ indicate the linear relationships between multichannel EEG data and sine-cosine waves.

Sine-cosine waves having different frequencies are uncorrelated and have no noise. Thus, the regressions coefficients computed in (7) can be treated equally. The related degrees (RDs) between the multichannel EEG data Y and the reference signals $S_{i(i=1,2,\cdots,K)}$ can be calculated using

$$RD_i = \sum_{j=1}^{2H} \sum_{m=1}^{M} \left|\beta^i_{j,m}\right|. \quad (8)$$

The frequency corresponding to the maximum RD value is recognized as the target frequency.

3 Experiment

Twelve healthy volunteers (seven female, five male; aged 18–28 years) participated in the study. Four white arrows flicking at 7.5, 8.57, 10, and 12 Hz on a black screen were used as visual stimuli. In each trial, after 1-s target cue presented in the middle of the screen, four white arrows flickered simultaneously for 5 s. In the experiment, three runs (40 trials for each run) separated by a 2-min break were recorded per participant.

EEG signals were recorded using a SynAmps2 (NeuroScan Inc., El Paso, TX, USA) from six commonly used parieto-occipital EEG channels P3, Pz, P4, O1, Oz, and O2 at 1000 Hz sampling rate, and were bandpass filtered between 0.1 and 100 Hz. Linked-mastoids were used as a reference, and a ground electrode was placed between FPz and Fz. A bandpass filter was used to filter the EEG between 6 and 28 Hz to improve the signal-to-noise ratio of SSVEP signals before further analysis.

A more detailed description of the experiment design can be found in [10].

4 Results

The effectiveness of LASSO for SSVEP detection has been validated in [8, 9], and [10]. In this study, to validate the performance of our proposed improved multiple LASSO model, we compared it with the traditional LASSO model using different EEG channel combinations and TW lengths. The results are shown in Table 1.

For 1-channel (Oz) EEG data, the performances of the improved multiple LASSO model and the traditional LASSO model were almost the same. When 3-channel (O1, Oz, and O2) and 6-channel (P3, Pz, P4, O1, Oz, and O2) EEG data were used, the improved multiple LASSO model achieved significantly higher classification accuracies than the traditional LASSO model (paired t-test, $p < 0.05$), and more EEG channels yielded higher performance. However, the average classification accuracies of the 6-channel EEG obtained by the traditional LASSO were lower than that of fewer channels, which was consistent with the research by Zhang et al. [9].

From the experiment results, we can conclude that our proposed improved multiple LASSO model is effective against dealing with multichannel EEG data.

5 Discussion

LASSO can only compute the regression of a single dependent vector on an independent matrix. For the traditional LASSO model, the linear regression of each EEG channel data on reference signals (sine-cosine waves) is calculated separately and treated without distinction.

Table 1 Averaged classification accuracies (%) for the 12 participants derived by the traditional and the improved multiple LASSO model

TW (s)	Traditional LASSO			Improved multiple LASSO		
	1-chl	3-chl	6-chl	1-chl	3-chl	6-chl
1.0	47.6	48.9	41.7	47.9	53.3	53.0
1.5	56.3	56.7	48.1	56.3	63.6	64.9
2.0	61.8	63.5	52.9	62.0	70.1	73.5
2.5	67.8	69.6	60.4	67.4	75.6	79.9
3.0	72.7	74.9	66.2	73.1	79.0	84.4
3.5	76.3	77.4	67.4	75.8	82.3	87.2
4.0	77.4	79.4	70.4	77.2	85.5	88.1
4.5	80.3	81.6	73.3	79.6	87.3	90.2
5.0	81.9	83.6	75.6	81.9	88.1	90.8

Note 1-chl: Oz; 3-chl: O1, Oz, O2; 6-chl: P3, Pz, P4, O1, Oz, and O2

However, multichannel EEG data are correlated and include different noise. So, for a multichannel BCI system, a channel selection process is indispensable to obtain high performance using the traditional LASSO model [10].

Instead of computing the regressions of EEG data on reference signals, the regressions of reference signals on EEG data are computed in the improved multiple LASSO model. In this way, all the EEG channels' data are combined effectively to model a sine or cosine wave, and the combined coefficients reflect their linear relationship to the sine or cosine wave. Unlike multichannel EEG data, the constructed sine-cosine waves are uncorrelated and have no noise, so they can be treated separately and equally.

From Table 1, we can see that the classification accuracies computed by the traditional and the improved multiple LASSO model were most the same for the 1-channel EEG. Both of the regression coefficients estimated by computing the LASSO regressions of EEG on reference signals (traditional LASSO) or by computing the LASSO regressions of reference signals on EEG (improved multiple LASSO) are effective features for SSVEP detection. However, the improved multiple LASSO model can extract useful information from multichannel data through combining them linearly, it has achieved significantly higher classification accuracies than the traditional model for a multichannel BCI system. In this sense, our proposed improved LASSO model can also be used for channel selection. Thus, no additional channel selection is needed for the improved multiple LASSO model.

As a shrinkage regression method, a penalty parameter needs to be set for LASSO [12]. In this study, to improve the system's robustness, the penalty parameter was set to 6 candidate values (0, 0.02, 0.04, 0.06, 0.08, 0.1), and the averaged results were used for analysis.

6 Conclusions

In this study, we proposed an improved multiple LASSO model for SSVEP detection without channel selection, which can achieve a significantly higher classification accuracy than the traditional LASSO model in a multichannel BCI system. This model does not need training, and its penalty parameter is set in a simple way. Thus, it is an effective approach to improve the SSVEP-based BCI systems' performance.

Acknowledgements This work was supported by the National Basic Research Program of China (2015CB351704), the Fundamental Research Funds for the Southeast University (CDLS-2015-01).

References

1. Ebrahimi T (2007) Recent advances in brain-computer interfaces, IEEE 9th Workshop on MMSP, Crete, Greece, 2007, p 17
2. Nicolas-Alonso LF, Gomez-Gil J (2012) Brain computer interfaces, a review. Sensors (Basel) 12:1211–1279
3. Cheng M, Gao X, Gao S, Xu D (2002) Design and implementation of a brain–computer interface with high transfer rates. IEEE Trans Biomed Eng 49:1181–1186
4. Liu Q, Chen K, Ai QS, Xie SQ (2014) Review: recent development of signal processing algorithms for SSVEP-based brain computer interfaces. J Med Biol Eng 34(4):299–309
5. Dyar MD, Carmosino ML, Breves EA et al (2012) Comparison of partial least squares and lasso regression techniques as applied to laser-induced breakdown spectroscopy of geological samples. Spectrochim Acta Part B 70:51–67
6. Xu C, Ladouceur M, Dastani Z, Richards JB, Ciampi A et al (2012) Multiple regression methods show great potential for rare variant association tests. PLoS ONE 7(8):e41694
7. Wang J, Xue F, Li H (2015) Simultaneous channel and feature selection of fused EEG features based on sparse group lasso. Biomed Res Int 2015:703768
8. Zhang Y, Jin J, Qing XY, Wang B, Wang XY (2012) LASSO based stimulus frequency recognition model for SSVEP BCIs. Biomed Signal Proces 7(2):104–111
9. Zhang YS, Dong L, Zhang R, Yao DZ, Zhang Y, Xu P (2014) An efficient frequency recognition method based on likelihood ratio test for SSVEP-based BCI. Comput Math Method M 2014:908719
10. Sheng G, Wang RM, Leng Y, Wang HX, Lin P, and Iramina KA double-partial least-squares model for the detection of steady-state visual evoked potentials. IEEE J Biomed Health Press
11. Lin ZL, Zhang CS, Wu W, Gao XR (2007) Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs. IEEE Trans Biomed Eng 54(6):1172–1176
12. Boulesteix AL, Richter A, Bernau C (2013) Complexity selection with cross-validation for lasso and sparse partial least squares using high-dimensional data. In: Lausen B, Van den Poel D, and Ultsch A (eds) Algorithms from and for nature and life. Springer International Publishing, Switzerland, vol 5517, pp 261–268