

基于 ICA 和极限学习机的模拟阅读脑电特征分类

官金安^{1,2}, 杨建华¹, 赵瑞娟²

(1 中南民族大学 生物医学工程学院, 认知科学国家民委重点实验室, 武汉 430074;

2 中南民族大学 医学信息分析及肿瘤诊疗湖北省重点实验室, 武汉 430074)

摘要 为了有效地提取 N2-P3 成分, 利用 ICA 对脑电数据进行盲源分离, 自动提取 N2-P3 成分; 同时为了克服传统方法如支持向量机、神经网络训练时间长、个别识别准确率不够高的缺点, 选择极限学习机作为分类器。在模拟阅读实验范式下, 记录了 7 名受试者的脑电数据, 利用 ICA 分别对每名受试者的高维脑电数据进行盲源分离, 提取出 N2-P3 成分, 以此作为靶特征, 并与非靶特征一起放入极限学习机分类器进行分类。训练得到 7 名受试者的训练时间和分类准确率, 并与支持向量机进行了比较。结果表明: 经过 ICA 特征提取后, 使用极限学习机进行分类, 该分类器学习速度快, 泛化能力强, 训练时间大大减少。在分类准确率上, ICA+ELM 的分类准确率较传统的最佳单通道+ SVM 有较大幅度的提升, 从后者平均的 82.4% 提升到了 97.7%。

关键词 模拟阅读; N2-P3 成分; 极限学习机

中图分类号 TP391.4 文献标识码 A 文章编号 1672-4321(2018)01-0085-05

EEG Feature Classification of Imitating-Reading Based on ICA and Extreme Learning Machine

Guan Jin'an^{1,2}, Yang Jianhua¹, Zhao Ruijuan²

(1 Key Laboratory of Cognitive Science, State Ethnic Affairs Commission, College of Biomedical Engineering, South-Central University for Nationalities, Wuhan 430074, China; 2 Hubei Key Laboratory of Medical Information Analysis and Tumor Diagnosis & Treatment, South-Central University for Nationalities, Wuhan 430074, China)

Abstract In order to effectively extract the N2-P3 components, ICA was introduced; At the same time, Extreme Learning Machine(ELM) was used as the classifier. The EEG data of 7 subjects were recorded, the ICA was used to separate the high-dimensional EEG data from each subject, and the N2-P3 component was extracted as the target Characteristics, and with untargeted features into the ELM for classification. The training time and classification accuracy of the seven subjects were trained and compared with the SVM. The results show that, after the ICA feature extraction, the training time is greatly reduced. In the classification accuracy, Compared with SVM, classification accuracy of ICA + ELM has a more substantial increase, from the latter average of 82.4% to 97.7%.

Keywords imitating-reading; N2-P3; ELM

近年来, 我们对“模拟阅读”脑-机接口(BCI)进行了持续的研究^[1-4], 其中, 脑电信号的特征提取和模式分类是两个重要研究内容, 对提高 BCI 的准确率起到至关重要的作用。

在特征提取方面, 先后采用了张量时频空模式^[5]、共空间模式(CSP)^[6]、最佳单通道^[7]等方法, 均取得了较好的效果。但我们发现, 有些受试者可以高效地使用 BCI, 也有许多人因错误率太高而无法

使用, 存在所谓“金牌受试者”的问题。这表明, 要使绝大多数人脑电信号的模式识别率都达到较高水平, 在特征提取方面还需进行新的探索。事实上, 可以把多道脑电信号的独立成分看成大脑中若干个“等效源”, 这些等效源的输出传播到不同的头皮电极位置, 就形成了所测得的多通道的脑电信号。此外, 脑电信号还包含 50Hz 工频干扰、眼电、肌电及心电等干扰信号。独立分量分析(ICA)是近年来发

收稿日期 2017-06-26

作者简介 官金安(1965-) 男, 教授, 博士, 研究方向: 生物医学信号检测与处理、脑-机接口, Email: guanja@tom.com.

基金项目 国家自然科学基金资助项目(91120017)

展起来的一种基于无监督统计学习的盲源分离方法,能够将复杂的混合信号分解成独立的信号分量.本文利用 ICA 方法分离其他干扰信号,达到消噪的效果,同时自动提取出混合在脑电信号中的 N2-P3 成分.实验结果表明,ICA 方法能够有效地提出 N2-P3 成分^[8],大幅提高分类正确率.

在 EEG 信号模式分类上,以支持向量机(SVM)^[9]和人工神经网络(ANNs)^[10]等最为常见,但是 SVM 和 ANN 等分类器会有训练速度慢、参数选择难、易产生局部最优解等缺点.Huang 于 2006 年提出了基于前馈神经网络的 ELM 学习算法^[11,12].近年来,该方法在大规模计算、高速信号处理、人工智能等领域应用越来越广泛^[13].与传统的神经网络和支持向量机器相比,ELM 拥有着显著的优势,如学习速度快、易于实施、和最小的人为干预等等.因此,选择 ELM 作为分类器来进行分类,在训练前只需设置合适的隐含层节点数和激活函数,然后为输入权值和隐含层偏差量进行随机赋值,最后引入矩阵广义逆的思想且通过最小二乘法获得输出权值,整个训练过程快速简单,无需繁琐地迭代和调整参数.

本文在“模拟阅读”实验模式下,对 7 名健康受试者用 ICA 进行 EEG 信号的盲源分离^[13],提取出 N2 和 P3 成分作为靶特征,采用 ELM 对靶与非靶特征进行训练和分类.为了验证 ICA 方法的有效性,与采用最佳单通道时域信号作为特征的分类效果进行了比较.此外,在模型训练所需的时间、分类准确率等方面,将 ELM 与 SVM 进行了对比.

1 实验模式以及数据获取

1.1 实验模式

本实验采用中南民族大学提出的模拟阅读(IR)诱发电位模式,该模式能够让受试者在实验中像平常阅读书本文字一样来获得视觉刺激,从而产生基于视觉诱发事件相关电位的 N2 和 P3 成分.用此模式在屏幕上构建一个虚拟键盘,就能实现人机交互.“模拟阅读”与通常阅读书本不同的是,实验过程中,受试者的视线相对视觉诱发界面保持静止状态,刺激符号串相对于视线移动,以减少视线移动引起的眼电对脑电信号的污染.在这种刺激方式中,将一个靶标符号(图 1(a))随机放到多个相似的非靶标符号(图 1(b))中,构成刺激符号串(图 1(c)).实验时,符号串匀速移过小视窗的速度是 5ms/pixel.一个试次(trial)的运行时间指的是符号串刚开始进入小视窗到符号串全部离开小视窗的时间间隔.图

(1(d))为一个试次的示意图.一旦靶标通过小视窗就会产生稳定的事件相关电位.其中,靶标符号、非靶标符号和小视窗的尺寸是 30*30 像素.靶标符号和非靶标符号具有相同的结构,两种符号的差别在于中间竖线是否被染成红色,靶标符号的中间竖线被染成红色,非靶标符号没有.

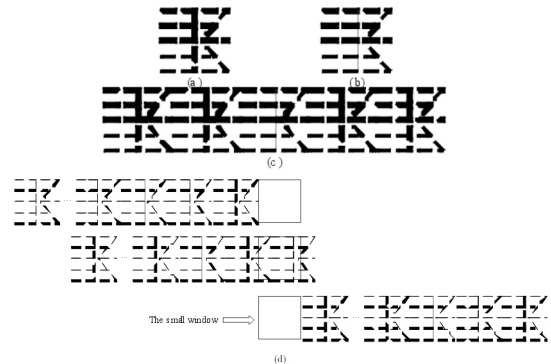


图 1 “模拟自然阅读”模式下的诱发字符串

Fig.1 Character String on the pattern of Imitating-Reading

1.2 实验数据获取

实验使用 Biosemi Active II 多导生理信号采集装置,采样频率为 2048Hz.实验采集了 7 名健康受试者在“模拟阅读”实验范式下的脑电信号,编号 S1-S7.每个受试者采集 4 组数据,每组 50 个试次,一共 200 个试次.数据的存储格式为:通道数×通道采样点数×实验试次数.实验采集的 EEG 会受到工频和肌电等噪声干扰,为了更加有效地提取出信号的特征,首先对 EEG 数据进行预处理,包括:去均值、20Hz 低通滤波及归一化.实验流程如图 2 所示.



图 2 实验过程示意图

Fig.2 Diagram of the experimental process

2 研究方法

2.1 独立分量分析(ICA)

本文分别对每个受试者的脑电数据进行 ICA 盲源分离.独立分量分析是近年发展起来的一种盲源分离方法(BSS).ICA 分离算法的关键在于如何度量分离结果的独立性,以随机变量的非高斯性和相互独立为分析目标,最终是为了从多通道观测数据中分离出相互独立的信号源.不仅可以消除噪声,而且可有效分离出相应的特征信号.ICA 算法的数学模型为:

$$\vec{X} = \vec{A}\vec{S}, \quad (1)$$

$$\vec{u} = \vec{W}\vec{X}, \quad (2)$$

公式(1)中:观测信号矢量 $\vec{X} = (\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n)^T$, 其余的每个分量 x_i 有相同的观测样本长度. $\vec{S} = (\vec{s}_1, \vec{s}_2, \dots, \vec{s}_n)^T$ 为未知源信号矢量 \vec{A} 为混合矩阵 \vec{S} 反映源信号在观测信号 \vec{x} 中的加权比重. 公式(2)中: $\vec{u} = (\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n)^T$ 为分离出的独立分量. \vec{W} 为待求的分离矩阵. ICA 算法的真正目的就是求得分离矩阵 \vec{W} , 它反映的是使输出的独立分量 \vec{u} 尽可能地逼近源信号 \vec{S} . 将独立分量 \vec{u} 尽可能地逼近源信号 \vec{S} . 根据公式(1)和(2)可得:

$$\vec{S} = \vec{W}\vec{X} = \vec{A}^{-1}\vec{X}. \tag{3}$$

但是由于混合矩阵 \vec{A} 是未知的, 因此不能直接得到最优的分离矩阵. 解决 ICA 问题方法要从寻优的角度出发, 首先随机给分离矩阵一个初始值, 然后根据优化准则调整分离矩阵, 使得到的分量尽可能相互独立. 用非高斯性描述 ICA 分离的分量的相互独立程度是非高斯性极大优化准则的基本思想.

2.2 极限学习机

2.2.1 单隐层前馈神经网络

设训练集 X 有 N 个样本, 表示为 $X = [x_1, x_2, \dots, x_N]$ 输出集为 T , 表示为 $T = [t_1, t_2, \dots, t_N]$ 激活函数为 $g(x)$ 的标准单隐层前馈神经网络^[14-16] 的数学模型可以表示为:

$$\sum_{i=1}^L \beta_i g_i(x_j) = \sum_{i=1}^L \beta_i g(\alpha_i x_j + b_i) = t_j, \quad j = 1, 2, \dots, N, \tag{4}$$

式中 x_j 和 t_j 表示第 j 个样本的输入和输出, α_i 是连接第 i 个隐层节点与输入神经元的权重向量, b_i 为第 i 个隐层节点的偏置, β_i 是第 i 个隐层节点和输出神经元的连接权值, $g(x)$ 是无线可微的激励函数.

式(4)可以表示为:

$$H\beta = T, \tag{5}$$

$$H(a_1, a_2, \dots, a_L, b_1, b_2, \dots, b_L, x_1, x_2, \dots, x_N) = \begin{bmatrix} g(a_1 x_1 + b_1) & g(a_2 x_1 + b_2) & \dots & g(a_L x_1 + b_L) \\ g(a_1 x_2 + b_1) & g(a_2 x_2 + b_2) & \dots & g(a_L x_2 + b_L) \\ \vdots & \vdots & \ddots & \vdots \\ g(a_1 x_N + b_1) & g(a_2 x_N + b_2) & \dots & g(a_L x_N + b_L) \end{bmatrix}_{N \times L}. \tag{6}$$

输出权值矩阵 β 可以表示为:

$$\begin{cases} \beta = [\beta_1, \beta_2, \dots, \beta_N]_{L \times m} \\ T = [t_1, t_2, \dots, t_N]_{N \times m} \end{cases} \tag{7}$$

2.2.2 极限学习机理论

根据黄广斌教授的理论^[10, 11], 如果训练集样本数比较大时, 隐含层神经元个数可以比样本数小, 此

时, 单隐层前馈神经网络的训练误差可以无限小,

$$\text{即: } \sum_{j=1}^N \|t_j - y_j\| < \varepsilon, \tag{8}$$

其中输出 $y_j = [y_{1j}, y_{2j}, \dots, y_{mj}]^T, j = 1, 2, \dots, n$.

极限学习机的理论思想可以总结为: 如果选择的激活函数 $g(x)$ 无限可微, 可以不用设置单隐层神经网络的输出层偏置, 同时, 训练前随机产生连接权值 a 和隐含层神经元阈值 b , 在训练过程中不用调整. 通过求解(9)式的最小二乘解可以得到输出权值矩阵 β :

$$\min \|H\beta - T\| \text{ s.t. } \min \|\beta\|, \tag{9}$$

其解为: $\beta = H^+ T$, 其中 H^+ 为 H 的 Moore - Penrose 广义逆.

标准的 ELM 算法过程为:

- (1) 设定隐含层神经元的个数 L , 给出连接权值 α 和隐含层神经元阈值 b ;
- (2) 选择合适的激活函数, 计算隐含层输出矩阵 H ;
- (3) 根据公式 $\beta = H^+ T$ 计算输出权值 β .

2.3 实验方法

本文结合 ICA 和 ELM 对 EEG 数据进行模式分类. 首先对单试次的 EEG 数据进行 ICA 分离, 实验是在 Matlab 平台下进行, 利用 eeglab 工具包下的 runica() 函数对数据进行 ICA 分解. 然后采用固定时间段样本方差最大来自动提取 N2 和 P3 成分, 样本方差如公式 $S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_j - \bar{x})^2$, 公式中: S^2 为样本方差, $x_i (i=1, 2, \dots, n)$ 为固定时间段信号的幅度, n 为固定时间段数据点数, \bar{x} 为其均值. 把提取的分量直接作为实验刺激诱发的特征进行模式分类.

对实验靶数据进行 ICA 分解, N2 成分的提取是自动查询 150~250ms 时间段的样本方差的最大的分量, P3 的提取是自动查询 250~450ms 的样本方差的最大的分量, ICA 分离可能会造成波形倒相, 所以要进行极性调整. 把提取的分量直接作为靶特征, 相对应的对非靶数据也进行上述特征提取, 提取的对应分量作为非靶特征. 对每个受试者 200 个 trial 的样本进行上述特征提取过程, 得到 200 个靶样本集和 200 个非靶样本集, 拿出一半的数据集组成训练样本, 另外一半作为测试样本.

将训练样本和测试样本导入 ELM 分类器进行训练. 根据标准的 ELM 算法过程, 首先要设定隐含层神经元个数 L , 随机产生连接权值 a 和隐含层神经元阈值. 选择合适的激活函数, 本文选择的是

“sigmoid”激活函数.因为神经元个数不同会对分类准确率有影响,所以在本文中对隐层神经元数进行寻优,经过多次训练,选择寻优空间为 $L = [100, 600]$,步长为 10.找出最佳的分类器模型,然后对测试样本分类,得到分类准确率.为了验证 ELM 的分类效果,本文利用 ICA+SVM 分类器对比.

3 实验结果及分析

本实验采集了 7 名受试者的脑电数据.首先将所有受试者的脑电数据分别进行 ICA,将经 ICA 得到的 N2 和 P3 分别作为靶特征,与非靶特征一起用 ELM 和 SVM 进行分类,分别对分类准确率和训练时间做了对比,同时还将比较最佳单通道与 ICA 两种方式下的分类准确率.本文所有程序及实验结果都是在 MATLAB R2016a 环境下运行得到.

3.1 分类准确率

表 1 记录了最佳单通道和 ICA 在 SVM 和 ELM 两种分类器的训练准确率.从表中可以看出,SVM 和 ELM 分别做最佳单通道的分类,两者的分类准确率差别很小,平均分类准确率分别为 82.4%和 83.1%.经过 ICA 之后的分类准确率相较传统的最佳单通道有较大幅度的提升,平均准确率由 82.4%和 83.1%上升到 94%和 96%以上,除了 S4 和 S7 这两名受试者原本准确率就很高,可见 ICA 的对那些准确率不高的受试者的效果非常明显.其次是在以 N2 和 P3 分别作为靶特征时的准确率比较,从图中可以看出,以 N2 成分和 P3 成分分别作为靶特征时,用 ELM 分类的准确率基本都比用 SVM 分类的准确率高,ELM 作为分类器达到了很好的效果.

表 1 分类准确率

Tab.1 The classification accuracies

受试者编号	分类准确率/%					
	单通道 SVM	单通道 ELM	ICA N2+SVM	ICA P3+SVM	ICA N2+ELM	ICA P3+ELM
S1	69.5	66.5	94.0	93.5	98.5	92.5
S2	76.0	70.5	95.5	98.5	96.5	99.5
S3	70.0	73.0	89.5	94.5	98.0	96.5
S4	96.0	96.5	97.0	94.5	96.5	97.0
S5	82.5	79.0	94.0	97.5	98.5	97.0
S6	84.0	97.0	95.0	94.5	99.0	98.5
S7	98.5	99.0	95.5	99.0	97.0	99.5
平均准确率	82.4	83.1	94.4	96.0	97.7	96.6

3.2 训练时间

本文将 ELM 和 SVM 的训练时间进行比较.在 SVM 分类器中,对惩罚因子 C 和 RBF 核函数 $\exp(-\text{gamma} * (u-v)^2)$ 的参数 gamma 通过以下的网格 $C = [2^{-5}, 2^{-4}, \dots, 2^5]$ 和 $\text{gamma} = [2^{-5}, 2^{-4}, \dots, 2^5]$ 进行训练.在 ELM 分类器中选择 sigmoid 激活函数,对隐含层神经元个数 $L = [100, 600]$ 步长为 10 进行训练.两种方法对应的训练时间如表 2 所示.

表 2 不同分类方法下的耗时

Tab.2 The training time of different method

受试者编号	分类耗时/s			
	N2+SVM	P3+SVM	N2+ELM	P3+ELM
S1	26.591	83.969	3.835	3.670
S2	26.872	73.305	3.684	3.900
S3	30.527	27.575	3.651	3.731
S4	85.243	73.431	3.989	3.893
S5	26.745	81.228	3.687	3.039
S6	27.638	82.772	3.411	3.408
S7	26.904	67.766	3.428	3.470

平均耗时 35.789 70.006 3.669 3.587

可以看出 ELM 的训练时间明显比 SVM 的训练时间少.ELM 平均训练时间只需 3.699s 和 3.587s,而 SVM 平均训练时间需要 35.789s 和 70.006s,两者差异巨大.这是因为 ELM 只需要设置网络的隐层节点个数,在算法执行过程中不需要调整网络的输入权值以及隐元的偏置,并且产生唯一的最优解,因此学习速度快;SVM 方法通过核函数把原空间映射到高维空间,使得非线性可分的问题转化为在特征空间中的线性可分的问题,升维这个过程会增加计算的复杂度,导致分类耗时增加.

4 结语

本文提出了一种将 ICA 与 ELM 结合的脑电信号特征分类方法,该方法使用 ELM 作为分类器,结合 ICA 作为特征选择算法,对 7 名受试者在“模拟阅读”范式下的电信号进行分析,同时对 ELM 隐含层的神经元个数进行寻优,以训练准确率为评价指

标,确定最佳神经元个数.为了验证该方法的有效性,与最佳单通道时域信号作为特征的分类正确率进行了比较.同时,还对SVM和ELM分类器进行了对比,以确定二者在我们的BCI应用上的效果.实验结果表明,使用ICA特征提取后,分类准确率有较大幅度的提高;使用ELM作为分类器,训练时间较SVM大幅度减少,平均准确率也有所提高.这一为构建在线BCI系统打下了良好的基础.

参 考 文 献

- [1] 官金安. 脑-机接口及其信号的单次提取[D]. 武汉:华中科技大学,2005.
- [2] 官金安,陈亚光. 脑控双页虚拟键盘的设计与性能分析[J]. 中国临床康复,2006,10(9):124-126.
- [3] 谢水清,杨 阳,杨仲乐. 脑-机接口中高性能虚拟键盘的实现[J]. 中南民族大学学报(自然科学版),2004,23(2):38-40.
- [4] 李 梅. 模拟阅读BCI信号空时特征提取与模式识别[D]. 武汉:中南民族大学,2013.
- [5] 邱静雯,官金安. 张量时频空模式在脑电信号特征提取中的研究[J]. 计算机与数字工程,2015,43(5):797-800.
- [6] 官金安,李 梅,荆汉娜等. 基于CSP的模拟阅读脑-机接口信号分类[J]. 华中科技大学学报(自然科学版),2013,41(11):123-127.
- [7] 官金安,陈亚光. 通道选择对诱发脑电单次提取精度影响的研究[J]. 计算机应用,2006,26(8):1932-1934.
- [8] Kachenoura A, Albera L, Senhadji L, et al. Ica: a potential tool for bei systems[J]. IEEE Signal Processing Magazine,2007,25(1):57-68.
- [9] Gao J, Lu L, Yang Y, et al. A novel concealed information test method based on independent component analysis and support vector machine [J]. Clinical EEG and Neuroscience,2012,43(1):54.
- [10] Wu T, Yang B, Sun H. EEG Classification Based on Artificial Neural Network in Brain Computer Interface [C]// Springer. Life System Modeling and Intelligent Computing. Berlin:Springer,2010:154-162.
- [11] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: Theory and applications [J]. Neurocomputing,2006,70(1-3):489-501.
- [12] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: a new learning scheme of feedforward neural networks [C]// IEEE. International Joint Conference on Neural Networks,2004. Proceedings. New Jersey: IEEE,2005:985-990
- [13] Huang G, Huang G B, Song S, et al. Trends in extreme learning machines: a review [J]. Neural Networks the Official Journal of the International Neural Network Society,2015,61(C):32.
- [14] Zhang W J. Feedforward Neural Networks [J]. Comprehensive Chemometrics,2009,13(4):27-31.
- [15] Zurada J M, Kang M J. Stationary points of single-layer feedback neural networks [C]// IEEE. International Symposium on Circuits and Systems,1992. ISCAS '92. Proceedings. New Jersey:IEEE,2002:57-60.
- [16] Hagan M T, Menhaj M B. Training feedforward networks with the Marquardt algorithm [J]. IEEE Transactions on Neural Networks,2002,5(6):989-993.

(责任编辑 颜昌学)