

基于表面肌电和位姿信息融合的手势动作识别^①

杨庆华^② 金圣权 都明宇 王志恒^③

(浙江工业大学机械工程学院 杭州 310023)

(特种装备制造与先进加工技术教育部/浙江省重点实验室 杭州 310023)

摘要 针对仅通过表面肌电信号(sEMG)进行手势识别难以应对复杂手势的问题,提出一种基于表面肌电和位姿信息融合的手势识别方法。通过双阈值方法对信号活动段进行分割,提取表面肌电信号、位姿信号的特征,使用核主成分分析方法(KPCA)对提取特征进行降维融合,使提取特征中的非线性信息得到较好保留,最后通过随机森林(RF)分类器进行分类识别。实验结果显示,该方法对 10 名受试者的 11 种不同手势的最佳平均识别率为 98.23%,单个动作的识别准确率均在 90% 以上,验证了提出方法的可靠性。

关键词 表面肌电信号(sEMG); 位姿信号; 多模融合; 核主成分分析方法(KPCA); 随机森林(RF)

0 引言

手势作为一种有效传递信息手段,其非接触式、多样性的特点,使其成为人机交互领域一种常用的控制手段,手势识别技术也成为了智能假肢^[1]、康复训练^[2]、手语识别^[3]、智能家居^[4]等领域的热点话题。根据获取信息方式的不同,手势识别可以分为基于机器视觉^[5]以及基于穿戴式传感器^[6]2类。基于机器视觉的手势识别,受环境影响较大,遮挡、光照、背景都会影响识别准确率。而基于穿戴式传感器的手势识别克服了以上问题。其中数据手套可以精确地采集各个手指的运动信息,但其本身笨重又无法应用于手部残疾的人群。表面肌电信号(surface electromyography, sEMG)因为其蕴含信息丰富,检测用的非侵入式传感器受限制较少、便于携带的特点,越来越成为手势识别领域关注的重点^[7-9]。

不少研究者通过对 sEMG 信号进行分析并提取特征。谢小雨和刘喆颖^[10]通过动态时间规整(dynamic time warping, DTW)算法、Omari 等人^[11]使用

广义回归神经网络(generalized regression neural network, GRNN)、都明宇^[12]使用 BP 神经网络(back propagation neural network, BPNN)和 Cheng 等人^[13]使用卷积神经网络(convolutional neural network, CNN)对手部手势进行分类识别,均取得了不错的效果。但是仅使用 sEMG 进行手势识别往往会忽略手部动作的位姿信息,而通过不同种类的传感器采集信息,可以改善这一问题。Kundu 等人^[4]利用肌电和惯性传感器采集信号,并使用支持向量机(support vector machines, SVM)对 7 个动作进行识别并控制轮椅,可以达到平均 90.5% 的识别率。韩志昕和隋修武^[14]利用表面肌电信号、加速度信号和角速度信号,通过极限学习机对 9 种手势可实现 94.7% 的识别率,并挑选其中 6 种用于控制轮椅运动,效果较好。鲍磊等人^[15]提取 4 路 sEMG 与 3 轴加速度信号的特征,输入 SVM 对 4 名受试者执行的 6 种中国手语手势的最高识别率为 91.2%。周非等人^[16]使用 sEMG 信号及加速度信号,通过线性判别式对 5 种手势动作进行分类识别,识别结果较仅使用 sEMG 的方式准确率提

① 国家重点研发计划(2018YFE0125600)和浙江省基础公益研究计划(LGG19E050023)资助项目。

② 男,1964年生,博士,教授;研究方向:智能装备,机器人;E-mail: 1024898522@qq.com。

③ 通信作者,E-mail: wzh232@zjut.edu.cn。

(收稿日期:2022-03-18)

升 14.8%, 并成功应用于飞行器的控制。虽然目前利用 sEMG 结合多传感器的手部动作识别总识别率较好, 但用于识别的手势种类数量少, 绝大多数在 10 种以下。手势识别种类越多, 识别准确率越低。并且手势识别动作大多以单一的手腕或者手指动作为主, 不利于使用范围的拓展。

1 手势识别系统

本文建立的基于表面肌电和位姿信息融合的手势识别系统如图 1 所示。首先是通过 4 路肌电传感器采集手部动作 sEMG 信号, 利用固定于手背的惯性传感器记录手势加速度和速度信号; 然后以速度信号为基准, 使用双阈值分割方法划分活动段, 再通过时间轴同步的方式分割 sEMG 信号和加速度信号; 提取 sEMG 信号和速度、加速度信号特征, 通过核主成分分析 (kernel principal component analysis, KPCA) 的方式对特征进行降维融合, 最后使用随机森林 (random forest, RF) 法完成分类识别。



图 1 多模态信息采集系统

1.1 数据采集

实验中, 肌电信号采集设备使用傲意科技的 gForceOct 无线肌电仪, 采集拇短伸肌、指浅屈肌、尺侧腕屈肌和桡侧腕伸肌 4 路 sEMG 信号, 其中拇短伸肌和指浅屈肌主要与手指动作有关, 尺侧腕屈肌和桡侧腕伸肌主要与手腕动作有关。位姿信号使用维特智能的 BWT901CL 惯性传感器, 为保证可以采集到仅有手指动作手势的位姿信息, 将惯性传感器设置于手背处, 传感器布置位置如图 2 所示。其中 sEMG 信号采样频率为 1000 Hz, 速度和加速度信号

的采样频率为 100 Hz。

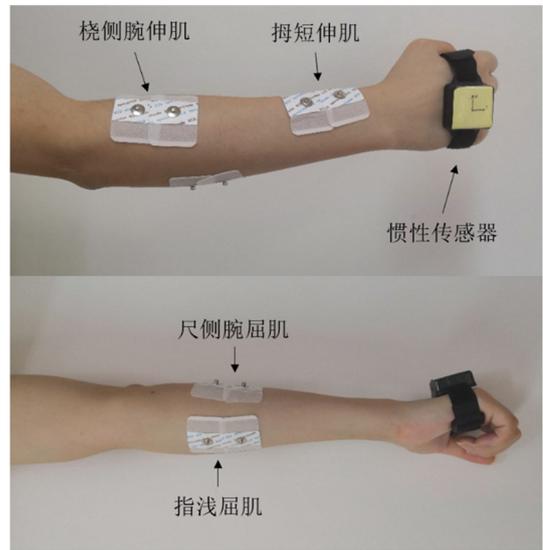


图 2 传感器布置位置

1.2 实验动作规范

实验参考 NinaPro 数据集的动作类型^[17], 从中选取 11 种手势、手腕动作进行实验, 具体形式如图 3 所示, 其中编号 0 为放松, 编号 1 ~ 11 分别为伸张、握拳、腕屈、腕伸、腕侧、右旋、左旋、指向、点赞、OK 和 V。每种动作执行过程中穿插 1 s 左右的放松时间, 每组动作共采集 2 min 的数据, 之后被试对象休息 5 min, 继续进行下一种动作的数据采集。

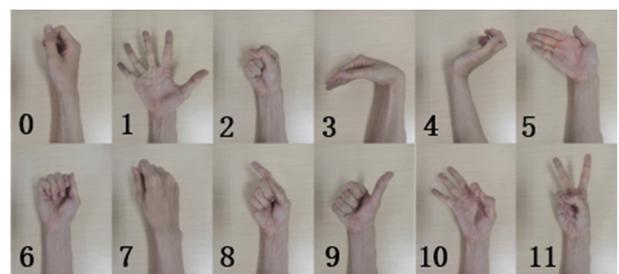


图 3 手势动作

本文被试对象共 10 人, 其中 8 人为男性, 2 人为女性, 年龄在 20 ~ 25 岁之间, 均为右利手且无任何肌肉神经疾病和残疾的健康人。为保证手势标准性及执行时间尽可能相同, 实验对象事先均进行过一定练习。

1.3 活动段分割

对采集信号手势动作段的分割会影响分类器训练的准确性^[18], 因此准确地对信号活动段进行提取

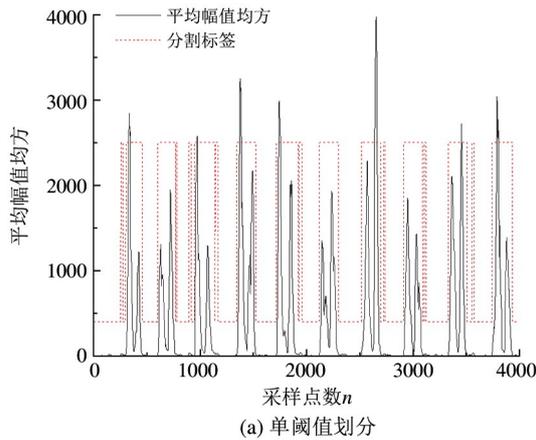
十分关键。本文采用平均能量双阈值分割的方式提取活动段,针对速度信号进行手势动作的活动段分割,再使用时间轴同步提取其余两信号活动段。首先通过滑动窗口方式计算速度信号总平均能量:

$$E(n) = \frac{1}{m} \sum_{i=n+1}^{n+m} v_{\text{sum}}(i)^2 \quad (1)$$

其中, m 为窗口宽度,其值为20; l 为滑动步长,其值为1; $v_{\text{sum}}(i)$ 为3轴速度信号绝对值之和。

设置阈值 $th1$,当速度信号总平均能量开始大于 $th1$ 时视其为待定活动段起点 A ,当其开始小于 $th1$ 时视其为待定活动段终点 B 。设置阈值 $th2$,待定活动段起点和终点之间的长度小于阈值 $th2$ 则视为干扰信号,将其舍去;满足双阈值限制条件的才视为动作活动段。其中 $th1$ 的取值由式(2)确定。

$$th1 = \alpha\sigma \quad (2)$$



其中, α 为增益系数,此处取2.5; σ 为放松状态下受试者速度信号总平均能量, $th1$ 的值向上取整。将阈值 $th2$ 均设置为15,即将持续时间少于150 ms的信号区段视为干扰信号。

将分割后的信号二值化,便于同步分割sEMG及ACC信号,二值化公式如式(3)所示。

$$y(n) = \begin{cases} 2500 & e(n) > th1, l_{A-B} > th2 \\ 400 & \text{其他} \end{cases} \quad (3)$$

其中, l_{A-B} 为两相邻待定活动段起点和终点之间的距离。

实验过程中受试者难免会因为不可控因素产生轻微抖动,对活动段分割产生干扰。以受试者 F 的指向动作为例,采用单阈值方法时很可能导致划分错误进而影响后续识别,如图4(a)所示;而双阈值方法则可以很好地改善这一问题,如图4(b)所示。

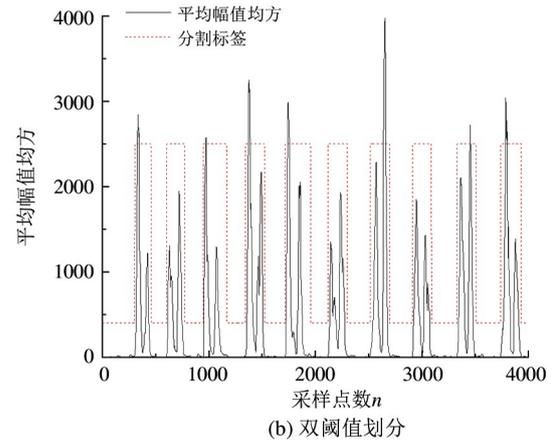


图4 划分动作段标签

1.4 特征提取

采集到的原始sEMG信号中存在较多干扰,特征提取如果不事先进行降噪处理会影响其信号质量,进而影响识别准确率。因此先用梳状陷波过滤50 Hz的工频干扰及其谐波,再用20~500 Hz的巴特沃斯带通滤波器,消去高频噪声及低频运动伪迹干扰,对经过降噪处理后的sEMG信号再进行特征提取。

针对sEMG信号,主要的特征提取方法有时域特征、频域特征、时频域特征、非线性特征等,其中时域特征实时性较好,提取简单,使用最为广泛。因此选用平均绝对值(absolute mean value, MAV)、均方

根(root mean square, RMS)、方差(variance, VAR)和过零点数(zero crossing, ZC)4个时域特征作为sEMG信号特征。特征提取方式采用滑动窗口法^[19],出于实时性考虑,滑动窗口宽度应在300 ms以内,本文选择200 ms为滑动窗口宽度,50 ms为滑动步长。

sEMG的MAV大小与其信号强度成正比,因此其可以反映选定肌肉在特定动作下的输出力度,MAV的表达式为

$$MAV = \frac{1}{N} \sum_{i=1}^N |x_i| \quad (4)$$

其中, x_i 为 i 时刻的肌电信号, N 为滑动窗口宽度,后续的时域特征中 x_i 和 N 具有的意义相同。

RMS 为信号平均能量的平方根,可以反映其能量水平,其公式为

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2} \quad (5)$$

VAR 可以反映信号与均值的差异程度,其公式为

$$VAR = \frac{\sum_{i=1}^N \left(x_i - \frac{1}{N} \sum_{i=1}^N x_i \right)^2}{N} \quad (6)$$

ZC 可以反映信号穿越零点的频率,其公式为

$$ZC = \sum_{i=1}^N \text{sgn}(-x_i x_{i+1})$$

$$\text{sgn}(x) = \begin{cases} 1 & x > 0, |x_i - x_{i+1}| \geq x_{th} \\ 0 & \text{其他} \end{cases} \quad (7)$$

其中, x_{th} 是为了防止噪声干扰导致的信号幅值在 0 值附近波动而设置的阈值,其值为标准差的 1%。

对 4 通道 sEMG 信号各提取 4 个时域特征,可以构成 16 维特征矩阵:

$$S_1 = [MAV_1, \dots, MAV_4, RMS_1, \dots, RMS_4, VAR_1, \dots, VAR_4, ZC_1, \dots, ZC_4] \quad (8)$$

针对速度和加速度信号,选用平均绝对值反映其总体输出力度。波形长度(waveform length, WL)是时间段中波形的累积长度,是信号复杂程度的度量,可以反映位姿信号的变化规律^[20],其公式为

$$WL = \sum_{i=1}^{N-1} |x_{i+1} - x_i| \quad (9)$$

对加速度和速度信号,各提取 3 维的 MAV 和 WL 特征,可以构成 12 维特征矩阵:

$$S_2 = [MAV_{x1}, MAV_{x2}, \dots, MAV_{z1}, MAV_{z2}, WL_{x1}, WL_{x2}, \dots, WL_{z1}, WL_{z2}] \quad (10)$$

1.5 降维融合

用于分类识别的特征维数过多会增加分类器的训练和识别时间,对特征进行降维可以改善这个问题。肌电信号和位姿信号中含有非线性的特征,而常用的主成分分析方法(principal component analysis, PCA)对线性数据降维处理有较好的效果,却无法兼顾非线性数据,而核主成分分析可以很好地解决这一点。以下是 KPCA 算法实现过程。

由于提取的信号不同特征尺度差异较大,所以需要先对特征值进行归一化处理,将特征值归一化为均值为 0、方差为 1 的数据,其公式为

$$x_0 = \frac{x - \mu}{\sigma} \quad (11)$$

其中, μ 为数据的均值, σ 为数据的方差。

利用高斯径向基核函数,将标准化处理后的数据由数据空间映射得到核矩阵,其中高斯径向基核函数公式为

$$K(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right) \quad (12)$$

其中, σ 为核参数,通过网格搜索选取使识别准确率最高的值; x 和 y 均为标准化后的数据。

而后中心化核矩阵,其公式为

$$K_c = K - I_N K - K I_N + I_N K I_N \quad (13)$$

其中, I_N 为 $N \times N$ 的矩阵,每一个元素都为 $1/N$, N 为样本数量; K 为核矩阵。

之后计算矩阵 K_c 的特征值,得到对应的特征向量为 $\lambda_1, \lambda_2, \dots, \lambda_n$, 因为特征值决定方差 v_1, v_2, \dots, v_n 的大小,特征值越大所含有效信息越多,按特征值降序对其进行排序,对应的特征向量也做出调整。通过施密特正交化方法,正交并单位化特征向量,得到 a_1, a_2, \dots, a_n 。

计算重新排序后的特征值累计贡献率 r_1, r_2, \dots, r_n , 取大于 95% 的部分,公式为

$$\frac{\sum_{i=1}^t \lambda_i}{\sum_{j=1}^n \lambda_j} \geq 95\% \quad (14)$$

选取 a_1, a_2, \dots, a_t 为降维融合后的数据,分别对 sEMG 时域特征及位姿特征使用 KPCA 降维后,将两者进行组合形成融合后的特征矩阵。

1.6 分类器分类

随机森林(RF)^[21]是一个包含多个决策树的分类器,并且其输出的类别是由所用树输出类别的众数而定,是 Bagging 算法的一种。它学习速度快,针对多种特征和大量数据处理效果都较好^[22]。

随机森林算法的基本思路如图 5 所示,假设训练集 S 中共有 n 个不同的样本,首先通过 Bootstrap 方法,有放回地从训练集 S 中抽取 n 次产生一个训练集;并将此步骤重复 K 次,随机产生 K 个训练集 S_1, S_2, \dots, S_K ; 随后通过分类回归树方法构建对应决策树 C_1, C_2, \dots, C_K , 对于每一棵决策树而言,各自从 M 个总特征中随机选择 m 个用于分类,利用每个决

策树对测试样本 X 进行分类得到 $C_1(X)$ 、 $C_2(X)$ 、 \dots 、 $C_K(X)$; 最后通过投票的方法, 将 K 个决策树中输出最多的类别作为测试样本 X 的分类结果。

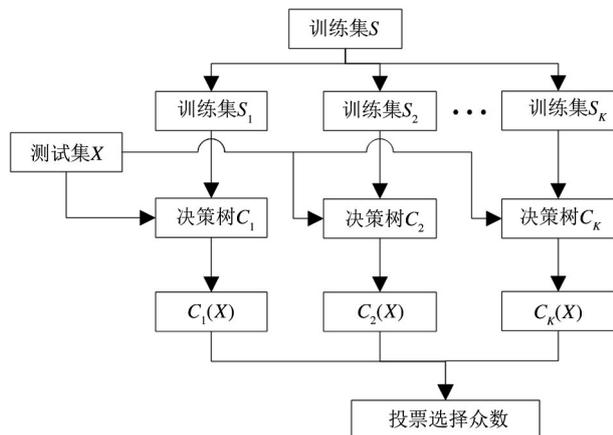


图5 随机森林分类器

2 实验与分析

2.1 随机森林参数确定

通过 RF 进行分类需要确定的参数有随机选择的特征个数 m 和构建的决策树总数 N , 对于随机选择的特征个数 m , 其决定方式为

$$m = \sqrt{M} \quad (15)$$

其中, M 为总特征数, m 向下取整。

对于决策树总数 N , 选用受试者 B 的数据进行分析, N 的取值范围在 100 ~ 1000 之间, 每隔 50 进行一次测试, 重复 20 次, 识别准确率取平均值, 其结果如图 6 所示。

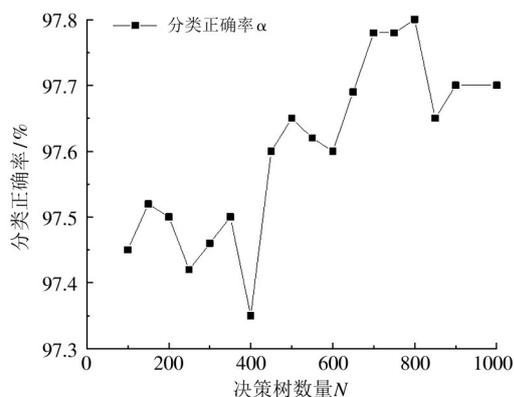


图6 决策树数量的影响

可以看出, 在决策树总数 N 为 500 时识别准确

率较好, 而随机森林分类器训练及识别时间会随着决策树数量增加而增加, 决策树棵树在 600 以上时, 对于准确率的提升不大, 但训练和识别所花的时间则会大大增加, 因此选用 500 作为决策树总数。

2.2 分类识别

表面肌电信号和位姿信号相结合的方式可以提高手势识别的正确率。为说明其各自的作用, 用不同的特征组合构成特征矩阵进行识别分析, 结果如表 1 所示。

表1 实验组特征向量列表

	sEMG 特征组 E1	位姿 特征组 E2	融合 特征组 E3
特征矩阵	S_1	S_2	$S_1 + S_2$

表 1 所示的 3 组实验均使用 RF 进行分类, 各实验人员的每个动作均取 500 组, 共 5500 组作为实验总样本作为实验总样本; 随机选取样本数据的 60% 作为训练样本对分类器进行训练, 40% 作为测试样本进行测试分析; 每组均进行 20 次, 识别准确率取平均值。实验所采用的计算机设备配置如下: 处理器为 Intel Corei5-1035G1, 主频为 1.00 GHz, 内存为 16 GB, 操作系统为 Windows10, 识别结果如图 7 所示。

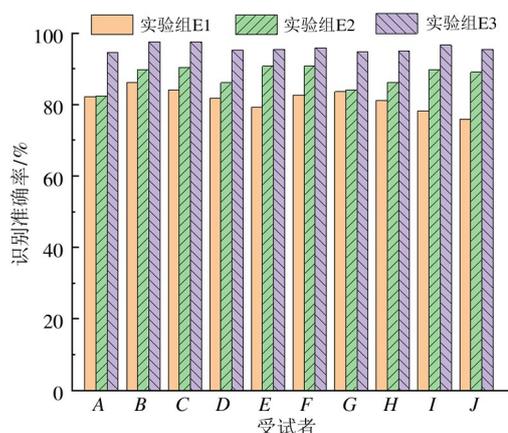


图7 不同实验组识别准确率

如图 7 所示, 对于所有受试者而言, sEMG 与位姿信号组成的特征向量识别准确率均高于两者单独用于识别, 其中实验中 $E1$ 、 $E2$ 、 $E3$ 中最高识别率分别为 86.10%、90.78% 和 98.23%。

以受试者 *B* 的数据为例,特征选用融合特征组 *E3*,选择文献[11]、[12]、[15]中使用的 BPNN、SVM 和 GRNN 分类器与本文的 RF 方法进行对比。其中 BPNN 的初始化权值阈值取(-1, 1)之间的随机数,训练后的最优模型用于识别;SVM 的惩罚系数及核参数和 GRNN 的扩展速度均通过网格搜索法寻找最优值。各分类方法的分类识别准确率如表 2 所示,通过对比可以看出使用 RF 进行分类识别准确率最高。

表 2 不同分类器识别准确率

	RF	BPNN	SVM	GRNN
准确率/%	98.15	93.12	95.21	85.54

通过 KPCA 降维可以减少 RF 的训练时间,将 sEMG 和位姿信号通过 KPCA 降维融合和未使用 KPCA 进行降维的各受试者识别准确率及识别所用分类器训练时间如表 3 所示。由结果可知,将 sEMG 和位姿信号特征通过 KPCA 降维融合后再通过 RF 分类识别在加快训练时间的同时还可以微略地提

表 3 使用 KPCA 降维前后识别准确率及分类器训练时间

	未使用 KPCA		KPCA 降维融合	
	准确率/%	时间/s	准确率/%	时间/s
受试者 <i>A</i>	94.45	7.41	96.23	3.79
受试者 <i>B</i>	96.83	4.34	98.15	2.55
受试者 <i>C</i>	97.59	6.45	98.23	3.71
受试者 <i>D</i>	95.17	5.49	95.78	3.28
受试者 <i>E</i>	95.21	3.55	97.58	2.19
受试者 <i>F</i>	94.97	4.23	97.69	2.51
受试者 <i>G</i>	95.71	3.97	96.81	1.97
受试者 <i>H</i>	94.81	4.42	95.35	3.82
受试者 <i>I</i>	96.77	6.75	97.85	4.27
受试者 <i>J</i>	96.03	5.42	97.72	3.02

升准确率,识别准确率为 $97.14 \pm 1.03\%$ 。

以受试者 *C* 为例,表 4 展示其识别结果的混淆矩阵,矩阵每一列数值表示真实数据被预测为该类的数目,每一行数值表示数据的真实归属类别。从表中可知,KPCA 降维融合的手势识别方法在保障总体识别率高的同时,单一动作的识别率均在 95% 以上。

表 4 受试者 *C* 手势动作识别混淆矩阵

真实类	预测类											实际准确率/%
	伸张	握拳	腕屈	腕伸	腕侧	右旋	左旋	指向	点赞	OK	V	
伸张	198	0	0	0	0	0	0	1	0	1	0	99.0
握拳	0	196	1	2	0	1	0	0	0	0	0	98.0
腕屈	0	0	199	0	0	0	0	0	0	1	0	99.5
腕伸	0	1	1	197	0	0	0	0	0	0	0	99.0
腕侧	0	0	0	1	194	1	1	0	3	0	0	97.0
右旋	0	0	0	0	0	195	5	0	0	0	0	97.5
左旋	0	0	0	0	2	0	198	0	0	0	0	99.0
指向	0	0	0	0	2	0	0	192	6	0	0	96.0
点赞	0	0	0	0	0	0	0	2	198	0	0	99.0
OK	0	0	0	2	0	0	0	2	0	195	1	97.5
V	0	0	0	0	0	1	0	0	1	0	198	99.0
样本准确率/%	100	99.5	99.0	97.5	98.0	98.5	97.1	97.5	95.2	99.0	99.5	98.2

对比传统 PCA 方法和 KPCA 方法的效果,图 8 展示了各动作的识别准确率。可以看到,PCA 方法的识别准确率相比于 KPCA 有很大程度的下降,这

是因为 KPCA 降维方法可以更好地保留数据的非线性特征。所以,使用 KPCA 降维的方式优于 PCA 降维方式。

2.3 疲劳状态效果验证

肌肉疲劳状态下 sEMG 信号的幅度以及频率会出现变化,提取特征会与正常状态下有差异,对手势识别算法是一个挑战。让受试者按 1.2 节实验规范完成正常数据的采集后,先无间歇地执行单个手势动作 1~3 min,直到受试者明显感觉到手部肌肉出现疲劳;再采集该手势数据,不断重复这一过程,直到全部 11 种手势数据均采集完毕。之后采用正常数据提取特征训练分类器,使用疲劳数据提取的特征进行验证,各手势的识别准确率如表 5 所示。可

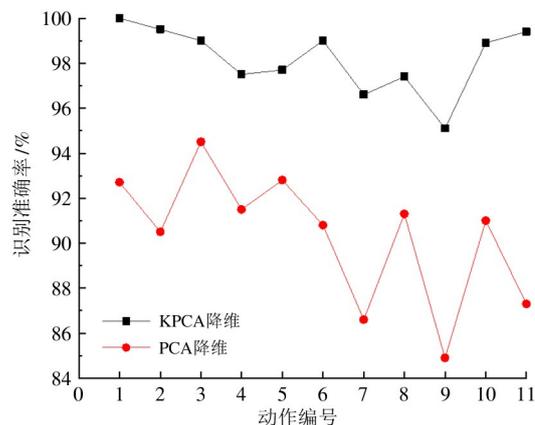


图 8 不同降维方式动作识别准确率

表 5 疲劳状态下各受试者识别准确率

受试者	A	B	C	D	E	F	G	H	I	J
准确率/%	93.25	94.56	95.76	92.95	93.82	94.13	92.38	91.87	93.67	93.51

以看出,所有受试者的手势识别准确率为 $93.59 \pm 1.10\%$,识别效果依旧较好,本文所提出方法对肌肉疲劳状态鲁棒性较好。

2.4 时效分析

本文所使用方法对于单个手势识别总延时小于 80 ms,其中滑动窗口延时为 50 ms,3 种信号特征提取平均时间为 1.5 ms,使用 KPCA 降维融合平均耗时为 27.4 ms,RF 分类识别耗时小于 1 ms。因此本文所提出方法可以满足人机交互领域中的实时性要求。

3 结论

本文提出一种基于 sEMG 和位姿信息融合的手势识别方法。通过速度信号双阈值分割方法对活动段进行划分,在提高准确率的同时降低了操作的复杂性;使用 KPCA 方法对提取的 sEMG、加速度和速度信号特性进行降维融合,在减少分类器负担的同时保留了信号的非线性特征,进而极大减少分类器训练耗时并略微提高识别准确率。相较于使用 PCA 方法进行特征降维,本文方法大幅提高了个别动作的识别准确率,证明了在手势识别中非线性特征的重要性。设计并使用 RF 对 11 种不同手势进行分类,并与不同分类器进行比较,RF 的识别效果最好,平均最佳识别率最高可达 98.23%。所提方

法在保证较高平均识别率的同时,对单一动作的识别准确率也保持在 90% 以上,具有良好的识别性能。该识别算法对单个手势的识别延时小于 80 ms,可满足实际应用的需求。

本文所提的手势识别算法针对 10 种以上的手指、手腕动作进行识别,可以得到较好的识别效果,这有利于手势识别在实际工程应用中的范围扩展,为如智能假肢动作、智能家居控制等提供更多的选择。单种动作识别的准确率均在较高的水平,也有利于康复训练等工程使用中安全性的提高。

参考文献

- [1] PAN L, ZHANG D, LIU J, et al. Continuous estimation of finger joint angles under different static wrist motions from surface EMG signals[J]. Biomedical Signal Processing and Control, 2014,14:265-271.
- [2] KHUSHABA R N, KODAGODA S, TAKRURI M, et al. Toward improved control of prosthetic fingers using surface electromyogram (EMG) signals[J]. Expert Systems with Applications, 2012,39(12):10731-10738.
- [3] 王震,张震,姚松丽,等.一种基于肌电信号的踝关节动作预测方法的研究[J].高技术通讯,2010,20(11):1173-1177.
- [4] LI Y, CHEN X, ZHANG X, et al. Assign-component-based framework for chinese sign language recognition using accelerometer and sEMG data[J]. IEEE Transactions on Biomedical Engineering, 2012,59(10):2695-2704.

- [5] KUNDU A S, MAZUMDER O, LENKA P K, et al. Hand gesture recognition based omnidirectional wheelchair control using IMU and EMG sensors[J]. *Journal of Intelligent and Robotic Systems*, 2018,91(3-4):529-541.
- [6] LU D L, YU Y L, LIU H P, et al. Gesture recognition using data glove: an extreme learning machine method [C]//2016 IEEE International Conference on Robotics and Biomimetics. Qingdao: IEEE, 2016:1349-1354.
- [7] CAO T, LIU D, WANG Q, et al. Surface electromyography-based action recognition and manipulator control[J]. *Applied Sciences-Basel*, 2020,10(17):5823.
- [8] 丁其川, 熊安斌, 赵新刚, 等. 基于表面肌电的运动意图识别方法研究及应用综述[J]. *自动化学报*, 2016, 42(1):13-25.
- [9] 李自由, 赵新刚, 张弼, 等. 基于表面肌电的意图识别方法在非理想条件下的研究进展[J]. *自动化学报*, 2021,47(5):955-969.
- [10] 谢小雨, 刘喆颖. 基于 DTW 算法的肌电信号手势识别方法[J]. *计算机工程与应用*, 2018,54(5):132-137.
- [11] OMARI F A, HUI J, MEI C, et al. Pattern recognition of eight hand motions using feature extraction of forearm EMG signal[J]. *Proceedings of the National Academy of Sciences, India-Section A*, 2014,84(3):473-480.
- [12] 都明宇, 王志恒, 荀一, 等. 多模式人手动作分类识别方法[J]. *中国机械工程*, 2019,30(12):1474-1479.
- [13] CHENG Y, LI G, YU M, et al. Gesture recognition based on surface electromyography-feature image [J]. *Concurrency and Computation: Practice and Experience*, 2021,33(6):1-13.
- [14] 韩志昕, 隋修武. 基于多源信息融合的肌电轮椅智能控制技术[J]. *现代制造工程*, 2020(9):136-144.
- [15] 鲍磊, 罗志增, 席旭刚, 等. 融合表面肌电和加速度的手势动作识别[J]. *传感技术学报*, 2019,32(12):1843-1848,1863.
- [16] 周菲, 蔡晨晓, 郑标. 基于多源信息融合的手势智能交互系统[J]. *信息与控制*, 2019,48(4):452-458, 468.
- [17] ATZORI M, GIJSBERTS A, HEYNEN S, et al. Building the Ninapro database: a resource for the biorobotics community[C]//2012 4th IEEE RAS and EMBS International Conference on Biomedical Robotics and Biomechanics. Rome: IEEE, 2012:1258-1265.
- [18] 段有康, 陈小刚, 桂剑, 等. 基于相位划分的下肢连续运动预测[J]. *浙江大学学报(工学版)*, 2021,55(1):89-95.
- [19] SHULTZ A H, LAWSON B E, GOLDFARB M. Variable cadence walking and ground adaptive standing with a powered ankle prosthesis[J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2016, 24(4):495-505.
- [20] PHINYOMARK A, PHUKPATTARANONT P, LIMSAKUL C. Feature reduction and selection for EMG signal classification[J]. *Expert Systems with Applications*, 2012,39(8):7420-7431.
- [21] 高清. 基于肌电信号的手势动作识别[D]. 南京: 东南大学, 2019.
- [22] BREIMA N. Random forests [J]. *Machine Learning*, 2001, 2001,45(1):5-32.

Gesture action recognition based on fusion of surface electromyography and pose information

YINAG Qinghua, JIN Shengquan, DU Mingyu, WANG Zhiheng

(College of Mechanical Engineering, Zhejiang University of Technology, Hangzhou 310023)

(Key Laboratory of Special Purpose Equipment and Advanced Processing Technology,

Ministry of Education and Zhejiang Province, Zhejiang University of Technology, Hangzhou 310023)

Abstract

Aiming at the problem of the difficulty of complex gestures recognition just by surface electromyography (sEMG) signals, a gesture recognition method based on the fusion of surface electromyography and pose information is proposed. First, the signal active segments are segmented by the dual-threshold method, and then the features of the surface electromyography signal and the pose signal are extracted. Second, the kernel principal component analysis (KPCA) is used to perform dimensionality reduction and fusion on the extracted features so that the non-linear information in the extracted features is preserved well. Finally, the random forest (RF) classifier is utilized for classification. The experimental results show that the highest average accuracy of the proposed method is 98.23% for 11 different gestures of 10 subjects, and every individual gesture is above 90%, which verifies the reliability of the proposed method.

Key words: surface electromyography (sEMG), position signal, multi-mode fusion, kernel principal component analysis (KPCA), random forest (RF)