

基于 HPSO-SVM 的多传感器手语识别方法研究

刘闯闯 朱正伟

(常州大学 微电子与控制工程学院 常州 213000)

摘要: 为了提高手语识别准确率,提出一种基于混合粒子群优化的支持向量机(HPSO-SVM)的多传感器手语识别方法。在原始数据采集阶段,利用 ZTEMG-2000 肌电传感器采集人体手臂表肌电信号、MPU6050 传感器采集右手加速度和角速度信号;在数据处理阶段,增加一个自适应容错长度,提高了短时能量法提取活动段的精度;在分类方法阶段,通过混合粒子群算法(HPSO)寻找出支持向量机(SVM)的惩罚因子和核函数参数的最优组合,优化了 SVM 模型。实验上,对每名受试者分别执行的 5 种中国手语进行识别,平均识别率达到了 96.78%。该方法利用数量较少的、经济实惠的传感器对手语进行识别,且识别准确率较传统 SVM 算法提高了 5%,展现了该方法在手语识别上的优越性。

关键词: 手语识别;肌电传感器;支持向量机;混合粒子群算法;特征融合

中图分类号: TP391.4 **文献标识码:** A **国家标准学科分类代码:** 520.2040

Research on multi-sensor sign language recognition method based on HPSO-SVM

Liu Chuangchuang Zhu Zhengwei

(School of Microelectronics and Control Engineering, Changzhou University, Changzhou 213000, China)

Abstract: In order to improve the accuracy of sign language recognition, this paper proposes a multi-sensor sign language recognition method based on hybrid particle swarm algorithm support vector machine (HPSO-SVM). In the raw data collection stage, the ZTEMG-2000 EMG sensor is used to collect the EMG signal of the human arm surface, and the MPU6050 sensor is used to collect the right-hand acceleration and angular velocity signals. In the pretreatment phase of the experiment, short-term energy method, optimized by altering the adaptive fault tolerance length, is introduced to improve the extraction accuracy of the active segment. In the classification method stage, the optimal combination of the penalty factor of the support vector machine (SVM) and the kernel function parameter is found through the hybrid particle swarm algorithm (HPSO), and the SVM model is optimized. In experiments, the five Chinese sign languages performed by each subject were recognized, and the average recognition rate reached 96.78%. This method uses a relatively small number of more economical sensors to recognize sign language, and the recognition accuracy is 5% higher than that of the traditional SVM algorithm, demonstrating the superiority of this method in sign language recognition.

Keywords: sign language recognition; electromyography sensor; support vector machine; hybrid particle swarm algorithm; feature fusion

0 引言

随着时代的进步,人们对弱势群体的关注度越来越高,尤其是残疾人士,为了实现聋哑人与普通人之间的正常沟通,本文提出了一种混合粒子群优化的支持向量机(hybrid particle swarm algorithm support vector machine, HPSO-SVM)算法来实现手语识别。根据采集数据方式的不同,通常将手语识别分为 3 类,分别为基于图像的手语识别、基

于数据手套的手语识别以及使用多传感器的手语识别^[1],三者的优缺点如下。

基于图像的手语识别技术。大多数实验者利用 Kinect 获取深度图像和骨骼信息来识别手势^[2-4],其优点是无需佩戴设备,且理论研究和算法的实现都比较成熟,但是基于图像的手语识别对环境的要求比较高,在识别过程中会受到光线、背景环境等因素的影响,因此未能得到广泛应用。

基于数据手套的手语识别^[5-7]。此方法虽然可以精准

获取每个手指时间及空间轨迹信息,但系统穿戴复杂、设备不易携带,而且价格昂贵,难以大量推广。

由于图像和数据手套的识别方法存在弊端,基于表面肌电和陀螺仪的多传感器手语识别方法更具优势;与基于图像的手语识别相比,其避免了环境的干扰;与基于数据手套^[8]手语识别相比,其在价格上更具优势。目前利用 MYO 臂环传感器采集数据的识别方法居多,但其价格昂贵,对于聋哑人家庭的负担较大^[9]。因此,需要设计一种更加实惠、有效的传感器组合装置对手语进行识别。

日前已有学者利用多传感器融合技术对手语识别进行了研究。张启忠等^[10]以关联维和分维数作为表面肌电信号的特征向量,采用 SVM 作为分类器对 4 种手势进行识别,识别率达到 91%,准确率有待提高。Wu 等^[11]使用 4 通道肌电信号采集数据,采用信息增益的方法选择特征,使用决策树算法和神经网络算法对 10 种手势进行识别,识别率分别为 85% 和 94%。Cheng 等^[12]使用 4 通道肌电信号和加速度计传感器对 6 种手势进行识别,准确率达到 94%,证实了多传感器融合的可行性。张永强等^[13]提出一种基于 10 通道肌电信号和三轴加速度计的方法,其对 16 种手语进行识别的准确率达到 95%,效果较好,但使用了较多的传感器,性价比不高。

为了平衡传感器^[14]价格和手语识别准确率的矛盾,本文使用两种价格更为低廉的传感器进行融合以实现高准确率手语识别。首先,利用 ZTEMG-2000 传感器采集肌电信号,在此基础上,增加了 MPU6050 传感器以完成对加速度

和角速度信号的采集,从而获取较为全面的手语数据信息。其次,在利用短时能量法对数据段进行提取时,通过增加了一个自适应容错长度,优化了该算法,从而有效地提取出活动段。此外,在特征选择上,对比不同传感器的数据特征分布,选取聚类效果好的特征,并对不同的特征融合进行分析,找出了最优特征组合,并将其组成特征矩阵。最后,选择分类效果较好的 SVM 作为分类器,利用混合粒子群算法优化其参数,对 5 种中国手语进行识别。实验结果表明,本文提出的 HPSO-SVM 在手语识别率上较传统 SVM 算法有 5% 的提高,为实现一种实惠通用的手语识别机器的研发提供了可能。

1 方法

1.1 系统设计方案

本文研究的动态手语识别系统原理如图 1 所示。该系统采集装置由 2 个肌电信号采集传感器和 1 个六轴陀螺仪传感器组成,通过 Arduino 开发板分别采集左右手肌电信号和主利手(一般为右手)的角速度、加速度信号,并利用均值滤波法对数据进行滤波;先用改进的短时能量法提取相应的数据活动段,之后再从中分别提取肌电信号绝对平均值、中值频率、小波系数绝对平均值等特征,加速度和角速度信号的均值、方差等特征,将其组成 26 维度的特征矩阵并构建数据集,最后将得到的数据集划分为测试集和训练集,利用混合粒子群算法优化的支持向量机对 5 种手语动作进行分类识别,实现对 5 种动作的精准分类。

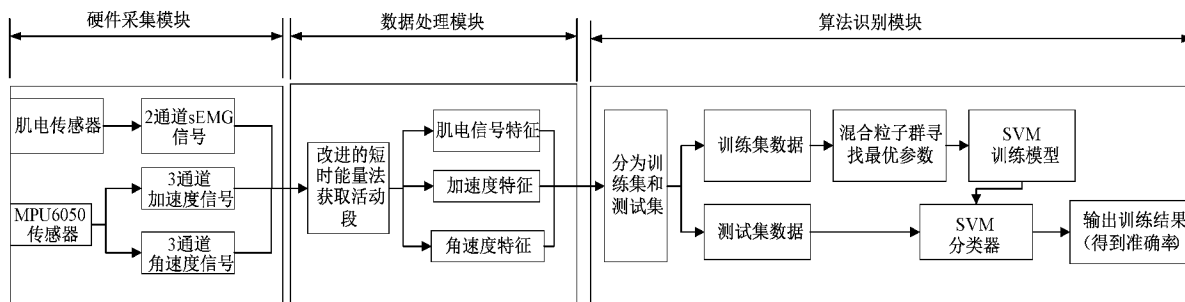


图1 动态手语识别系统原理

1.2 活动段检测

由于皮肤表面肌电信号非常微弱,在采集的过程中容易受到工频干扰、白噪声等干扰^[15],为了提高信号的信噪比,需要对采集的加速度、角速度,肌电信号进行均值滤波。在实际手语表达中,两个手语动作之间有明显的活动间隙,在此期间,肌肉处于静息状态,表现为一段低幅值的波形。在执行某个手语动作时,人们的手指、手掌和手臂都会挥动,这时肌肉会进行收缩,肌电信号会有强烈的幅值变化。表面肌电信号可以反映肌肉收缩的情况,因此可以用来提取活动段。

图2显示的是利用短时能量法检测活动段的波形图,其原理如下。

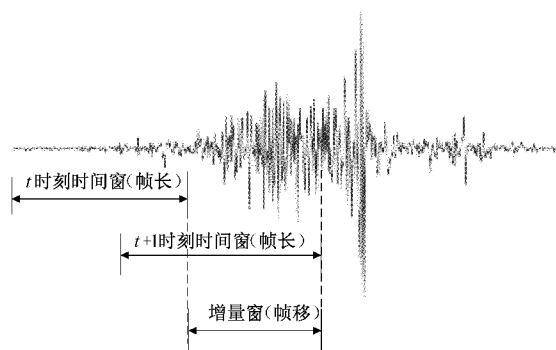


图2 活动段检测波形图

根据采样频率和采样数据长度设置一个帧长为 N 的滑动窗口和帧移 d , 因为肌电信号在静息状态下的短时能量较小, 在运动状态下的短时能量较大, 所以设置一个阈值, 当某一帧的能量大于最大能量的 5% 时, 记录下该帧为起始帧; 当某一帧的能量小于最大能量的 5% 时记录下该帧为终止帧。但是该方法存在缺陷, 在一些复杂的手语动作中不能正确地提取出有效活动段。

为了能提取复杂动作的完整活动段, 本文对此方法进行了改进, 改进后的短时能量法流程如下。

1) 初始化时间窗口长度 N , 增量窗 d , 帧移距离 len , 短时能量门限值 $energy_1, energy_2$ (本文根据采集数据频率设置 $N = 32, d = 10, len = n/d, n$ 为数据长度, $energy_1$ 为起始门限值, $energy_2$ 为终止门限值)。

2) 使用分帧函数, 将活动段分为 n 帧, 计算每一帧的短时能量 E 如式(1)所示。

$$E(i) = \frac{1}{N \times k} \sum_{j=1}^N \sum_{i=1}^k x_{ij}^2, i = 1, 2, \dots, n \quad (1)$$

式中: k 为肌电通道数。

3) 判断起始点位置, 与设置的短时能量进行比较, 当 $E(i) \geq energy_1$ 时, 记录此时的位置 x_1 。

4) 某些动作中间会出现短暂的间隔, 这样会导致识别到的活动段长度和真实不符合, 所以设置一个容错长度 l , 本文取 $l = len/6$, 其判断过程如下: 如果 $E(x_1 + j) \leq energy_1, x_2 = x_1 + j, j \in 1, \dots, n - x_1$, 则执行步骤 5)。

5) 如果 $E(x_2 + m) \leq energy_2, m \in 1, \dots, l$, 则输出起始点 x_1 和终止点 x_2 , 结束; 否则返回步骤 4)。

因为手语动作的数据都是同步采集的, 5 种手语动作都有右手参与, 本文先对右手表面肌电信号进行活动段的分割, 并根据得到的活动段位置映射到左手肌电信号, 以及右手的加速度和角速度信号, 这样就能获得所有数据的活动段。

以动作“对不起”为例, 其动作在运动过程中有短暂的停息, 如果用传统的活动段检测法则不能完整的提取整个活动段, 但在改良之后的提取算法能精确的提取到活动段, 为后续提取精确特征提供了保障。图 3(a) 展示的是原始肌电信号, 从中可以看出该信号既包含静息状态下的肌电信号也包含运动状态下的肌电信号, 图 3(b) 表示的原始肌电信号每一帧对应的短时能量, 图 3(c) 展示的是利用改进的短时能量法提取出的活动段, 可以看出利用该方法能有效地提取活动段数据。

1.3 特征提取

在获取完活动段数据后, 需要对肌电信号和加速度、角速度信号进行特征提取。肌电信号能够反映手指屈伸的强度, 对于手指的细微动作有较好的判断, 但肌电信号比较微弱, 传感器的位置要求比较高, 六轴传感器可以检测手的大尺度运动, 对于一些手臂运动类的动作有较好的区分^[16], 因此将两者的特征组合, 会对动作区分有很大的

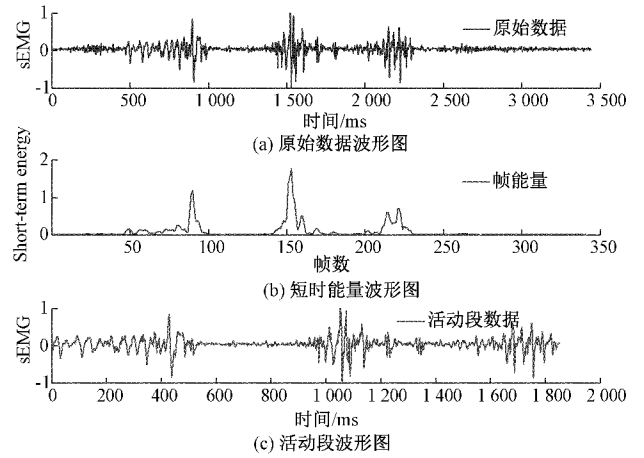


图3 改进短时能量法提取活动段图

帮助。经实验研究, 肌电信号的时域特征、频域特征、时频特征, 以及加速度的时域特征都对动作区分有帮助, 其中部分特征如下。

1) 平均绝对值: 平均绝对值是通过判断肌电信号活动强度来判断幅度变化, 从而区分不同的动作, 如式(2)所示。

$$MVA = \frac{1}{N} \sum_{i=1}^N |X_i| \quad (2)$$

式中: X_i 为肌电数据, N 为数据长度。

2) 中值频率: 中值频率是将功率谱求和之后取其 1/2 的值, 也就是总功率的 1/2 对应的频率, 如式(3)所示。

$$F_z = \frac{1}{2} \sum_{k=1}^N |fft(X_k)| \quad (3)$$

式中: X_k 为肌电数据, N 为数据长度。

3) 小波系数绝对平均值: 将信号进行 3 层小波分解, 得到小波系数, 然后对其求平均值, 如式(4)所示。

$$\begin{cases} c_i = (cD_1, cD_2, cD_3, cA_3) \\ MAV_{c_i} = \frac{\sum_{i=1}^n |c_i|}{n_i} \end{cases} \quad (4)$$

式中: c_i 为小波系数。

1.4 算法设计

1) 支持向量机

为确保手语识别的分类精度, 本文选用鲁棒性较好的支持向量机^[17-18]作为分类器。通过混合粒子群算法寻找支持向量机最优参数, 从而提高识别率。

最早提出的支持向量机是用于求解二分类问题, 对于线性可分性问题, 假定存在样本集 $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, 我们的目标是寻找一个超平面, 通过构造分隔超平面, 将数据分为 1, -1 两个类别, 超平面方程如式(5)所示。

$$\omega \cdot x_i + b = 0 \quad (5)$$

找出一个超平面, 使距离超平面最近的点之间的距离

达到最大值,并由此得到一个有约束的优化问题,如式(6)所示。

$$\begin{cases} \min & \frac{\|\omega\|^2}{2} \\ \text{s. t.} & y_i(\omega \cdot x_i + b) \geq 1, i = 1, 2, \dots, m \end{cases} \quad (6)$$

考虑到容错性的问题加入了惩罚因子 c 和松弛变量 ξ_i 。其约束条件如式(7)所示。实际上,松弛变量 ξ_i 的值指示了对应的点到群的距离,数值越大,距离越远;惩罚因子 c 用来衡量分类损失和最大间隔之间的关系,也就是用来控制目标函数中两个(“寻找分类间隔最大的超平面”和“保证数据点偏差量最小”)之间的权重^[19]。

$$\begin{cases} \min \left\{ \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n \xi_i \right\} \\ \text{s. t.} & y_i(\omega \cdot x_i + b) \geq 1 - \xi_i, i = 1, 2, \dots, m \end{cases} \quad (7)$$

在对偶问题上,可用拉格朗日乘法解,由此可得到拉格朗日方程如式(8)所示。

$$L(\omega, b, \alpha) = \frac{\|\omega\|^2}{2} + \sum_{i=1}^n \alpha_i (1 - y_i(\omega x_i + b)) \quad (8)$$

式中: $\alpha_i \geq 0, i = 1, 2, \dots, n, \alpha_i$ 为拉格朗日乘子。

方程对 ω, b 求偏导,可以确定线性可分的最优分类面如式(9)所示。

$$f(x) = \text{sgn}(\alpha_i y_i (x_i \cdot x) + b) \quad (9)$$

对于线性不可分的问题,则需要加入核函数,一般的核函数有多项式核函数、径向基核函数、柯西核函数等,本文选择使用径向基核函数,其表达式如式(10)所示。

$$k(x_i, x) = \exp\left(-\frac{\|x - x_i\|^2}{\sigma^2}\right) \quad (10)$$

式中: σ 为核函数参数。

在引入核函数后,分类判别函数如式(11)所示。

$$\begin{cases} \max J(a) = \max \left\{ \sum_{i=1}^l a_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l a_i a_j y_i y_j K(X_i \cdot X_j) \right\} \\ f(X) = \text{sgn} \left(\sum_{i=1}^n a_i y_i K(X \cdot X_i) + b^* \right) \end{cases} \quad (11)$$

通过上式可以看出,惩罚因子 c 和核函数参数 σ 对支持向量机识别准确率有很大的影响,为了使系统获得最优的参数,本文使用混合粒子群算法对这两个参数寻优,以获得最好的支持向量机模型,提高手语识别的准确度。

2) 混合粒子群算法优化的支持向量机

标准的粒子群算法具有快速逼近最优解的能力,可以有效的对系统的参数进行优化^[20]。标准的粒子群算法的速度和位置如式(12)所示。

$$\begin{cases} v_{id}^{k+1} = \omega \cdot v_{id}^k + c_1 r_1 \cdot (p_{id}^k - x_{id}^k) + c_2 r_2 \cdot (p_{gd}^k - x_{id}^k) \\ x_{id}^{k+1} = x_{id}^k + v_{id}^{k+1} \end{cases} \quad (12)$$

式中: ω 为惯性权重, v_{id}^k 为上次速度, p_{id}^k 为个体极值位

置, p_{gd}^k 为群体极值位置, c_1 为认知学习因子, c_2 为社会学习因子, r_1, r_2 为(0,1)的随机值。

$\omega \cdot v_{id}^k$ 为一个惯性系数,下一次迭代包含了先前迭代的速度信息。 $c_1 r_1 \cdot (p_{id}^k - x_{id}^k)$ 为自身最优因子,可以理解为历史上自身离最优解最近的位置^[21]。 $c_2 r_2 \cdot (p_{gd}^k - x_{id}^k)$ 为社会因子,因为这是所有粒子与最优解之间最接近的位置。

但标准的粒子群算法容易产生早熟的问题,容易陷入局部最小值,混合粒子群算法(HPSO)将社会因子分解为了两个部分:局部社会因子和全局社会因子,融合两种社会因子可以降低陷入局部最优解的情况,其公式如式(13)所示。

$$\begin{cases} v_{id}^{k+1} = \omega \cdot v_{id}^k + c_1 r_1 \cdot (p_{id}^k - x_{id}^k) + v_{pd}^k \\ v_{pd}^k = c_2 r_2 (q \cdot (p_{gd}^k - x_{id}^k) + (1-q)(p_l^k - x_{id}^k)) \\ x_{id}^{k+1} = x_{id}^k + v_{id}^{k+1} \end{cases} \quad (13)$$

式中: p_l^k 为局部最优解, $q \in (0,1)$ 为全局变量所占权重。混合粒子群算法流程如图4所示。

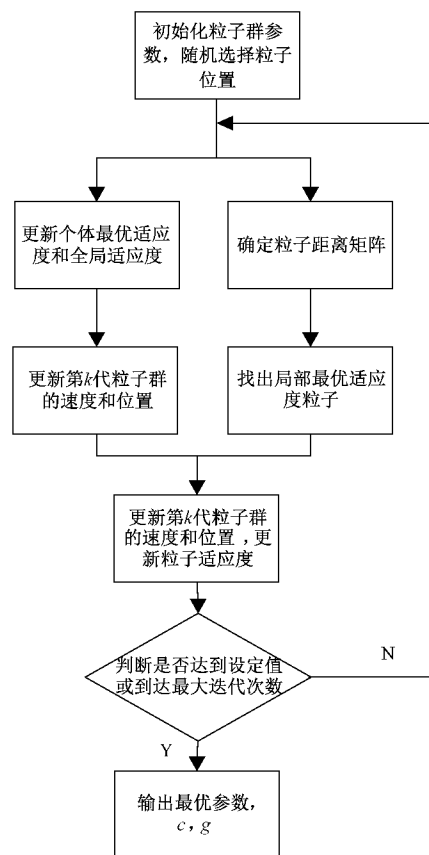


图4 混合粒子群优化算法流程

HPSO-SVM算法步骤如下。

(1)初始化粒子群算法参数:加速因子 c_1, c_2 , 惯性权重 ω , 初始速度 v , 种群数量 $sizepop$, 最大进化代数 n , 距离矩阵 D , 局部因子作用半径 R , 全局变量所占权重 q , 初始速度设置为惩罚因子 c 和核函数参数 σ 的变化范围, 每

个粒子的位置坐标为 (c, σ) 。

(2) 随机生成初始群体和初始速度, 计算初始群体适应度, 将动作识别率作为适应度函数, 并根据该函数求取全局极值, 记录全局极值位置。

(3) 计算距离矩阵 D , 距离公式为 $d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$, 根据适应度, 更新局部最优解, 记录局部最优解位置, 根据混合粒子群算法速度更新公式, 更新粒子速度, 更新粒子适应度, 全局最优解; 距离公式如式(14)所示。

$$D = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{bmatrix} \quad (14)$$

(4) 判断是否达到最大迭代次数, 达到则结束, 否则返回步骤(3)。

(5) 输出最优惩罚因子 c 和核函数参数 σ , 以及识别准确率。

2 实验与结果分析

2.1 数据采集

实验对象均为右手为主利手的学生, 而且没有任何肌肉神经疾病史。其中男性 3 名, 女性 1 名, 年龄均在 23~26 岁之间, 所有受试者均自愿参加此次测试, 且测试前 1 天未做剧烈运动。实验前每名测试对象先用酒精对测试部位肌肉进行消毒, 然后将肌电传感器对称佩戴到左右手上^[22]。受试者均以舒适的姿态保持站立, 所有动作都以双手竖直下垂为起始状态, 然后要求受试者依次完成动作, 每个动作做 2~4 s, 每个动作间隔 3 s, 每做一组休息 30 s, 每组动作做 30 次。做动作时尽量保证开始前处于平稳状态, 执行动作时保持的力度和速度适中。本文主要研究了 5 种中国手语动作, 分别为实、在、对不起、麻烦、打, 每次采集 5 种动作, 每组动作重复 30 遍, 每次能得到 150 组数据, 4 位同学每人采集两次, 共组成 1 200 组数据样本。

算法运行环境在 MATLAB R2016b 上完成, 计算机处理器是 Intel(R) Core(TM) i5-4210M CPU@2.6 GHz, 安装内存是 8.00 G(7.53 G 可用)。实际采集电路放置位置如图 5 所示。本文试验的 5 种中国手语动作如图 6 所示。

实际测试动作肌电信号波形如图 7 所示, 从图 7 中可以看出动作“实”和动作“打”都只有短暂的挥动, 动作“对不起”和“麻烦”都是 3 次波动, 而且前一次都是从下垂状态到额头, 区分难度较大。

2.2 特征聚类分析

为了减少算法运算时间, 提高识别效率, 需要选取聚类效果明显的特征值。本文对 5 个手语动作共选取 150 组实验数据来观察聚类效果, 加速度信号和肌电信号的特征分布如图 8 所示。图 8(a) 为合成加速度的均方根、合成加速



图 5 实际采集实物图



图 6 测试手语图

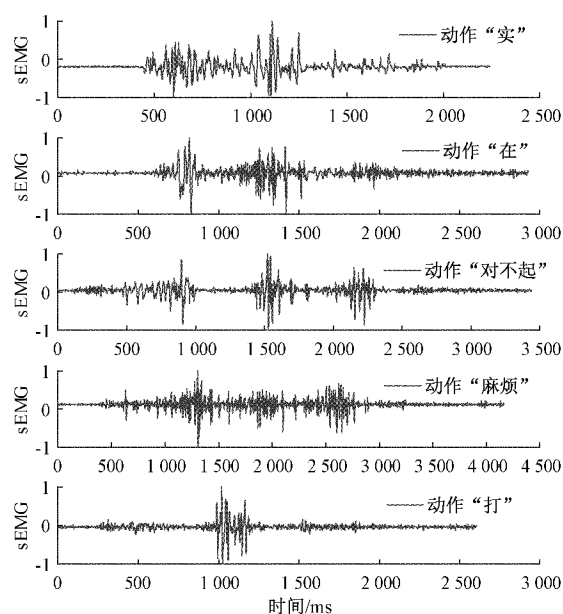


图 7 5 种肌电信号原始数据图

度均值和方差的三维图, 从图中可以看出动作“对不起”和动作“麻烦”均有两下敲击, 因此两个动作之间的运动轨迹有一部分重叠, 导致两种动作的区分度不大。图 8(b) 为肌电信号频域特征三维分布图, 从图中可以看出 5 种动作在肌电信号频域特征上展现出较好的区分度, 因此将该特征与其他特征结合可能会有更好的分类效果。图 8(c) 为肌电信号时频域特征, 该特征对动作“打”区分度比较好, 对其他动

作区分度不太明显。图8(d)为肌电信号的时域特征,取其中的部分特征,均方根、方差、SSC,可以看出该特征对动作“实”、“在”和“打”区分度较明显,对其他两种动作区分度不

明显。通过比较以上不同传感器的特征值,可以发现不同特征对不同动作的区分度有所不同,因此可以选取不同传感器聚类效果较好的特征,以便提高算法识别准确率。

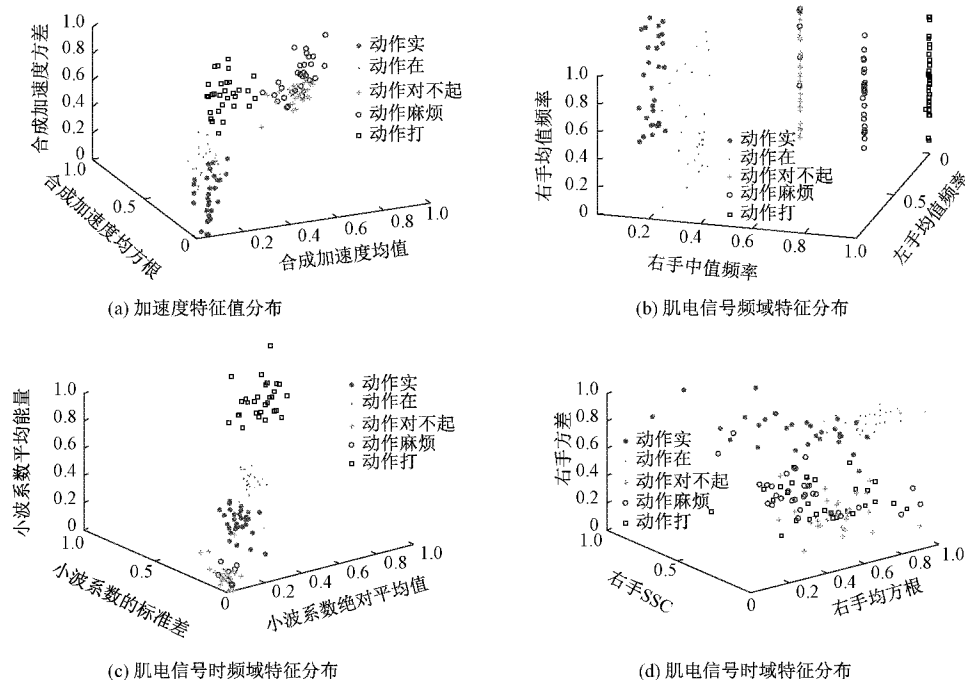


图8 不同特征分布图

2.3 实验结果与分析

为了探究手语识别的准确率,本文分别从不同特征组合和不同算法两个方面对识别准确率进行了探究。在探究不同特征组合对手语识别的影响时,首先将提取的特征分为4类,分别为加速度和角速度特征、肌电信号的时域特征、肌电信号频域特征和时频域特征,对不同的特征进行组合,把手语加速度特征的均方根、方差等特征组成特征矩阵 Λ_1 ,手语的加速度和角速度特征组合为 Λ_2 ,在 Λ_2 特征组合的基础上增加肌电信号频域特征中的中值频率、均值频率等组合为 Λ_3 ,在 Λ_3 特征组合的基础上增加肌电信号时域特征中的均方根、方差等特征组成特征矩阵 Λ_4 ,在 Λ_4 特征组合的基础上增加肌电信号时频域特征的小波系数均值、小波系数方差等特征组成特征矩阵 Λ_5 ,分别设计4种分类算法进行分类,支持向量机(SVM)算法核函数选择为径向基核函数;决策树(decision tree, DT)算法以C4.5算法作为分类标准;BP神经网络(back propagation neural network, BPNN)算法^[29]使用3层神经网络进行预测;混合粒子群优化的支持向量机(HPSO-SVM)算法选取种群数量为20的粒子群进行参数寻优。对所选特征识别率进行验证。在不同特征组合下,利用上述4种算法得到的手语识别率结果如图9所示。

从图9可以看出,虽然不同算法在不同特征下所得识别率有所不同,但大体走势很相似,且本文提出的HPSO-

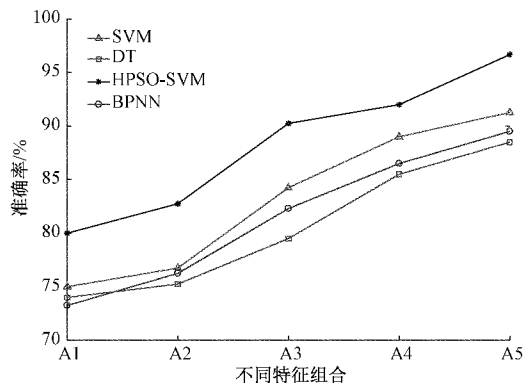


图9 4种算法下的不同特征识别率比较图

SVM算法在识别率上高于其他算法。此外,随着不同特征的不断加入,算法的识别率越来越高;加入肌电信号频域特征后识别率提高10%左右,加入肌电信号时频域特征提高5%左右,证实了多种传感器特征组合有助于提高手语识别率。

为了验证混合粒子群优化的支持向量机的手语识别效果,本文设置了对比试验,通过肌电传感器和六轴陀螺仪共采集了1200组实验数据,将其中300组作为训练集、剩余900组作为测试集,分别用支持向量机(SVM)算法、DT算法、BPNN算法和混合粒子群优化的支持向量机(HPSO-SVM)算法在MATLAB中构建分类器对每个动

作的 180 组数据进行预测,在特征组合为 A6 的情况下,4 种算法模型的识别时间和平均识别率如表 1 所示。在训练时间上,SVM 算法训练时间最短,其次是 HPSO-SVM 算法,HPSO-SVM 算法训练时间优于 DT 和 BPNN 算法。在预测时间上,HPSO-SVM 算法同样优于 DT 和 BPNN 算法,略长于 SVM 算法;在识别率上,采用 SVM 算法的识别率为 91.11%,采用 DT 算法的识别率为 88.56%,采用 BPNN 算法的识别率为 89.56%,而 HPSO-SVM 算法识别率能达到了 96.78%,其识别率优于前 3 种算法,证明该算法可以有效提高手语识别率。

表 1 不同算法效果对比

算法	训练时间/s	预测时间/s	平均识别率/%
SVM	0.007 8	0.004 7	91.11
DT	0.618 3	0.023 5	88.56
BPNN	0.608 3	0.083 2	89.56
HPSO-SVM	0.009 4	0.013 6	96.78

在特征组合为 A6 的情况下,不同算法下不同测试人员的手语识别率如图 10 所示。从图中可以看出不同测试人员在不同算法下识别率相差不大,但 HPSO-SVM 算法的分类效果优于其他算法,有更好的分类效果。

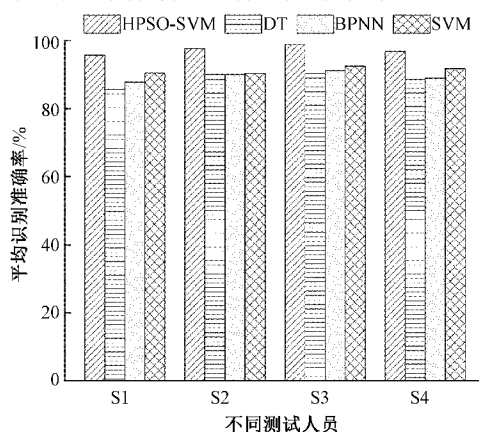


图 10 4 种算法下的不同测试人员识别率比较图

本文在不同算法下对不同动作的识别率进行了对比实验,对每个动作的 180 组实验数据分别进行测试,实验结果分别如表 2~5 所示。

表 2 SVM 算法动作识别结果统计

手语动作	实/次	在/次	对不起/次	麻烦/次	打/次	识别率/%
实	165	7	1	1	6	91.67
在	3	170	2	3	2	94.44
对不起	5	3	158	13	1	87.78
麻烦	3	2	15	156	4	86.67
打	3	5	0	1	171	95.00

表 3 DT 算法动作识别结果统计

手语动作	实/次	在/次	对不起/次	麻烦/次	打/次	识别率/%
实	166	6	0	2	6	92.22
在	5	167	3	3	2	92.77
对不起	6	2	155	16	1	86.11
麻烦	7	2	19	146	6	81.11
打	7	5	1	4	163	90.55

表 4 BPNN 算法动作识别结果统计

手语动作	实/次	在/次	对不起/次	麻烦/次	打/次	识别率/%
实	167	7	0	0	6	92.77
在	8	164	1	1	4	91.11
对不起	5	4	157	14	0	87.22
麻烦	5	2	17	153	3	85.00
打	6	7	0	2	165	91.66

表 5 HPSO-SVM 算法动作识别结果统计

手语动作	实/次	在/次	对不起/次	麻烦/次	打/次	识别率/%
实	177	1	0	0	2	98.33
在	2	176	0	0	2	97.78
对不起	1	2	171	5	1	95.00
麻烦	1	1	6	172	0	95.55
打	2	1	0	2	175	97.22

表 2 展示的是在 SVM 模型下对 5 种动作的识别结果,从表中可以看出对 180 组动作“实”的能准确识别出 165 组数据,识别率能达到 91.67%,对其他 4 个动作“在”、“对不起”、“麻烦”和“打”的识别率分别能达到 94.44%、87.78%、86.67% 和 95%,可以发现在该算法下对动作“在”和“打”的识别率相对较高。表 3 展示的是在 DT 算法下对 5 种动作的识别结果,可以发现在该算法下对动作“实”、“在”和“打”的识别率相对较好,其识别准确率分别能达到 92.22%、92.27% 和 90.55%。表 4 展示的是在 BPNN 算法下识别结果,可以发现该算法对动作“实”、“在”和“打”的识别率同样比动作“对不起”、“麻烦”略高。表 5 展示的是在 HPSO-SVM 模型下对 5 种动作的识别结果,从中可以看出在该算法下对 5 种动作识别准确率都较高,且高于另外 3 种算法,展现出 HPSO-SVM 算法有更好的分类效果。从表 2~5 可以看出,动作“实”、“在”和“打”的识别率要高于动作“对不起”和“麻烦”,原因在于动作“对不起”和动作“麻烦”的相似度较高,且动作操作不规范,导致识别率较低。

表 6 所示为在 HPSO-SVM 算法下,不同测试人员的 5 种动作识别准确率,不同测试人员的测试结果略有不同,

因为每个人在执行手语动作时所出现的误差都不相同。整体来看测试人员 S4 略好于其他测试人员,可能主要是采集时动作更加规范。其他人的平均动作识别率也都达到了 96% 以上,说明此方法具有较高的实用价值。

表 6 动态手语识别结果 %

手语动作	实	在	对不起	麻烦	打	识别率
S1	100.00	95.55	95.55	93.33	97.77	96.44
S2	97.77	97.77	93.33	95.55	100.00	96.88
S3	95.55	97.77	93.33	97.77	95.55	96.00
S4	100.00	100.00	97.77	95.55	95.55	97.77

3 结 论

本文在传统 SVM 的基础上提出了一种 HPSO-SVM 优化算法,通过融合表面肌电特征和加速度特征以及角速度特征,从而捕获了较为完备的手语数据信息,实现了对 5 种手语“实”、“在”、“对不起”、“麻烦”、“打”的有效识别。同时,在数据预处理过程中,增加了一个自适应容错长度以优化短时能量法,从而有效地提取出活动段;通过选取聚类效果好的特征,分析并找到最优特征组合,得到特征矩阵;之后,用 HPSO-SVM 算法、SVM 算法、DT 算法和 BPNN 算法对 5 种动作的识别准确率进行比较分析。最后得出结论,本文使用更加实惠的两种传感器采集数据,降低了数据采集成本,且提出的 HPSO-SVM 算法识别率能达到 96.78%,优于其他算法,提高了手语识别的准确度,且算法鲁棒性较好,为实现一种实惠通用的手语识别机器研究提供了方法。然而本文只做了单个手语的识别,对于连续手语动作识别需要进一步研究。

参考文献

- [1] 姚登峰,江铭虎,鲍泓,等. 手语计算 30 年:回顾与展望[J]. 计算机学报, 2019, 42(1): 113-137.
- [2] LI C S, ZHOU Y Q, LI Y. The signal processing and identification of upper limb motion based on sEMG[J]. Wireless Personal Communications, 2018, 103(1): 887-896.
- [3] 刘传乐,魏士松,贾峻山,等. 基于 Kinect v2 的呼吸运动监测方法研究[J]. 国外电子测量技术, 2020, 39(12): 6-10.
- [4] 万晨晖,杨开明,汪伟,等. 基于 Kinect 的自适应跑步机行走速度估计方法[J]. 仪器仪表学报, 2019, 40(10): 219-226.
- [5] ZHANG Y, LIU B, LIU Z. Recognizing hand gestures with pressure-sensor-based motion sensing [J]. IEEE Transactions on Biomedical Circuits and Systems, 2020, 13(6): 1425-1436.
- [6] 吴常斌,曹青青,费飞,等. 一种可穿戴指间角度测量系统设计[J]. 电子测量与仪器学报, 2020, 34(12): 94-100.
- [7] 孟录,高恒上,张含光,等. 基于全连接神经网络的三维人体姿态估计[J]. 仪器仪表学报, 2020, 41(10): 165-177.
- [8] LU Z, CHEN X, LI Q, et al. A hand gesture recognition framework and wearable gesture-based interaction prototype for mobile devices [J]. IEEE Transactions on Human-Machine Systems, 2017, 44(2): 293-299.
- [9] 齐琦,钱慧芳. 基于融合 3DCNN 神经网络的行为识别[J]. 电子测量技术, 2019, 42(22): 140-144.
- [10] 张启忠,席旭刚,罗志增. 基于表面肌电信号形态特征的多模式识别研究[J]. 传感技术学报, 2012, 25(12): 1636-1642.
- [11] WU J, SUN L, JAFARI R. A wearable system for recognizing american sign language in real-time using IMU and surface EMG sensors[J]. IEEE Journal of Biomedical and Health Informatics, 2016, 20(5): 1281-1290.
- [12] CHENG J, CHEN X, LIU A, et al. A novel phonology- and radical-coded Chinese sign language recognition framework using accelerometer and surface electromyography sensors[J]. Sensors, 2015, 15(9): 23303-23324.
- [13] 张永强,陈香,王从政,等. 基于 FPGA 的实时手势识别系统[J]. 传感技术学报, 2011, 24(11): 1653-1657.
- [14] 孔冬荣,朱杰. 基于表面肌电和加速度信息融合的手势识别[J]. 电子测量技术, 2019, 42(5): 91-95.
- [15] SHULL P B, JIANG S, ZHU Y, et al. Hand gesture recognition and finger angle estimation via wrist-worn modified barometric pressure sensing [J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2019, 27(4): 724-732.
- [16] YANG X, CHEN X, CAO X, et al. Chinese sign language recognition based on an optimized tree-structure framework[J]. IEEE Journal of Biomedical & Health Informatics, 2017, 21(4): 994-1004.
- [17] XIE W, LIANG G, YUAN P. Research on the incremental learning SVM algorithm based on the improved generalized KKT condition[J]. Journal of Physics Conference Series, 2019, 1237(2): 022150.
- [18] 张熙,许浩,李沁程. 基于 SVM 的飞机驾驶舱头顶板字符识别方法研究[J]. 电子测量技术, 2020, 43(22): 82-87.
- [19] LI G, WU H, JIANG G, et al. Dynamic gesture

- recognition in the internet of things[J]. IEEE Access, 2019, 7: 23713-23724.
- [20] 陈家瑞, 陈忠孝, 秦刚, 等. 基于 PSO 算法与 SVR 算法在企业直流配电网短期负荷预测的研究[J]. 国外电子测量技术, 2020, 39(12): 70-73.
- [21] LV Z, WANG L, HAN Z, et al. Surrogate-assisted particle swarm optimization algorithm with pareto active learning for expensive multi-objective optimization[J]. IEEE/CAA Journal of Automatica Sinica, 2019, 6(3): 838-849.
- [22] 王文会, 陈香, 阳平, 等. 基于多传感器信息检测和融合的中国手语识别研究[J]. 中国生物医学工程学报, 2010, 29(5): 665-671.
- [23] 涂文涛, 刘韬, 刘浩炜, 等. BP 与 AR 模型在轴承性能退化评估和预测中的应用[J]. 电子测量与仪器学报, 2019, 33(11): 79-88.

作者简介

刘闯闯, 硕士研究生, 主要研究方向为嵌入式系统开发。

E-mail: lcc1685136818@163.com

朱正伟, 博士, 教授, 主要研究方向为智能检测技术及应用、嵌入式系统及应用。

E-mail: zhuzw@cczu.edu.cn