

Research Article

Improving Classification by Feature Discretization and Optimization for fNIRS-based BCI

Baolei Xu^{1,2*}, Yunfa Fu³, Gang Shi¹, Xuxian Yin^{1,2}, Zhidong Wang^{1,4} and Hongyi Li^{1,5}

¹State Key Laboratory of Robotics, Shenyang Institute of Automation (SIA), Chinese Academy of Sciences (CAS), Shenyang 110016, P. R. China ²University of Chinese Academy of Sciences, Beijing 100049, P. R. China ³School of Automation and Information Engineering, Kunming University of Science and Technology, Kunming 650500, P. R. China

⁴Department of Advanced Robotics, Chiba Institute of Technology, Chiba 2750016, Japan ⁵School of Mechanical Engineering & Automation, Northeastern University, Shenyang, China

Abstract

In this paper, we present a signal discretization and feature selection method to improve classification accuracy for fNIRS based brain computer interface (BCI) system, which can classify right hand clench force motor imagery and clench speed motor imagery at an accuracy of 69%-81% through 5 fold cross validation in 6 subjects. Difference between oxyhemoglobin and deoxyhemoglobin (we abbreviate this difference as HbD) is proposed as a new feature type and shows outstanding performance in some subjects. Linear kernal support vector machine (SVM) classification between clench force motor imagery and clench speed motor imagery using four concentration feature types (oxyhemoglobin, deoxyhemoglobin, totalhemoglobin, and HbD) is implemented. Our results demonstrate that feature discretization using Chi2 algorighm and feature optimization using 'MIFS' (Mutual Information Feature Selection) criterion can improve the classification accuracy by more than 35%. Except total hemoglobin, all the other three feature types can be used as the optimum feature for different subjects. The results of this paper can also be used in online BCI applications.

Keywords: Brain computer interface; BCI; fNIRS; Feature discretization; Feature selection

Introduction

Brain Computer interface (BCI) is a technology to provide device control using brain signals directely without the participation of the peripheral nervous system and muscles. It was first funded by the Defense Advanced esearch Projects Agency (DARPA) in 1970s [1] and has got great developments since 1990s [2-4]. Many signals can be used to build a BCI system, including electroencephalography (EEG), electrocorticography (ECoG), magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI) and functional near infrared spectroscopy (fNIRS) [5]. At present, EEG is the most wildly used non invasive signal for its high temporal resolution, portability and low cost. fNIRS is another promising modality for BCI purpose. This technology can provide the metabolism information during the cognitive process, and can be used with EEG simultaneously to achieve higher classification accuracy for BCI systems [6].

fNIRS is used in BCI research since 2004 by Coyle [7], and BCI reseachers have focused on this modality increasingly in the following years [8-10]. fNIRS can obtain concentration changes of deoxyhemoglobin (Hb) and oxyhemoglobin (HbO) calculated by the modified Beer-Lamber Law from the detected otpical signals [11]. Both concentration signals and optical signals can be used as features in BCI research. Our previous research shows that concentration signals can get better results than optical signals [12]. Early researches often use the amplitude of concentration signals as features directly [7,13]. Cui et al. compares the classification accuracy of different feature spaces, including signal amplitude, signal history, history gradient and spatial pattern. Their research indicate that combining multi channel signal history can get the best results [14]. The slope of the concentration signal can also be used as the feature for classification and get desirable results [15]. Classification accuracy can be further improved by using both 'filter' and 'wrapper' feature selection methods [16]. Herff et al. take the mean hemodynamic concentration difference between task period and rest period as features, use Mutual Information based Best Individual Feature (MIBIF) to decrease feature space, getting a

classification accuracy of 61% between three speaking modes using support vector machines (SVM) [17]. Power et al. take the slope of Hb and HbO concentration as features, and use sequential feed-forward feature selection algorithm to choose the optimal feature subset with the Fisher criterion, getting a classification accruracy of 71.1% between the Mental Arithmetic (MA) task and the No Control (NC) state using Linear Discriminant Analysis (LDA) [18].

In this paper, we apply motor imagery paradigm in the experiment. A new signal type of Difference between HbO and Hb (HbD) is proposed as a feature. The continous feature is discretized first, then the Mutual Information Based Feature Selection (MIFS) is used to get the optimal feature subset, and the Support Vector Machine (SVM) with linear kernal is used to classify the task of clench speed motor imagery from the task of clench force motor imagery on the same hand. The comparison of classification accuracy between signal types of Hb, HbO, HbT(total hemoglobin) and HbD shows that the best results relies on individuals. For some subjects HbD can get comparable results with HbO, and for some subjects HbD can get better results than all the other three feature types.

Experiment Design

Paradigm

In our experiment, the motor imagery of clench force and clench

*Corresponding author: Baolei Xu, State Key Laboratory of Robotics, Shenyang Institute of Automation (SIA), Chinese Academy of Sciences (CAS), Shenyang 110016, P. R. China, Tel: +86-15002426919; E-mail: blxu@sia.cn

Received September 28, 2013; Accepted November 18, 2013; Published December 26, 2013

Citation: Xu B, Fu Y, Shi G, Yin X, Wang Z, et al. (2014) Improving Classification by Feature Discretization and Optimization for fNIRS-based BCI. J Biomim Biomater Tissue Eng 19: 119. doi: 10.4172/1662-100X.1000119

Copyright: © 2013 Xu B, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

speed of right hand is used as the paradigm. We measure each subject's Maximum clench Force (MF) using hand dynamometers, and told them to do real clench force practice at the level of 20%, 50% and 80% of MF. The clench speed task practice request the subject to clench their right hand at the frequency of 0.5Hz, 1Hz and 2Hz according to a metronome. We divide both task into three sub-types because we want to average the effect of different task parameters. During the experiment, the subjects are told to recall the feeling of doing real movement task. We adopt this motor parameters imagery paradigm based on two considerations: First, motor imagery is more natural for device control than paradigms using mental arithmetic and singing. Second, this paradigm can provide more direct control commands than paradigms that do not distinguish the hand motor imagery types, and such discrimination is especially important to improve the information transfer rate for BCI applications if both hands are used.

Six healthy subjects (three males and three females, averaged age 26.8 years) participate in the experiment. All the subjects are right handed. Three of them are trained three times before the experiment (subject number 1, 2, and 6), while the other three subjects are not trained (subject number 3, 4, and 5). All of them give written informed consent to participate in the experiment, and the experiment is approved by the Ethnical Committee of the Shenyang Institute of Automation (SIA), Chinese Academy of Sciences (CAS).

A single trial of the experiment consists of four parts: the base line period, the cue period, the task period and the rest period, as shown in Figure 1. The baseline period is used to produce a baseline for task period. The cue period is used to inform the motor imagery type to implement in the following task period that lasts 10 seconds. The rest peroid is used to make the hemodynamic level return to a normal level. The duration of a single trial in fNIRS experiment is much longer than EEG experiment due to the intrinsic time lag of hemodynamic signals [11]. Each subject take part in 3 sessions, and each session contains 60 trials (30 trials of clench force motor imagery and 30 trials of clench speed motor imagery).

Data aquisition

We acquire the hemodynamic signals using ETG-4000 produced by the Hitachi Co., Ltd. This device is a continuous wave equipment,

and it has two types of optode. One type is the illuminator and the other type is the detector, as is shown in Figure 2. The middle point between those two optodes is the meaurement channel. The illuminator optode can emit near infrared light at wavelength of 695 nm and 830 nm simultaneously. The detector receives the output light that has been modulated by the oxyhemoglobin and deoxyhemoglobin concentration changes, and transmits it to an Avalanche Photo Diode (APD) that converts optical changes to voltage changes. The voltage changes is digitalized by the analog to digital converter (ADC) and then used to calculate the concentraion changes by the modified Beer-Lamber Law as shown in Eq. (1), where $\mathcal{E}_{\lambda_1/\lambda_2,Hb/HbO}$ is the extinction coefficient of Hb/HbO under the corresponding wavelength, $I_{\lambda_1/\lambda_2}(t_1/t_2)$ is the light intensity, and DPF is the ratio of optical photons' actual path length and the illuminator detector distance [19]. To cover the sensorimotor area at both hemisphere, we choose two 3×3 optode layout helmets, which contain 24 channels altogether. The channel 7 and channel 21 are adjusted to directly above the C4 and C3 position in the standard 10-20 international system used for EEG recordings. The sampling frequency is 10Hz, which is sufficient enough for fNIRS based BCI system.

Page 2 of 5

In our experiment, the EEG and fNIRS data are acquired simultaneously because we want to research the multi modality feature for BCI application. However, in this paper, we only focus on the fNIRS signal, and the combined features will be studied in the future.

$$\Delta C = \begin{pmatrix} \Delta c_{Hb} \\ \Delta c_{HbO} \end{pmatrix} = \begin{pmatrix} \varepsilon_{\lambda_1,Hb} & \varepsilon_{\lambda_1,HbO} \\ \varepsilon_{\lambda_2,Hb} & \varepsilon_{\lambda_2,HbO} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\log \frac{I_{\lambda_1}(t_1)}{I_{\lambda_1}(t_2)}}{DPF_{\lambda_1}d} \\ \frac{\log \frac{I_{\lambda_2}(t_1)}{I_{\lambda_2}(t_2)}}{DPF_{\lambda_2}d} \end{pmatrix}$$
(1)

Data Processing Methods

Preprocessing

The original concentration data is linearly detrened first to remove low frequency drift due to optode movements, and then filtered by a Chebyshev II low pass IIR filter at a cutoff frequency of 0.1 Hz to remove high frequency artifacts such as heart rate and breathing rate. This is



reasonable because our previous research shows that the frequency component between 0.02 Hz and 0.08 Hz plays important role for data classification [20]. The time course average of the four different preprocessed fNIRS signals is shown in Figure 3. The data is then downsampled to 1 Hz to decrease feature space. The downsampling process loss no information on the low passed data according to the Shannon's sampling theorem.

Discretization

Discretization is the process of partitioning continous variables into categories. It is very useful for many machine learning (ML) algorithms to get better classification results [21]. Based on whether the class information is taken into account, discretization methods can classified into supervised methods and unsupervised methods. Generally, the supervised methods outperforms the unsupervised methods [22]. The equal width and equal frequency discretizaions are two representative algorithms of the unsupervised methods, but their performance is relatively poor. In this paper, we choose the Chi2 algorithm to discretize the continous concentration sigal. In Chi2 algorithm, the significance on relationship between the values of a feature and a class type is tested using the χ^2 statistic as shown in Eq. (2), where k is the number

of classes, A_{ii} is the number of patterns in the *ith* interval of *jth* class,

 $R_i = \sum_{i=1}^k A_{ij}$ is the number of patterns in the *ith* interval, $C_j = \sum_{i=1}^2 A_{ij}$ is the number of patterns in the *jth* class, $N = \sum_{i=1}^{2} R_i$ is the total number of

patterns [23], and E_{ii} is the expected frequency of $A_{ii} = R_i * C_i / N$.

$$\chi^{2} = \sum_{i=1}^{2} \sum_{j=1}^{k} \frac{(A_{ij} - E_{ij})^{2}}{E_{ij}}$$
(2)

Feature selection

Currently, there are three types of concentration signal used in fNIRS based BCI application: Hb, HbO and HbT. In this paper, we proposed a new concentration signal HbD, which refers to the difference between HbO and Hb concentration. HbO and Hb are typically strongly negatively correlated with each other under normal circumstances [24], and we propose that HbD may amplify the amplitude of concentration changes due to cognitive tasks.





The orignal feature space consists of 24 channels of the downsampled amplitude data within the task start point and 2s after the task end point, resulting in a feature set of 312 values. The 2s time lag is set to take into consideration of the intrinsic time lag of hemodynamic response. The four signal types (Hb, HbO, HbT, and HbD) are researched respectively.

Page 3 of 5

The 'wrapper' methods and the 'filter' methods are two brodly grouped approaches in feature selection techniques [16]. The former method takes the training accuracy of a paticular classifier as the criterion when searching for the best feature subsets. This method can get better results, while the subsets are overly fit to the classifier used, and the computational cost is considerably expensive. To make the optimized subset suitable for more classification methods, we apply the 'filter' method in our research.

Eight types of information based criteria, including 'MIFS', 'mRMR', 'CMIM', 'JMI', 'DISR', 'CIFE', 'ICAP', and 'CondRed' [16], are tested using one subject's data to choose the best criterion for our research. The 'MIFS' (Mutual Information Feature Selection) criterion is chooed in the following research because it's performance is best in all the tested criteria. The details feature selection method using 'MIFS' criterion can be found in [25].

Classification

Support vector machine (SVM) is a widely used efficient method in machine learning research, especially for classification with small data set [26,27]. Unlike other classification method that use all the sample data to train the model, SVM only selects the samples (the support vectors) that can produce the largest margin between two classes. Also, SVM maps the input vectors into a high dimensional feature space using a kernal function, and the pproblems that can not linearly separable in low dimensional space becomes possible in high dimensional feature space. SVM can get better performance by solving the optimization problem in Eq. (3) [26].

$$\min_{\substack{w,b,\xi\\w}} \frac{1}{2}w^Tw + C\sum_{i=1}^{l}\xi_i$$
(3)
subject to
$$y_i(w^T\phi(x_i) + b) \ge 1 - \xi_i, \quad \xi_i \ge 0, \ i = 1, ..., l$$

Where *x* is the training vevtors, $\phi()$ is the kernal function, *w* is the weight, b is the bias, ξ is the soft margin and C is a constant. The linear kernal $K(xi, xj) = x_i^T x_j$ is used in our reseach because it requires lesser parameters to optimize, and can achieve much higher classification results compared with other kernals when their parameters are not been optimized. The 5 fold cross validation is used to get the results more credible.

Results

To demonstrate the effectiveness of feature discretization and feature selection, we first compared classification accuracy of four different feature sets. The first feature set consists of the continuous data from the channel above C3. The second feature set consists of the discrete data from C3 channel. The third feature set consists of the discrete data from all the 24 channels, and the fourth feature set is the optimum subset chosen from the third feature set using the 'MIFS' cirterion. The comparison of these four different feature sets are shown in Figure 4 (left). The error rate of continuous C3 feature set decreases by 9%, 5%, 2%, and 6% when the data is discretized for signal type of Hb, HbO, HbT and HbD respectively, which implies that feature discretization can improve the classification accuracy. The error rate of the feature set containing discrete data from all the 24 channels reaches

Page 4 of 5

nearly 60%, and can be decreased to 25% when the optimum subset is chosen. The relationship between error rate and the feature number chosen by the 'MIFS' feature selection algorithm is shown in Figure 4 (right). The optimum feature number of most of the subjects is between 20 to 50.

The classification error rate using the optimized feature subset of four types of concentraion signal is shown in Table 1 and Figure 5 (left). Similar to EEG, fNIRS signal's characteristic varies between subjects. For subject 1, the mean error rate of Hb feature type for three sessons shows the lowest error rate than other three feature types, but the difference is not significant. Significant difference is observed olny in 3 subjects: for subject 2, HbD feature type has a lower error rate than Hb and HbT feature type at a significant level of 0.01 and 0.05; for subject 3, Hb feature has a lower error rate than HbO feature at a significant level of 0.05; for subject 6, HbO feature type has a lower error rate than HbT feature type at a significant level of 0.05. HbT feature type shows no significant lower error rate than other three feature type in all the subjects.

The trained subjects has lower mean error rate than no trained subjects, as is shown in Figure 5 (right). Though the difference level for



Figure 4: The error rate comparison of four different feature sets (left), and the error rate curve corresponding to increased discrete feature number chosen using 'MIFS' criterion (right).

Signal type		Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Subject 6
Hb	Session 1	0.25	0.28	0.27	0.25	0.27	0.24
	Session 2	0.15	0.27	0.20	0.28	0.25	0.20
	Session 3	0.17	0.27	0.27	0.27	0.25	0.25
	Mean ± std	0.19 ± 0.05	0.27 ± 0.01	0.24 ± 0.04	0.27 ± 0.02	0.26 ± 0.01	0.23 ± 0.03
HbO	Session 1	0.30	0.17	0.28	0.32	0.25	0.15
	Session 2	0.17	0.22	0.34	0.30	0.32	0.23
	Session 3	0.22	0.23	0.30	0.22	0.23	0.20
	Mean ± std	0.23 ± 0.07	0.21 ± 0.03	0.31 ± 0.03	0.28 ± 0.05	0.27 ± 0.04	0.19 ± 0.04
HbT	Session 1	0.17	0.20	0.30	0.23	0.32	0.27
	Session 2	0.22	0.32	0.23	0.25	0.30	0.27
	Session 3	0.27	0.25	0.27	0.27	0.28	0.38
	Mean ± std	0.22 ± 0.05	0.26 ± 0.06	0.27 ± 0.03	0.25 ± 0.02	0.30 ± 0.02	0.31 ± 0.07
HbD	Session 1	0.23	0.13	0.35	0.32	0.27	0.19
	Session 2	0.15	0.18	0.26	0.20	0.32	0.17
	Session 3	0.25	0.20	0.30	0.28	0.18	0.25
	Mean ± std	0.21 ± 0.05	0.17 ± 0.03	0.30 ± 0.04	0.27 ± 0.06	0.26 ± 0.07	0.20 ± 0.04





Figure 5: The error rate comparison of all the subjects using the optimized featue (left), and the comparison of training effets on classification results (right).

J Biomim Biomater Tissue Eng ISSN: 1662-100X Biochem, an open access journal

feature type of Hb and HbT is not significant, the error rate of trained subjects is lower than the no trained subjects at a significant level of 0.01 for both feature type of HbO and HbD.

Discussions and Conclusions

In this paper, we present a signal discretization and feature selection method to improve classification accuracy for fNIRS based BCI system, which can classify right hand clench force motor imagery and clench speed motor imagery at an accuracy of 69%-81% through 5 fold cross validation. HbD (Difference between HbO and Hb) is proposed as a new feature type and shows outstanding performance in some subjects. Although feature discretization has been used in many machine learning areas, we notice that it has not been used in BCI applications before. Our research shows that discretization using Chi2 algorithm can improve classification accuracy by 2%-9% for different feature types using signal from the channel above C3. Further more, the optimum feature subset chosen by 'MIFS' criterion can improve the classification accuracy by 35% compared with feature set containing all the discrete channels.

Due to the specificity of fNIRS signals, it is hard to conclude which concentration signal type is best for all the subjects. Our research shows that the best signal type depends on subject. In our study, Hb, HbO and HbD can get optimum results for different subjects, but HbT does not perform best in all the subjects. This may be explained by the strongly negatively correlated relationship between Hb and HbO. Our study also shows that subjects can improve their performance through training using feature type of HbO and HbD at a significant level of 0.01, which means that sufficient training before experiment is necessary to improve the classification accuracy.

Our research shows that classification of clench force motor imagery and clench speed motor imagey of the same hand is possible at a resonable accuracy, which can be used to increase the command number for BCI control. Also, the feature discretization and feature optimization method of this paper can be used in online BCI applications.

References

- 1. Vidal J (1973) Toward direct brain computer communication. Annu Rev Biophys Bioeng 2: 157-80.
- Pfurtscheller G, Doris F, Kalcher J (1993) Brain computer interface—A new communication device for handicapped persons. J Microcomput Applicat 16: 293-299.
- Sellers EW (2013) New horizons in brain computer interface research. Clinical Neurophysiology 124: 2-4.
- An J, Lee SH, Jin SH, Abibullaev B, Gwanghee J, et al. (2013) The beginning of neurohaptics: Controlling cognitive interaction via brain haptic interface. in Brain Computer Interface (BCI), 2013 International Winter Workshop on, Gangwo, 103-106.
- Nicolas Alonso LF, Gomez Gil J (2012) Brain computer interfaces, a review. Sensors (Basel, Switzerland) 12: 1211-79.
- Fazli S, Mehnert J, Steinbrink J, Curio G, Villringer A, et al. (2012) Enhanced performance by a hybrid NIRS EEG brain computer interface. Neuroimage 59: 519-529.
- Coyle S, Ward T, Markham C, McDarby G (2004) On the suitability of near infrared (NIR) systems for next generation brain computer interfaces. Physiological Measurement 25: 815-822.
- Abibullaev B, An J, Lee S, Jin S, Moon JI (2013) Minimizing inter subject variability in fNIRS based Brain Computer Interfaces via multiple kernel support vector learning. Medical Engineering Physics.
- 9. Abibullaev B, An J (2012) Classification of frontal cortex haemodynamic

responses during cognitive tasks using wavelet transforms and machine learning algorithms. Med Eng Phys 34: 1394-410.

- Abibullaev B, An J, JI Moon (2011) Neural Network Classification of Brain Hemodynamic Responses from Four Mental Tasks. International Journal of Optomechatronics 5: 340-359.
- Matthews F, Pearlmutter BA, Ward TE, Christopher S, Markham C (2008) Hemodynamics for brain computer interfaces. IEEE Signal Processing Magazine 25: 87-94.
- Xu B, Fu Y, Shi G, Yin X, Miao L, et al. (2013) Comparison of optical and concentration feature used for fNIRS based BCI system using HMM. Applied Mechanics and Materials 385-386: 1443-1448.
- Sitaram R, Zhang HH, Guan CT, Thulasidas M, Hoshi Y, et al. (2007) Temporal classification of multichannel near infrared spectroscopy signals of motor imagery for developing a brain computer interface. NeuroImage 34: 1416-1427.
- Cui X, Bray S, Reiss AL (2010) Speeded Near Infrared Spectroscopy (NIRS) Response Detection. PLoS One 5: e15474.
- Power SD, Kushki A, Chau T (2011) Towards a system paced near infrared spectroscopy brain computer interface: differentiating prefrontal activity due to mental arithmetic and mental singing from the no control state. J Neural Eng 8.
- Brown G, Pocock A, Zhao MJ, Luján M (2012) Conditional Likelihood Maximisation: A Unifying Framework for Information Theoretic Feature Selection. Journal of Machine Learning Research 13: 27-66.
- Herff C, Putze F, Heger D, Guan CT, Schultz T (2012) Speaking Mode Recognition from Functional Near Infrared Spectroscopy, in 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society 1715-1718.
- Power SD, Chau T (2013) Automatic single trial classification of prefrontal hemodynamic activity in an individual with Duchenne muscular dystrophy. Developmental Neurorehabilitation 16: 67-72.
- Delpy DT, Cope M, van der Zee P, Arridge S, Wray S, et al. (1988) Estimation of optical pathlength through tissue from direct time of flight measurement. Phys Med Biol 33: 1433-42.
- Xu B, Fu Y, Miao L, Wang Z, Li H (2011) Classification of fNIRS Data Using Wavelets and Support Vector Machine during Speed and Force Imagination.
 2011 IEEE International Conference on Robotics and Biomimetics (ROBIO). Phuket, Thailand.
- Kotsiantis S, Kanellopoulos D (2006) Discretization Techniques: A recent survey. International Transactions on Computer Science and Engineering 32: 47-58.
- Dougherty J, Kohavi R, Sahami M (1995) Supervised and Unsupervised Discretization of Continuous Features. The Twelfth International Conference on Machine Learning. Tahoe City, California, USA.
- Liu H, Setiono R (1997) Feature selection via discretization. leee Transactions on Knowledge and Data Engineering 9: 642-645.
- 24. Cui X, Bray S, Reiss AL (2010) Functional near infrared spectroscopy (NIRS) signal improvement based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics. NeuroImage 49: 3039-3046.
- Battiti R (1994) Using mutual information for selecting features in supervised neural net learning. IEEE Transactions on Neural Networks 5: 537-550.
- Chang CC, Lin CJ (2011) LIBSVM: A Library for Support Vector Machines. Acm Transactions on Intelligent Systems and Technology 2.
- 27. Cortes C, Vapnik V (1995) Support Vector Networks. Machine Learning 20: 273-297.

Citation: Xu B, Fu Y, Shi G, Yin X, Wang Z, et al. (2014) Improving Classification by Feature Discretization and Optimization for fNIRS-based BCI. J Biomim Biomater Tissue Eng 19: 119. doi: 10.4172/1662-100X.1000119