

基于表面肌电和加速度信息融合的手势识别

孔冬荣 朱 杰

(河海大学 计算机与信息学院 南京 211100)

摘 要: 针对现有手势识别方法中人为提取特征具有局限性的问题,考虑到 MYO 臂环稀疏多通道的特点,提出一种融合表肌电信号(sEMG)和加速度(ACC)信息的手势识别方案。首先由 5 位受试者佩戴 MYO 臂环,同步采集 8 组不同手势的多通道 sEMG、ACC 数据;其次采用一种新的数据预处理方法,使用高能量滑动窗将数据进行分割,得到短时有效活动段;最后设计出一种并行多层的 LSTM 网络,对两类数据进行特征提取和融合,实验准确率达 96.87%。结果表明,所提出的手势识别方案简便可行。

关键词: MYO 臂环;肌电信号;加速度;手势识别

中图分类号: TP391.4; TN911.7 **文献标识码:** A **国家标准学科分类代码:** 520.2040

Gesture recognition based on fusion of surface electromyography and acceleration information

Kong Dongrong Zhu Jie

(Computer and Information College, Hohai University, Nanjing 211100, China)

Abstract: Aiming at the limitation of the existing hand gesture recognition methods, considering the sparse multi-channel feature of MYO armband, this paper proposes a hand gesture recognition scheme which combines surface electromyogram (sEMG) and acceleration (ACC) information. Firstly, 5 participants wore MYO armbands, and simultaneously collected 8 groups of sEMG and ACC data of different gestures. Secondly, a new data preprocessing method is proposed, which will segment multi-channel sEMG and ACC data through high-energy sliding windows to obtain short-term effective active parts. Finally, a parallel sequence LSTM network is designed to extract and fuse the features of the two types of data. The accuracy of gesture recognition is 96.87%. The results show that the proposed scheme is simple and feasible.

Keywords: MYO armband; myoelectric signal; acceleration; gesture recognition

0 引 言

随着人工智能的飞速发展,手势动作识别成为了人机交互、手语识别、康复训练、运动医学等领域的研究热点^[1]。根据数据采集设备的不同,手势识别技术^[2]主要分为基于机器视觉的手势识别和基于可穿戴设备的手势识别。其中,可穿戴设备又分为数据手套、表面肌电信号(surface electromyography, sEMG)^[3]传感器、惯性测量单元(inertial measurement unit, IMU)传感器等。基于机器视觉的手势识别无需用户穿戴任何设备,人机交互体验较好,但其识别结果易受到光照强度、复杂背景及视角遮挡等客观因素的影响;数据手套可以精准获取每个手指时间及空间轨迹信息,但缺点在于价格相对昂贵,在实际应用中用户体验舒适度不佳。

Gieser 等^[4]证实加拿大 Thalmic^[5]推出的 MYO 臂环,相较于其他自制传感器贴片以及 Delsys Trigno Lab 等采集设备在相似识别效果前提下具有成本低廉、佩戴舒适、更具实用性等优点。

考虑到表肌电信号是一种不稳定的信号,每个受试者的肌电信号具有明显的差异性。Lu 等^[6-7]提出了一种采用融合加速度和表面肌电信号进行手势识别的算法框架,并设计了包含一个可穿戴的姿态传感装置(嵌入式三轴加速度计和四导肌电信号传感器),平均识别率为 89.6%。Li 等^[8]开发了一种采用便携式加速度计和 EMG 传感器识别中国手语的系统,对于 120 个手语的词汇表总体分类精度为 96.5%,证明了融合 ACC 和 sEMG 数据进行手势识别的可行性。谢小雨等^[9]对 MYO 臂环采集得到的 sEMG 数据采用滑动窗口分割方法提取平均绝对值(average absolute value, MAV)

等时域特征,使用衡量两个时间序列之间相似度的经典算法 DTW 对肌电信号进行手势识别,其动作识别的准确率达到了 96.09%。

为解决传统方法中手工提取特征存在局限性以及不同受试者在采集数据时速度不同导致有效活动段长短不一的问题,本文利用深度学习算法自主学习特征的优势,提出短时局部有效活动段的预处理方法,同时设计一种并行多层的 LSTM 神经网络结构分别对 sEMG 和 ACC 数据进行训练,最终实现对 MYO 多传感器手势数据的特征融合与识别。

1 手势识别方法

本文使用 MYO 臂环采集数据,其有 8 片独立传感器,

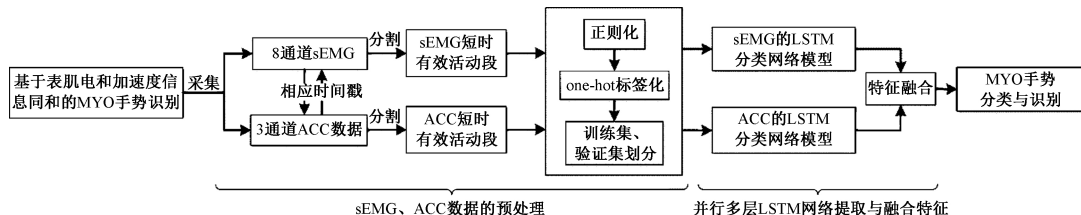


图 2 基于 MYO 臂环多传感器融合的手势识别方法实现流程

1.1 数据预处理

本文使用时间序列表示用户佩戴 MYO 臂环后采集到的手势动作数据。假设一个用户在时间 T 完成某个手势动作 X 。

MYO 臂环中,每个时间戳标记的数据记为 1 帧,同步采集得到的表肌电信号序列为 X_{emg} , 加速度信息序列为 X_{acc} 。

$$X_{\text{emg}} = (X_{\text{emg}}^1, X_{\text{emg}}^2, \dots, X_{\text{emg}}^t, \dots, X_{\text{emg}}^M) \quad (1)$$

$$X_{\text{acc}} = (X_{\text{acc}}^1, X_{\text{acc}}^2, \dots, X_{\text{acc}}^t, \dots, X_{\text{acc}}^N) \quad (2)$$

式中: M 表示 T 时间内表肌电传感器采集的总帧数,每帧为一维 8 通道 sEMG 数据,分别表示为 $\text{emg}1, \text{emg}2, \dots, \text{emg}8$; N 表示 T 时间内加速度传感器采集的总帧数,每帧为一维 3 通道 ACC 数据,分别表示为 x, y, z 。

由于 MYO 臂环的 8 块独立的表肌电传感器贴片的采样频率为 200 Hz, 惯性测量单元 IMU 的采样频率为 50 Hz, 因此帧数 M 与 N 比例近似为 4:1。考虑表肌电信号与加速度数据帧数之间的关系,为了便于后面使用 LSTM 深度学习网络进行同步特征提取与融合,本文对加速度序列进行加帧处理,即序列中每帧 X_{acc}^t 扩充成原来的 4 倍。

$$X_{\text{acc}}^{\text{new}} = (\dots, X_{\text{acc}}^t, X_{\text{acc}}^t, X_{\text{acc}}^t, X_{\text{acc}}^t, \dots) \quad (3)$$

传统的预处理方法通常采用“移动平均法”、“滑动能量平均”等阈值判断方法,利用有预设窗口增量大小的滑动窗探测出动作开始 X_{start} 和结束的位置 X_{end} 进而分割出有效活动段,但不同受试者做相同手势的速度具有一定的差异,导致有效活动段长度差异大难以直接使用深度学习网络进行训练。

每片传感器对应的编号如图 1 所示。本文提出的基于 MYO 臂环多传感器信息融合的手势识别方法,其实现流程如图 2 所示。在采集数据后,大致分为两个主要步骤:对 sEMG 和 ACC 数据进行预处理以及设计一种并行多层的 LSTM 神经网络用于提取并融合特征,最终实现手势的分类与识别。



图 1 MYO 臂环传感器贴片顺序

本文提出一种分割短时有效序列片段的预处理方式,如图 3 所示。首先,在所有用户采集完手势后,对每个用户每个手势对应的 8 通道的 sEMG 和 3 通道的 ACC 信息,对表面肌电信号进行全波整流,采用固定长度大小为 W 的能量采集滑动窗对有效活动的片段进行采样,计算每个划分窗口内所有信号绝对值之和的平均,按照计算得到的平均能量值由大到小排序取前 E 个能量值较大的短时有效序列,每个窗口划分的连续短时序列片段视为每个手势的短时有效活动段。

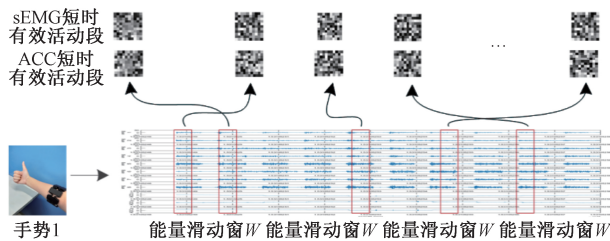


图 3 短时有效活动段的探测与分割

能量计算公式如下:

$$E_{\text{energy}} = \frac{1}{W \cdot k} \sum_{j=0}^{W-1} \sum_{i=0}^{k-1} |X_{ij}| \quad (4)$$

式中: k 指的是数据的维数, sEMG 数据预处理中 k 取 7, ACC 数据中 k 取 2。本方法选用了有效活动段中的局部片段,将数量有限的手势数据有效活动段扩充成数量可观的短时序列片段,可以近似看作不同大小滑动窗获取到的肌电数据片段,扩充了训练集和测试集的样本数量,更有助于网络学习样本数据的多样性;同时本方法解决了完整的手

势有效活动段之间的高相似度易产生训练过拟合的问题。

数据分割后分别形成 E 个 $k \times W$ 大小的短时 sEMG、ACC 有效活动段作为数据集。在送入网络进行训练之前,还需要使用 L2 范数进行正则化以及人为打上标签。本文的数据标记采用 *one-hot* 方法,将 0~7 共 8 个数字标签转化成 *one-hot* 标签。例如数字标签“6”转化为标签为[00000010]。将已经处理好的数据集随机打乱后,按照 9:1 比例划分训练集、测试集,送入神经网络进行特征提取与分类识别,数据预处理阶段至此完成。

1.2 并行多层 LSTM 网络的设计

1) 长短期记忆网络

传统的前馈神经网络例如多层感知机(MLP)是一种从输入到输出的单向映射,难以利用输入数据在时间序列上的相互依赖关系。Hochreiter 等^[10]提出递归神经网络(recurrent neural networks, RNN),在输入到输出的过程中存在环状的递归反馈机制,在处理时序数据时具有较好的优势。

设 RNN 的输入序列 $(x^1, x^2, \dots, x^t, \dots, x^T)$, 隐层序列 $(h^1, h^2, \dots, h^t, \dots, h^T)$ 。当前时刻 t 的隐层输出 O^t 共分为两个部分:一部分是当前时刻 t 的输入 x^t , 另一部分则是 $t-1$ 时刻的隐层输出 h^{t-1} 。

$$h^t = H(W_h g[x^t, h^{t-1}] + b_h) \quad (5)$$

式中: H 为激活函数; W_h 、 b_h 分别为隐层的权重和偏置值。可以看出求解隐层输出是一个迭代的过程,但基于梯度下降的 RNN 算法会产生梯度爆炸或梯度消失的问题^[11]。而 LSTM 网络弥补了 RNN 的不足,其隐层中加入了状态 C 以及 3 个重要的“门”概念:遗忘门 f 、输入门 i 和输出门 o , 调节单元状态中遗忘、输入以及输出的比例,使其可以学习较长时间间隔的帧序列之间的关系。

$$f_t = \sigma(W_f[h^{t-1}, x^t] + b_f) \quad (6)$$

$$i_t = \sigma(W_i[h^{t-1}, x^t] + b_i) \quad (7)$$

$$O_t = \sigma(W_o[h^{t-1}, x^t] + b_o) \quad (8)$$

$$\tilde{C}_t = \tanh(W_c[h^{t-1}, x_t] + b_c) \quad (9)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (10)$$

$$h_t = O_t \tanh(C_t) \quad (11)$$

式中: σ 表示 sigmoid 激活函数; $\odot$ 表示向量元素乘。最终当前时刻 t 的 LSTM 隐层状态 h^t 由单元状态 C 决定,易知隐层状态 h^t 同时受到 f 、 i 、 o 3 个门的影响。

2) 并行多层神经网络的设计

鉴于 MYO 臂环采集到的手势信号具有时序性^[12],采用深度学习 LSTM 算法能较好地学习肌电信号的特征。多层神经网络能够更准确地拟合数据的特征信息,但随着网络的深度的增加,网络训练易产生过拟合问题。因此,本文设计了层数适中的并行多层 LSTM 网络,其结构如图 4 所示。

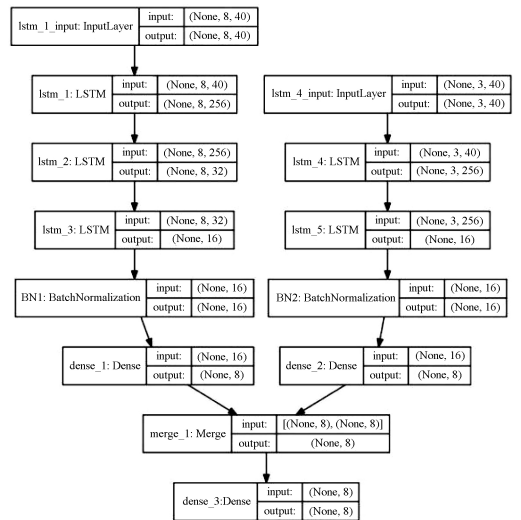


图 4 并行多层神经网络结构

本文设计的并行多层网络共分为 sEMG 分类网络模型和 ACC 分类网络模型两部分,分别将预处理好的数据集送入各自对应的模型网络进行训练。在网络设计中,添加批量归一化层 BN^[13]可以控制过拟合加速网络收敛,所有全连接层使用 softmax 作为激活函数。sEMG 分类网络模型有 3 个 LSTM 层 lstm_1、lstm_2、lstm_3,一个批量归一化层 BN1,一个全连接层 dense_1。ACC 分类网络模型有两个 LSTM 层 lstm_4、lstm_5,一个批量归一化层 BN2,一个全连接层 dense_2。两个模型最后的输出结果通过一个合并层 merge_1 进行融合,加上最后一个全连接层 dense_3,得到网络训练后的最终分类结果。

2 实验结果与分析

2.1 数据采集

鉴于基准数据集的先验性和权威性,本实验使用 MYO 臂环采集原始 sEMG、ACC 信号模仿 NinaPro 肌电信号基准数据集^[9]里与手指动作相关的 13~20 号手势,其分别为伸大拇指、伸食指中指、伸拇指食指中指、大拇指内扣、五指伸展、五指握拳、伸食指、五指伸展内扣等共 8 个动作,如图 5(a)~(h)所示。为避免不同受试者采集时数据传感器通道错杂对后续特征提取与识别工作带来干扰,要求受试者佩戴时统一将 MYO 臂环戴前臂靠近肘关节上方肌肉凸起位置,臂环 LED 标识与手心方向相反。

实验对象为 23~25 岁健康受试者共 5 名。在采集过程中,考虑到 sEMG 信号存在较大的个体差异,同一个体的特性受采集电极安放位置、肌肉疲劳程度、惯用手不同等因素的影响,要求每位受试者佩戴上 MYO 臂环采集完全部数据后摘下,每个手势动作大致 5 s,随后伸展五指放松休息 3 s,每个动作重复执行 10 次,轮换其左右两只手分别进行采集,不同手势之间休息 15~20 min。8 个动作总共执行 $5 \times 10 \times 2 \times 8 = 800$ 次,数据通过蓝牙实时传输至电脑端保存。

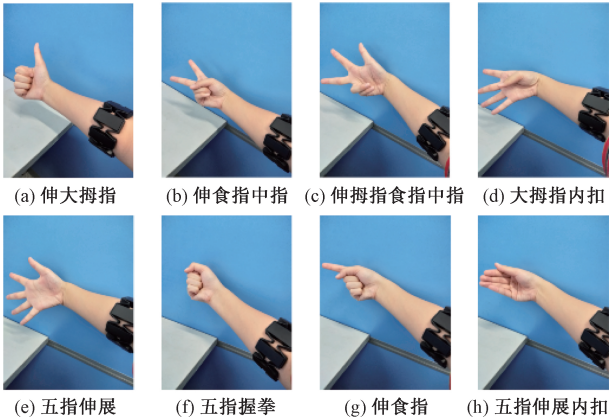


图 5 采集的 8 组手势动作

2.2 实验结果及分析

本文实验环境为 Ubuntu 16.04.3 LTS,使用 Python 语言以及 Keras 深度学习库。为验证 LSTM 网络更适合处理 MYO 臂环采集到的手势信号数据,本文在 8 通道 sEMG 信号数据上对比了人工神经网络(ANN)^[14]、卷积神经网络(CNN)^[15]、VGG-16 以及长短期记忆网络(LSTM)这 4 种不同网络结构的识别准确率,如表 1 所示。无论在训练集还是测试集上,LSTM 网络在准确率方面的明显优于其他 3 种算法。实验结果表明,使用 LSTM 算法处理 MYO 臂环的时序数据具有优越性。

表 1 4 种网络识别准确率的对比表 (%)

算法	训练准确率	测试准确率
ANN	0.916 5	0.850 2
CNN	0.905 1	0.880 8
VGG-16	0.970 6	0.836 8
LSTM	0.998 1	0.906 2

为确定 MYO 臂环数据预处理中最佳的高能量滑动窗口大小,本实验将采集到的 sEMG、ACC 数据分别通过本文提出的方法进行预处理。高能量滑动窗口 W 分别设置长度为 5、10、20、30、40、60、80、100、120、140,取能量窗口个数 $E=300$,送入本文设计的并行多层结构 LSTM 网络进行实验。如图 6 所示,在测试样本上得出窗口大小与准确率的关系曲线,当窗口 $W \geq 40$ 时,准确率基本稳定在 96%~98%。因此,设定高能量滑动窗口大小 $W=40$ 。

为了进一步说明 sEMG 和 ACC 两类传感器信息融合的效果,实验对比了仅采用 sEMG 信息、仅采用 ACC 信息和 sEMG-ACC 数据特征融合这 3 种方法。取每个采样信号的有效活动段个数 $E=300$,样本依然按 9:1 划分,每 300 批次计算一次准确率和损失值,对样本的一次迭代需要 72 次,整个迭代过程进行 200 个周期。

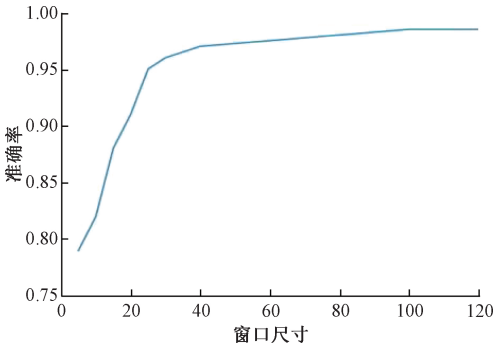


图 6 窗口大小与准确率关系曲线

采用单一传感器数据和两类传感器数据训练的对比结果如图 7 所示。在训练集上迭代到第 60 个周期时准确率均已经基本收敛,但 sEMG-ACC 数据特征融合的结果明显高于仅采用 sEMG 信息和仅采用 ACC 信息的方法。

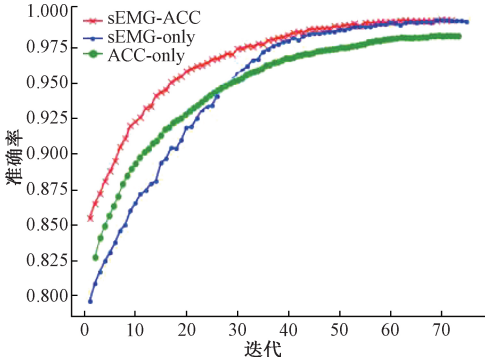


图 7 训练集准确率对比曲线

可以发现 sEMG 特征有更好的表征能力,ACC 特征的加入能够加快信号拟合速度。在初始阶段,加入 ACC 特征后的网络比原始 sEMG 网络高出 5% 的准确率,但迭代 70 个周期后,ACC 准确率不再提高,基本保持在 97%。通过对比可知,融合 sEMG-ACC 两种传感器数据特征的方法最佳。此方法在测试集上 8 组手势分类结果的评估指标结果如表 2 所示,使用并行多层 LSTM 结构对融合 sEMG-ACC 两种传感器数据分类的平均准确率为 96.87%。此结果证实了本文将两类传感器信息融合方法的有效性。

表 2 融合 sEMG 和 ACC 测试集结果

手势	sEMG-ACC		
	准确率	召回率	f1 值
1	0.97	0.97	0.97
2	0.97	0.96	0.95
3	0.96	0.96	0.96
4	0.97	0.95	0.96
5	0.97	0.99	0.98
6	0.99	0.97	0.98
7	0.94	0.98	0.96
8	0.98	0.96	0.97
平均	0.968 7	0.967 5	0.966 2

3 结 论

本文结合手势数据传感器数据具有较强的时序性,提出“短时表肌电加速度 sEMG-ACC 帧序列”方法,采用高能量滑动窗口分割整个手势动作有效活动段内的局部高能量片段作为样本数据。同时,本文设计了并行多层 LSTM 网络进行训练,实现了基于 MYO 臂环的手势分类和识别。实验结果验证,使用本文提出的短时有效活动段分割的预处理方法和分类网络具有较为理想的分类结果。

基于 MYO 臂环的手势动作是一种自然的信息传递方式,可方便地用于正常人和计算机之间的信息交互,具有直接性、可靠性以及低廉便携的特点。接下来将会对动态手语进行深入学习和研究,相信在实际生活中具有实用意义以及市场价值。

参考文献

- [1] FARINA D, JIANG N, REHBAUM H, et al. The extraction of neural information from the surface EMG for the control of upper-limb prostheses: emerging avenues and challenges[J]. IEEE Transactions on Neural Systems & Rehabilitation Engineering, 2014, 22(4):797-809.
- [2] 赵小川. 手势识别技术研究综述与应用展望[J]. 计算机工程与应用, 2017, 53(S2):294-297.
- [3] CHOWDHURY R, REAZ M, MOHD A, et al. Surface electromyography signal processing and classification techniques[J]. Sensors, 2013, 13(9):12431-12466.
- [4] GIESER S N, KANAL V, MAKEDON F. Evaluation of a low cost EMG sensor as a modality for use in virtual reality applications [C]. International Conference on Virtual, Augmented and Mixed Reality. Springer, Cham, 2017:97-110.
- [5] THALMIC LABS. Myo User Guide[EB/OL]. 2013. <https://developer.thalmic.com/docs/api.reference/platform/gettingstarted.html>.
- [6] LU Z, CHEN X, LI Q, et al. A hand gesture recognition framework and wearable gesture based

interaction prototype for mobile de-vices[J]. IEEE Transactions on Human Machine Systems, 2014, 44(2):293-299.

- [7] LIU J H, CHEN W ZH, LI M Y, et al. Continuous recognition of multifunctional finger and wrist movements in amputee subjects based on sEMG and accelerometry[J]. The Open Biomedical Engineering Journal, 2016, 10(1):101-110.
- [8] LI Y, CHEN X, ZHANG X, et al. A sign-component-based framework for Chinese sign language recognition using accelerometer and sEMG data[J]. IEEE Trans Biomed Eng, 2012, 59(10):2695-2704.
- [9] 谢小雨, 刘喆颀. 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.
- [10] HOCHREITER S, SCHMIDHUBER J. Long Short-Term Memory[J]. Neural Computation, 1997, 9(8):1735-1780.
- [11] BENGIO Y, SIMARD P, FRASCONI P. Learning long-term dependencies with gradient descent is difficult[J]. IEEE Transactions on Neural Networks, 1994, 5(2):157-166.
- [12] GRAVES A. Supervised Sequence Labelling with Recurrent Neural Networks [M]. Springer Berlin Heidelberg, 2012.
- [13] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[C]. International Conference on International Conference on Machine Learning, 2015:448-456.
- [14] SCHMIDHUBER, JÜRGEN. Deep learning in neural networks: An overview[J]. Neural Networks, 2015, 61:85-117.
- [15] 刘嘉莹, 张孙杰. 融合视频时空域运动信息的 3D CNN 人体行为识别[J]. 电子测量技术, 2018, 41(7):43-49.

作者简介

孔冬荣, 硕士, 主要研究方向为神经网络、机器视觉等。

朱杰, 硕士, 主要研究方向为文本分析、数据挖掘与深度学习等。

E-mail:815335003@qq.com