

DOI:10.19651/j.cnki.emt.2208864

面向康复训练的多通道 mRMR-PSO 肌电特征选择算法

胡蓉华 姚圣曾 成

(南昌大学先进制造学院机械工程系 南昌 330031)

摘要: 表面肌电信号的产生超前于肢体运动的发生,具有预测肢体运动的能力,常常辅助患者的康复训练。针对单通道表面肌电信号难以有效进行关节角度预测的问题,提出了一种基于多通道肌电特征采集下的最大相关最小冗余结合粒子群优化的特征选择算法。通过与采用 mRMR 算法、主成分分析法对关节角度预测精度进行实验对比,验证 mRMR-PSO 算法的性能。实验结果表明,基于多通道的 mRMR-PSO 特征选择算法相比 mRMR 和 PCA 算法在关节角度预测精度分别提高了 32.6% 和 14.9%,从而验证了算法的有效性,并将该算法应用于实际场景。

关键词: 表面肌电信号;关节角度预测;特征选择;互信息;PSO

中图分类号: TP391 **文献标识码:** A **国家标准学科分类代码:** 510.4030

Multi-channel mRMR-PSO sEMG feature selection algorithm for rehabilitation training

Hu Ronghua Yao Sheng Zeng Cheng

(Department of Mechanical Engineering, School of Advanced Manufacturing, Nanchang University, Nanchang 330031, China)

Abstract: The generation of surface electromyography is ahead of the occurrence of body movement and has the ability to predict body movement, which often assists patients in rehabilitation training. To solve the problem that single channel sEMG signal is difficult to predict people's joint angles effectively, this paper proposed a maximum relevance minimum redundancy based on multi-channel EMG feature acquisition and particle swarm optimization feature selection algorithm. The performance of mRMR-PSO algorithm was verified by comparing with that of mRMR algorithm and principal component analysis algorithm for joint angle prediction accuracy. Experimental results show that the joint angle prediction accuracy of mRMR-PSO based on multi-channel feature selection algorithm is 32.6% and 14.9% higher than that of mRMR and PCA, respectively, which verifies the effectiveness of the algorithm, and the algorithm is applied to actual scenarios.

Keywords: surface electromyography; joint angle prediction; feature selection; mutual information; PSO

0 引言

因为疾病造成的肢体瘫痪可以通过康复训练来逐步恢复肌肉功能^[1]。进行康复训练的过程,需要对肌肉进行刺激使其收缩或舒张并带动骨骼绕关节转动。在康复训练初期,患者肌张力改变,需要外部力量辅助才能将肢体运动到有效位置。表面肌电信号(surface electromyography, sEMG)的产生超前于肢体运动的发生,具有预测肢体运动的能力,因此利用 sEMG 信号来辅助患者进行康复训练的研究,受到很多学者的关注^[2]。

随着基于表面肌电信号的康复训练研究不断发展,涌现出了愈来愈多的研究成果。其关键在于建立 sEMG 信号和关节角度回归模型进行运动信息预测的研究,国内外

学者已有不少相关研究进展,如 Zhang 等^[3]应用 BP 神经网络(back propagation neural networks, BPNN)模型估计下肢踝关节、膝关节、髋关节角度。张松等^[4]提出应用多通道 sEMG 分解与 LSTM 神经网络(long short term memory neural networks, LSTMNN)模型相结合的手势识别方法,比分解信号的 SVM 高 4.17%。Pradhan 等^[5]采集同一肌群下不同位置的信息,布置多通道采集的方法来提高 sEMG 识别的准确性。石欣等^[6]提出了一种基于改进能量核的下肢表面肌电信号特征提取方法,相比于传统的两种能量核特征提取方法,运算效率分别提升了 13 倍和 9 倍。哈尔滨工业大学设计了一种基于肌电信号的腕关节术后康复外骨骼系统^[7],能够发挥患者主观能动性,提高患者康复治疗的积极性。合肥工业大学李舒怡^[8]设计了一种基

收稿日期:2022-01-17

于表面肌电信号的腰部康复机器人控制系统,能够更加智能化帮助腰部损伤患者恢复腰部运动能力。

由上可知,虽然基于 sEMG 的康复研究越来越多,但依旧面临着许多挑战,如对关节角度连续运动估计的识别精度仍有待提高。

因此,本文针对肘关节动作提出一种基于多通道的最大相关最小冗余(maximum relevance-minimum redundancy, mRMR)结合粒子群优化(particle swarm optimization, PSO)的特征选择方法来提高关节角度连续运动估计的识别精度。本文使用表贴式电极片布置多采集通道,使用 LSTM 神经网络作为回归算法,使用 mRMR-PSO 进行特征选择,进行验证实验并与 mRMR 和主成分分析法(principal component analysis, PCA)对比,验证该方

法对提高肘关节角度预测精度具有可行性。

1 研究方案

1.1 实验设计

实验流程如图 1 所示,在同一肌群采集多通道的 sEMG 信号,受试者正视 Kinect 相机,与 Kinect 相机水平距离保持为 100 cm。实验分为离线训练和在线控制两部分。图 1 左侧为离线训练,其主要流程为采集 sEMG 信号和肘关节角度信号,对训练数据进行处理,通过 LSTM 算法建立处理后角度信号和 sEMG 信号之间的关系,得到回归器。图 1 右侧为在线控制部分,采集受试者上肢的 sEMG 信号,对 sEMG 信号处理后通过回归器预测受试者关节角度,使用预测结果控制上肢康复机器人的运动。

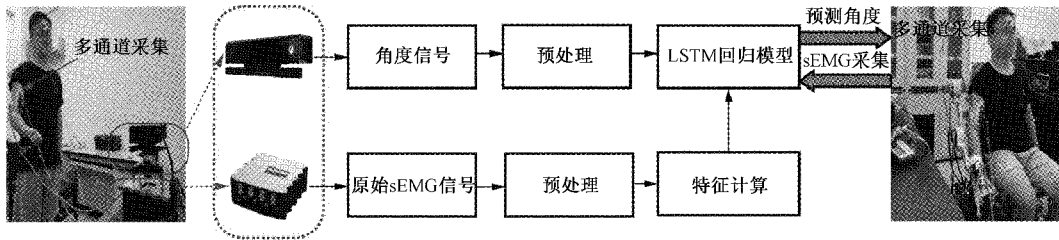


图 1 基于 sEMG 信号的上肢康复机器人控制流程

sEMG 信号采样频率为 1 200 Hz,预处理采用 5~500 Hz 的巴特沃斯带通滤波器、50 Hz 的巴特沃斯陷波器和小波阈值去噪算法去除噪声。Kinect 相机采样频率为 30 Hz, Kinect 相机采集受试者手臂的骨骼点坐标,通过余弦定理获得肘关节角度值, Kinect 相机采集到的角度信号作为真实角度信号。

特征的好坏对泛化能力有至关重要的影响,良好的特征选择能够帮助减少数据冗余和输入数据带宽,对预测的精度能够有所提高。使用 mRMR、PCA 和 mRMR-PSO 3 种降维算法的特征子集分别进行特征计算,并比较 3 种算法的优劣。

1.2 多采集通道布置

根据肌电图实验设计共识^[9](the consensus for experimental design in electromyography, CEDE),选取尺寸为 10 mm 的圆形 Ag/AgCl 一次性表面电极片在肱二头肌和肱三头肌肌腹处^[10]采集 sEMG 信号,电极间距^[11]为 20 mm,采集通道数由双通道变为六通道,如图 2 所示。四通道采集肱二头肌,两通道采集肱三头肌。肱二头肌和肱三头肌之间放置参考电极,额头处为接地电极。

1.3 LSTM 回归算法

特征选择后的 sEMG 信号是一个大小为 $N \times M$ 的矩阵 \mathbf{X} ,角度信号为一个大小为 $N \times 1$ 的矩阵 \mathbf{O} 。 N 为信号样本数即样本长度, M 为特征选择后的特征数。回归指研究矩阵 \mathbf{X} 和矩阵 \mathbf{Y} 之间的统计分析方法,通过回归算法可以建立 sEMG 信号和角度信号之间的关系,得到回归器。

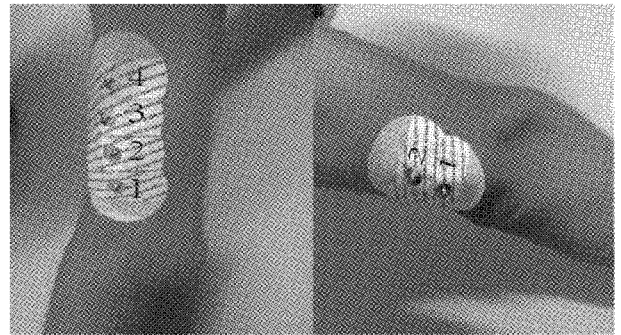


图 2 采集通道布置

LSTM 是一种特殊的循环神经网络(recurrent neural network, RNN)类型,通过添加门,包括输入门 i_t 、遗忘门 f_t 和输出门 o_t , LSTM 的重要信息能保持较长时间,在 sEMG 信号回归问题中表现出良好的性能。

$$\begin{cases} h_t = o_t \odot \varphi(c_t) \\ o_t = \sigma(\mathbf{W}_o x_t + \mathbf{U}_o h_{t-1} + \mathbf{V}_o c_{t-1}) \\ c_t = f_t \odot c_{t-1} + i_t \odot \varphi(\mathbf{W}_c x_t + \mathbf{U}_c h_{t-1}) \\ f_t = \sigma(\mathbf{W}_f x_t + \mathbf{U}_f h_{t-1} + \mathbf{V}_f c_{t-1}) \\ i_t = \sigma(\mathbf{W}_i x_t + \mathbf{U}_i h_{t-1} + \mathbf{V}_i c_{t-1}) \end{cases} \quad (1)$$

式中: i_t 、 f_t 、 c_t 、 o_t 、 h_t 分别为输入门、遗忘门、记忆单元、输出门和隐藏状态; \mathbf{W} 、 \mathbf{U} 和 \mathbf{V} 分别为对应连接的的权值矩阵; x_t 为当前输入; σ 为 sigmoid 函数; \odot 表示元数相乘。

2 基于 mRMR-PSO 的特征选择算法

六通道下采集 sEMG 信号的特征空间维数较高,后续处理过程难度大,因此我们需要对特征空间进行优化。同一肌群下不同位置的采集通道其特征组合相同,这些特征集进行特征选择更具有统计学规律。针对这一特性,本文提出一种基于 mRMR-PSO 算法的特征选择方法,如图 3 所示。这是一个两阶段的特征选择算法,第 1 阶段的 mRMR 算法^[12]采用统计学方法对特征集里的冗余信息进行初步排除,得到候选特征集。第 2 阶段的 PSO 算法^[13]以候选特征集作为搜索的初始范围,以神经网络模型的预测精度作为适应度函数,不断更新粒子和种群最优位置,迭代选出最好的特征子集。

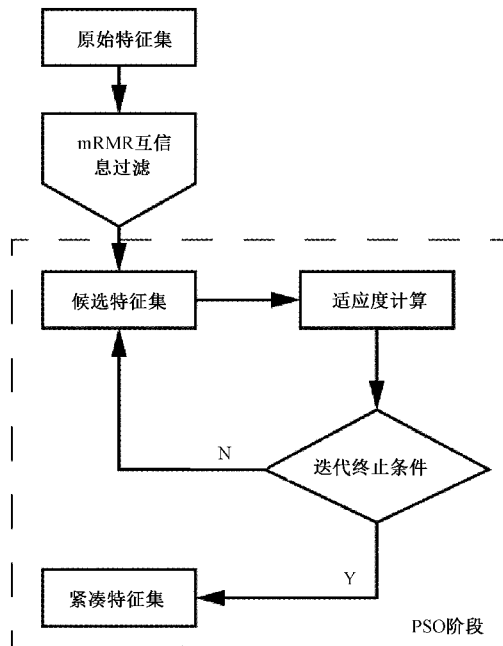


图 3 mRMR-PSO 特征选择算法

mRMR 算法采用互信息度量指标。这种无参、非线性的测度指标正适用于衡量 sEMG 信号特征变量间的非线性关系。

在进行特征选择时,依次计算 sEMG 信号特征值 X_i 与关节角度值 c 的互信息,用最大互信息 $I(x_i; c)$ 表示该特征值与关节角度值的最大相关性。

$$\max D(S, c), D = \frac{1}{|S|} \sum_{x_i} I(x_i; c) \quad (2)$$

式中: S 为 sEMG 信号的特征数据集。虽然式(2)根据最大相关性选出了特征子集,但选出的特征变量之间可能存在较大的冗余性。为了消除特征变量之间的相互冗余,最小冗余准则被引入。

$$\min R(S), R = \frac{1}{|S|^2} \sum_{x_i, x_j \in S} I(x_i; x_j) \quad (3)$$

将式(2)和(3)进行组合得到特征选择的最大相关最

小冗余(mRMR)算法:

$$\max \Phi(D, R) = \max_{x_j \in X-S_{m-1}} \left[I(x_j; c) - \frac{1}{m-1} I(x_j; x_i) \right] \quad (4)$$

mRMR 大大缩短了连续特征的特征选择时间,并显著提高了预测精度。对于具有大量样本的数据集, mRMR 的预测精度具有优越性。不足之处在于 mRMR 对于局部细化具有局限性。

PSO 算法源于对鸟群捕食行为的模拟。在实际应用中, sEMG 信号的特征选择是一个典型的 0~1 组合优化问题,需要将连续的 PSO 进行离散化。

本文中采用 sigmoid 函数对粒子位置进行转换,通过 mRMR 得到一个 D 维的特征空间,群体中第 i 个粒子的位置的每一维表示一个特征变量。每次粒子位置更新后将转换成由 0、1 组成的一个向量。0 表示该变量未被选中, 1 表示该变量被选中。粒子位置转换如下:

$$x_{ij} = \begin{cases} 1, & \text{rand}() \leq S(V_{ij}) \\ 0, & \text{其他} \end{cases} \quad (5)$$

$$S(V_{ij}) = \frac{1}{1 + E^{-v_{ij}}} \quad (6)$$

其中, V_{ij} 表示第 i 个粒子在第 j 维搜索空间的速度。

PSO 算法的搜索性能取决于其全局探索和局部细化的平衡,算法通用性强,需要调整的参数少,易实现。不足之处在于搜索耗时长,样本维数越高,耗时越长。

mRMR 算法和 PSO 算法进行结合,对于多通道采集的大样本,通过 mRMR 算法能够快速在全局上进行初步筛选,得到候选特征集。通过 PSO 算法对候选特征集进行局部细化,得到紧凑特征集。

3 实验结果与分析

3.1 实验结果

主要设备组成包含:奥地利 G. Tec 公司生产的 80 通道系统主机放大器、G. Hiamp 脑电/肌电两用采集设备和 Ag/AgCl 一次性电极片,微软公司推出的 Kinect V2.0 深度相机,进行信号存储和处理的计算机,作为下位机的 STM32F1ZET6 芯片和四自由度上肢康复机器人。

实验选择了 8 名年龄在 22~24 岁的健康男性为受试者。受试者在实验之前已知悉了实验内容。在放置电极之前,电极片位置使用酒精擦拭。sEMG 信号采集的一个完整周期时长为 10 s,前 4 s 手臂保持放松为休息段,后 6 s 肘关节做屈曲运动为动作段。10 个完整周期为 1 组,1 组过后缓解肌肉疲劳,受试者休息 3 min,共做 10 组。实验随机选取 2 组作为测试集,另外 8 组作为训练集。

初始选择^[14]肱二头肌每个采集通道特征组合为波长(waveform length, WL)、斜率变化率(slope sign change, SSC)、差绝对标准偏差值(difference absolute standard deviation valuc, DASDV)、第五时间矩的绝对值^[15](the

fifth moments, TM5), 肱三头肌每个采集通道特征组合为方差 (variance, VAR)、脉搏百分比^[16] (myopulse percentage rate, MYOP)。电极片的设计如 1.2 节所示, 通过 mRMR-PSO 算法对六通道的特征集进行特征选择, 选出最优的特征组合。以 LSTM 神经网络模型的预测精度作为适应度函数, 将均方误差 (mean-square error, MSE) 作为预测精度的指标。

以 1 号受试者的数据进行特征选择。首先通过 mRMR 获得候选特征集, 其步骤如下:

1) 使用 mRMR 的增量选择方法, 得到 n 个顺序特征子集, 并满足 $S_1 \subset S_2 \subset \dots \subset S_{n-1} \subset S_n$; 2) 以神经网络模型的预测精度作为特征子集评价指标, 比较 n 个顺序特征子集的 MSE, 找到 MSE 较小的特征个数范围, $S_1 \subset \dots \subset S_k \subset \dots \subset S_n (1 \leq k \leq n)$, 即得到 k 的取值范围, 该范围称为 Ω ; 3) 在 Ω 内, 找到最小的分类误差 e^* 。该分类误差所对应的特征集为候选特征集。

如图 4 所示, 特征子集为 11 个其均方误差最小, 再增加特征数量其均方误差趋于稳定, 选择 11 个特征数量作为候选特征数量。将特征排名前 11 个的特征作为候选特征集。候选特征集作为 PSO 搜索的初始范围获得紧凑特征集, 其具体步骤如下: 1) 将候选特征集作为 PSO 搜索的初始范围, 进行种群及参数初始化; 2) 以神经网络模型的预测精度作为适应度函数, 评价粒子的适应度值, 进行粒子位置及速度更新; 3) 将更新后的粒子位置进行离散化, 更新粒子和种群最优位置, 检查是否满足终止条件。若满足则得到紧凑特征集; 反之跳转至步骤 1)。

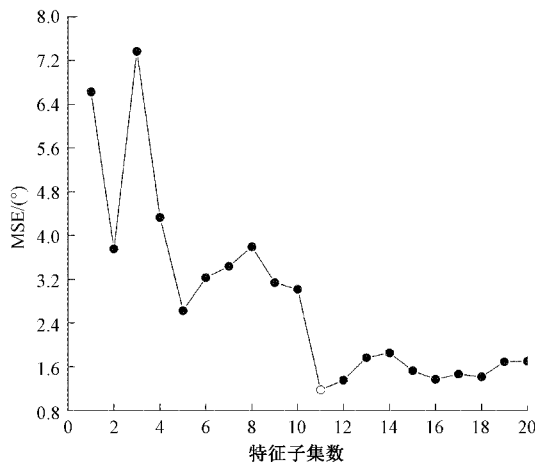


图 4 mRMR 候选特征集数量

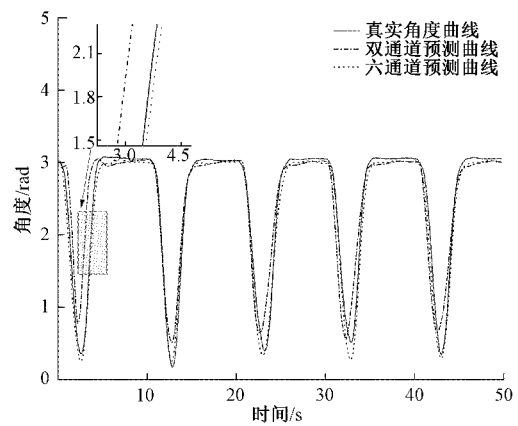
通过 mRMR-PSO 选出的特征组合为肱二头肌采集通道 1 的 WL、采集通道 2 的 DASDV、采集通道 3 的 SSC 和采集通道 4 的 DASDV 以及肱三头肌采集通道 1 的 VAR、采集通道 2 的 VAR、MYOP。

3.2 结果分析

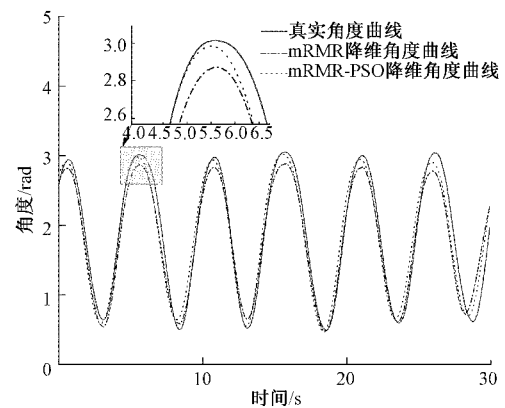
针对 8 名受试者在肘关节屈曲过程中关节角度预测精度结果进行了统计, 每位受试者预测数据进行 20 次计

算, 每次计算时重新随机选取训练集数据和测试集数据, 取 20 次中精度的平均值作为统计的预测精度结果。

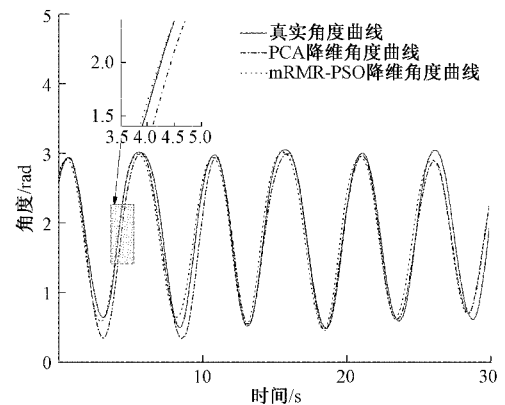
图 5(a) 中为肘关节屈曲动作的真实角度曲线、双通道预测曲线和六通道预测曲线。在肘关节屈曲动作段, 六通道预测曲线比双通道预测曲线更优。在肘关节屈曲休息段预测曲线拟合情况近似。表 1 中以传统的双通道采集和改进的六通道采集方式进行了对比。8 名受试者的预测精度提高了 29.0%。由表 2 可知, 六通道采集的平均耗时比双通道采集的平均耗时时间更长。从数据预处理到预测出关节角度平均用时 73.60 ms。sEMG 信号的产生超



(a) 不同通道数量角度预测曲线



(b) mRMR和mRMR-PSO降维下动作段角度预测曲线



(c) PCA和mRMR-PSO降维下动作段角度预测曲线

图 5 角度预测曲线图

表 1 通道采集方式均方误差对比

受试者 编号	双通道采集 MSE/(°)	六通道采集 MSE/(°)	精度 提高率/%
1	0.900	0.658	26.9
2	0.882	0.648	26.5
3	0.342	0.252	26.3
4	0.324	0.248	23.5
5	0.648	0.424	34.6
6	0.485	0.343	29.3
7	0.378	0.259	31.4
8	0.465	0.308	33.7
平均	0.553	0.393	29.0

表 2 通道采集方式实时性对比 ms

	预处理	特征处理	角度预测	总时间
双通道采集	6.04	9.88	42.25	58.17
六通道采集	15.85	16.21	41.54	73.60

前于肢体运动,此方案所消耗的时间小于 100 ms,使用者不会感到明显滞后现象。六通道采集的实时性能够满足康复机器人的时间要求。

通过对比图 5(b)中 mRMR 和 mRMR-PSO 两种降维方式动作段的角度预测曲线发现,肘关节屈曲动作中屈肘和伸肘动作段预测曲线区别不大,但在波峰段采用 mRMR 降维的方法出现较大误差。因为只采用 mRMR 进行特征选择,特征子集还存在潜在的冗余特征,这些冗余特征使得 mRMR 的精度不如 mRMR-PSO 精度。通过对比图 5(c)中 PCA 和 mRMR-PSO 两种降维方式动作段的角度预测曲线发现,肘关节屈曲动作中屈肘和波峰波谷段区别不大,但在肘关节伸肘采用 mRMR-PSO 进行降维的角度预测曲线效果更好。因为采用 PCA 降维的方法产生的子集还基本保留原来的信息,对无效的特征没有起到滤除作用,相比而言 mRMR-PSO 的方法产生的特征子集预测结果更优。表 3 对比 8 位受试者的 MSE 发现基于 mRMR-PSO 的角度预测精度平均比基于 mRMR 的角度预测精度高 32.6%。和基于 PCA 的精度对比中发现,基于 mRMR-PSO 的角度预测精度比基于 PCA 的角度预测精度略高 14.9%,在对比 4 号受试者时,两种降维方式的预测精度基本持平。表 4 在实时性上对比了 3 种算法的运算时间,mRMR 的特征子集数量比采用 mRMR-PSO 的特征子集数量更多,所消耗的时间也更长。采用 PCA 降维的方法需要计算全部特征,消耗的时间比采用 mRMR-PSO 消耗的时间更长。

3.3 算法应用

为更直观表现算法应用效果,采用实验室已有的六自由度双臂机器人为控制对象,采用 mRMR-PSO 进行特征选择的 LSTM 回归模型进行示教,实验应用展示如图 6 所示。

表 3 降维方式均方误差对比 (°)

受试者编号	PCA MSE	mRMR MSE	mRMR-PSO MSE
1	2.9	2.7	1.9
2	3.3	4.7	3.1
3	3.2	3.8	2.8
4	1.0	1.8	1.0
5	1.1	1.3	1.0
6	2.6	3.4	2.1
7	1.9	2.7	1.6
8	1.7	1.9	1.5
平均	2.21	2.79	1.88

表 4 降维方式实时性对比 ms

	预处理	特征处理	角度预测	总时间
PCA	14.18	11.93	45.98	72.09
mRMR	14.73	7.96	52.17	74.86
mRMR-PSO	14.82	1.61	39.61	56.04

注:所使用计算机相关配置为 inter Core i5 7th Gen, 8 G RAM,表中列出时间为非首次运行时间。

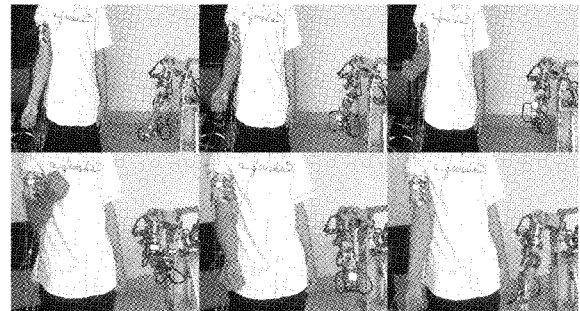


图 6 实验应用展示

4 结 论

本文提出了基于多通道的 mRMR-PSO 特征选择算法。多通道下采集的特征空间维数较高处理过程难度大。根据多通道采集特点,该算法先通过 mRMR 算法快速在全局上进行初步筛选,再通过 PSO 算法对候选特征集进行局部细化。实验结果表明基于多通道的 mRMR-PSO 特征选择方法可以有效提升肘关节角度预测精度,同时该方法能够应用于实际场景。在辅助患者进行康复训练方面具有一定的理论和实用价值。然而多通道的采集方式比传统的双通道采集方式耗时较多,如何减少预处理和特征计算的处理时间将会是下一步工作的重点。

参考文献

[1] 徐瑞,李志才,王雯婕,等.基于肌电的人机交互控制策略及其应用与挑战[J].电子测量与仪器学报,2020,

- 34(2):1-11.
- [2] 侯增广, 赵新刚, 程龙, 等. 康复机器人与智能辅助系统的研究进展[J]. 自动化学报, 2016, 42(12): 1765-1779.
- [3] ZHANG F, LI P F, HOU Z G, et al. sEMG-based continuous estimation of joint angles of human legs by using BP neural network[J]. Neurocomputing, 2012, 78(1): 139-148.
- [4] 张松, 李江涛, 别东洋, 等. 一种基于单通道 sEMG 分解与 LSTM 神经网络相结合的手势识别方法[J]. 仪器仪表学报, 2021, 42(4): 228-235.
- [5] PRADHAN A, HE J, JIANG N. Performance optimization of surface electromyography (sEMG) based biometric sensing system for both verification and identification[J]. IEEE Sensors Journal, 2021.
- [6] 石欣, 朱家庆, 秦鹏杰, 等. 基于改进能量核的下肢表面肌电信号特征提取方法[J]. 仪器仪表学报, 2020, 41(1): 121-128.
- [7] 罗洋. 基于肌电信号控制的腕关节术后康复外骨骼系统研究[D]. 哈尔滨: 哈尔滨工业大学, 2018.
- [8] 李舒怡. 基于表面肌电信号的腰部康复机器人控制系统设计及试验研究[D]. 合肥: 合肥工业大学, 2019.
- [9] BESOMI M, HODGES P W, VAN D J, et al. Consensus for experimental design in electromyography(CEDE) project: Electrode selection matrix [J]. Journal of Electromyography and Kinesiology: Official Journal of the International Society of Electrophysiological Kinesiology, 2019, 48: 128.
- [10] 袁林涛. 基于肌肉主动力响应的肘关节稳定性研究[D]. 天津: 天津科技大学, 2020.
- [11] 于祥, 赵翠莲. 基于 sEMG 和肌肉深度的多通道阵列电极间距研究[J]. 电子测量技术, 2021, 44(11): 16-21.
- [12] MESA I, RUBIO A, TUBIA I, et al. Channel and feature selection for a surface electromyographic pattern recognition task [J]. Expert Systems with Applications, 2014, 41(11): 5190-5200.
- [13] 李炜, 巢秀琴. 改进的粒子群算法优化的特征选择方法[J]. 计算机科学与探索, 2019, 13(6): 990-1004.
- [14] 宋岩. 基于 sEMG 的上肢康复机器人关节角度在线预测算法研究[D]. 南昌: 南昌大学, 2021.
- [15] PHINYOMARK A, PHUKPATTARANONT P, LIMSAKUL C. Feature reduction and selection for EMG signal classification [J]. Expert Systems with Applications, 2012, 39(8): 7420-7431.
- [16] PHINYOMARK A, NUIDOD A, PHUKPATTARANONT P, et al. Feature extraction and reduction of wavelet transform coefficients for EMG pattern classification [J]. Elektronika Ir Elektrotechnika, 2012, 122(6): 27-32.

作者简介

胡璐华, 博士, 教授, 主要研究方向为智能机器人与视觉。

E-mail: 330209066@qq.com

姚圣, 硕士, 主要研究方向为肌电信号处理。

E-mail: 664776685@qq.com

曾成(通信作者), 博士, 讲师, 主要研究方向为目标检测、机器视觉和遥操作。

E-mail: chengzeng@ncu.edu.cn