

DOI: 10.19650/j.cnki.cjsi.J2209228

基于全视域特征表征与 ELM-Adaboost 方法的情绪辨识研究^{*}

潘礼正¹, 王顺超¹, 丁 忆¹, 宋爱国², 曾 洪²

(1. 常州大学机械与轨道交通学院 常州 213164; 2. 东南大学仪器科学与工程学院 南京 210096)

摘 要:情绪在人类的行为和认知中具有不可或缺的地位,开展情绪辨识研究具有非常重要的实际意义。为提高四类情绪多被试交叉辨识的准确性,提出了基于全视域特征表征与 ELM-Adaboost 的情绪辨识方法。首先,提出了基于融合信息的数据处理策略,交叉融合的多类生理信号构成的样本数据,有助于从全视域角度提取样本特征。其次,采用最大相关性最小冗余度的特征选择方法对融合后的特征进行选择,进而实现有效的情绪状态表征。最后,通过 ELM-Adaboost 方法构建的强分类器实现情绪状态辨识。多角度的实验分析研究充分证明了所提方法的有效性,四类情绪多被试交叉辨识准确率达到了 83.06%。

关键词:情绪辨识;特征选择;极限学习机;分类器强化

中图分类号: TP391 TH7 **文献标识码:** A **国家标准学科分类代码:** 510.40

Emotion recognition research based on the full-view feature representation and ELM-Adaboost

Pan Lizheng¹, Wang Shunchao¹, Ding Yi¹, Song Aiguo², Zeng Hong²

(1. School of Mechanical Engineering and Rail Transit, Changzhou University, Changzhou 213164, China;

2. School of Instrument Science and Engineering, Southeast University, Nanjing 210096, China)

Abstract: Emotion plays an indispensable role in human behavior and cognition. It is of great practical significance to carry out research on emotion recognition. To improve the accuracy of cross-subject identification of four kinds of emotional states, the feature representation method based on the full-view and ELM-Adaboost method is proposed. Firstly, a data processing strategy based on the fused information is proposed. The sample data consisted of multiple types of physiological signals are cross-fused to help extract sample features from the perspective of full view. Secondly, the feature selection method with the maximum correlation and the minimum redundancy is used to select the fused features to achieve effective emotional state representation. Finally, the strong classifier constructed by the ELM-Adaboost method is employed to realize emotion recognition. Experimental results from multi-perspective evaluate the effectiveness of the proposed method. The identification accuracy of cross-subject with four types of emotional states reaches 83.06%.

Keywords: emotion recognition; feature selection; extreme learning machine; classifier reinforcement

0 引 言

情绪在日常生活中扮演着重要角色,已经深深地渗透在人类的行为和认知中。情绪不仅可以由外在身体的感知或周边环境渲染所引起,也可以由内在的推论、记忆、信念或者遐想所触发^[1]。近年来,对情绪的探讨和应

用引起了学者广泛的兴趣和关注。研究领域已经覆盖在康复医疗、智能驾驶、脑机接口等^[2]。

目前,研究者已通过多样的途径开展了情绪辨识的相关研究。外在在信号源的选择中,常见的有语音、姿势、面部表情等。Kacem 等^[3]提出了一种基于时空几何表征和推导工具的情绪状态辨识方法,通过对人体部位的动力学分析来实现情绪状态识别。Su 等^[4]利用面部

收稿日期:2022-01-20 Received Date: 2022-01-20

^{*} 基金项目:国家自然科学基金(61773078, 62173089)、江苏省研究生科研与实践创新计划(KYCX20_2533)项目资助

表情和言语反应特征,通过细胞耦合的 L-skip 融合机制的长短时记忆网络来表示 6 类具体情绪状态,获得了 76.9% 的最佳准确率。然而,情绪作为一种内在的感知,一些细微的变化难以在表面察觉,因此使用外在的信号源进行情绪状态辨识存在一定弊端。相反,利用内在的信号源,使用生理信号开展情绪状态辨识能够起到很好的效果,其中脑电信号一直是众多学者研究的焦点之一。Gannouni 等^[5]提出了基于零时间窗口的纪元估计和相关电极识别的方法,实现基于脑电分析的情绪识别。Huang 等^[6]提出了用于脑电情绪识别的双半球差异卷积神经网络模型,该模型可以有效地学习大脑左右半球的不同反应模式并设计成具有 3 个卷积神经网络层的三输入单输出网络结构,基于主体依赖的实验结果在效价和唤醒度的二分类情绪辨识准确率分别为 94.38% 和 94.72%。脑电信号复杂性、易受干扰、穿戴不便等问题是制约辨识效率的潜在因素。最近,一些学者探索利用其他外围生理信号描述情绪状态,取得了很好表征效果。Ganapathy 等^[7]使用皮肤电活动信号与多尺度卷积神经网络来实现效价和唤醒状态的识别。Santamaria-Granados 等^[8]采用深度卷积神经网络结合心电和皮电从唤醒和效价角度进行情绪检测,辨识准确率分别为 76% 和 75%。相比脑电,其他外围的生理信号复杂度较低,采集和穿戴也较为便捷。因此,应用其他外围的生理信号用于表征情绪,同样是一条有效途径。

数据处理是情绪辨识中重要的工作之一,包括对数据降噪、样本划分、特征提取等。样本划分大多依据经验,但是样本划分的方式对情绪辨识具有一定的影响。特征作为样本的一种属性,是其突出性质的表现。一个有效的特征表征对于模型的构建、运行学习算法效率的提升具有事半功倍的效果。实际应用中采集的信号往往较杂乱且非线性,对于微弱的、易受干扰的生理信号,提取单一的特征是难以全面充分地表征样本信息。因而融合不同类型、不同维度的特征具有更好的情绪表征效果^[9]。特征的融合或者信号通道的融合通常会导致特征维度的升高,从而使得计算量增加,模型结构趋向复杂化。想要获得更佳的情绪辨识效果,特征选择是不可或缺的一个重要环节。目前已有较多特征选择方法被应用于特征评估中,例如成对样本 T 检验、Fisher 距离、ReliefF、FCBF 等。这些方法中有的只适用于判别线性特征,有的只考虑特征与目标类别的相关性而忽视了特征间的冗余性。因此,采用有效的特征选择方法较为关键,既要考虑特征与类别相关问题,又要考虑特征之间冗余问题,其目的是为了减少特征间不必要的噪声,降低模型复杂度。

辨识分类是情绪识别中另一个至关重要的环节。目前,被广泛应用的分类器有决策树、支持向量机、随机森

林、神经网络、深度学习等。如何构建一个强有力的分类器一直是众多学者追求的目标。有的学者致力于在单分类器上进行传统的参数寻优,然而一个最优的单分类器在复杂的分类环境下,应用场合是相对有限的。因此,有的学者则将不同作用的分类器进行融合进而发挥各自性能的优势。在前期研究中,本团队曾提出一种融合非线性特征和团队协作辨识的方法^[10],用于提高最终情绪状态辨识的决策效率。Chen 等^[11]提出一种基于模糊积分的多分类器融合方法,可以融合各个 LIBSVM 分类器在分类中得到的结果,以此反映融合过程中每个 LIBSVM 分类器的重要性,从而改善辨识的效率。自适应增强算法是近年来比较流行的一种强化方法,目的是为了强化多个弱分类器的最终分类性能。Li 等^[12]提出一种基于支持向量机 (support vector machine, SVM) 的自适应增强 (adaptive boosting, Adaboost) 算法,通过设定 SVM 的惩罚参数 c 构建不同的分类器以此组建强分类器。相同的方法被 Wan 等^[13]采用,用于提高铣削过程中颤振分类的精度。但是自适应增强算法只适用于二分类,使得应用的范围受限。另外,为了得到更有效的样本训练,如何使训练样本的分布更加差异化也是值得探讨的问题。

综上所述,针对目前情绪辨识研究的现状,本文利用外围的生理信号开展基于全视域特征表征与 ELM-Adaboost 的情绪辨识方法研究,采用基于融合信息的数据处理策略,使样本包含的信息更丰富,有助于提取更具全视域性的特征。为更有效地进行情绪状态表征,运用最大相关性最小冗余度的特征选择方法对融合后的特征进行选择。采用自适应增强算法作为评估手段,训练多个极限学习机弱分类器并整合为一个强二分类器以此强化情绪辨识性能;并融合多组强二分类器通过投票原则实现四类情绪状态的辨识。

1 情绪辨识方法

1.1 整体方法框架

为提高情绪辨识精度,本文提出了一种基于全视域特征表征及 ELM-Adaboost 的情绪辨识方法。如图 1 所示,首先针对多路生理信号,采用全视域特征表征方法,利用融合的数据处理策略对数据进行交叉融合划分和特征提取。随后应用 mRMR 最大相关-最小冗余算法进行特征选择和优化。最后采用 ELM 作为基分类器,Adaboost 算法作为评估手段构建 ELM-Adaboost 强二分类器模型,通过投票原则实现多类的情绪辨识。

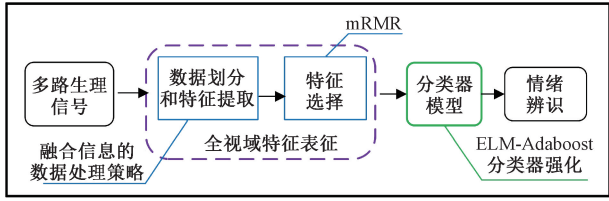


图1 整体方法原理示意图

Fig. 1 Schematic diagram of overall method principle

1.2 基于融合信息的全视域特征表征

1) 融合信息处理策略

本文采用预处理后的 DEAP 数据集,采样频率已由 512 Hz 下采样至 128 Hz,滤波并去除了眼电伪迹。在本研究中,有 6 路外围的生理信号被选用,分别为血压 (blood volume by plethysmograph, BVP)、皮肤电 (galvanic skin response, GSR)、呼吸 (respiration amplitude, RSP)、肌电 (electromyograms of zygomaticus, zEMG) 以及两路在不同位置上采集的眼电 (electrooculograms of horizontal and vertical, hEOG, vEOG)。

常规的数据样本划分方法如图 2 所示,采集一路去除了 3 s 基线后生理信号的试验数据,采样频率 128 Hz, 1 min 数据总共 7 680 个数据点 (128×60)。每个样本被划分成 6 s 一段并伴有 3 s 的滑窗,包含 768 个数据点 (128×6),一次试验的每路信号被划分成 19 个样本。对每路样本提取特征,最终 6 路信号进行特征融合用于全面表征同类型的试验情绪状态。这种常规的样本数据处理方式,数据信息维度较单一,不利于样本特征的表征。为了提取更具有全视域性能的特征,本文提出一种基于融合信息的数据处理策略,如图 3 所示。一次试验在去除 3 s 基线后,6 路信号依次循环交叉取点被整合为一个融合的信号段。融合的信号段被等分 12 份,每等份是基于 5 s 全视域信息的数据源 (3 840 数据点)。每个数据源再被等分为 6 段,每段 (640 数据点) 同样划分 19 个样本,每个样本的数据长度为 64 个点并伴有 32 个点的滑窗。这样处理划分的样本融有多路生理信号,所包含信息更丰富,更能多维度展现所在情绪状态时的生理信息。对每段的样本提取特征,6 段样本特征融合用于全方位地表征 5 s 全视域的情绪状态。

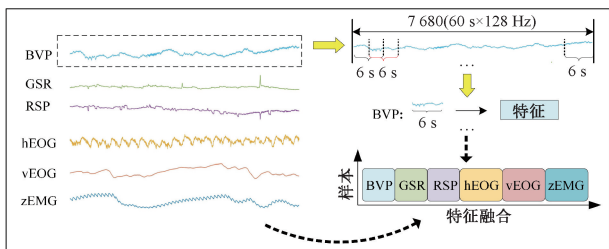


图2 常规的数据划分

Fig. 2 The routine data partitioning

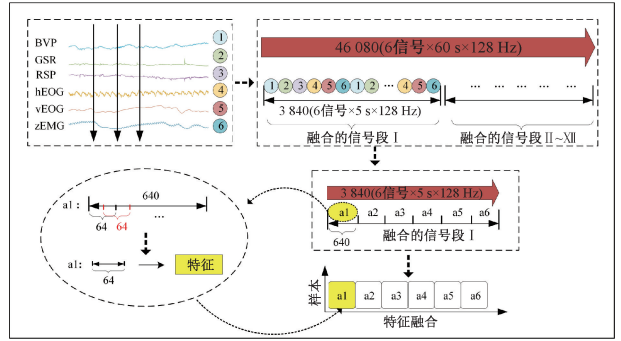


图3 基于融合信息的数据处理策略

Fig. 3 Data processing strategy based on fused information

2) mRMR 特征选择

本文提取了常规的时域特征和熵特征,包括最大值、最小值、平均值、中值、范围、方差、标准差、调和平均、众数、斜度、近似熵、样本熵、模糊熵、小波包熵以及排列熵。另外,从不同的时域和频域维度对信号样本进行预处理,再同样提取以上特征,可以增大对样本额外的表征力度。对信号向量的预处理方式包括:离散傅立叶变换和一阶差分,相关公式如式 (1)、(2) 所示。一阶差分用于信号的平滑处理,反映信号是否平稳及奇异点的检测。离散傅立叶变换用于反映信号的频域特性。

$$\begin{cases} Y(k) = \sum_{j=1}^n X(j) W_n^{(j-1)(k-1)} \\ X(j) = \frac{1}{n} \sum_{k=1}^n Y(k) W_n^{-(j-1)(k-1)} \\ W_n = e^{(-2\pi i)/n} \end{cases} \quad (1)$$

式中:对于长度为 n 的向量 X 和 Y , Y 和 X 分别实现向量的离散傅里叶变换和离散傅里叶逆变换,其中 W_n 为 n 次单位根之一。

$$1\text{diff} = [X_2 - X_1, X_3 - X_2, \dots, X_n - X_{n-1}] \quad (2)$$

式中:1diff 返回长度为 $n-1$ 的向量。1diff 的元素是 X 相邻元素之间的差分。

本文采用最大相关-最小冗余 (max-relevance and min-redundancy, mRMR) 算法用于筛选最优的特征维度。mRMR 是 Peng 等^[14]提出的一种特征选择算法,其目的是在原始特征集中找到与类别标签相关性最大,且特征彼此之间相关性最小的一组特征。mRMR 算法速度快,估计结果鲁棒性较好,是所有互信息的一阶最优估计。mRMR 原理定义如下:

(1) 特征集 S 与类别 c 的相关性 $D(S, c)$ 由各个特征 f_i 和类 c 之间的所有互信息值的平均值定义:

$$D(S, c) = \frac{1}{|S|} \sum_{f_i \in S} I(f_i, c) \quad (3)$$

$$I(f_i, c) = \iint p(f_i, c) \log \frac{p(f_i, c)}{p(f_i)p(c)} dx dy \quad (4)$$

其中, $p(f_i)$, $p(c)$, $p(f_i, c)$ 是变量 f_i 和 c 的概率密度函数。

(2) 特征集 S 中所有特征的冗余性是由特征 f_i 和特征 f_j 之间的所有互信息值的平均值定义:

$$R(S) = \frac{1}{|S|^2 \sum_{f_i, f_j \in S} I(f_i, f_j)} \quad (5)$$

(3) 权衡最大相关性和最小冗余性, mRMR 的定义如下:

$$\text{mRMR} = \max_S [D(S, c) - R(S)] \quad (6)$$

使用 mRMR 进行特征选择后, 每个特征会得到一个重要性排名, 结合交叉验证来选择结果性能最好的特征子集。即根据特征重要性排名对特征子集从 $\text{top1} \sim \text{topn}$ 一个个进行交叉验证测试, 然后选择结果最好的一组特征集即可。

1.3 基于 ELM-Adaboost 的分类器自适应增强算法

极限学习机 (extreme learning machine, ELM) 是一种单隐层前馈型神经网络^[15]。ELM 的结构简单, 由输入层、隐含层和输出层构成, 如图 4 所示。与传统的神经网络相比, 它克服了传统算法 (如 BP 神经网络) 局部最小化及过度拟合的问题。另外, ELM 的隐层参数 (输入权值和隐层偏置) 随机产生, 且无需迭代更新, 模型训练速度快, 适用于维度较高数据样本量大的场合。其算法原理如下:

对于 N 个随机样本 $\{(x_i, y_i) \mid x_i \in \mathbf{R}^n, y_i \in \mathbf{R}^m, i = 1, 2, \dots, N\}$, 假定隐含层节点个数为 L , 输出函数可表示为:

$$\sum_{j=1}^L \beta_j g(\mathbf{w}_j \cdot \mathbf{x}_i + \mathbf{b}_j) = y_i, \quad i = 1, 2, \dots, m \quad (7)$$

式中: β_j 为第 j 个隐层节点的输出权值, \mathbf{w}_j 和 \mathbf{b}_j 分别表示为第 j 个隐层节点的输入权值和偏置, $g(\cdot)$ 表示激励函数, $\mathbf{w}_j \cdot \mathbf{x}_i$ 表示 \mathbf{w}_j 与 \mathbf{x}_i 的内积。

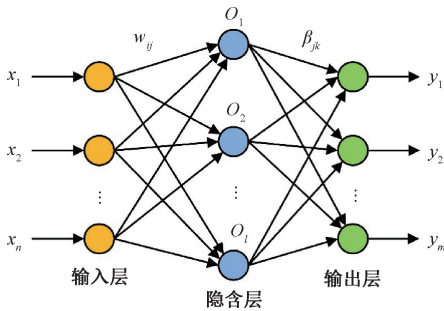


图 4 极限学习机结构

Fig. 4 The structure of ELM

对于隐含层来说, 节点数应设置在合理的范围内, 节点过多或过少一定程度上会影响分类性能。在隐含层节点数选择上, 结合参考文献^[16-17]采用经验公式:

$$X_1 = \sqrt{N(C+P)} + Q \quad (8)$$

式中: N 为样本数量, C 为分类类别数, P 为特征维度, Q 为常数 ($0 \sim 1\,000$, 默认为 0)。通过式 (8) 可以初步得出较为适宜的节点, 本文为符合实际应用的需求, 需进一步地进行节点数寻优。寻优的节点范围设置在式 (8) 所得节点的 $\pm(30 \sim 40)\%$, 并依据范围大小设置合适的步长。

自适应增强算法 (Adaboost) 是由 Freund 等^[18]提出的一种重要的集成学习技术。它的自适应体现在被前一个基分类器错误分类的样本, 其权值会增大, 而正确分类的样本的权值会减小, 然后更新权值的全体训练样本被再次用来训练下一个基分类器。在每轮迭代中, 加入一个新的基分类器并依据训练结果对弱分类器作出评估。通过设置分类器的迭代次数或预定的足够小的错误率, 最终整合为一个强分类器。

本文采用 ELM 作为基分类器, 整合 Adaboost 的集成学习技术, 构成基于极限学习机的自适应增强算法 (ELM-Adaboost)。同时, 打破 Adaboost 算法只能应用于二分类的局限性, 本文采用一对一方法, 在任意两类样本之间设计一个强二分类器, 采用投票原则最终实现四分情绪辨识问题。如图 5 所示, ELM-Adaboost 算法的步骤如下:

1) 将四分类数据随机划分为训练集和测试集, 80% 用于训练, 20% 用于测试。再将四分类的训练集细分为二分类的数据集, 两两任意组合共有 6 种类型的二分类数据集, 分别为 $\{(1\&2), (1\&3), (1\&4), (2\&3), (2\&4), (3\&4)\}$ 。

2) 每种二分类数据集构建该类型的强分类器。以 1&2 数据集为例, 将数据集进一步划分为自训练集和自测试集, 训练测试比为 8:2。设置构建弱分类器的迭代次数 T ($T=10$), 对二分类数据集的所有样本赋予初值权重 D_i , 定义如下:

$$D_i(i) = 1/m \quad (i = 1, 2, \dots, m) \quad (9)$$

其中, t 表示当前迭代次数, m 是该二分类数据集的总样本数。

3) 将自训练集输入 ELM 分类器进行训练并寻优最佳节点, 参照式 (8) 节点范围选择在 $[800:1\,600]$, 步长为 20。随后, 将当前训练好的模型用自测试集进行测试。

4) 统计当前弱分类器自训练和自测试的总错误率 e_t , 并计算当前弱分类器的权重 a_t 。

$$e_t = \sum_i D_t(i), (g_t \neq y) \quad (10)$$

$$a_t = \frac{1}{2} \ln\left(\frac{1 - e_t}{e_t}\right) \quad (11)$$

其中, g_t 是预测结果, y 是预期结果。

5) 调整下一轮二分类数据集全体样本的权重 D_{t+1} , 并依据权重由大到小排序, 重新按比例挑选自训练和

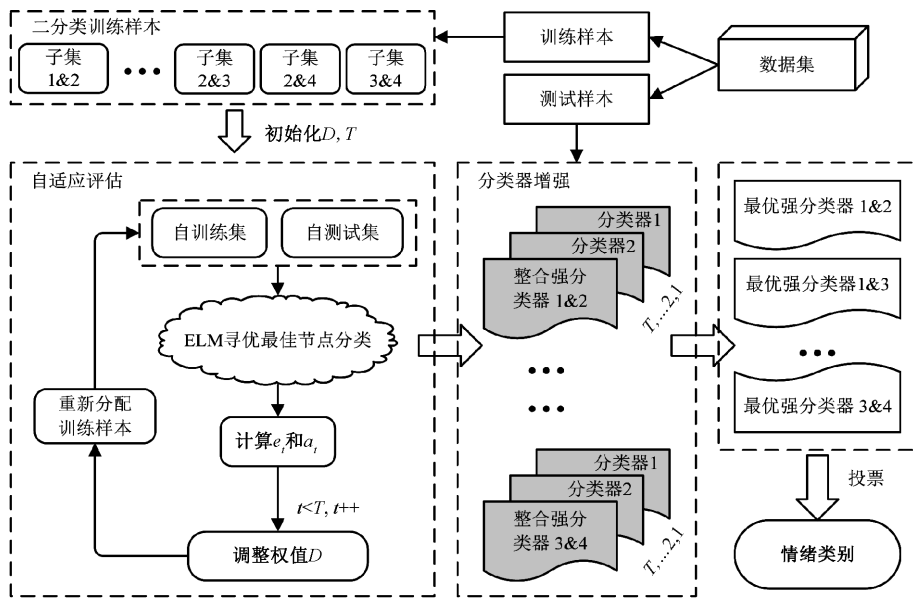


图 5 基于极限学习机的自适应增强算法

Fig. 5 The flow chart of ELM-Adaboost

自测试集。权重大的,即易分错的样本将优先被挑选出作为自训练集,剩余作为自测试集。

$$D_{t+1}(i) = \frac{D_t(i)}{B_t} \exp[-a_t y_i g_t(x_i)] \quad (12)$$

其中, B_t 是归一化因子,目的是在权重比例不变的情况下使分布权值和为 1。

6) 重复步骤 3)~5), 迭代至最后一个弱分类器。整合所有弱分类器, 构建一个强分类器, 输入测试集。训练 T 轮后得到 T 组弱分类器函数 $f(g_t, a_t)$, 由 T 组弱分类函数 $f(g_t, a_t)$ 组合得到强分类函数 $h(x)$ 。

$$h(x) = \text{sign} \left[\sum_{i=1}^T a_i \times f(g_i, a_i) \right] \quad (13)$$

7) 选择出各二分类中最优的强分类器并整合作为一个整体, 测试样本分别进入 6 类强分类器后依据投票原则最终实现四分类的情绪辨识。

2 实验分析

2.1 实验数据

本文采用国际公开的生理信号情绪数据集(a database for emotion analysis using physiological signals, DEAP)作为情绪辨识模型的数据来源^[19]。在 DEAP 数据集中, 参与试验的被试共有 32 名, 其中男性和女性的比例各占 1/2, 年龄范围在 19~37 岁之间。实验时, 每个被试被要求观看一段长达 1 min 的音乐视频剪辑, 每段视频都是经过精心筛选以此确保被试的情绪能够被合理的诱发。同时被试需要在实验后, 对视频在唤醒度和效价等情绪维度进行适当

地自我评估。另外, 在实验期间, 每位被试身上共有 40 路信号被同时采集, 分别包括 32 路脑电 (electroencephalogram, EEG) 和 8 路其它的外围生理信号, 信号采集的原始频率为 512 Hz。每人分别进行了 40 次实验, 在每次实验开始时都会有 3 s 的基线时间做准备。

2.2 情绪类别划分

如何恰当地对情绪进行定性分析是研究情绪识别的基础。目前, 有两种主流的评估方法。一种是离散的情绪表达, Ekman 等^[20]提出了 6 种基本情绪, 即快乐、悲伤、惊讶、厌恶、愤怒和恐惧。目前, 情绪的研究被划分越来越细, 虽然比较直观, 但是情绪是一种复杂的产物, 划分的越细, 对应诱发的场景及要求也越高。另一种是连续的情绪表达, Russell^[21]采用效价 (valence) 和唤醒度 (arousal) 的二维平面来泛化对情绪的描述, 使其约束于一定空间以便进行较为合理的情绪分析。

本文采用 Russell 的情绪维度评估理论将情绪定性地表征在效价和唤醒度所组成的二维空间。如图 6 所示, 由效价和唤醒度构成二维平面, 其刻度范围都是 [1, 9]。效价反映愉悦的程度, 表示从不开心到开心; 唤醒度反映兴奋的程度, 表示从不激动到激动, 刻度越小意味着程度越弱。参考同类研究方法, 本文把刻度大于 5 定义为高, 小于 5 定义为低, 因而, 效价-唤醒度平面被等分为 4 个象限 (四分类), 即高唤醒度高效价 (high-arousal high-valence, HAHV)/低唤醒度高效价 (low-arousal high-valence, LAHV)/低唤醒度低效价 (low-arousal low-valence, LALV)/高唤醒度低效价 (high-arousal low-valence, HALV)。

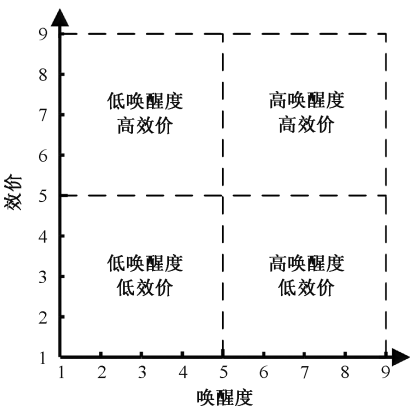


图 6 情绪维度评估

Fig. 6 The assessment of emotion

2.3 单被试情绪识别分析

针对每个单被试的四分类情绪识别,采用融合信息的数据处理策略对其试验数据进行处理划分。每个被试包含 40 个试验,每个试验基于 5 s 全视域被划分 19 个样本,单人总共 760 个样本,训练和测试比为 8:2。对每类信号段的每个样本提取共 45 维特征,将 6 段信号的特征进行融合以便多角度的描述同一状态下的情绪特性,特征总维度为 270 维。分类器选用传统的极限学习机,依据经验公式节点范围选择在[900,2 100],步长为 50。将常规的数据处理方式和融合信息的数据处理策略在 32 位被试上分别进行应用比较,实验选取 5 倍交叉验证并取平均结果作为最终的四分类情绪辨识结果,实验结果如图 7 所示。

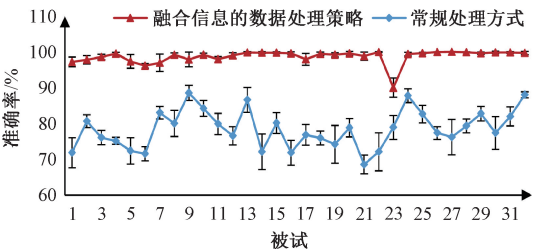


图 7 融合信息的数据处理策略和常规处理方式比较

Fig. 7 Comparison of the routine method and data processing strategy based on fused information

图 7 中就每位被试四类情绪辨识精度而言,本文所提的融合信息的数据处理策略皆显著地优于常规的数据处理方式。常规的数据处理方式不仅因个人的差异性而波动较大,而且整体的稳定性也较差。融合信息的数据处理策略整体精度高且相对稳定。在融合信息的数据处理策略中虽然第 23 个被试出现一定的波动,但同比结果仍好于常规的数据处理方式。如表 1 所示,在常规的数据处理方式中前 22 人的单人辨识整体平均水平要低于

后 10 人的整体平均水平。研究分析表明引起这种波动的主要原因在于试验(s01-s22、s23-s32)是在两个不同国家的地区进行采集。不同国家的人文、环境、文化差异甚至同一型号设备的细微干扰都有可能影响情绪辨识实验的效果。但用融合信息的数据处理策略,该类影响较小。

表 1 单被试的整体平均辨识度

Table 1 Overall average recognition of one-subject experiments

| 方法 | 实验被试 | | |
|----------------|------------|------------|------------|
| | s01-s22 | s23-s32 | s01-s32 |
| 常规 (Acc %) | 77.14±2.97 | 80.41±2.90 | 78.42±2.87 |
| 本文融合信息 (Acc %) | 98.66±0.86 | 98.90±0.46 | 98.70±0.75 |

从表 1 可看出,在 32 个单被试的整体情绪辨识中,使用融合信息的数据处理策略可以获得 98.70% 的辨识精度,相比于常规处理方式要高 20.28%,表明融合信息的数据处理策略更有助于从全视域的角度提取有效的特征,从而改善情绪辨识准确性。

2.4 多被试交叉情绪识别分析

本文选用 32 人全体数据进行四分类的多被试交叉情绪识别研究。相比于基于单被试的情绪辨识,多被试交叉情绪辨识计算更为复杂,模型辨识难度大,然而构建一个具有较好泛化能力的多被试交叉的情绪识别模型更具有实用性和普适性意义。

在确定数据处理策略和所选特征维度后,针对由于特征维度偏高而导致模型结构复杂、训练时间较长、辨识精度低的情况,本文采用 mRMR 特征选择方法对特征的相关性与冗余性进行分析,进而筛选出最优的特征维度来实现情绪状态表征。

通过使用 mRMR 进行特征选择,每个特征会得到一个重要性排名,选用 5 倍交叉验证的方法来选择结果性能最好的特征子集。总特征维数为 270 维,特征选择范围在[30,270],维度搜索梯度下降步长为 30,以此初步分析选择较优的特征子集范围,如图 8 所示。

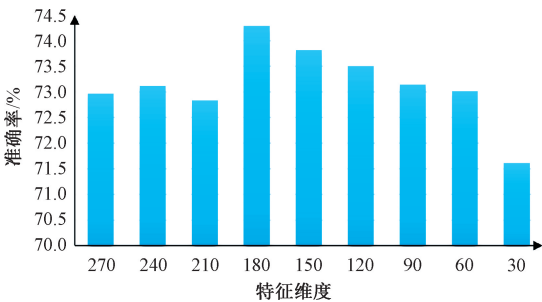


图 8 mRMR 粗选特征维度

Fig. 8 Roughly selecting the feature dimension with mRMR

依据图 8,较优的特征维度选择范围可以被粗略地确定在 $[120,180]$ 维。在该维度范围下进一步精选特征维度,维度搜索梯度下降步长为 5,结果如图 9 所示。

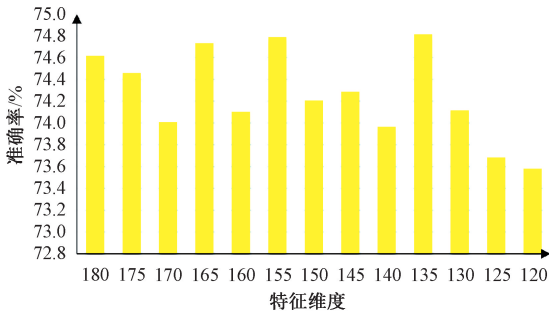
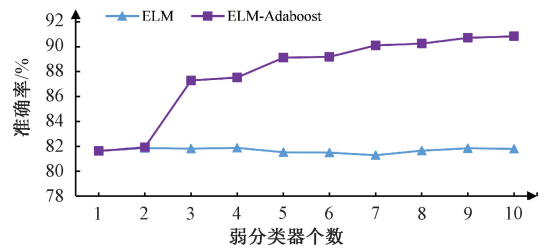


图 9 mRMR 精选最优特征维度

Fig. 9 Selecting the optimal feature dimension with mRMR

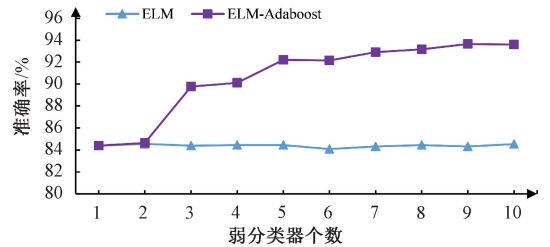
由图 9 可见,最优维度的特征子集最终确定在 135 维,该维度下的特征相比于原始的特征维度实现了较大幅度的降低,减少了模型训练量,且从特征子集的辨识精度来看,该维度下特征子集情绪辨识效果最好。

本文采用 mRMR 特征选择方法来优化特征表征,结合基于 ELM-Adaboost 的强化分类器方法,从而改善多被试交叉的四分类情绪辨识准确性。将四分类的数据,两两组合拆分为二分类的数据集,共计分为 6 类情况,分别是 1&2、1&3、1&4、2&3、2&4、3&4,弱分类器的迭代次数 T 选为 10。对每类的二分类数据集,分别进行 ELM-Adaboost 和 ELM 的分类结果比较,如图 10 所示。在图 10 中,ELM 的横坐标表示第 t 个弱分类器所对应的精度,而 ELM-Adaboost 的横坐标表示整合前 t 个弱分类器后所得的强化后的分类精度。不难发现,所提的



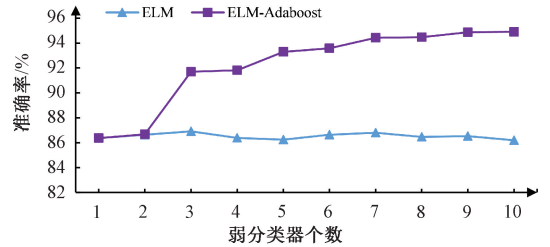
(c) 基于数据集1&4的辨识结果比较

(c) Comparison of identification results based on datasets 1&4



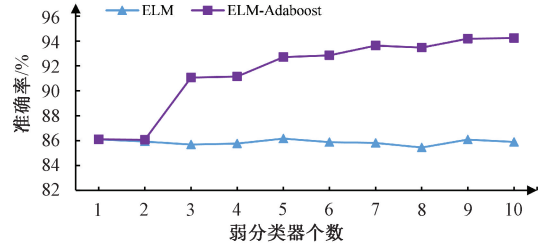
(d) 基于数据集2&3的辨识结果比较

(d) Comparison of identification results based on datasets 2&3



(e) 基于数据集2&4的辨识结果比较

(e) Comparison of identification results based on datasets 2&4



(f) 基于数据集3&4的辨识结果比较

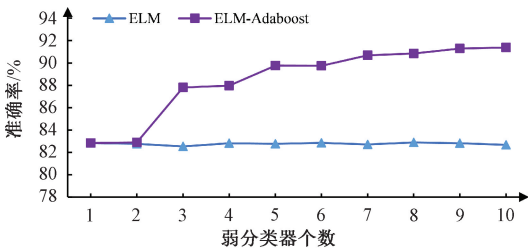
(f) Comparison of identification results based on datasets 3&4

图 10 不同弱分类器数的 ELM-Adaboost 和 ELM 的辨识精度比较

Fig. 10 Accuracy comparison between ELM-Adaboost with different number of weak classifiers and ELM

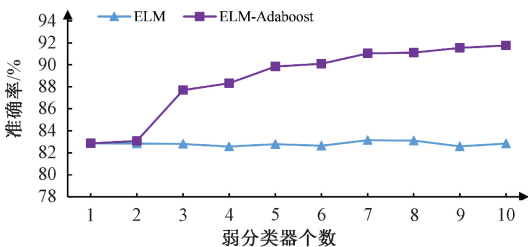
ELM-Adaboost 分类方法,当组合弱分类器数低于 3 时,所得结果和 ELM 差异不大;然而组合弱分类器数大于 3 时,强化后的分类结果得到明显的提升并逐渐趋于平稳。

挑出 ELM-Adaboost 方法下最佳迭代的强二分类器作为最终四类情绪辨识决策的最优二分类器。如表 2 所示,取迭代次数大于 3 的所有强分类器的整体平均辨识结果与 10 次传统的 ELM 分类器的整体平均结果进行比较,观测强化后的分类器对整体辨识性能提升的影响。



(a) 基于数据集1&2的辨识结果比较

(a) Comparison of identification results based on datasets 1&2



(b) 基于数据集1&3的辨识结果比较

(b) Comparison of identification results based on datasets 1&3

表 2 中 Acc 表示准确率 (accuracy), Pe 表示精度 (precision), Re 表示召回率 (recall)。

表 2 ELM 与 ELM-Adaboost 整体辨识性能比较
Table 2 Comparison of overall identification performances between ELM and ELM-Adaboost

| 类别 | ELM | | | | ELM-Adabost | | | |
|-----|------|------|------|------|-------------|------|------|------|
| | Acc | Pe | Re | F1 | Acc | Pe | Re | F1 |
| 1&2 | 0.83 | 0.79 | 0.73 | 0.76 | 0.91 | 0.92 | 0.83 | 0.87 |
| 1&3 | 0.81 | 0.76 | 0.71 | 0.73 | 0.91 | 0.91 | 0.84 | 0.87 |
| 1&4 | 0.80 | 0.77 | 0.72 | 0.74 | 0.91 | 0.91 | 0.84 | 0.87 |
| 2&3 | 0.87 | 0.82 | 0.82 | 0.82 | 0.93 | 1.00 | 0.93 | 0.97 |
| 2&4 | 0.85 | 0.86 | 0.86 | 0.86 | 0.94 | 0.95 | 0.94 | 0.94 |
| 3&4 | 0.85 | 0.85 | 0.86 | 0.86 | 0.94 | 1.00 | 0.94 | 0.97 |

从表 2 可以看出,每类 ELM-Adaboost 依次迭代所得的强二分类器整体平均辨识结果明显好于传统的 ELM 分类器的整体平均精度。每类强二分类器获得的平均准确率增益效果依次为 8% (1&2)、10% (1&3)、11% (1&4)、6% (2&3)、9% (2&4)、9% (3&4)。

通过整合 6 类最优的强二分类器模型,依据投票决策原则,实现四分类的情绪辨识。多被试交叉四分类辨识精度的比较如表 3 所示。

在表 3 中,所提的融合 6 类最优的二分类 ELM-Adaboost 模型针对四分类交叉情绪辨识具有很好的辨识性能,辨识精度达到了 83.06%;相比 mRMR+ELM 的方法提升了 8.25%。

表 4 本文方法与其它研究工作的比较
Table 4 Comparison between the proposed method and other works

| 研究 | 被试 | 信号 | 情绪类别 | 方法 | 准确率/% |
|-------------------|----|---------------------------------|------------------------------|-------------------|----------|
| 文献[22] (2019) | 32 | EDA | 4 (HAHV, LAHV, LALV, HALV) | CNN | 85 * /82 |
| 文献[23] (2020) | 16 | EEG, GSR, RB, EOG | 4 (HAHV, LAHV, LALV, HALV) | ECNN | 82.92 * |
| 文献[24] (2019) | 32 | EEG | 4 (HAHV, LAHV, LALV, HALV) | GELM | 69.67 |
| 文献[25] (2020) | 32 | EEG | 4 (HAHV, LAHV, LALV, HALV) | DFC-DNN-SVM | 81.3 |
| 文献[26] (2020) | 32 | EEG | 4 (HAHV, LAHV, LALV, HALV) | HOS-LSTM-Softmax | 82.01 |
| 本文 | 32 | BVP, vEOG, RSP, zEMG, GSR, hEOG | 4 (HAHV, LAHV, LALV, HALV) | mRMR-ELM-Adaboost | 83.06 |

注:有 * 号表示对单被试进行训练和测试,最后取总体平均值。没有 * 标志,意味着多被试交叉训练和测试。

3 结 论

在本文研究中,利用 6 种生理信号用于情绪状态辨识,为提高情绪辨识的准确性,提出了基于全视域特征表

表 3 不同方法辨识精度比较
Table 3 Comparison of the identification accuracy with different methods

| 被试 | 准确率/% | | |
|---------|------------|------------|---------------------|
| | ELM | mRMR + ELM | mRMR + ELM-Adaboost |
| s01-s32 | 72.96±1.04 | 74.81±0.17 | 83.06±1.25 |

2.5 讨论

本文提出了基于融合信息的全视域特征表征和分类器强化的情绪识别方法,以实现四类交叉情绪辨识。在 2.4 节中,分别从不同角度分析了所提方法的有效性。近年来,较多研究者也利用 DEAP 数据库进行情绪识别的研究。本文提出的方法与近期相关研究工作的比较如表 4 所示。

从表 4 分析可知,研究者从不同的角度针对情绪辨识开展了大量研究工作并取得了不错效果。目前,单人的四分类相比于多人交叉的四分类情绪辨识,整体精度明显要高。通常情况下,情绪类型越多,识别越困难;交叉辨识比单被试辨识难度大。参考文献[22-23],单人整体四分类情绪辨识精度分别达到了 85%、82.92%。在生理信号的选择中,大多研究者利用脑电开展情绪辨识研究^[24-26],而相比于脑电,利用其他的生理信号同样也获得不错的辨识效果^[22-23]。此外,利用各种神经网络开展的深度学习方法也越来越多的受到关注^[22-23,25]。

本文探索了四分类的交叉情绪辨识方法,情绪识别类型情境与文献[22,24-26]相同。所提的融合信息的全视域特征表征,结合 mRMR 特征优化及 ELM-Adaboost 的分类方法,实现四分类交叉情绪辨识准确率达到了 83.06%,同比相关研究方法具有更好的辨识性能。

征与 ELM-Adaboost 的情绪辨识方法。运用所提的融合信息的数据处理策略,由于样本中交叉融合了不同类型的生理信号,能够从全视域角度提取更有效的样本特征。采用基于最大相关性最小冗余度的特征选择评估方法对融合后的特征进行选择,去除冗余特征,实现情绪状态的

有效表征。运用 ELM-Adaboost 的分类器强化算法,有效避免传统分类器的局限性,提升情绪辨识分类性能。基于 DEAP 数据库多角度的实验比较分析,充分证明了本文所提方法的有效性。

参考文献

- [1] LI W, ZHANG ZH, SONG A G. Physiological-signal-based emotion recognition: An odyssey from methodology to philosophy[J]. *Measurement*, 2021, 172: 108747.
- [2] 苗敏敏, 徐宝国, 胡文军, 等. 基于自适应优化空频微分熵的情感脑电识别[J]. *仪器仪表学报*, 2021, 42(3): 221-230.
- MIAO M M, XU B G, HU W J, et al. Emotion EEG recognition based on the adaptive optimized spatial-frequency differential entropy [J]. *Chinese Journal of Scientific Instrument*, 2021, 42(3): 221-230.
- [3] KACEM A, DAOUDI M, BEN A B, et al. A novel geometric framework on gram matrix trajectories for human behavior understanding[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(1): 1-14.
- [4] SU M H, WU C H, HUANG K Y, et al. Cell-coupled long short-term memory with l-skip fusion mechanism for mood disorder detection through elicited audiovisual features[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2020, 31(1): 124-135.
- [5] GANNOUNI S, ALEDAILY A, BELWAFI K, et al. Emotion detection using electroencephalography signals and a zero-time windowing-based epoch estimation and relevant electrode identification[J]. *Scientific Reports*, 2021, 11(1): 7071.
- [6] HUANG D M, CHEN S T, LIU CH, et al. Differences first in asymmetric brain: A bi-hemisphere discrepancy convolutional neural network for EEG emotion recognition[J]. *Neurocomputing*, 2021, 448: 140-151.
- [7] GANAPATHY N, VEERANKI Y R, KUMAR H, et al. Emotion recognition using electrodermal activity signals and multiscale deep convolutional neural network [J]. *Journal of Medical Systems*, 2021, 45(4): 49.
- [8] SANTAMARIA-GRANADOS L, MUNOZ-ORGANERO M, RAMIREZ-GONZALEZ G, et al. Using deep convolutional neural network for emotion detection on a physiological signals dataset (AMIGOS) [J]. *IEEE Access*, 2019, 7: 57-67.
- [9] 赵紫宁, 李智, 张绍荣. 运动想象脑机接口中两种改进的脑电共空域模式特征提取方法[J]. *电子测量与仪器学报*, 2019, 33(12): 64-70.
- ZHAO Z N, LI ZH, ZHANG SH R. Two improved methods for EEG common spatial pattern feature extraction in motor imagery based brain-computer interface[J]. *Journal of Electronic Measurement and Instrumentation*, 2019, 33(12): 64-70.
- [10] PAN L ZH, YIN Z M, SHE SH G, et al. Emotional state recognition from peripheral physiological signals using fused nonlinear features and team-collaboration identification strategy[J]. *Entropy*, 2020, 22(5): 511.
- [11] CHEN T, JU S H, REN F J, et al. EEG emotion recognition model based on the LIBSVM classifier[J]. *Measurement*, 2020, 164: 108047.
- [12] LI X CH, WANG L, SUNG E. AdaBoost with SVM-based component classifiers [J]. *Engineering Applications of Artificial Intelligence: The International Journal of Intelligent Real-Time Automation*, 2008, 21(5): 785-795.
- [13] WAN SH K, LI X H, YIN Y J, et al. Milling chatter detection by multi-feature fusion and Adaboost-SVM[J]. *Mechanical Systems and Signal Processing*, 2021, 156: 107671.
- [14] PENG H C, LONG F H, DING C. Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2005, 27(8): 1226-1238.
- [15] 袁琦, 周卫东, 李淑芳, 等. 基于 ELM 和近似熵的脑电信号检测方法[J]. *仪器仪表学报*, 2012, 33(3): 514-519.
- YUAN Q, ZHOU W D, LI SH F, et al. Approach of EEG detection based on ELM and approximate entropy[J]. *Chinese Journal of Scientific Instrument*, 2012, 33(3): 514-519.
- [16] SARTORI M A, ANTSAKLIS P J. A simple method to derive bounds on the size and to train multilayer neural networks[J]. *IEEE Transactions on Neural Networks*, 1991, 2(4): 467-471.
- [17] TAMURA S, TATEISHI M. Capabilities of a four-layered feedforward neural network: Four layers versus three[J]. *IEEE Transactions on Neural Networks*, 1997, 8(2): 251-255.
- [18] FREUND Y, SCHAPIRE R E. A decision-theoretic

- generalization of on-line learning and an application to boosting[J]. Journal of Computer and System Sciences, 1997, 55(1): 119-139.
- [19] KOELSTRA S, MUHL C, SOLEYMANI M, et al. DEAP: A database for emotion analysis using physiological signals[J]. IEEE Transactions on Affective Computing, 2012, 3(1): 18-31.
- [20] EKMAN P, FRIESEN W V, O'SULLIVAN M, et al. Universals and cultural differences in the judgments of facial expressions of emotion[J]. Journal of Personality and Social Psychology, 1987, 53(4): 712-717.
- [21] RUSSELL J A. A circumplex model of affect[J]. Journal of Personality and Social Psychology, 1980, 39(6): 1161-1178.
- [22] AL MACHOT F, ELMACHOT A, ALI M, et al. A deep-learning model for subject-independent human emotion recognition using electrodermal activity sensors [J]. Sensors, 2019, 19(7): 1659.
- [23] HUANG H P, HU ZH CH, WANG W M, et al. Multimodal emotion recognition based on ensemble convolutional neural network[J]. IEEE Access, 2020, 8: 3265-3271.
- [24] ZHENG W L, ZHU J Y, LU B L. Identifying stable patterns over time for emotion recognition from EEG[J]. IEEE Transactions on Affective Computing, 2019, 10(3): 417-429.
- [25] ASGHAR M A, KHAN M J, RIZWAN M, et al. An innovative multi-model neural network approach for feature selection in emotion recognition using deep feature

clustering[J]. Sensors, 2020, 20(13): 3765.

- [26] SHARMA R, PACHORI R B, SIRCAR P. Automated emotion recognition based on higher order statistics and deep learning algorithm [J]. Biomedical Signal Processing & Control, 2020, 58: 101867.

作者简介



潘礼正 (通信作者), 2013 年于东南大学获得博士学位, 现为常州大学副教授、硕士生导师, 主要研究方向为康复机器人、人机交互技术。

E-mail: panlz@cczu.edu.cn

Pan Lizheng (Corresponding author) received his Ph. D. degree from Southeast University in 2013. He is now an associate professor and master supervisor at Changzhou University. His research interests include rehabilitation robot and human computer interaction technology.



宋爱国, 1993 年于南京航空航天大学获得硕士学位, 1996 年于东南大学获得博士学位, 现为东南大学教授。主要研究方向为遥操作技术、人机交互技术、康复机器人技术。

E-mail: a. g. song@seu.edu.cn

Song Aiguo received his M. Sc. degree from Nanjing University of Aeronautics and Astronautics in 1993 and received his Ph. D. degree from Southeast University in 1996. Now, he is a professor in Southeast University. His main research interests include teleoperation, human computer interaction, and rehabilitation robot.