

文章编号:1006-3080(2018)05-0730-07

DOI:10.14135/j.cnki.1006-3080.20170715001

## 基于禁忌搜索的混合算法在驾驶压力识别中的应用

叶朋飞, 陈兰岚, 张 傲

(华东理工大学化工过程先进控制和优化技术教育部重点实验室, 上海 200237)

**摘要:**驾驶员在压力状态下行车会对驾驶安全产生很大影响,严重时甚至会造成交通事故。为准确检测驾驶员的压力状态,提取了驾驶员生理信号的多模态特征并提出了一种基于多种过滤式算法(Multi-filter, MF)与禁忌搜索算法(Tabu Search, TS)相结合的混合算法来选择有效特征向量。该算法首先采用多种过滤式算法的综合评分对原始特征集进行排序和过滤,有效降低特征维度;然后利用禁忌搜索算法进一步选出最优特征组合;最后采用支持向量机对3种不同驾驶压力水平进行分类。实验结果表明,本文提出的混合算法不仅有效地消除了高维特征向量中的冗余信息,还提升了分类准确率。

**关键词:**生理信号; 过滤算法; 禁忌搜索; 特征选择; 驾驶压力识别

中图分类号:TP391

文献标志码:A

## Hybrid Algorithm Based on Tabu Search for Drive Stress Recognition

YE Peng-fei, CHEN Lan-lan, ZHANG Ao

(Key Laboratory of Advanced Control and Optimization for Chemical Processes, Ministry of Education, East China University of Science and Technology, Shanghai 200237, China)

**Abstract:** Driving under stress status has a great impact on drive safety and even causes traffic accidents in severe cases. It is important to develop appropriate approaches for early detecting the driving stress and then improving the awareness and performance of drivers. Due to its strong objectivity, robustness and real-time property, the physiological signal analysis has been validated as an effective approach for reflecting physiological conditions of human operators. In order to accurately detect the stress status of drivers, multi-modal features are extracted from time, spectral, wavelet, and nonlinear domains based on physiological signals, i. e., electrocardiogram, galvanic skin response and respiration of drivers. And then, a hybrid algorithm based on the multi-filters (MF) and tabu search (TS) is proposed in this work to select effective features, since there are many redundant features and invalid features in the feature set. Firstly, in order to effectively reduce the feature dimension, the original features are sorted and filtered according to the comprehensive score calculated by multiple filtering algorithms. Three different filter methods including Chi test, Wilcoxon test and manual information are employed to generate a new feature subset, which had smaller feature dimensional. And then, the tabu search algorithm is adopted to further select the optimal feature combination from the optimal feature subset. Finally, three different drive stress levels are classified by means of support vector machine (SVM). Some widely-used feature selection methods, including genetic algorithm (GA) and random tabu search algorithm, are also

收稿日期:2017-07-15

基金项目:国家自然科学基金(61201124);中央高校基本科研业务费重点科研基地创新基金(222201717006)

作者简介:叶朋飞(1992-),男,硕士生,研究方向为生理信号处理与机器学习。E-mail: sky\_ecust@163.com

通信联系人:陈兰岚,E-mail:llchen@ecust.edu.cn

implemented for comparisons. The proposed hybrid algorithm not only eliminates redundant information in high dimensional eigenvectors effectively, but also improves the classification accuracy.

**Key words:** physiological signal; filter algorithm; tabu search; feature selection; drive stress recognition

驾驶员在压力状态下行车会对驾驶安全产生很大影响,不仅会导致驾驶员危险驾驶,甚至还有可能造成致命的交通事故<sup>[1]</sup>,因此,对驾驶员的压力状态进行检测,减少由压力引发的事故有着重要的现实意义。近年来,不少学者提出应用各种生理信号诸如皮电、血容、呼吸、脑电等检测驾驶员的压力状态,因为生理信号具有很强的客观性、鲁棒性和实时性,可以有效反映人的生理和情绪状态,故而在驾驶员压力和情绪识别方面取得了不少的研究成果<sup>[2-3]</sup>。

通常生理信号中蕴含了包括时间相位、波形、功率谱等多维度的特征信息,甚至还有学者从小波变换和非线性分析的角度来挖掘生理信号中能反映人体生理、心理状态的特征变量。多模态特征提取和分析在一定程度上可以提高识别准确率,但也引入了很多冗余和干扰信息,因此有效的特征降维和特征选择手段在驾驶员压力识别系统中非常关键,且有助于消除噪声影响、降低特征维度、缩短训练时间和提高分类准确率。

特征选择方法分为过滤式(Filter)和封装式(Wrapper)<sup>[4]</sup>。过滤式方法基于特征集的内在联系筛选特征,该方法简单灵活、运算速度快,但相较于封装式算法分类性能较差。封装式方法通过内嵌分类器有效筛选出分类性能好的特征子集,但在原始特征集维数较高时运算时间较长。禁忌搜索算法作为一种启发式算法,最初应用于各种组合优化问题,近年来也被作为封装式算法用于特征选择,并且由于其在初始解、搜索策略、禁忌长度、目标函数等方面可以根据具体问题优化设置,相比其他算法更具灵活性和优越性。本文提出了一种新颖的特征选择混合算法,有效集成了过滤式和封装式算法两者的特点。首先采用3种过滤式算法对所有原始特征进行综合评分从而生成初始解;随后用禁忌搜索算法进一步选择产生冗余度低且分类识别性能好的最优特征子集;最后采用支持向量机(SVM)对不同驾驶压力状态进行识别。

## 1 实验数据

实验采用美国国家卫生研究院 PhysioBank 数

据库中的实验数据<sup>[5]</sup>,该数据集记录了9位驾驶员在特定路线上完成27次驾驶任务中的4种生理信号:心电信号(Electrocardiogram,ECG)、左手和左脚上的皮电(Galvanic Skin Response,GSR)以及呼吸信号(Respiration,RESP)。实验中的驾驶路况分为休息、高速公路和城市街道3种,并通过视频打分和问卷调查验证了这3种路况分别对应驾驶员低压、中压和高压状态。

分别从每组实验数据的低压、中压和高压周期中提取10 min数据段,并将数据划分为时长为100 s的片段,低压、中压和高压3种状态分别产生了388个样本,总计1 764个样本。

## 2 特征提取

对原始数据进行预处理后,从每个数据片段提取了时域、频域、小波域和非线性分析4类总共154个特征。

(1) 时域特征。对4种生理信号分别计算了6个统计特征:均值、标准差、方差、均方根、最大值和最小值。另外,基于皮电在波形上的特性,特别提取了12个基于信号一阶差分序列和二阶差分序列的统计量特征以及4个波形形态特征<sup>[5]</sup>。

(2) 频域特征。考虑到各种生理信号的差异性,对每种信号分别计算特定频段的功率谱特征。呼吸信号提取了0~0.1 Hz、0.1~0.2 Hz、0.2~0.3 Hz、0.3~0.4 Hz和全频带总计5个能量特征;心电信号提取了11个心率变异性(HRV)相关特征,主要是0~0.04 Hz、0.04~0.15 Hz、0.15~0.5 Hz和全频带的功率谱及相关的组合特征;皮电信号则提取了5个0~0.5 Hz频带功率谱统计量特征。

(3) 小波特征。采用Daubechies小波函数将各生理信号分解为1个近似信号和5个细节信号<sup>[6]</sup>,进而提取了各子带小波系数的均值和标准差总计12个特征。

(4) 非线性特征。考虑到生理信号具有高度的非线性和非平稳性,对4种信号分别计算了6个非线性特征:样本熵<sup>[7]</sup>、C0复杂度<sup>[8]</sup>、近似熵<sup>[9]</sup>、Hurst指数<sup>[10]</sup>、排列熵<sup>[11]</sup>和分形维数<sup>[12]</sup>。

### 3 特征选择

#### 3.1 MFTS 混合算法架构

针对高维空间特征选择问题,充分考虑多种过滤式算法和禁忌搜索算法的优缺点后,本文提出了一种新颖的基于多种过滤式算法和禁忌搜索算法相结合的混合型特征选择算法(Multi Filter and Tabu

Search Algorithm, MFTS)。首先采用多种过滤式算法对原始特征集中的各个特征进行排序和综合评分产生高质量的新特征子集,极大地降低了特征维数;随后采用禁忌搜索算法对新特征集进行搜索,得到最优特征子集后将其送入支持向量机,最终实现对3种不同驾驶压力状态的分类识别。MFTS 算法总体架构如图1所示。

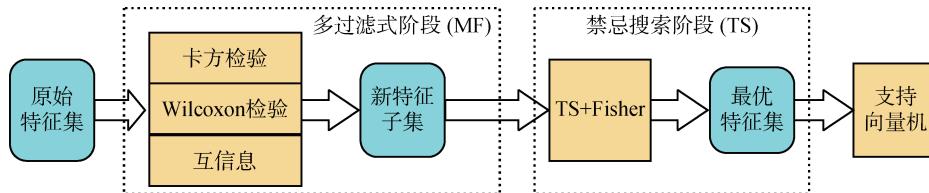


图1 MFTS 混合算法架构

Fig. 1 Framework of MFTS hybrid algorithm

#### 3.2 多过滤式阶段

3.2.1 卡方检验 卡方分布(Chi-square distribution)是概率论与统计学中常用的一种概率分布,是统计推断中应用最为广泛的概率分布之一。卡方检验是参照卡方分布来求取概率和临界值的统计检验,是一种非常有效的检验方法。作为独立性统计检验方法,卡方检验可以用来确定两个变量的依赖关系和独立性,故而可以用来进行特征选择。本文采用皮尔森卡方检验进行特征选择。

设有两个变量  $X$  和  $Y$ ,分别有  $m$  和  $n$  种状态,共有  $l$  个样本,检验原假设  $H_0: X$  和  $Y$  相互独立。根据式(1)可以计算变量的卡方值:

$$\chi^2 = l \left[ \sum_{i=1}^m \sum_{j=1}^n \frac{f_{ij}^2}{\sum_{j=1}^n f_{ij} \cdot \sum_{i=1}^m f_{ij}} - 1 \right] \quad (1)$$

其中  $f_{ij}$  表示变量  $X$  和  $Y$  分别处于第  $i$  种和第  $j$  种状态时的频数。

对于给定的显著性水平  $\alpha$ ,自由度  $(m-1)(n-1)$ ,查卡方分布表得分位数为  $\chi^2_{1-\alpha}((m-1)(n-1))$ ,使  $P\{\chi^2 > \chi^2_{1-\alpha}((m-1)(n-1))\} = \alpha$ ,当  $\chi^2 > \chi^2_{1-\alpha}[(m-1)(n-1)]$  时拒绝  $H_0$ ,否则接受  $H_0$ 。

3.2.2 Wilcoxon 检验 Wilcoxon 检验(Wilcoxon Test, WT)将观测值和零假设的中心位置之差的绝对值的秩分别按照不同的符号相加作为其检验统计量。这种检验方法适用于  $t$  检验中的成对比较,但并不要求成对数据之差服从正态分布,只要求对称分布即可。

假设有  $N$  对样本数据  $(X, Y)$ ,对于  $i = 1, \dots, N$ ,  $x_i$  和  $y_i$  表示具体测量值。

(1) 对于  $i = 1, \dots, N$ ,计算  $|y_i - x_i|$  和  $\text{sgn}(y_i - x_i)$ ,  $\text{sgn}$  为符号函数。

(2) 去除  $|y_i - x_i| = 0$  的数据对,将样本量降低为  $N_r$ 。

(3) 将筛选出的  $N_r$  个样本按照  $|y_i - x_i|$  从小到大排序,并找出它们的秩,绝对值相同的样本点将它们的秩的平均值作为自身的秩,记为  $R_i$ 。

(4) 计算检验统计量  $W = \sum_{i=1}^{N_r} [\text{sgn}(y_i - x_i) \cdot R_i]$ 。

(5) 对于给定的显著性水平  $\alpha$ ,查分布表可得临界值  $W_{\text{critical}, N_r}$ ,若  $|W| > W_{\text{critical}, N_r}$  则拒绝  $H_0$ ,否则接受  $H_0$ 。

3.2.3 互信息 互信息(Mutual Information, MI)是两个随机变量间的相互依赖性的度量。不同于相关系数,互信息并不局限于实值随机变量。当一个随机变量改变时,可以通过互信息来衡量另一个随机变量的不确定性。例如,如果两个变量  $X$  和  $Y$  相互独立,则知道  $X$  不对  $Y$  提供任何信息,反之亦然。

假设有两个离散随机变量  $X$  和  $Y$ ,则互信息可以定义为

$$I(X;Y) = \sum_{y \in Y} \sum_{x \in X} p(x,y) \lg \left( \frac{p(x,y)}{p(x)p(y)} \right) \quad (2)$$

其中:  $p(x,y)$  为  $X$  和  $Y$  的联合概率分布函数;  $p(x)$  和  $p(y)$  是  $X$  和  $Y$  的边际概率分布函数。

互信息是计算语言学模型分析的常用方法,它度量两个变量之间的相互性,在选择问题中用于度量特征对类别的区分度。

3.2.4 多种过滤式算法 过滤式算法根据统计量来进行特征选择。由于原理不同,每一种过滤式算

法都会选择出不同的特征集,且或多或少都会遗漏一些重要特征。为了克服这个不足之处,本文提出了一种新颖的基于多种过滤式算法(Multi-Filter, MF)的信息融合策略,基于多种过滤式算法得到一系列重要的特征,进而将这些特征组成的集合作为第二阶段的搜索集。MF 算法的流程如图 2 所示。

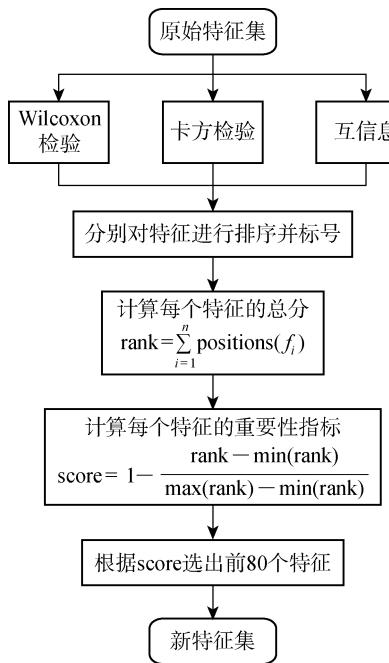


图 2 多种过滤式算法的流程框图

Fig. 2 Flow chart of multiple filter algorithm

MF 算法的基本思路如下:

(1) 利用 3 种不同的过滤式特征选择算法(卡方检验、Wilcoxon 检验、互信息)分别计算 154 个原始特征的相关性,并按相关性由强到弱对特征进行排序并标号。

(2) 针对每种过滤式算法,将特征的排序编号作为分数,并计算综合 3 种过滤式算法的分数后所得到的各个特征的总分数:

$$\text{rank} = \sum_{i=1}^n \text{positions}(f_i) \quad (3)$$

其中: $n$  为所使用的过滤式特征选择算法的个数; $\text{positions}(f_i)$  表示特征在第  $i$  个过滤式算法中的排列位置。

(3) 根据步骤(2)中的评分结果,按照式(4)计算各个特征的重要性指标:

$$\text{score} = 1 - \frac{\text{rank} - \min(\text{rank})}{\max(\text{rank}) - \min(\text{rank})} \quad (4)$$

其中  $\max$  和  $\min$  分别对应所有特征综合评分的最大值和最小值。

(4) 根据步骤(3)计算得出的重要性指标,将所

选特征由高到低排序,并选出前 80 个特征组成新的特征集,用于第二阶段的禁忌搜索算法。

为便于理解,本文通过举例来阐述多种过滤式算法,如图 3 所示。

特征排序	卡方检验	Wilcoxon 检验	互信息
	41	136	107
92	58	136	
136	35	58	
35	140	35	
...	...	...	
17	122	70	

特征	排序位置	总分
136	3, 1, 2	6

顺序	分数	特征
1	1	136
2	0.998 6	35
...	...	...

图 3 多种过滤式算法的实现

Fig. 3 Achievement of multiple filter algorithm

### 3.3 禁忌搜索阶段

禁忌搜索算法(TS)是一种模拟人的智力和记忆过程的智能优化算法,因为其极具特色的记忆功能,自从问世以来在组合优化领域取得了一系列的成就,有效地解决了局部最优问题。而特征选择作为一个典型的组合优化问题,同样可以采用禁忌搜索算法处理。

本文在特征选择第二阶段采用禁忌搜索算法,与第一阶段的 MF 算法结合组成特征选择混合算法。首先,根据 MF 算法生成的新特征集,将数据集分为训练集、测试集。然后,将训练集进一步划分为训练集和验证集,在训练集上采用禁忌搜索算法进行搜索,并用验证集来检验所选特征的性能,从而得到最优特征组合。考虑到原始禁忌搜索算法不完善,本文提出了部分改进。禁忌搜索算法的流程图如图 4 所示,其算法步骤如下:

(1) 对算法进行初始化,随机产生初始解,初始化禁忌表,设定搜索策略、邻域构造方法、禁忌对象、禁忌长度、适应度函数、特赦准则以及终止条件等。

(2) 根据当前解以及邻域构造方法产生邻域空间,随机选择其中部分元素作为候选解。

(3) 根据适应度函数选出候选集中的最好解作为邻域最优解,并判断当前最优解是否优于邻域最

优解。

(4) 如果邻域最优解优于当前最优解, 将邻域最优解作为新的当前最优解和当前解; 如果当前最优解优于邻域最优解且满足特赦准则, 将邻域最优解作为当前解。

(5) 禁忌邻域最优解, 并更新禁忌表。

(6) 判断是否达到终止条件。如果满足, 终止算法并输出最优解; 否则转至步骤(2)。

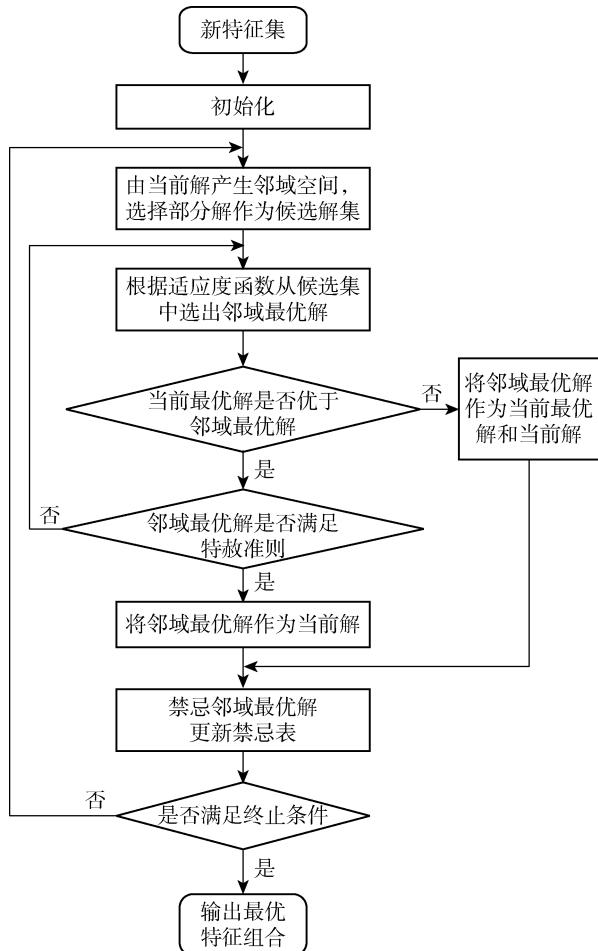


图4 禁忌搜索算法的流程图

Fig. 4 Flow chart of tabu search algorithm

在保持禁忌搜索算法基本框架不变的前提下, 本文针对具体的研究对象, 对禁忌搜索算法几个关键环节的特定设置如下:

(1) 特征编码。采用由0和1组成的二进制编码方式对特征向量进行编码, 每一种特征组合都有一个唯一的与之对应的二进制序列。序列中的0表示所对应的特征被舍弃, 1代表对应的特征被选中。

(2) 初始解。初始解的质量对禁忌搜索算法最终搜索的结果有很大的影响, 目前绝大部分的研究采取随机方法产生初始解。为了得到一个优质的初始解, 本文利用过滤式特征选择算法产生禁忌搜索

算法的初始解空间。

(3) 搜索策略。禁忌搜索算法在构造候选集时有两种搜索策略, 分别为集中性和多样性搜索策略。集中性搜索策略可以对当前的邻域空间进行充分搜索, 有利于找到最优解。多样性搜索策略扩大了搜索领域, 丰富了候选解, 能够防止算法陷入局部最优。考虑到本文算法事先通过过滤式特征选择算法产生了优质的特征子集, 极大提升了了解的质量, 减少了冗余性, 因此采用集中性搜索策略, 以期对解的邻域空间进行更充分的搜索。

(4) 禁忌长度。禁忌长度代表着在不使用特赦准则的情况下某个解不能被选中的迭代次数。从算法的计算量、内存空间以及搜索效率考虑, 禁忌长度应该较小; 但较大的禁忌长度才能帮助禁忌搜索算法克服局部最优困境。本文采用固定禁忌长度, 并将其设定为15。

(5) 适应度函数。智能优化算法根据适应度值来判断解的优劣, 故而特征选择的目的是寻找到一个最优特征组合, 实现用较少的特征维数来获得较高的分类准确率。定义适应度函数如式(6)所示, 包含了反映特征个数和分类准确率两个部分。

$$\text{rate} = \text{wrong\_rate}_1^2 + \text{wrong\_rate}_2^2 + \text{wrong\_rate}_3^2 \quad (5)$$

$$\text{Fitness} = k_1 \cdot \text{rate} + k_2 \cdot \text{num} \quad (6)$$

$$K = k_1 / k_2 \quad (7)$$

其中:  $\text{wrong\_rate}_1$ 、 $\text{wrong\_rate}_2$  和  $\text{wrong\_rate}_3$  分别表示3个类的分类错误率;  $\text{num}$  是选中的特征个数;  $k_1$  和  $k_2$  是相关系数, 分别代表分类准确率和特征个数对适应度函数的重要性;  $K$  为比例系数, 本文选取  $K = 500$ , 同时选择 Fisher 分类器作为禁忌搜索算法的嵌入分类器来计算每一个解的分类准确率。

(6) 特赦准则与终止条件。禁忌搜索在每次迭代后都会禁忌一个解, 但有时会在迭代搜索过程中得到一个比过往所有候选解还要好的解, 如果这个解被禁忌, 将会导致算法无法朝更好解的方向搜索, 甚至可能陷入局部最优。为了解决这个问题, 需要设立特赦准则, 用来释放某些不应该被禁忌的解。根据问题性质的不同, 禁忌搜索算法有多种特赦准则, 本文采用了基于适应度值的特赦准则, 根据解的适应度值来判断此解是否应该被特赦。

终止条件用来强制结束算法的搜索进程。智能优化算法很难自主停止搜索, 而无限搜索显然不切实际, 因此需要人为地设立终止条件来节省算法搜索时间, 本文设定最大迭代步数为50来强制终止算法搜索进程。

## 4 实验结果与分析

### 4.1 实验方法

针对过滤式算法和禁忌搜索算法选择出的最优特征组合,本文采用 SVM 作为最终的分类器。基于 MFTS 算法的结果,用训练集训练 SVM 得到最终的模型,并用测试集验证模型的有效性。为了防止训练模型时发生过拟合,考虑到一共有 14 条驾驶记录,故而采用弃一法进行 14 折交叉验证,并将 14 次交叉验证的平均结果作为最终的结果。

### 4.2 过滤式算法选择结果

在分析组合式过滤算法的性能之前,首先评估和比较单一过滤式算法的有效性。利用 3 种过滤式算法,即卡方检验(Chi)、Wilcoxon 检验(WT)和互信息(MI)分别选出各自最优的前 80 个特征,并采用 SVM 识别最终结果。

如图 5 所示,3 种过滤式特征选择算法均可以从原始高维特征集中选出对识别贡献度较高的特征。随着所选特征的增加,识别准确率有非常明显的提升,尤其前 20 维特征的识别贡献率较高,超过 20 维之后识别准确率的增长趋势有所放缓。此外,对比 3 条曲线可见,Wilcoxon 检验所选特征比卡方检验和互信息所选特征效果更好,准确率更高。

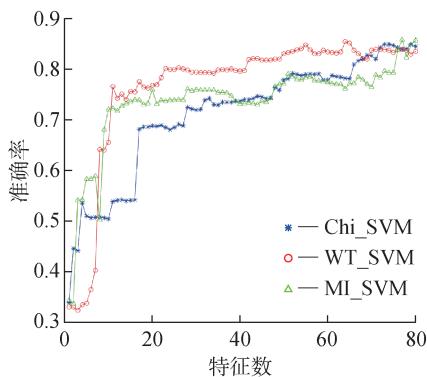


图 5 单一过滤式算法的性能比较

Fig. 5 Performance of single filter algorithm

### 4.3 混合算法选择结果

本文在单一过滤式算法的基础上,采用多种过滤式算法(MF)对特征进行综合评分,选择出排名前 80 个特征组成新的特征集,随后采用禁忌搜索算法对新的特征集进行搜索。将混合算法(MFTS)与基于随机初始解的禁忌搜索算法(Random\_TS)和遗传算法(GA)进行比较。设定算法终止条件为迭代 50 次,50 次迭代的结果如图 6 所示。由图 6 可知,局部搜索算法(禁忌搜索算法)的准确率变化比较平缓,而全局搜索算法(遗传算法)的准确率波动

较大,两种算法各有优势。MFTS 算法与 Random\_TS 和 GA 算法相比效果要好,搜索效率高,且准确率高。

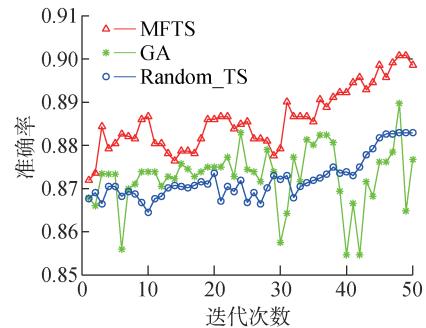


图 6 不同算法的识别准确率

Fig. 6 Accuracy curves of different algorithms

采用 MFTS 算法进行 50 次迭代后选择出的最优特征有 20 个左右,这些特征主要包括皮电信号的时域特征、小波域特征及非线性特征,其次为心率信号的小波特征和呼吸信号的功率谱特征,这与其他学者的研究结果基本一致。Healey 等<sup>[6]</sup>的研究曾证明相较于其他生理信号,皮电和心率相关特征对压力状态的识别效果最佳。Wang 等<sup>[13]</sup>用心率变异率相关的趋势和统计特征来识别驾驶者的压力状态取得了不错的结果。

为了进一步验证 MFTS 算法的有效性,本文还分析了单一过滤式算法和禁忌搜索相结合的算法(Chi\_TS、MI\_TS、WT\_TS)性能,比较结果如表 2 所示。由表 2 可知,与几种 Single Filter\_TS 算法的结果相比,MFTS 算法在具有最少特征的同时,中压分类准确率接近,而低压和高压的分类准确率、3 类的平均准确率都是最高。虽然中压的分类准确率低于 Random\_TS 算法,但 MFTS 算法选择的特征最少,不到 Random\_TS 算法的一半。而这一半减少的特征势必导致一部分信息的流失,但是相较于减少的特征数,中压状态分类准确率的下降幅度还是可以接受的,因此 MFTS 算法无论是准确率还是特征个数都明显更优。

将本文算法与采用 PhysioBank 数据库中相同数据集的其他算法进行比较,结果如表 3 所示。文献[13]将分别对应休息、高速公路和城市街道 3 种路况的低压、中压和高压 3 种状态进行了合并,分成了低压和中高压两类,即将实验驾驶路况分为了休息和驾驶两大类,而这两类的差异无疑非常明显,所以当文献[14]采用 PBFG+KBCS+LDA+PCA 方法对特征降维并进行二分类时,最终得到了很高的准确率;文献[15]采用了 SBL+PCA 的算法,最终对 3 种状态的识别率达到了 0.877;文献[16]采用

表2 基于禁忌搜索的各种特征选择算法效果比较

Table 2 Comparison of feature selection algorithms based on tabu search

方法	每类平均准确率			平均准确率	特征个数
	低压	中压	高压		
Single Filter_TS	Chi_TS	0.931	0.829	0.849	0.869
	MI_TS	0.927	0.827	0.876	0.877
	WT_TS	0.913	0.826	0.872	0.870
Random_TS		0.926	0.842	0.871	0.879
MFTS		0.945	0.828	0.886	0.886

组合融合(Combinatorial Fusion, CF)的方式降低特征维度,最终平均准确率达到了0.857。相较于其他作用于三分类的算法,MFTS算法取得了较好的平均准确率,验证了本文算法的有效性。

表3 同类研究的对比结果

Table 3 Comparison with similar studies

算法	平均准确率	分类
文献[15]	0.877	3
文献[16]	0.857	3
文献[14]	1.000	2
MFTS	0.886	3

## 5 结 论

本文采用皮电、心电和呼吸3种生理信号来识别驾驶员在3种不同路况下的压力状态。为了充分挖掘生理信号中所蕴含的与驾驶压力相关的信息,提取了包含时域、频域、小波域和非线性分析在内的多模态特征,提出了一种基于多种过滤式算法与禁忌搜索算法相结合的混合算法(MFTS)进行特征选择。实验结果表明该混合算法不仅有效地减少了特征冗余,消除了噪声,还提升了分类准确率。

## 参考文献:

- [1] JI Q, ZHU Z, LAN P. Real-time nonintrusive monitoring and prediction of driver fatigue[J]. IEEE Transactions on Vehicular Technology, 2004, 53: 1052-1068.
- [2] CONATI C. Probabilistic assessment of user's emotions in educational games[J]. Journal of Applied Artificial Intelligence, 2002, 16(7/8): 555-575.
- [3] FU R, WANG H, ZHAO W. Dynamic driver fatigue detection using hidden Markov model in real driving condition[J]. Expert Systems with Applications, 2016, 63(C): 397-411.
- [4] SAHAYADHAS A, SUNDARAJ K, MURUGAPPAN M, et al. Physiological signal based detection of driver hypovigilance using higher order spectra[J]. Expert Systems with Applications, 2015, 42(22): 8669-8677.
- [5] ASIR D, APPAVU S, JEBAMALAR E. Literature review on feature selection methods for high-dimensional data[J]. International Journal of Computer Applications, 2016, 136(1): 9-17.
- [6] HEALEY J A, PICARD R W. Detecting stress during real-world driving tasks using physiological sensors[J]. IEEE Transactions on Intelligent Transportation Systems, 2005, 6(2): 156-166.
- [7] SUBASI A. Automatic recognition of alertness from EEG by using neural networks and wavelet coefficients[J]. Expert Systems with Applications, 2005, 28(4): 701-711.
- [8] CHEN L L, ZHANG J, ZOU J Z, et al. A framework on wavelet-based nonlinear features and extreme learning machine for epileptic seizure detection[J]. Biomedical Signal Processing and Control, 2014, 10: 1-10.
- [9] SHEN E H, CAI Z J, GU F J. Mathematical foundation of C0 complexity[J]. Applied Mathematics and Mechanics, 2005, 26(9): 1083-1090.
- [10] GU Q, DING Y S, ZHANG T L. An ensemble classifier based prediction of G-protein-coupled receptor classes in low homology [J]. Neurocomputing, 2015, 154: 110-118.
- [11] HENRÍQUEZ P, ALONSO J B, FERRER M A, et al. Non-linear dynamics characterization of emotional speech[J]. Neurocomputing, 2014, 132: 126-135.
- [12] ACHARYA U R, FUJITA H, SUDARSHAN V K, et al. Application of entropies for automated diagnosis of epilepsy using EEG signals: A review[J]. Knowledge-Based Systems, 2015, 88(C): 85-96.
- [13] ACHARYA U R, SREE S V, SWAPNA G, et al. Automated EEG analysis of epilepsy: A review[J]. Knowledge-Based Systems, 2013, 45(3): 147-165.
- [14] WANG J S, LIN C W, YANG Y T C. A k-nearest-neighbor classifier with heart rate variability feature-based transformation algorithm for driving stress recognition[J]. Neurocomputing, 2013, 116: 136-143.
- [15] CHEN L L, ZHAO Y, YE P F, et al. Detecting driving stress in physiological signals based on multimodal feature analysis and kernel classifiers [J]. Expert Systems with Applications, 2017, 85(C): 279-291.
- [16] DENG Y, HSU D F, WU Z. Feature selection and combination for stress identification using correlation and diversity [J]. International Symposium on Pervasive Systems, 2013, 7363(1): 37-43.