

文章编号: 1671-7104(2020)04-0283-05

基于多种生理信号的情绪识别研究

【作者】 陈沙利^{1,2,3}, 张柳依⁴, 江锋^{1,2,3}, 陈婉琳^{1,2,3}, 缪家骏^{1,2,3}, 陈杭^{1,2,3}

1 浙江大学 生物医学工程与仪器科学学院, 杭州市, 310027

2 生物医学工程教育部重点实验室, 杭州市, 310027

3 浙江省心脑血管检测技术与药效评价重点实验室, 杭州市, 310027

4 浙江大学 心理与行为科学系, 杭州市, 310027

【摘要】 情绪是由某种特定的对象或情景所引发的一系列反应, 会影响人的生理状态, 因此可通过生理信号进行识别。提出一种融合脉搏波、皮肤电反应、呼吸、皮肤温度等多种信号的特征, 结合SVM-RFE-CBR(基于支持向量机可减少相关性偏差的递归特征消除)特征排序算法进行特征选择, 利用支持向量机进行分类的情绪识别模型, 并通过DEAP数据集验证该模型在愉悦度、唤醒度、优势度上的二分类效果, 分别获得了73.5%、81.3%、76.1%的准确率。结果表明, 利用多种生理信号融合可有效地进行情绪识别。

【关键词】 情绪识别; 多生理信号; 支持向量机

【中图分类号】 R318

【文献标志码】 A

doi: 10.3969/j.issn.1671-7104.2020.04.001

Emotion Recognition Based on Multiple Physiological Signals

【Writers】 CHEN Shali^{1,2,3}, ZHANG Liuyi⁴, JIANG Feng^{1,2,3}, CHEN Wanlin^{1,2,3}, MIAO Jiajun^{1,2,3}, CHEN Hang^{1,2,3}

1 College of Biomedical Engineering and Instrument Science, Zhejiang University, Hangzhou, 310027

2 Key Laboratory of Biomedical Engineering of Ministry of Education, Zhejiang University, Hangzhou, 310027

3 Zhejiang Provincial Key Laboratory of Cardio-Cerebral Vascular Detection Technology and Medicinal Effectiveness Appraisal, Zhejiang University, Hangzhou, 310027

4 Department of Psychology and Behavioral Science, Zhejiang University, Hangzhou, 310027

【Abstract】 Emotion is a series of reactions triggered by a specific object or situation that affects a person's physiological state and can, therefore, be identified by physiological signals. This paper proposes an emotion recognition model. Extracted the features of physiological signals such as photoplethysmography, galvanic skin response, respiration amplitude, and skin temperature. The SVM-RFE-CBR(Recursive Feature Elimination-Correlation Bias Reduction-Support Vector Machine) algorithm was performed to select features and support vector machines for classification. Finally, the model was implemented on the DEAP dataset for an emotion recognition experiment. In the rating scale of valence, arousal, and dominance, the accuracy rates of 73.5%, 81.3%, and 76.1% were obtained respectively. The result shows that emotional recognition can be effectively performed by combining a variety of physiological signals.

【Key words】 emotion recognition, multiple physiological signals, support vector machine

0 引言

情绪是由某种特定的对象或情景所引发的一系列反应, 通常由主观感受、肢体表达和生理状态的协调变化组成^[1]。研究表明, 情绪会影响人的解释、判断、决策、推理及风险认知等高层次的认知过程^[2], 在人的日常生活及沟通交流过程中扮演着非常重要的角色。近年来, 情绪识别技

术逐渐成为研究热点, 在人机交互、医疗健康、社交媒体、游戏、智能教育、认知负荷^[3]等领域拥有广泛的应用前景。

詹姆士(James)-兰格(Lange)^[4]的情绪理论认为, 情绪是自身对由刺激引发的循环系统、消化系统、内分泌系统、交感神经系统等机体变化的一种体验和感觉。因此, 可以通过分析与每种情绪相关的生理变化的模式来检测情绪。同时, 由于伴随情绪的生理变化具有自主性, 不受主观意

收稿日期: 2019-11-19

作者简介: 陈沙利, E-mail: 21815049@zju.edu.cn

通信作者: 陈杭, E-mail: ch-sun@263.net

念控制,因此与直接观察面部表情和身体姿态等外部行为表现相比较,基于生理信号的情绪识别能获得更加客观真实的结果^[5]。

目前常用的情绪识别方案多使用脑电信号来表征大脑皮层的神经活动,以反映情绪状态,已获得了较好的结果。但脑电信号采集困难,并且由于采集装置的复杂性,会对受试者造成不适,影响其情绪,最终影响信号采集效果。而外周生理信号采集简便,同时也能在一定程度上反映情绪状态,可作为一种有效的补充方案。

本研究提取了脉搏波、皮肤电反应、呼吸、体温等多种信号的时域、频域、非线性特征,采用一种基于支持向量机的可减少相关性偏差的递归消除特征排序(SVM-RFE-CBR)算法进行特征选择,并使用DEAP数据库进行验证,利用支持向量机分别对愉悦度、唤醒度和优势度进行二分类,实现了基于多种外周生理信号的情绪识别。

1 数据来源

DEAP是一个用于人类情绪分析的多模态数据库^[6]。数据库共有32名参与者,每人观看40个情绪内容不同的音乐视频,其中17个视频具有来自last.fm网站的情感标签,每个视频持续1 min。采用512 Hz的采样率采集脑电、肌电、呼吸、皮肤电反应、眼电、脉搏波和皮肤温度等生理信号,同时记录了1~22号参与者观看视频时的面部表情。参与者在每次试验后对视频进行Valence(愉悦度,范围从消极到积极,表明情绪从令人不愉快到愉快)、Arousal(唤醒度,范围从被动到活跃,表明情绪的激烈程度)、Dominance(优势度,范围从受控到自主,反映了人在某种情绪中的控制能力)几个维度的评分,范围为1~9。

每位被试的生理数据按照观看不同视频时的状态被分为40段,每段包含3 s基线状态和60 s测试状态(观看情绪视频时的状态)。对于采集了面部表情视频数据的参与者,选取观看带有情感标签的视频时采集的脉搏波、皮肤电反应、呼吸和皮肤温度4种信号的数据构建子数据集,便

于由人工判断受试者是否被激发了相应的情绪。分别使用参与者所打出的Valence、Arousal、Dominance评分作为分类标准,将1~3分和7~9分分为两类,使用0和1进行标记。

2 情绪识别方案

情绪识别方案包含特征提取、特征选择及情绪分类三个部分。

2.1 特征提取

对选取的每种信号,进行低通滤波、去除基线漂移等预处理,取60 s测试状态下的数据进行特征提取。得到的序列记为PPS(PP interval sequence)。对预处理后脉搏波(photoplethysmography, PPG)、皮肤电反应(galvanic skin response, GSR)、呼吸(respiration amplitude, RSP)和皮肤温度(skin temperature, SKT)信号分别提取多种特征,具体如表1所示。

表1 提取的特征
Tab.1 Extracted features

生理信号	提取特征
PPS	低频(0.04~0.15 Hz)、高频(0.15~0.4 Hz)能量之比FFT_ratio; FFT绝对值全积分L1_Norm; 一阶差分的最值diffPP_max
PPG	PPG、PPS及脉率信号PR的Picard参数和最大最小值之比PPG_r、PPS_r、PR_r; 脉动周期的标准差、脉动周期差值的标准差及它们的比值 ^[7] ; 采用“db8”小波对PPG进行四层小波分解后的各频段能量比,各近似系数和细节系数的最大值、最小值、均值、中值、标准差及总和; PSI ₃₍₁₀₎ 指数 ^[8]
GSR	衰变期间的平均下降率AOND; 局部最小值的个数NOLM; 0~2.4 Hz频段的能量FP及频谱功率SP; 低频信号SR(0~0.2 Hz)和极低频信号VSR(0~0.08 Hz)的正负向过零率Pnum、Nnum
RSP	低频(0.05~0.25 Hz)和高频(0.25~5 Hz)段的能量SP_L、SP_H及频谱功率的比值ER_FFT; 信号均值ARS、变化率MOD; 呼吸幅度和频率的均值MB、MBR; 呼吸峰值间期的均值APT
SKT	信号均值MSKT、方差STDSKT; 低频(0~0.1 Hz)和高频(0.1~0.2 Hz)段的能量SP_L、SP_H及频谱功率的比值ER_FFT

表1中, Picard参数^[9]包括6个统计学参数:原始信号均值 μ_s 、标准差 σ_s 、一阶差分绝对值的均值 δ_s 、二阶差分绝对值的均值 γ_s , 标准化后信号的一阶及二阶差分绝对值的均值 $\tilde{\delta}_s$ 、 $\tilde{\gamma}_s$ 。小波分解后,用ca1表示低频(近似)系数,cd1~cd4分别表示1~4层高频(细节)系数,最大值、最

小值、均值、中值、标准差及总和分别用max、min、mean、median、std、sum表示。此外,对每种信号均提取了谱熵SpeEn、小波熵WavEn、样本熵SampEn、近似熵ApEn等特征。

为了消除不同个体间的差异,提高数据可比性,对所提取的每个特征均采用Z-score标准化方法进行处理。计算式如下:

$$\tilde{X} = \frac{X - \mu}{\sigma} \quad (1)$$

其中 \tilde{X} 为标准化后数据, μ 和 σ 分别为数据的均值和标准差。经处理后的数据符合标准正态分布,即均值为0,标准差为1。

2.2 特征选择

所提取的多个特征中,有些特征之间可能存在相关性,会产生冗余,容易导致以下后果:

①特征分析及模型训练时间过长;②引起过度拟合,泛化能力差,识别率低;③引发“维数灾难”,导致模型性能下降^[10]。因此,进行特征选择十分必要。利用SVM-RFE-CBR方法进行特征排序后,采用步进法进行特征选择。

SVM-RFE (recursive feature elimination) 是一种基于支持向量机 (support vector machine, SVM) 进行特征排序的方法。对于线性分类问题,进行特征选择时理想的目标函数是计算误差的期望值,即在无限个特征组合情况下计算误差率。在SVM-RFE中,这个目标函数由在训练集上计算的成本函数进行替代^[11]。因此SVM-RFE的基本思想为,计算由于删除特征 x_i 而导致的成本函数 $DJ(i)$ 的变化。

对于训练集 $\{(\mathbf{x}_i, y_i)\}_{i=1}^N$,其中 $\mathbf{x}_i \in R^D$, $y_i \in \{+1, -1\}$, \mathbf{x}_i 为第 i 个样本, N 为样本总量, D 为样本特征数, R^D 为特征空间, SVM的最优分类超平面为:

$$\mathbf{w}^T \cdot \mathbf{x}_i + b = 0 \quad (2)$$

可求得 \mathbf{w} 的解为:

$$\mathbf{w} = \sum_{i=1}^N \alpha_i y_i \mathbf{x}_i \quad (3)$$

α_i 为拉格朗日乘子,则成本函数^[12]:

$$DJ(i) = \mathbf{w}_i^2 \quad (4)$$

通过构建的SVM分类器训练样本,计算每个特征的成本函数并对其进行排序,将得分最小的特征移入无效特征集 F_{out} 后,用剩余有效特征集 F_{in} 再次训练并进行下一次迭代,直到对所有特征完成排序。具体迭代步骤如下:①初始化有效特征集 F_{in} 为所有提取出的特征,无效特征集 F_{out} 为空;②利用特征集 F_{in} 训练SVM分类器,得到 \mathbf{w} ;③计算 $DJ(i)$,找出 $\min\{DJ(i)\}$, $i \in F_{in}$,将该特征移入 F_{out} ;④更新特征集 F_{in} 并进行下一次迭代;⑤得到排序后的特征集。

虽然SVM-RFE是一种有效的特征选择方法,在特征数量较大时,可有效地避免过度拟合。但是研究表明,当特征空间中存在大量高度相关的特征时,这种方法会产生较大偏差。即:若在一次迭代中,某一组高度相关的特征中有一部分特征的成本函数因得分较低而被移除,那么这整组特征都可能会被移出 F_{in} 。

因此采用了一种可减少相关性偏差 (correlation bias reduction, CBR) 的方法^[13]。这种方法旨在将 F_{out} 中潜在的有效特征移回 F_{in} ,避免某些特征的重要性被低估而得出错误的特征排序,被证明可有效地提高特征选择的表现。具体过程为:①在每次RFE迭代完成后,找出 F_{out} 中得分最高的特征并标记为K;②找到所有与K的相关系数较高的特征;③若这些特征没有一个处在 F_{in} 中,就将K移回 F_{in} 并进行下一次迭代。

在利用SVM-RFE-CBR对所有特征进行排序后,采用SPSS (版本15.0) 中的步进法选择特征,在每个步骤中,输入可以使总体威尔克Lambda最小化的变量,再将最终选出的特征作为分类器的输入进行情绪分类。

2.3 情绪分类

SVM是目前情绪识别中最常用的一种方法,它在高维的小样本数据处理方面有独特的优势^[14],且拥有优秀的泛化能力。SVM为有监督的学习模型,基本思想是:线性可分的情况下,在特征空间 R^D 中构建一个超平面 $f(\mathbf{x}) = \mathbf{w}^T \cdot \mathbf{x}_i + b$,将不

同种类的样本区分开来。对于预测类别 y_i , 有:

$$y_i = \begin{cases} 0, & \text{if } f(x_i) < 0 \\ 1, & \text{if } f(x_i) \geq 0 \end{cases} \quad (5)$$

为了尽可能地增大样本间的区分度, 提高模型的分类准确率, 需要将样本点到超平面的最小距离最大化。因此目标函数为:

$$\max_{w, b} \left\{ \min_i \frac{|w^T \cdot x_i + b|}{\|w\|} \right\} \quad (6)$$

即 $\min_i \frac{1}{2} \|w\|^2$ (约束条件为: $y_i(w^T x_i + b) \geq 1, i=1, 2, \dots, n$)。

线性不可分的情况下, 可通过一个非线性映射, 把特征空间映射到一个具有核函数的高维空间, 使得在原来的特征空间中非线性可分的问题转化为在高维空间中线性可分的问题。核函数可以降低维数增加带来的计算复杂性^[10]。

采用K折交叉验证^[15]进行检验以得到稳定可靠的分类模型, 即将原始数据随机分成K份, 保留其中一份作为测试集, 其余K-1份作为训练集, 重复K次并将分类准确率的结果取平均以得到一个单一的估测值。本研究中K=5。

3 结果

3.1 特征选择结果

经过特征排序和步进法进行特征选择后, 最终选出分别用于Valence、Arousal、Dominance的分类, 提取自不同信号的特征集如表2所示:

表2 用于愉悦度、唤醒度和优势度的分类特征
Tab.2 Classification features for Valence, Arousal and Dominance

分类维度轴	分类特征
Valence (6个)	PPG: PR_r、diffPP_max、cd3_median cd3_sum; GSR: AOND; SKT: SP_L
Arousal (9个)	PPG: L1_Norm、PPG_r、PP4_r、cd4_std; RSP: ER_FFT、MB; SKT: MSKT、STDSKT、SKT_WavEn
Dominance (11个)	PPG: FFT_ratio、PPG_μs、PPG_γs、cd1_median、cd4_min; GSR: AOND、VSR_Nnum; RSP: SP_H、ARS、MB、RSP_WavEn

从表2中可知:

(1) 特征集中对Valence轴进行分类的有效特征较少, 表明Valence的区分难度更大;

(2) 相比原始PPG信号, 脉率变异性等特征

包含了更多情绪方面的信息;

(3) PPG高频信号隐含的情绪信息比低频更多, 因此更多的高频特征被选入特征集;

(4) 特征集中包含了多个小波变换相关特征, 表明小波变换是一种对于情绪特征的有效提取方法;

(5) 与Valence有关的特征几乎来自于脉搏波信号; Arousal与脉搏波、呼吸、皮肤温度等信号有更大的关联; 而Dominance则与脉搏波、皮肤电反应、呼吸等信号有更大的关联。

3.2 情绪识别结果

对于每个数据样本, 根据参与者所打出的Valence、Arousal、Dominance评分, 可分别标记不同的真实标签0或1。通过分类器进行分类后, SVM模型会判断该样本所属的类别, 并输出结果标签0或1。使用TN(true negative)表示真实标签为0, 结果标签为0的样本数量; FN(false negative)表示真实标签为1, 结果标签为0的样本数量; FP(false positive)表示真实标签为0, 结果标签为1的样本数量; TP(true positive)表示真实标签为1, 结果标签为1的样本数量。

根据TN、FN、FP、TP构建分类结果矩阵, 展示在每个类别中分类正确和错误的比例, 结构如表3所示。此外, 分类准确率(accuracy, ACC)被定义为: $(TP+TN)/(TP+TN+FP+FN)$, 表示正确分类的样本数占总体样本数的比例。阴性预测值(negative predictive value, NPV)和阳性预测值(positive predictive value, PPV)数值越高, 分类准确率越高, 表明分类效果越好。

表3 分类结果矩阵结构
Tab.3 The structure of classification result matrix

分类结果		真实标签	
		0	1
结果标签	0	$\frac{TN}{TN+FN}$	$\frac{FN}{TN+FN}$
	1	$\frac{FP}{TP+FP}$	$\frac{TP}{TP+FP}$

表4展示了使用SVM分类器对Valence、Arousal、Dominance进行二分类的分类结果, 表

5展示了三种分类的准确率。

表4 支持向量机的分类结果(%)
Tab.4 Result of classification by support vector machine

分类结果		Valence		Arousal		Dominance	
		0	1	0	1	0	1
SVM	0	41.8	58.2	78.7	21.3	72.6	27.4
	1	10.3	89.7	16.4	83.6	20.3	79.7

表5 支持向量机的分类准确率(%)
Tab.5 Classification accuracy of support vector machine

分类维度	Valence	Arousal	Dominance
分类准确率	73.5	81.3	76.1

(1) 在分类结果矩阵中, 对Arousal进行分类的NPN、PPV以及ACC最高, 表明对Arousal所提取的特征有最好的分类效果, 对Valence和Dominance所提取的特征分类效果较差。这可能是由于Arousal所代表的情绪激烈程度的变化会更多地引起生理方面的变化; 也可能是由于视频材料更多地诱发了在Arousal轴上区分度较大的情绪。

(2) 整体的分类准确率不够高, 原因可能是数据的不平衡。不同类别中的数据量相差较大导致构建的分类模型不够稳定, 分类器倾向于将样本分到数据量较大的类别中。同时由于总体样本量较少导致模型性能下降, 识别率低。

4 讨论

本研究提供了一种使用外周生理信号进行情绪识别的方案, 采用DEAP数据库构建的子数据库进行验证, 对Valence、Arousal、Dominance三个轴分别进行二分类, 获得了最高81.3%的识别准确率。在没有使用脑电信号的情况下, 获得了与使用脑电和多种外周生理信号进行情绪识别时近似的识别效果, 为今后发展更加便捷有效的情绪识别方案提供了可能。

本研究所提出的情绪识别方案分类准确率不够高, 主要有以下两个原因。首先, 只使用了DEAP数据库的部分子数据集, 因此用于模型训练的数据量不够大, 导致构建的模型不够完善。其次, 情绪识别过程受人体差异的影响较大, 通常来说对于单个人不同情绪的识别准确率要远高于对不同人的情绪识别准确率。因此, 在后续研

究中需要增大数据量, 以优化分类模型。也可将生理信号与语音、面部表情等与情绪有关的信息相结合, 形成多模态、多特征融合的情绪识别方案, 以进一步提高情绪识别的准确性和鲁棒性。

参考文献

- [1] MAUSS I B, LEVENSON R W, MCCARTER L, et al. The tie that binds coherence among emotion experience, behavior, and physiology[J]. Emotion, 2005, 5(2): 175-190.
- [2] BLANCHETTE I, RICHARDS A. The influence of affect on higher level cognition: a review of research on interpretation, judgement, decision making and reasoning[J]. Cognit Emot, 2010, 24(4): 561-595.
- [3] CALVO R A, D'MELLO S K. Affect detection: an interdisciplinary review of models, methods, and their applications[J]. IEEE Trans Affect Comput, 2010, 1(1):18-37.
- [4] JAMES W. What is an emotion[J]. Mind, 1884, 9(34):188-205.
- [5] WU N, JIANG H, YANG G, et al. Emotion recognition based on physiological signals[C]. International Conference on Brain Inspired Cognitive Systems. Springer Berlin Heidelberg, 2012: 311-320.
- [6] KOELSTRA S, MUHL C, SOLEYMANI M, et al. DEAP: A database for emotion analysis; using physiological signals[J]. IEEE Trans Affect Comput, 2012, 3(1): 18-31.
- [7] 陈继华, 李岚, 钱坤喜. 基于多生理信号的情绪初步识别[J]. 生物医学工程研究, 2006, 25(3): 141-146.
- [8] 陈婉琳, 江锋, 陈新忠, 等. 脉搏波形态学研究及其在子痫前期的应用[J]. 天津大学学报(自然科学与工程技术版), 2019(52): 83-87.
- [9] TAKAHASHI K. Remarks on SVM-based emotion recognition from multi-modal bio-potential signals[C]. IEEE International Workshop on Robot & Human Interactive Communication, 2004: 95-100.
- [10] LIN S, JINYAN X, MINGYUE Y, et al. A review of emotion recognition using physiological signals[J]. Sensors, 2018, 18(7): 2074-2115.
- [11] TORRES C A, OROZCO A A, ALVAREZ M A. Feature selection for multimodal emotion recognition in the arousal-valence space[C]. Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2013: 4330-4333.
- [12] 黄晓娟, 张莉. 改进的多类支持向量机递归特征消除在癌症多分类中的应用[J]. 计算机应用, 2015, 35(10): 2798-2802.
- [13] YAN K, ZHANG D. Feature selection and analysis on correlated gas sensor data with recursive feature elimination[J]. Sensor Actuat B-Chem, 2015, 212: 353-363.
- [14] 雷沛, 王静, 周昕炜, 等. 基于Relief特征选择算法与多生理信号的焦虑状态识别[J]. 中国医疗器械杂志, 2014(3): 186-189.
- [15] 赵国联, 宋金晶, 葛燕, 等. 基于生理大数据的情绪识别研究进展[J]. 计算机研究与发展, 2016, 53(1): 80-92.