

基于数据空间自适应与共空间模式的脑电情感分类

陈景霞^{1,2}, 郑 茹¹, 张鹏伟¹, 贾小云¹

(1. 陕西科技大学 电气与信息工程学院, 西安 710021; 2. 西北工业大学 计算机学院, 西安 710129)

摘 要: 为缓解日间脑电信号波动和差异导致情感分类性能下降的问题, 在数据空间自适应(DSA)与共空间模式(CSP)迭代的基础上, 提出一种用于脑电情感分类的特征提取算法。针对 12 个受试者连续 5 d 的情感脑电信号, 采用 DSA 算法对脑电信号进行空间线性变换, 再使用 CSP 将脑电信号变换到最优子空间, 提取日间差异最小且类间差异最大的脑电功率谱密度特征及微分偏侧与差异因果特征。实验结果表明, 该算法能提高脑电信号情感分类的准确率和稳定性。

关键词: 脑电; 数据空间自适应; 共空间模式; 迭代; 情感分类

中文引用格式: 陈景霞, 郑茹, 张鹏伟, 等. 基于数据空间自适应与共空间模式的脑电情感分类[J]. 计算机工程, 2019, 45(4): 296-301.

英文引用格式: CHEN Jingxia, ZHENG Ru, ZHANG Pengwei, et al. Electroencephalogram emotion classification based on data space adaptation and common spatial pattern[J]. Computer Engineering, 2019, 45(4): 296-301.

Electroencephalogram Emotion Classification Based on Data Space Adaptation and Common Spatial Pattern

CHEN Jingxia^{1,2}, ZHENG Ru¹, ZHANG Pengwei¹, JIA Xiaoyun¹

(1. College of Electrical and Information Engineering, Shaanxi University of Science and Technology, Xi'an 710021, China;

2. College of Computer Science, Northwestern Polytechnical University, Xi'an 710129, China)

[Abstract] In order to alleviate the problem of degraded emotional classification performance caused by fluctuation and difference of day-to-day Electroencephalogram (EEG) signals, an EEG feature extraction algorithm for EEG emotion classification based on the combination of Data Space Adaptation(DSA) and Common Spatial Pattern(CSP) iteration is proposed. For 5 days of EEG data from 12 subjects, the DSA algorithm is first used to perform spatial linear transformation to minimize the difference between cross-day EEG signals. Then, the CSP algorithm is used to transform it into an optimal subspace to maximize the variance between two emotional classes. The power spectral density features and differential lateral and causal features are extracted. Experimental results show that the proposed algorithm improves the accuracy and stability of emotional classification on EEG signals.

[Key words] Electroencephalogram(EEG); Data Space Adaptation(DSA); Common Spatial Pattern(CSP); iteration; emotion classification

DOI: 10.19678/j.issn.1000-3428.0050948

0 概述

情感包括人对外界或自身刺激的心理反应, 也伴随伴随这种心理反应的生理反应, 而且涉及到人的自身需求和主观态度^[1]。近年来的研究表明, 情感的产生或活动与大脑皮层的活动有着密切关系, 这为通过脑电(Electroencephalogram, EEG)信号研究情感识别提供了依据。基于脑电的情感识别已经逐渐成为情

感研究领域的热点^[2], 对于实现高级人机交互、脑-机接口以及人工智能系统具有重要意义。

目前, 基于 EEG 的情感识别研究所面临的挑战^[3-4]之一是如何有效缓解不同受试者以及同一受试者在不同时间段诱发脑电信号之间的波动和差异对情感分类的影响。对此, 文献[5]采用自适应 Lempel-Ziv 复杂度算法提取 EEG 特征, 平均分类准确率高于传统的二值 LZC 算法和多尺度 LZC 算法。

基金项目: 国家自然科学基金青年基金(61806118)。

作者简介: 陈景霞(1979—), 女, 副教授、硕士, 主研方向为脑电信号处理、机器学习、模式识别; 郑 茹, 硕士研究生; 张鹏伟、贾小云, 副教授、硕士。

收稿日期: 2018-03-26 **修回日期:** 2018-05-10 **E-mail:** chenjx_sust@foxmail.com

文献[6]使用独立成分分析(Independent Component Analysis, ICA)等预处理步骤去除脑电图的伪迹,有效缓解部分脑电信号的日间差异。文献[7]采用 EEG 的微分偏侧(Differential Laterality, DLAT)特征与差异因果(Differential Causality, DCAU)特征,将 EEG 空间模式与隐性情感状态相关联,并通过实验证明采用跨日 EEG 数据集进行情感分类可能得到较好的分类性能。文献[8]采用鲁棒的主成分分析(Robust Principal Component Analysis, RPCA)缓解 EEG 信号的日间差异,并略微改善了情感分类的效果。文献[9]利用 RPCA 分解的稀疏信号捕获多日情感反应相对一致的 EEG 振荡,结果显示 EEG 特征数量和跨日情感分类的精度较之前的最好结果分别提高了 8% 和 6%。

数据空间自适应(Data Space Adaptation, DSA)^[10]算法的目的是实现数据源域与目标域的自适应。文献[10]采用监督和非监督的数据空间适应技术,缓解 EEG 数据间歇的非平稳性导致脑-机接口性能恶化的问题。共空间模式(Common Spatial Pattern, CSP)^[11-12]可使信号矩阵的方差最大化。文献[13]采用离散小波变换对原始脑电数据进行滤波预处理,利用样本熵算法提取 EEG 信号的非线性特征,并验证了该特征在睡眠分期分类实验中具有较高的准确率和泛化能力。文献[14]针对多类运动想象 EEG 信号在脑-机接口方面存在的分类识别率低和被试者差异性问题的,提出一种基于 ERS/ERD 现象的二级共空间模式特征提取方法。

受到上述研究的启发,针对如何缓解同一受试者日间及实验间脑电信号差异对分类性能的影响问题,本文提出一种 DSA 与 CSP 相结合的 EEG 情感特征提取算法(DSA-CSP)。该算法采用 DSA 算法缩小日间 EEG 信号的差异,使用 CSP 算法找到最优空间投影使日间 2 类情感对应 EEG 信号的差异最大,以此估计出 2 个空间滤波器。将 CSP 处理后的数据作为下一次 DSA 算法的输入进行多次迭代,从而提取日间差异最小、类间差异最大的判别性特征,同时去除任务不相关成分和噪声。

1 DSA-CSP 算法

1.1 DSA 算法

DSA 是一种典型的自适应算法,通过缓解源空间和目标空间数据分布的差异,提高分类模型的空间自适应性。当目标空间数据的联合分布 $P(\mathbf{X}, \mathbf{Y})$ 不同于源空间的联合分布 $P(\mathbf{M}, \mathbf{N})$ 时,可以通过固定 \mathbf{Y} 改变 \mathbf{X} 来调整目标空间的联合分布,使其与源空间的联合分布尽可能接近。基于该思路,假设存在一个变换函数 $g: \mathbf{X} \rightarrow \mathbf{Z}$, 将目标空间 \mathbf{X} 中的 EEG 数据 \mathbf{x} 变换到另一个空间 $\mathbf{Z}: z = g(\mathbf{x}) \in \mathbf{Z}$, 使得目标空间和源空间具有相同的联合分布,即

$P(\mathbf{Z}, \mathbf{Y}) = P(\mathbf{M}, \mathbf{N})$, 则源空间分布 $P(\mathbf{M}, \mathbf{N})$ 上的最优分类模型对于近似分布 $P(\mathbf{Y}|\mathbf{Z})$ 仍然是最佳。由此, DSA 算法提出一种线性变换函数, 定义为:

$$\mathbf{Z} = \mathbf{V}^T \mathbf{X} \quad (1)$$

其中, $\mathbf{V} \in \mathbb{R}^{n \times n}$ 表示转换矩阵, n 表示 EEG 通道的数量, \mathbf{T} 表示转置运算符。计算转换矩阵 \mathbf{V} , 使得目标空间和源空间之间的分布差异最小化。

本文假设不同天单次实验获取的 EEG 数据之间存在一定的非稳定性。为简化计算, 首先将第 1 天的单次 EEG 数据作为源空间, 将第 2 天的单次 EEG 数据作为目标空间, 通过最小化源空间与目标空间 EEG 平均分布之间的差异来计算一个线性变换矩阵。本文基于 12 个受试者连续 5 d 每天 1 次实验采集的 2 类(开心/悲伤)情感 EEG 数据进行实验。令 $\mathbf{X} = \mathbf{d}h_i \cup \mathbf{d}s_i (1 \leq i \leq 5)$ 代表某一受试者 5 d 的 EEG 情感数据, 其中, $\mathbf{d}h_i$ 代表第 i 天开心的 EEG 矩阵, $\mathbf{d}s_i$ 代表第 i 天悲伤的 EEG 矩阵, 矩阵的行表示不同的通道, 矩阵的列表示采样时间点。由于带滤波的 EEG 信号分布的均值近似为 0, 其分布的标准协方差矩阵可以估计为:

$$\boldsymbol{\Sigma} = \frac{\mathbf{x}\mathbf{x}^T}{\text{trace}(\mathbf{x}\mathbf{x}^T)} \quad (2)$$

其中, $\mathbf{x} \in \mathbf{X}$ 代表单次实验获取的 EEG 信号数据, $\text{trace}(\mathbf{x})$ 表示 \mathbf{x} 的对角线元素之和, 即 \mathbf{x} 的迹。

基于最大熵原理, 采用鲁棒的高斯函数对具有零均值和协方差矩阵的 EEG 分布进行建模, 采用 KL 距离来衡量 2 个高斯分布之间的差异^[10]。当类概率分布均匀时, 通过最小化 KL 距离可以得到最优的线性变换矩阵 \mathbf{V} 。因此, 首先利用式(2)求出 $\mathbf{d}h_1, \mathbf{d}s_1, \mathbf{d}h_2, \mathbf{d}s_2$ 在目标空间的标准协方差矩阵 $\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, \boldsymbol{\Sigma}_3, \boldsymbol{\Sigma}_4$, 然后计算:

$$\mathbf{V}_{\text{sup}}^* = \sqrt{2} (\boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_{12}^{-1} \boldsymbol{\Sigma}_2)^{-0.5} \quad (3)$$

其中, $\boldsymbol{\Sigma}_{11}^{-1}$ 是对 $\boldsymbol{\Sigma}_{11}$ 求倒数, $\boldsymbol{\Sigma}_{11}$ 是 $\mathbf{d}h_1$ 在源空间的标准协方差矩阵。为计算最优线性变换矩阵 \mathbf{V} (详细推导过程见文献[10]), 对 $\mathbf{d}h_1, \mathbf{d}s_1, \mathbf{d}h_2, \mathbf{d}s_2$ 利用式(4)求其映射 $o_{h_1}, o_{s_1}, o_{h_2}, o_{s_2}$ 矩阵:

$$\mathbf{o}_i = \mathbf{V}^T \mathbf{d}_i \quad (4)$$

其中, i 代表第 i 天, \mathbf{d}_i 代表第 i 天的开心或者悲伤的 EEG 矩阵, \mathbf{o}_i 代表 \mathbf{d}_i 的映射矩阵。

将 $o_{h_1}, o_{s_1}, o_{h_2}, o_{s_2}$ 连接起来作为新的源空间的 $\mathbf{d}h$ 和 $\mathbf{d}s$, 使用第 3 天的 $\mathbf{d}h_3, \mathbf{d}s_3$ 作为目标空间数据, 按照如上方法求第 3 天的映射, 再通过前 3 天的映射连接求第 4 天的映射。最后连接前 4 天的映射求第 5 天的映射, 直到将 5 d 的 EEG 数据转换为具有相同的分布为止。

1.2 CSP 算法

共空间模式^[15]利用线性变换将多通道脑电信号数据通过映射矩阵投影到低维子空间, 其中映射矩阵每一行由通道权值组成。这种变换可以使 2 类

信号矩阵的差异最大化。CSP算法基于2类协方差矩阵的同时对角化来实现。

若 X 代表每一天某个受试者开心的EEG二维矩阵, Y 代表每一天该受试者悲伤的EEG二维矩阵,EEG数据的归一化空间协方差矩阵可以表示为:

$$\begin{aligned} R_x &= \frac{XX^T}{\text{trace}(XX^T)} \\ R_y &= \frac{YY^T}{\text{trace}(YY^T)} \end{aligned} \quad (5)$$

其中, X^T 代表矩阵 X 的转置, $\text{trace}(X)$ 计算矩阵 X 的对角线元素的和。通过对每组的所有通道进行平均来计算平均归一化协方差矩阵 E_x 和 E_y 。复合空间协方差可以被分解为:

$$R = E_x + E_y = U_0 \Sigma U_0^T \quad (6)$$

其中, U_0 是特征向量矩阵, Σ 是特征值的对角矩阵。白化转换矩阵和转换平均协方差矩阵分别为:

$$P = \sqrt{\Sigma} U_0^T \quad (7)$$

$$S_x = PE_x P^T$$

$$S_y = PE_y P^T \quad (8)$$

其中, S_x 和 S_y 共享共同的特征向量,2个矩阵对应的特征值之和为1。

$$S_x = U \Sigma_x U^T$$

$$S_y = U \Sigma_y U^T \quad (9)$$

对于 S_x 具有最大特征值的特征向量对于 S_y 具有最小的特征值,反之亦然。将白化的EEG转化为对应于 Σ_x 和 Σ_y 中最大特征值的特征向量,使2类信号协方差矩阵的KL差异最大^[16]。映射矩阵 W 定义为:

$$W = U^T P \quad (10)$$

根据映射矩阵,原始EEG数据可按照式(11)进行转换:

$$Z_x = WX$$

$$Z_y = WY \quad (11)$$

其中, Z_x 和 Z_y 可以看作是脑电信号源的转换矩阵,包括不同任务的共同和特定的组成部分。原来的脑电信号矩阵 X 和 Y 可以通过式(12)重构:

$$X = W^{-1} Z_x$$

$$Y = W^{-1} Z_y \quad (12)$$

其中, W^{-1} 是 W 的逆阵。 W^{-1} 的列可以认为是EEG源向量分布的空间模式, W^{-1} 的第一列和最后一列分别对应一个类别的最大方差和另一个类别的最小方差的空间模式。

采用上述算法步骤,分别求出EEG开心数据 X 与EEG悲伤数据 Y 的映射矩阵,通过映射矩阵将2类EEG数据投影到带有映射矩阵的低维子空间中。

1.3 DSA-CSP算法迭代

为进一步减小脑电信号的日间差异,提取日间

更具判别性的EEG特征,本文将对上述DSA和CSP算法进行多次迭代计算。将经过一次DSA和CSP处理的EEG数据作为下一次DSA的输入,再进行CSP处理,直至迭代次数达到预设值,迭代步骤见图1。通过多次DSA-CSP算法迭代,提取出日间差异最小、2类差异最大且与情绪状态相关的脑电特征。分析比较每一次迭代后提取EEG特征的分类性能,找到其中的规律以及最优的迭代次数。

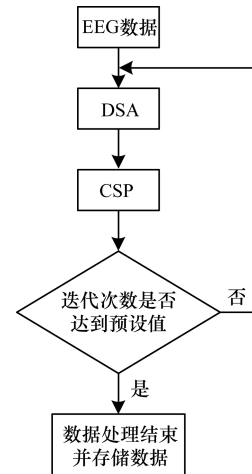


图1 DSA-CSP算法迭代过程

2 实验设计

本文实验环境为Win7 64位操作系统,Matlab R2016a软件,CPU为Core(TM)i7,主频4.2 GHz,运行内存32 GB。

2.1 实验数据集

本文采用的数据集来源于美国加利福尼亚大学圣地亚哥分校人类研究保护计划项目收集的数据集。该数据集采集了12个受试者连续5 d的脑电数据,每个受试者都参加了5 d相同的音乐诱发的脑电情感检测实验。实验采用14通道(T7和T8通道为参考电极)的Emotiv头戴便携式EEG采集设备进行EEG数据的采集,采样频率为128 Hz,带宽0.16 Hz~43 Hz。在每个实验中,受试者被要求根据其感受对每个音乐片段标注一种情感标签(高兴、悲伤或中性),共24个音乐片段,每个片段平均时长为37 s。每个受试者通过自报告的方式在每个实验最后的24 s~37 s标注情感标签。

2.2 数据预处理与特征提取

首先用1 Hz~43 Hz高通有限脉冲响应滤波器处理原始EEG信号,去除低频漂移。以128 Hz为采样频率进行EEG原始信号采样,并去除T7和T8通道的无关数据。为便于计算,从每个受试者每天的开心数据和悲伤数据中分别提取13 824个时间点,于是得到维度都是12通道乘以13 824个采样点的每个受试者每天的开心EEG数据和悲伤EEG数据,

再分别为每个通道取平均,然后采用本文提出的 DSA-CSP 迭代算法对预处理后的 EEG 数据进行线性变换与空间滤波处理。

由于功率谱密度 (Power Spectral Density, PSD) 特征和 DLAT&DCAU 特征在 EEG 情感分类上具有更好的判别性^[16],因此本文对于 DSA-CSP 滤波处理过的 EEG 数据,将每个受试者每天的开心数据与悲伤数据连接在一起,以 128 Hz 为采样频率进行 EEG 数据采样,对采样后的所有数据 (12 通道 × 1 080 个采样点) 使用与输入数据等长的矩形窗口,估算其 PSD 特征,再使用与输入数据等长的汉明窗计算其周期谱特征。由于 EEG 的 5 个频带,即 δ (1 Hz ~ 3 Hz)、 θ (4 Hz ~ 7 Hz)、 α (8 Hz ~ 13 Hz)、 β (14 Hz ~ 30 Hz) 和 γ (31 Hz ~ 43 Hz) 包含与情感相关的 EEG 特征,因此提取这 5 个 EEG 频带的平均功率。DLAT 和 DCAU 特征可以表示 EEG 波动与情感状态的联系,其原理是对每个通道上的 EEG 数据与其相邻通道上的 EEG 数据求差。因此,分别对每个 EEG 频带的 12 个通道求差提取 DLAT 和 DCAU 特征。

2.3 分类与验证

基于多日采集的带有 2 类情感标签的脑电数据集,本文使用简单机器学习算法中目前 EEG 分类性能最好的 4 种分类算法集成决策树 (Bagging Tree, BT)、线性判别分析 (Linear Discriminative Analysis, LDA)、贝叶斯线性判别分析 (Bayesian Linear Discriminative Analysis, BLDA) 和支持向量机 (Support Vector Machine, SVM)^[17],对没有使用和已使用 DSA-CSP 滤波处理 2 种情况下提取的 PSD 特征和 DLAT&DCAU 特征进行 2 类情感 (开心/悲伤) 分类实验,以验证 DSA-CSP 算法的适用性和有效性。

BT 算法是一种组合分类算法,通过降低基分类器的方差,改善泛化误差。其性能依赖于基分类器的稳定性,由于训练数据的随机波动会导致分类误差,如果基分类器不稳定,则分类误差会增大。LDA 是一种监督学习方法,其核心思想是把高维的样本映射到低维空间,若要分成 2 类则映射到一维平面,若要分成 3 类则映射到二维平面。LDA 映射的标准就是让同类的样本尽量靠近,而不同类的尽量分开。BLDA 利用已知的先验概率去推证将要发生的后验概率,就是计算每个样本的后验概率及其判错率,用最大后验概率来划分样本的类别并使得期望损失达到最小。SVM 的核心思想是找到不同类别之间的分界面,使得 2 类样本尽量落在面的两边,而且离分界面尽量远。SVM 对问题真实模型的逼近与问题真实解之间的累积误差小,并且其最终决策函数只由少数的支持向量确定,计算的复杂性取决于支持

向量的数目,而不是样本空间的维数,在某种意义上避免了维数灾难。

本文采用留一验证方法验证为期 5 d 的 EEG 情感分类的性能,选择其中 4 d 的 EEG 数据用于训练分类模型,使用余下 1 d 的 EEG 数据对训练模型进行测试。通过这种方式得到 5 组交叉验证集,如表 1 所示。

表 1 交叉验证集设置

交叉验证集	训练数据集	测试数据集
第 1 组	第 1、2、3、4 天的 EEG 数据	第 5 天的 EEG 数据
第 2 组	第 1、2、3、5 天的 EEG 数据	第 4 天的 EEG 数据
第 3 组	第 1、2、4、5 天的 EEG 数据	第 3 天的 EEG 数据
第 4 组	第 1、3、4、5 天的 EEG 数据	第 2 天的 EEG 数据
第 5 组	第 2、3、4、5 天的 EEG 数据	第 1 天的 EEG 数据

3 实验结果与分析

本文验证 DSA-CSP 迭代算法的有效性,并找出最优迭代次数。首先使用 DSA 算法对 5 d 的 EEG 数据进行空间线性变换,以最小化每个受试者日间脑电信号的不一致性和非平稳性。再应用 CSP 算法将其映射到一个最优子空间,以最大化 2 个类别之间的差异。然后将每次 CSP 映射的结果作为下一次 DSA 变换的输入,预设 5 次迭代并保存每次迭代预处理后的 EEG 数据。在此基础上,进一步提取 PSD 特征和 DLAT&DCAU 特征,然后使用 4 种分类算法对不同迭代次数处理的 EEG 数据的 PSD&DLAT&DCAU 特征进行情感分类,其精度如表 2 所示。

表 2 DSA-CSP 迭代算法分类精度

分类器	迭代次数为 1	迭代次数为 2	迭代次数为 3	迭代次数为 4	迭代次数为 5
BT	0.630 6	0.613 8	0.612 4	0.605 3	0.585 8
SVM	0.630 7	0.601 5	0.604 3	0.603 4	0.603 1
LDA	0.674 4	0.642 8	0.645 6	0.624 0	0.631 0
BLDA	0.683 4	0.661 8	0.640 8	0.618 2	0.597 0

由表 2 可知,经过 1 次 DSA-CSP 算法迭代处理后产生的 EEG 情感分类精度最高,其中最优的 BLDA 分类器精度达到 0.683 4,高于其他迭代次数的分类结果。分类精度没有随着迭代次数增加而提高的原因可能是 DSA 算法和 CSP 算法没有同步优化,导致使用 DSA 算法进行日间差异最小化变换时,削减了某些类间差异较大的判别性特征。同样地,在使用 CSP 算法进行类间差异最大化空间映射时,可能削减了某些日间差异较小的判别性特征。因此,并非迭代次数越多越好,而是要找到 2 种算法分步优化时的均衡点,才能得到较好的空间滤波效果。

本文还比较了未经过 DSA-CSP 滤波处理时 PCA 特征、PSD 特征、PSD&DLAT&DCAU 特征和 DSA&PSD&DLAT&DCAU 特征与经过 DSA-CSP 滤波处理后提取的 PSD 特征和 PSD&DLAT&DCAU 特征上 EEG 情感分类的精度,实验结果如表 3、图 2 所示。其中,PCA 也是一种常用的 EEG 特征提取算法,通过线性变换将原始 EEG 数据转化为每个维度上的一组线性无关向量,提取 EEG 数据的主要相关性特征。在图 2 中,特征 1 为 PCA,特征 2 为 PSD,特征 3 为 PSD&DLAT&DCAU,特征 4 为 DSA&PSD&DLAT&DCAU,特征 5 为 DSA-CSP(1-Iteration) & PSD,特征 6 为 DSA-CSP(1-Iteration) & PSD&DLAT&DCAU。

表 3 不同特征上 EEG 情感分类精度

特征	BT	SVM	LDA	BLDA
PCA	0.538 0	0.550 0	0.551 0	0.552 0
PSD	0.575 8	0.569 4	0.587 6	0.571 1
PSD & DLAT & DCAU	0.580 0	0.571 2	0.578 0	0.569 2
DSA & PSD & DLAT & DCAU	0.544 9	0.556 2	0.559 2	0.550 4
DSA-CSP(1-Iteration) & PSD	0.618 3	0.615 2	0.662 7	0.678 6
DSA-CSP(1-Iteration) & PSD & DLAT & DCAU	0.630 6	0.630 7	0.674 4	0.683 4

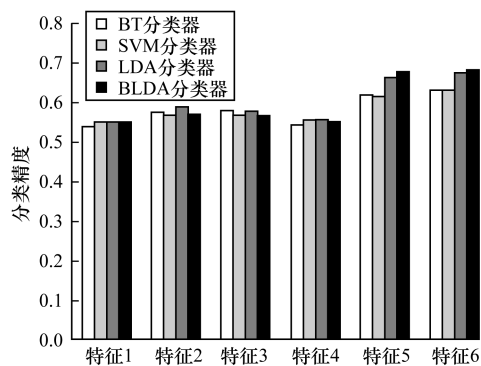


图 2 不同特征上 EEG 情感分类精度对比柱状图

实验结果显示,4 种分类器在 PSD 特征上的情感分类精度均高于 PCA 特征上的分类精度。BT 与 SVM 2 种分类器在 PSD&DLAT&DCAU 特征上的分类精度均高于 PSD 特征上的分类精度。而单纯使用 DSA 提取的 PSD&DLAT&DCAU 特征的分类精度比直接提取的 PSD 特征及 PSD&DLAT&DCAU 特征的分类精度都要低,究其原因这是由于 DSA 算法虽然减小了跨日 EEG 数据之间的差异,但同时使 2 类情感之间的差异也缩小了,因此单独使用 DSA 算法并不能有效地提高日间 EEG 情感分类的性能。CSP 算法能够提取使 2 类情感之间差异最大化的判别性特征。因此,本文提出基于 DSA-CSP 迭代的滤波算法来提取日间差异最小、类间差异最大的判别性特征进行情感分类识别。

4 种分类器在经过 DSA-CSP 算法 1 次迭代滤波后提取的 PSD 特征上的分类精度,均高于不使用

DSA-CSP 滤波处理时 4 种特征的分类精度。当使用经过 DSA-CSP 算法 1 次迭代滤波后提取的 PSD&DLAT&DCAU 特征进行分类时,其分类精度均高于前 5 种特征的分类精度,其中 BLDA 分类精度最高为 0.683 4,这一结果比单纯使用 PCA 特征的最优分类精度 0.552 0 要高 0.131 4,比单纯 PSD 特征的最优分类精度 0.587 6 高 0.095 8,比 PSD&DLAT&DCAU 特征的最优分类精度 0.580 0 高 0.103 4,比 DSA&PSD&DLAT&DCAU 特征的最优分类精度 0.559 2 高 0.124 2,比 DSA-CSP 处理后只提取 PSD 特征的最优分类精度 0.678 6 高 0.004 8。由此可见,本文提出的 DSA-CSP 迭代空间滤波算法比传统的 PCA 算法在 EEG 情感分类精度上提高了 23.8%,比直接在原始的 EEG 信号上提取的 PSD 特征的分类精度提高了 16.3%,比直接在原始 EEG 信号上提取的 PSD&DLAT&DCAU 特征的分类精度提高了 17.8%。实验结果表明,通过 DSA-CSP 算法对原始 EEG 数据进行线性变换和空间滤波,有助于提取出与 EEG 信号情感相关的判别性特征,同时缓解日间 EEG 信号的差异性和非平稳性,从而提高同一受试者日间和不同实验间的脑电信号情感分类的精度和稳定性。

4 结束语

本文提出一种基于 DSA-CSP 迭代的滤波算法用于脑电情感分类与识别。该方法运用 DSA 算法和 CSP 算法串行迭代处理 12 个受试者连续 5 d 采集的 2 类情感相关的 EEG 数据,在此基础上,提取 PSD 特征和微分偏侧与因果差异特征,缓解受试者日间及实验间 EEG 数据的非平稳性和差异性,同时提取出与 EEG 2 类情感相关的判别性特征。实验结果表明,本文算法相对于传统 PCA 算法在 EEG 情感分类精度上提高了 23.8%,相比于直接对原始 EEG 信号提取 PSD 特征和 PSD&DLAT&DCAU 特征的算法分类精度分别提高了 16.3%、17.8%。虽然本文提出的 DSA-CSP 空间滤波算法在一定程度上提高了多日 EEG 信号情感分类的精度和稳定性,但分类精度仍然不理想,下一步可将 DSP 和 CSP 统一到同个框架中进行优化,从而提高日间 EEG 情感分类性能。

参考文献

- [1] 蒋静芳,曾颖,林志敏,等. 基于脑电信号的情绪评估研究综述[J]. 信息工程大学学报,2016,17(6): 686-693.
- [2] 陆苗,邹俊忠,张见,等. 基于 IMF 能量熵的脑电情感特征提取研究[J]. 生物医学工程研究,2016,35(2): 71-74,80.
- [3] KASHIHARA K. A brain-computer interface for potential non-verbal facial communication based on EEG signals related to specific emotions[J]. European Journal of Clinical Investigation,2014,29(5):395-403.

- [4] TSIOURIS K M, MARKOULA S, KONITSIOTIS S, et al. A robust unsupervised epileptic seizure detection methodology to accelerate large EEG database evaluation[J]. *Biomedical Signal Processing and Control*, 2018, 40: 275-285.
- [5] 张栋, 陈东伟, 游雅, 等. 基于自适应 Lempel-Ziv 复杂度的情感脑电信号特征分析[J]. *计算机应用与软件*, 2014, 31(9): 162-165.
- [6] LIN Y P, JUNG T P. Exploring day-to-day variability in EEG-based emotion classification[C]//*Proceedings of IEEE International Conference on Systems, Man and Cybernetics*. Washington D. C., USA: IEEE Press, 2014: 2226-2229.
- [7] LIN Y P, HSU S H, JUNG T P. Exploring day-to-day variability in the relations between emotion and EEG signals[C]//*Proceedings of Foundations of Augmented Cognition*. Berlin, Germany: Springer, 2015: 461-469.
- [8] JAO P K, LIN Y P, YANG Y H, et al. Using robust principal component analysis to alleviate day-to-day variability in EEG based emotion classification[C]//*Proceedings of Engineering in Medicine and Biology Society*. Washington D. C., USA: IEEE Press, 2015: 570-573.
- [9] LIN Y P, JAO P K, YANG Y H. Improving cross-day EEG-based emotion classification using robust principal component analysis[J]. *Frontiers in Computational Neuroscience*, 2017, 11: 64.
- [10] ARVANEH M, GUAN C, ANG K K, et al. EEG data space adaptation to reduce intersession nonstationarity in brain-computer interface[J]. *Neural Computation*, 2013, 25(8): 2146-2171.
- [11] CHENG M M, LU Z H, WANG H X. Regularized common spatial patterns with subject-to-subject transfer of EEG signals[J]. *Cognitive Neurodynamics*, 2017, 11(2): 173-181.
- [12] SONG X M, YOON S C. Improving brain-computer interface classification using adaptive common spatial patterns[J]. *Computers in Biology and Medicine*, 2015, 61(1): 150-160.
- [13] 吕甜甜, 王心醉, 俞乾, 等. 基于脑电和肌电多特征的自动睡眠分期方法[J]. *计算机工程*, 2017, 43(10): 283-288.
- [14] 尚允坤, 段锁林, 潘礼正. 基于 ERS/ERD 的二级共空间模式的运动想象脑电信号特征提取[J]. *计算机工程与科学*, 2017, 39(7): 1385-1390.
- [15] 吴林彦, 鲁昊, 高诺, 等. 基于 CSP 算法与小波包分析方法的运动想象脑电信号特征提取性能的比较[J]. *生物医学工程研究*, 2017, 36(3): 224-228.
- [16] SAMEK W, KAWANABE M, MULLER K R. Divergence-based framework for common spatial patterns algorithms[J]. *IEEE Reviews in Biomedical Engineering*, 2014, 7: 50-72.
- [17] LIU Y, SOURINA O, NGUYEN M K. Real-time EEG-based emotion recognition and its applications[M]//GAVRILOVA M L, KENNETH C J. *Transactions on Computational Science XII*. Berlin, Germany: Springer, 2011: 256-277.

编辑 陆燕菲

(上接第 295 页)

- [20] GHANNAY S, BARRAULT L. Using hypothesis selection based features for confusion network MT system combination[C]//*Proceedings of the 3rd Workshop on Hybrid Approaches to Translation*. Stroudsburg, USA: Association for Computational Linguistics, 2014: 1-5.
- [21] RIKTERS M, SKADINA I. Combining machine translated sentence chunks from multiple MT systems[C]//*Proceedings of International Conference on Intelligent Text Processing and Computational Linguistics*. Berlin, Germany: Springer, 2016: 27-37.
- [22] YIMAM S M, ALONSO H M, RIEDL M, et al. Learning paraphrasing for multiword expressions[C]//*Proceedings of the 12th Workshop on Multiword Expressions*. Stroudsburg, USA: Association for Computational Linguistics, 2016: 1-10.
- [23] PAVLICK E, RASTOGI P, GANITKEVITCH J, et al. PPDB2.0: better paraphrase ranking, fine-grained entailment relations, word embeddings, and style classification[C]//*Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*. Stroudsburg, USA: Association for Computational Linguistics, 2015: 425-430.
- [24] DENKOWSKI M, LAVIE A. Meteor universal: language specific translation evaluation for any target language[C]//*Proceedings of the 9th Workshop on Statistical Machine Translation*. Stroudsburg, USA: Association for Computational Linguistics, 2014: 376-380.
- [25] KOEHN P, HOANG H, BIRCH A, et al. Moses: open source toolkit for statistical machine translation[C]//*Proceedings of the 45th Annual Meeting of the ACL on Interactive Poster and Demonstration Sessions*. Stroudsburg, USA: Association for Computational Linguistics, 2007: 177-180.
- [26] TURSUN E, GANGULY D, OSMAN T, et al. A semisupervised tag-transition-based markovian model for uygur morphology analysis[J]. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 2016, 16(2).
- [27] STOLCKE A. SRILM—an extensible language modeling toolkit[C]//*Proceedings of International Conference on Spoken Language Processing*. Stroudsburg, USA: Association for Computational Linguistics, 2002: 901-904.
- [28] OCH F J. Giza ++ : training of statistical translation models[EB/OL]. [2017-12-25]. <http://www.fjoch.com/GIZA++.html>.
- [29] PAPINENI K, ROUKOS S, WARD T, et al. BLEU: a method for automatic evaluation of machine translation[C]//*Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*. Stroudsburg, USA: Association for Computational Linguistics, 2002: 311-318.

编辑 赵辉