

DOI: 10.19650/j.cnki.cjsi.J1904759

穿戴式钢琴弹奏手套智能感知与手势识别技术*

叶素芬^{1,2}, 赖际舟^{1,2}, 吕 品^{1,2}, 朱超群^{1,2}

(1. 南京航空航天大学自动化学院 南京 211106; 2. 物联网与控制技术江苏省高校重点实验室 南京 211106)

摘 要:钢琴弹奏手套是一种新兴的智能可穿戴式设备,通过手套中的多惯性传感器对钢琴弹奏者的手势状态进行实时感知和分析,可以让钢琴学习者实时了解弹奏手势是否规范,从而有效提高钢琴学习效率和兴趣,并降低学习成本。区别于其他应用领域中的手势,钢琴弹奏手势具有多样性、快速性、大动态以及强时变的特点,设计了一套基于惯性数据手套与红外检测杆的钢琴弹奏手势识别系统,并提出一种基于机器学习的钢琴弹奏手势识别方法,以惯性数据手套与红外检测杆的输出作为数据样本,针对钢琴弹奏手势特性,进行多模态手势特征的提取并采用分层识别算法改善识别效果。实验表明,所提出的方法能够较好的适应钢琴弹奏手势识别需求,识别准确率达到99%以上。

关键词: 惯性数据手套;多模态特征;手势识别;机器学习;分层识别

中图分类号: TH89 TP274⁺.2 **文献标识码:** A **国家标准学科分类代码:** 460.40 510.80

An intelligent perception and gesture recognition technology for wearable piano-playing glove

Ye Sufen^{1,2}, Lai Jizhou^{1,2}, Lyu Pin^{1,2}, Zhu Chaoqun^{1,2}

(1. College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China;

2. Jiangsu University Key Laboratory of Internet of Things and Control Technology, Nanjing 211106, China)

Abstract: Piano-playing glove is one kind of emerging intelligent wearable equipment. By using the multi-inertial sensors in glove, the gesture of piano player can be real-time perceived and analyzed. The learners can know in real-time whether the playing gesture is right. Thus, the efficiency and interest of piano learning can be improved and the cost of learning can be reduced effectively. Different from gestures in other application fields, piano playing gestures have the characteristics of diversity, rapidity, large dynamics and strong time-varying. In this study, the piano playing gesture recognition system based on inertial data glove and infrared detecting rod is designed. A method of gesture recognition for piano playing based on machine learning is proposed. The output of inertia data gloves and infrared detection rods are used as data sample. According to the characteristics of piano playing gestures, multi-modal gesture features are extracted. Hierarchical recognition algorithm is adopted to improve the recognition effectiveness. Experimental results show that the proposed recognition method can better meet the needs of gesture recognition in piano playing. The recognition accuracy rate is better than 99%.

Keywords: inertial data gloves; multimodal features; gesture recognition; machine learning; hierarchical recognition

0 引 言

钢琴教育是一项专业性强、学习周期长的教学活动。传统的钢琴教学基于教师与学生的一对一教学,时间及经济成本较高。学生通过穿戴式传感器,对手部弹奏数

据进行采集与辨识,可以进行自主练习。教师可以通过对训练数据的分析,对学生练习情况进行全面掌握,并对学生进行反馈。

手势识别的实现方法主要包括视觉识别与数据手套两种^[1]。在基于视觉的手势识别中,用户一般不需要佩戴采集设备,但容易受到光照遮挡、相机放置等环境因素

收稿日期:2019-02-22 Received Date:2019-02-22

* 基金项目:国家自然科学基金(61703207)、江苏省自然科学基金(BK20170801)、航空基金(2017ZC52017)、中央高校基本科研业务费专项资金(NG2019001,NT2019008)项目资助

的影响。相比之下,基于数据手套的识别技术通过与用户之间的物理交互,可以得到更可靠的手部运动数据,识别精度相对较高^[2]。

目前,基于数据手套的手势识别被广泛应用于手语识别、游戏娱乐、医疗康复等场景^[3-7]。文献[8]采用具有5个弯曲传感器的数据手套,提出一种基于模板匹配和神经网络的联合机器学习算法,对数字1~6和字母A/T/W一共9种静态手势进行识别。文献[9]采用弯曲传感器和惯性传感器,运用线性判别分析法对手部五个不同按键动作进行识别。手势识别技术在很多领域都开展了研究,但在钢琴手势识别领域开展较少,它需要有非常精确的方法来衡量复杂的手部动作。基于视觉的手势识别技术已被应用于钢琴弹奏手势的识别^[10-11],但均基于MIDI键盘,识别的手势简单且识别效果不稳定,实用性及真实感较差。

相较于传统的手势识别领域,钢琴弹奏手势存在多样性、快速性、大动态以及强时变的特点,使得传统的手势识别方法在应用时存在以下问题:1)传统的手势识别大多为静态手势或规范性较强的手势,手势特征较为明显,但是本文研究的钢琴弹奏手势具有快速性、大动态以及强时变的特点,传统的单模态特征不适用于钢琴手势的识别;2)钢琴弹奏复杂手势多样且部分手势之间存在较大的相似性,传统的手势识别方法识别精度较低。针对以上问题,本文提出了一种基于惯性数据手套与红外检测杆相结合的钢琴弹奏手势识别技术。通过红外检测杆对琴键按下状态进行判断,协助对大动态的惯性数据手套进行有效数据截取。同时,研究钢琴弹奏相应的手势特征模型,并采用分层识别算法对弹奏手势进行训练,继而对弹奏手势进行识别,提高系统的实用性及可靠性。

1 钢琴弹奏手势识别系统设计

本文的钢琴弹奏手势识别系统整体框图如图1所示,该手势识别系统主要由数据采集模块、人机交互模块两个部分组成。首先,同步采集能够反映钢琴弹奏信息的手势数据以及琴键数据;其次,对采集的数据进行同步数据处理并进行手势识别;最后,运用辨识的手势进行手部弹奏动作的校正。

本文采用的惯性数据手套如图2所示。该数据手套由10个MPU6500型微电子机械系统(micro-electro-mechanical system, MEMS)惯性传感器(包括三轴陀螺仪和三轴加速度计)组成,10个惯性传感器分别分布于手背以及手指各关节处便于对手部各关键部位数据的采集,具体的惯性传感器分布如图2中序号1~10所示。采用手套中嵌入的ARM芯片(图2中方框标识)读取惯性传感器测量信息并通过姿态解算算法进行姿态解算。

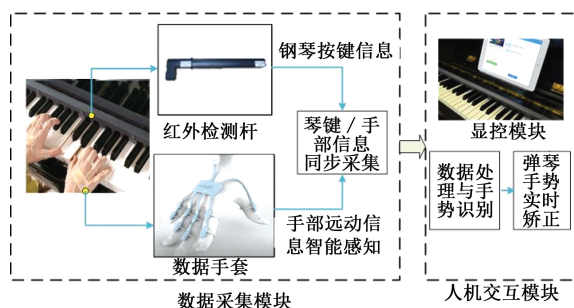


图1 钢琴弹奏手势识别系统

Fig.1 The gesture recognition system for piano playing

通过数据手套捕获的运动数据包括传感器所在的10个关键部位的运动信息及姿态信息,因此一帧手型数据可以表示为:

$$data_1 = [ax_i, ay_i, az_i, gx_i, gy_i, gz_i, pitch_i, roll_i, yaw_i] \quad (1)$$

$$i = 1, 2, \dots, 10$$

式中: ax_i, ay_i, az_i 分别指 x, y, z 三轴的加速度信息; gx_i, gy_i, gz_i 分别指三轴的角速度信息; $pitch_i, roll_i, yaw_i$ 分别指俯仰角、横滚角与偏航角; i 指钢琴弹奏手套上的惯性传感器标号1~10(如图2所示)。如图2所示,佩戴时令手背传感器1($x_1 - y_1 - z_1$) z 轴沿手背竖轴向上, y 轴沿手背纵轴向前, x 轴与 y 轴、 z 轴构成右手直角坐标系;手指的传感器以第9个为例($x_9 - y_9 - z_9$):传感器的 z 轴沿手指竖轴向上, y 轴沿手指纵轴向前, x 轴与 y 轴、 z 轴构成右手直角坐标系。

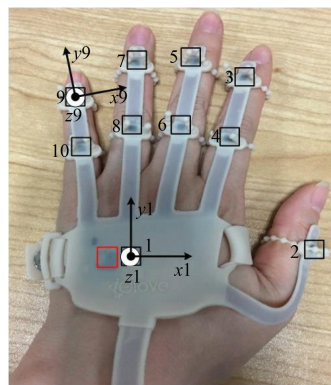


图2 本文惯性数据手套及传感器分布

Fig.2 The inertial data glove and sensor distribution

本文采用的红外检测杆具有88个红外传感器,传感器的分布与钢琴琴键一一匹配,使得红外传感器通过检测与琴键的距离继而判断琴键是否被按下或松开。将红外检测杆放置在琴键上方(具体放置方式如图1数据采集模块所示),通过红外检测杆获取钢琴琴键的按键信息,其输出的数据格式可以表示为:

$$data_2 = [number, 0 \text{ or } 1] \quad (2)$$

式中: *number* 代表按下的琴键编号(1 ~ 88);0 or 1 代表琴键被松开或按下。

2 基于机器学习的钢琴弹奏手势识别技术

为提高钢琴弹奏手势的识别精度,本文提出一种基于多模态特征的分层识别算法的钢琴弹奏手势识别技术。具体的算法流程图如图3所示。

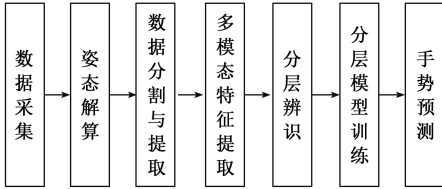


图3 钢琴弹奏手势分层识别算法流程
Fig.3 The flow chart of hierarchical recognition algorithm for piano playing gesture

1) 利用本文采用的惯性数据手套及红外检测杆采集包括各类钢琴弹奏手势的数据,并对惯性传感器数据进行姿态解算获得手部各关键部位的姿态信息。

2) 通过红外检测杆的数据利用固定宽度滑动窗口截取采集的手套数据获取有效的活动段数据,通过对数据的分析,并对窗口内数据进行预处理,提取能够反映钢琴弹奏手势的多模态特征。

3) 通过红外检测杆的数据特征对提取的多模态特征进行初步分类,并针对不同的分类结果将提取的手势特征作为各类极限学习机模型的输入进行建立各类手势识别分类模型,通过对模型进行学习与训练实现钢琴弹奏手势的识别。

2.1 基于互补滤波的手势姿态解算方法

惯性姿态解算算法是基于惯性器件数据手套手势识别的关键部分,其解算精度直接影响手部动作识别精度。本文惯性数据手套采用 MPU6500 六轴惯性传感器来获取手部运动信息并进行数据融合解算得到手部的姿态信息。加速度计低频特性较好,但在运动状态中会引入线性加速度,造成水平倾角计算不准确^[12]。陀螺仪动态特性较好,但长时间会造成积分误差的累加。因此采用互补滤波算法将加速度计长期稳定性和陀螺仪瞬时精度高的特征结合进行手部姿态估计以保证姿态更新速率及精度。由于在钢琴弹奏手势识别中对于航向没有要求,因此本文不对航向角进行校正。

对重力加速度在载体系下的投影 $\mathbf{g}^b$ 以及加速度计测得的重力加速度后的结果 $\vec{\mathbf{a}}$ 进行叉乘即可得到误差在各坐标轴上的投影:

$$\vec{\mathbf{e}} = \vec{\mathbf{a}} \times \mathbf{g}^b \quad (3)$$

通过 PI 控制器来修正陀螺仪误差,得:

$$\dot{\mathbf{e}}(\omega) = K_p \mathbf{e}(\omega) + K_i \int_0^t \mathbf{e}(\omega) dt \quad (4)$$

式中: $\dot{\mathbf{e}}(\omega)$ 是负反馈给陀螺仪进行校正补偿的值; $K_p \mathbf{e}(\omega)$ 是比例控制项; $K_i \int_0^t \mathbf{e}(\omega) dt$ 是积分控制项,在程序中采用离散累加计算。

考虑到钢琴弹奏过程中手部运动的特征,选取加速度的模在 0.9 g ~ 1.1 g 的情况下采用加速度计校正。

根据上述算法,可以估计每个 MEMS 传感器的姿态并确定由各手指和手掌的定向整合的手势。通过佩戴的惯性数据手套进行手部运动信息的智能感知,然后利用 10 个传感器的感知信息识别手势。

2.2 基于红外检测杆/惯性数据手套的手势特征建模与提取方法

1) 数据分割与截取

有效活动段数据截取可以提高算法运行效率并提高辨识精度,是手势识别关键的技术。手势截取的方式有很多种,文献[13]通过神经网络分类器实现手势的运动状态的检测;文献[14-15]通过设置判断相关阈值来判断手势的开始与结束;文献[16]通过添加按钮获取开始及结束时间戳。使用分类器分类需要实时监控数据流,带来较大的计算量;基于简单的阈值判断会导致阈值设置不合适时出现干扰数据;基于按钮的分割不能实现自动识别,需要额外增加步骤,用户体验感降低。

本文设计了一种基于红外检测杆的按键检测方法进行手势数据的精确自动截取,同步接收惯性数据手套及红外检测杆的两类信息并进行组合,无需添加额外步骤且能够实现手势的精准检测。接收端在判断收到红外检测杆发出的按键信号后即可获得数据的有效活动段时间继而进行有效活动段惯性数据手套数据的截取。由于完成不同的弹奏手势所需时间不同,根据弹奏动作的特性,弹奏手势在进行钢琴按键前后的一段数据具有较明显的特征。经过对钢琴弹奏手势所需时长的验证,普通人弹奏钢琴速度最快为 7~9 次/s,在钢琴练习中按键速度小于 7 次/s,本文的惯性数据手套器件输出频率为 42 Hz,由于钢琴弹奏手势在按键前后都有较为明显的特征,为了截取的数据能够完整的包含钢琴弹奏的手势动作,需要截取数据手套按键前后均为 7 组的数据。因此,本文设计了一种固定长度为 15 的滑动窗口在红外检测杆收集到按键信息时截取惯性数据手套在按键前后各七组数据作为一组样本。有效数据段截取流程如图4所示,其中 H 为连续手势动作序列, w 为滑动窗口大小, t 为获取的琴键按键时刻。截取的数据格式为:

$$\begin{cases} data_{select} = (a[t-7], \dots, a[t], \dots, a[t+7]) \\ a[n] = (ax[n], ay[n], az[n], gx[n], gy[n], \\ \quad gz[n], pitch[n], roll[n], yaw[n]) \end{cases}$$

(5)

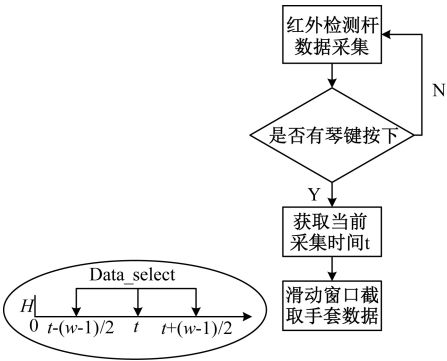


图4 有效活动段数据截取流程

Fig.4 The data interception flow chart of valid active segment

2) 多模态特征提取

惯性数据手套实时感知手部的运动变化信息,传达出时间维度信息。钢琴弹奏手势复杂多样,且具有较大的动态,因此采用单一特征参数难以表征手部的弹奏特性。为了从原始传感器测量数据中获取更多信息,针对钢琴弹奏手部特性本文提出一种多模态特征的手部运动特征提取方法,针对惯性数据手套的输出数据提取基于时间、空间及手指之间耦合特征,同时提取红外检测杆的数据信息作为辅助信息提高手势的识别精度。根据钢琴弹奏手势的特性,具体提取的特征如下。

基于时域的统计特征:由于弹奏的手指会具有较大的动态信息并且手背也会因为弹奏的手指不同而表现出不同的特性,因此提取数据的标准差/极差以及在按键前后的差值来体现手部的动态变化信息。针对手背以及不同手指的弹奏特性提取的手势特征有所不同,具体提取的基于时域手势特征如表1所示。

表1 基于时域的统计特征

Table 1 The statistical features based on time domain

提取特征		手指				
		拇指	食指	中指	无名指	小指
手背	俯仰角标准差	√	√	√	√	√
	横滚角标准差	√	√	√	√	√
	横滚角最大值	√	√	√	√	√
	俯仰角极差	√	√	√	√	√
	横滚角极差	√	√	√	√	√
所辨识手指下关节	y、z 轴加速度标准差	√	√	√	√	√
	x 轴加速度标准差	√				
	x/y 轴角速度标准差	√	√	√	√	√
	x/y 轴角速度前后差值		√	√	√	√
	x/y 轴角速度极差		√	√	√	√
	俯仰角按键前后差值	√	√	√	√	√
	俯仰角极差			√	√	
	俯仰角标准差			√	√	
所辨识手指上关节	y、z 轴加速度标准差		√	√	√	√
	x/y 轴角速度标准差		√	√	√	√
	x/y 轴角速度按键前后差值		√	√	√	√
	x/y 轴角速度极差		√	√	√	√
	俯仰角按键前后差值		√	√	√	√
	俯仰角标准差		√	√	√	√
	俯仰角标准差极差		√	√	√	√

基于空间特性的特征:在进行钢琴弹奏的过程中,手指的上下关节之间以及手指与手背之间姿态角会具有较

大的差值,因此提取按键时刻手指以及手背的信息。提取的表示空间特性特征如表2所示(*i,j*表示惯性器件的

编号,如图 2 编号所示)。

表 2 基于空间特性的特征
Table 2 The spatial-based features

特征描述	特征定义
手指上下关节弯曲度	$pitch_i - pitch_{i-1} (i = 4, 6, 8, 10)$
手指与手背空间弯曲度	$pitch_j - pitch_1 (j = 2, 4, 6, 8, 10)$

手指间的耦合特征:进行钢琴弹奏过程中,不同的弹奏手势导致手指之间的相对信息不同。如相邻手指之间的相对姿态角以及相对加速度、角速度等信息会发生变化。提取的手指间的耦合特征如表 3 所示。

表 3 手指间的耦合特征
Table 3 The coupling feature between fingers

特征描述	特征定义
相邻手指 x 轴相对加速度	$ax_i - ax_{i-2} (i = 4, \cdots, 10)$
相邻手指 y 轴相对加速度	$ay_i - ay_{i-2} (i = 4, \cdots, 10)$
相邻手指 x 轴相对角速度	$gx_i - gx_{i-2} (i = 4, \cdots, 10)$
相邻手指 y 轴相对角速度	$gy_i - gy_{i-2} (i = 4, \cdots, 10)$
相邻手指相对弯曲度	$pitch_i - pitch_{i-2} (i = 4, \cdots, 10)$

辅助特征:由于红外检测杆精准的检测了琴键的按键信息,按键信息也可以反映出当前弹奏的手指个数,因此本文采用了一种基于固定时间宽度的滑动窗口对红外检测杆的数据处理。考虑钢琴弹奏手部同时按下多个键所需的时间间隔,本文采用 100 ms 的窗口宽度对红外检测杆数据进行特征提取,根据一个窗口内的得到的按键数据个数即可确定同时弹奏的手指个数作为手势的辅助特征。

上述不同类型的特征构成了钢琴弹奏手势辨识的多模态特征,更充分的展示了钢琴弹奏手势的特性。为去除量纲的影响,对上述特征进行归一化处理:

$$f_{new-i} = \frac{f_{raw-i} - f_{\min}}{f_{\max} - f_{\min}} \tag{6}$$

2.3 基于分层识别模型的钢琴弹奏手势识别

极限学习机 (extremely learning machine, ELM) 是一种快速的新型学习算法,其结构模型如图 5 所示,通过随机初始化单隐层神经网络的输入权重和偏置得到相应的输出权重,具有学习速度快、泛化性能好,以及最少的人为干预的优点^[17]。

由于本文钢琴弹奏手势的类别较多,若采用单一模型进行辨识无法对手势进行高精度的识别。考虑本文手势类别较多且手势之间存在相似性易造成混淆,如单手指与多手指同时弹奏的情况容易混淆,因此,为了适应钢琴弹奏的手部动作特点以提高手势识别精度,本文提出

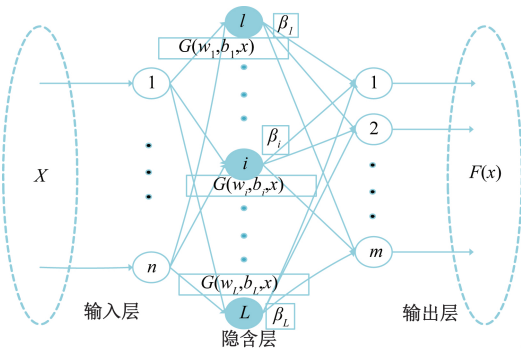


图 5 极限学习机网络结构模型
Fig.5 The network structure model of ELM

一种基于极限学习机的钢琴弹奏手势分层识别算法,该分层识别算法具体流程如图 6 所示。

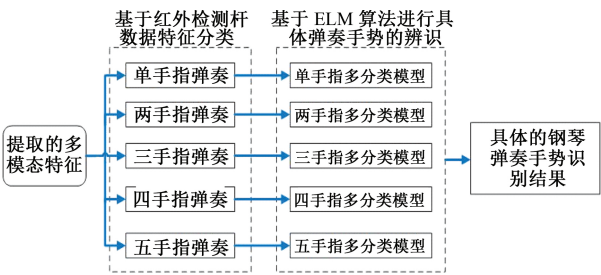


图 6 钢琴弹奏手势分层识别算法流程
Fig.6 The flow chart of hierarchical recognition algorithms for piano playing gesture

本文的分层识别算法包括两层分类处理,两层分类依次衔接,构成完整的手势识别方案。首先选择基于红外检测杆数据提取的辅助特征作为第 1 级特征从而将待识别手势初步分为单手指弹奏/二手指弹奏/三手指弹奏/四手指弹奏/五手指弹奏,然后针对不同的手指弹奏类型采用相应的 ELM 分类模型进行具体的弹奏手势的识别,最后输出弹奏手势识别的最终结果。

根据提出的分层识别算法,将采集的数据样本初步分为 5 个类别,因此需要训练 5 个独立的分类模型。

对于上述的每一个类别,给定采集的 N 个训练样本 $(\mathbf{X}, \mathbf{T}) = \{(\mathbf{x}_i, \mathbf{t}_i), i = 1, 2, \cdots, N\}$,此处 $\mathbf{X}$ 为提取的手势特征矩阵,为对应的期望输出, N 为该类别输入的样本个数, $\mathbf{x}_i = [x_{i1}, x_{i2}, \cdots, x_{in}]^T$ 表示 $\mathbf{X}$ 中的第 i 个样本, n 为输入的样本维, $\mathbf{t}_i = [t_{i1}, t_{i2}, \cdots, t_{im}]^T$ 表示第 i 个样本的对应输出, m 表示输出的向量维(与钢琴弹奏手势类别相同)。在本文中,期望输出向量采用 1-0 标记方式标定,即样本属于第几类手势,则将期望输出向量位置中的第几位标记为‘1’,其余位全均记为‘0’。设单隐层神经网络的隐含层有 L 个节点,其输出为:

$$\mathbf{o}_j = \sum_{i=1}^L \beta_i g_i(\omega_i, b_i, x_j) = \beta \cdot h(x_j), j = 1, 2, \dots, N \quad (7)$$

式中: ω_i 为连接输入神经元与第 i 个隐含层节点的权向量; b_i 为第 i 个隐含层节点的偏置; β_i 是第 i 个隐含层节点的输出权向量; $g(\cdot)$ 为激活函数。因此, 可将式(7) 简化成:

$$\mathbf{H} \cdot \beta = \mathbf{O} \quad (8)$$

式中: $\mathbf{H}$ 是隐层节点的输出; β 为输出权重; $\mathbf{O}$ 为单隐层神经网络的输出。

ELM 的学习目标是通过最小化预测误差损失函数之和来求解输出权重, 其运用最小范数得到最小误差解:

$$\|\mathbf{H} \cdot \hat{\beta} - \mathbf{T}\| = \min_{\beta} \|\mathbf{H} \cdot \beta - \mathbf{T}\| \quad (9)$$

因此根据上述公式可以推导出隐含神经元与输出神经元的输出权值矩阵最优解:

$$\hat{\beta} = \mathbf{H}^+ \mathbf{T} \quad (10)$$

式中: $\mathbf{H}^+$ 是矩阵 $\mathbf{H}$ 的摩尔-彭罗斯广义逆。

在 ELM 的训练过程中, 由于输入权重和隐层节点阈值是随机赋值的, 这会导致 ELM 模型的不稳定。为解决该问题并进一步提高 ELM 模型的泛化能力, 本文引入惩罚系数和核函数。通过引入核函数将输入样本映射到高维空间, 替代了 ELM 的随机隐层输出矩阵, 同时在计算输出权重时在对角线上添加正值 $1/C$, 计算得到输出权值为:

$$\beta = \mathbf{H}^T \left(\frac{I}{C} + \mathbf{H}\mathbf{H}^T \right)^{-1} \mathbf{T} \quad (11)$$

因此针对输入的弹奏手势训练样本, 可以得:

$$f(x) = h(x)\beta = h(x)\mathbf{H}^T \left(\frac{I}{C} + \mathbf{H}\mathbf{H}^T \right)^{-1} \mathbf{T} \quad (12)$$

构造核函数代替 $\mathbf{H}\mathbf{H}^T$:

$$\Omega_{\text{ELM}} = \mathbf{H}\mathbf{H}^T; \Omega_{\text{ELM},j} = h(x_i) \cdot h(x_j) = k(x_i, x_j) \quad (13)$$

因此基于核函数的 ELM 输出可以表示为:

$$f(x) = \begin{bmatrix} k(x, x_1) \\ \vdots \\ k(x, x_N) \end{bmatrix} \left(\frac{I}{C} + \Omega_{\text{ELM}} \right)^{-1} \mathbf{T} \quad (14)$$

对于给定的输入钢琴弹奏手势样本集 $(\mathbf{X}, \mathbf{T})$ 以及确定的核函数, 即可得到训练的每一类别的 ELM 钢琴弹奏手势分类器。

目前使用广泛的 ELM 核函数有: 多项式核函数、RBF 核函数、Sigmoid 核函数等。RBF 核函数能够实现非线性映射、模型复杂度较低且没有数值困难, 因此本文选用 RBF 核函数。引入 RBF 核函数:

$$k_{\text{RBF}}(x, z) = \exp\left(-\frac{\|x - z\|^2}{2\delta^2}\right) \quad (15)$$

3 钢琴弹奏手势感知与识别实验与分析

为验证本文所提出方法的有效性, 采用如图 1 所示的钢琴弹奏手势识别系统数据采集模块进行实验数据的采集。共采集《美丽的梦神》、《G 大调奏鸣曲》等钢琴曲在内的 11 首钢琴曲弹奏数据, 每首各采集四遍, 一共采集 44 首钢琴曲子的弹奏数据, 包括惯性数据手套的数据信息以及红外检测杆的数据信息。

首先, 为验证采用的惯性数据手套对手部钢琴弹奏动作的智能感知能力, 利用上述采集的数据进行验证。以《G 大调奏鸣曲》弹奏数据为例, 弹奏输出的各手指上下关节俯仰角数据如图 7 所示, 可以看出采用的姿态解算算法可以较好地对手部动作的捕获且惯性数据手套可以较好地感知手部的动态信息。图 8 所示为采用本文设计的有效数据段截取方法截取《G 大调奏鸣曲》弹奏数据, 曲线表示的是截取的 50 次有效按键数据俯仰角曲线, * 表示按键时刻, 其对应所在的曲线表示当前弹奏的手势类别, 可以看出截取的数据较好地反映了手指的弹奏动作。

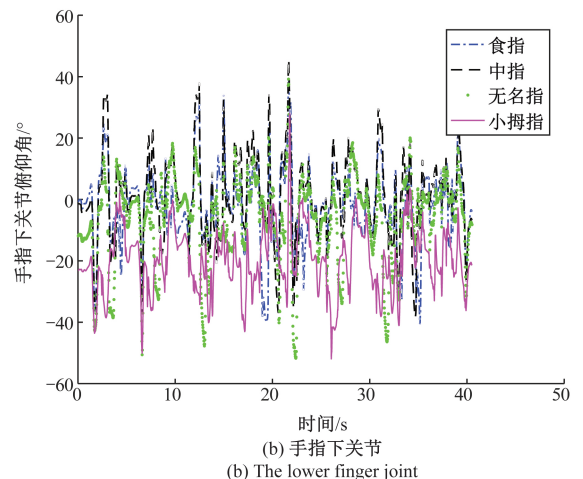
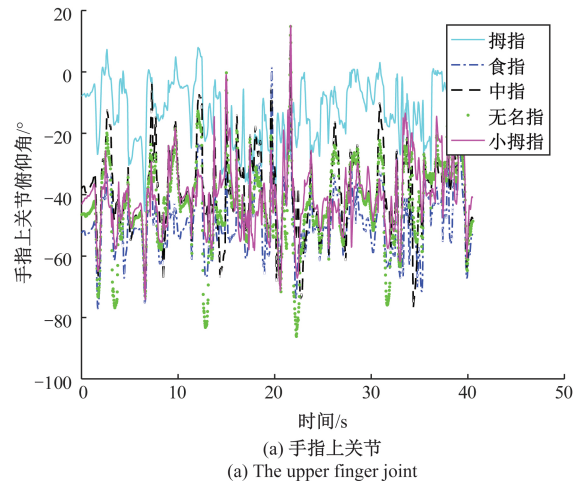


图7 钢琴弹奏手部上、下关节完整俯仰角变化曲线
Fig.7 The complete-data curve of the pitch angle of the upper and lower joints of the piano player

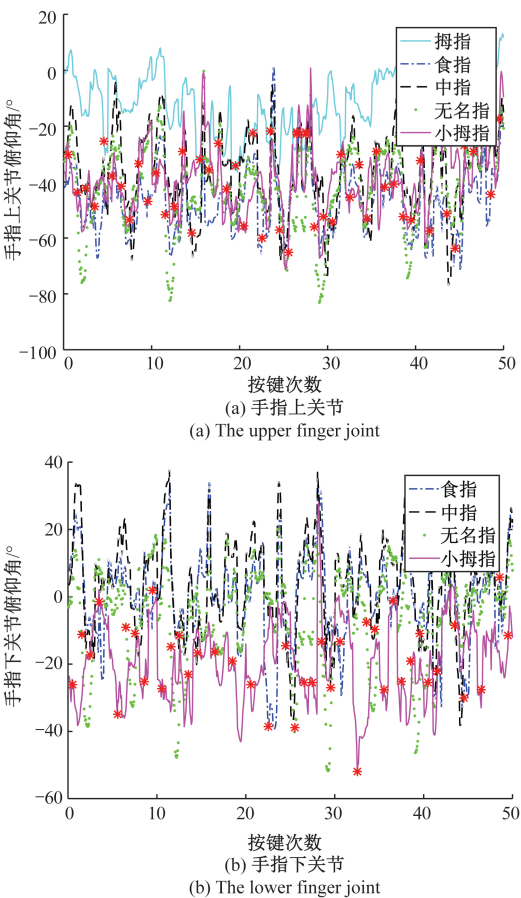


图 8 截取按键时刻数据钢琴弹奏手部的上、下关节俯仰角变化曲线

Fig.8 The valid-data curve of the full pitch angle of the upper and lower joints of the piano player

钢琴弹奏手势识别算法的性能评估分析,运用本文提出的手势识别方法进行模型的训练及分类测试。本文采用 4 折交叉验证法来评估本文所设计识别算法的性能,将采集的 44 首钢琴弹奏数据按照一致性平均分成 4 组(每组 11 首不同的曲子),共进行 4 组测试实验,每次取其中 3 组作为训练集,剩余一组作为测试集。

表 4 所示为不同方法的手势识别精度结果,传统方法为只采用时域的统计信息为特征且未采用提出的分层识别算法,使用单一模型进行辨识;本文改进方法为采用本文提出的多模态特征,同时采用本文提出的分层识别算法,使用 5 个不同的模型进行辨识。从表 4 可以看出,相比于传统方法,本文提出的方法具有更好的识别性能,提高了手势识别的精度。图 9 和 10 所示为传统方法与本文改进方法所有 14 个类别的混淆矩阵,混淆矩阵的每一列代表预测值,每一行代表的是实际的类别理想情况下。若模型的预测效果好,即所有样本的观测值和预测值一致,则所有样本均出现在混淆矩阵的对角线位置,即对角线上为 1,非对角线全为 0;若模型的预测效果较差,

非对角线位置的数值较大。图 9 和 10 中 1/2/3/4/5 分别表示大拇指/食指/中指/无名指/小指,一个类别中的多个数字则表示不同手指同时弹奏的情况。通过对 14 类混淆矩阵的比较发现,本文改进识别方法具有较好的识别精度,较好地减小了手势之间的误判。

表 4 不同方法钢琴弹奏手部动作识别精度
Table 4 The accuracy of piano-playing gesture recognition by different methods

组数	传统方法	本文改进方法
第 1 组	89. 83	99. 66
第 2 组	90. 22	99. 66
第 3 组	91. 06	99. 50
第 4 组	88. 72	99. 83
平均辨识精度/%	89. 85	99. 66

1	0.883	0.017	0.037	0.015	0.024	0.024	0.031	0.000	0.000	0.000	0.000	0.000	0.000	0.016	0.000
2	0.029	0.941	0.020	0.015	0.019	0.098	0.000	0.125	0.036	0.029	0.219	0.000	0.000	0.000	0.000
3	0.021	0.029	0.899	0.015	0.019	0.000	0.117	0.000	0.000	0.000	0.000	0.000	0.000	0.008	0.000
4	0.011	0.000	0.000	0.949	0.000	0.000	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
5	0.005	0.000	0.014	0.000	0.913	0.000	0.000	0.000	0.161	0.000	0.156	0.000	0.000	0.000	0.006
1/2	0.000	0.012	0.017	0.005	0.014	0.875	0.016	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1/3	0.019	0.002	0.014	0.000	0.000	0.000	0.813	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.000
1/4	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.875	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1/5	0.000	0.000	0.000	0.000	0.005	0.000	0.000	0.000	0.786	0.000	0.000	0.000	0.000	0.000	0.000
2/4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.971	0.000	0.000	0.000	0.000	0.000
2/5	0.000	0.002	0.000	0.000	0.005	0.000	0.000	0.000	0.000	0.000	0.994	0.000	0.000	0.000	0.000
1/2/3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000
1/3/5	0.027	0.000	0.000	0.000	0.010	0.000	0.008	0.000	0.018	0.000	0.000	0.000	0.976	0.025	0.000
1/2/3/5	0.000	0.000	0.000	0.000	0.000	0.000	0.008	0.000	0.000	0.000	0.031	0.000	0.000	0.000	0.969
	1	2	3	4	5	1/2	1/3	1/4	1/5	2/4	2/5	1/2/3	1/3/5	1/2/3/5	

图 9 传统方法混淆矩阵

Fig.9 The confusion matrix of traditional method

1	0.997	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.003	0.995	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	0.000	0.003	1.000	0.005	0.020	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4	0.000	0.000	0.000	0.995	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000	0.980	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1/2	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1/3	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1/4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1/5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000
2/4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
2/5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000
1/2/3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000
1/3/5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
1/2/3/5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000
	1	2	3	4	5	1/2	1/3	1/4	1/5	2/4	2/5	1/2/3	1/3/5	1/2/3/5	

图 10 本文改进方法混淆矩阵

Fig.10 The confusion matrix of improved method in this paper

4 结 论

本文提出了一种基于惯性数据手套与红外检测杆相结合的钢琴弹奏手势识别技术,并结合钢琴弹奏的特点,对弹奏手势数据分割与截取、特征建模、机器学习策略进

行了研究。通过实验表明,提出的方法可适应钢琴弹奏弱规范性、大动态性以及强时变性的手势运动特点,其识别精度高于99%,可以满足钢琴学习需求。本文所提出的方法可有效地识别人体手部动作,同时也可应用于手部医疗康复、单兵作战手势识别等领域。

参考文献

- [1] BERMAN S, STERN H. Sensors for gesture recognition systems [J]. IEEE Transactions on Systems Man & Cybernetics Part C, 2012, 42 (3): 277-290.
- [2] FANG B, SUN F C, LIU H P, et al. 3D human gesture capturing and recognition by the IMMU-based data glove[J]. Neurocomputing, 2018(277): 198-207.
- [3] SADDIQU N, CHAN R H M. A Wearable hand gesture recognition device based on acoustic measurements at wrist [C]. IEEE Engineering in Medicine and Biology Society, 2017.
- [4] KALPATTU S A, LEE C F Q, DEREK H. Glove-based hand gesture recognition sign language translator using capacitive touch sensor [C]. IEEE International Conference on Electron Devices and Solid-State Circuits (EDSSC), 2016.
- [5] CAVALLO F, ESPOSITO D, ROVINI E, et al. Preliminary evaluation of sens hand V1 in assessing motor skills performance in parkinson disease [C]. IEEE 13th International Conference on Rehabilitation Robotics, 2013: 1-6.
- [6] SANTOS L, CARBONARO N, TOBNETTI A, et al. Dynamic gesture recognition using a smart glove in hand-assisted laparoscopic surgery [J]. Technologies, 2018, 6(1): 8.
- [7] 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): 85-89.
KONG D R, ZHU J. Gesture recognition based on fusion of surface electromyography and acceleration information [J]. Electronic Measurement Technology 2019, 42(5): 85-89.
- [8] YIN S, YANG J, QU Y, et al. Research on gesture recognition technology of data glove based on joint algorithm [C]. 2nd International Conference on Mechanical, Electronic, Control and Automation Engineering, 2018.
- [9] SUN Q, GONZALEZ E, ABADINES B. A wearable sensor based hand movement rehabilitation and evaluation system [C]. 11th International Conference on Sensing Technology, 2017.
- [10] AKIYA O, MANABU H. Marker-less piano fingering recognition using sequential depth images [C]. 19th Korea-Japan Joint Workshop on Frontiers of Computer Vision, 2013.
- [11] GORODNICHY D. O, YOGESWARAN A. Detection and tracking of pianist hands and fingers [C]. 3rd Canadian Conference on Computer and Robot Vision, 2006.
- [12] 赵裕沛, 费飞, 杨德华等. 面向手部康复应用的数据手套技术研究 [J]. 仪器仪表学报, 2017, 38(10): 2415-2421.
ZHAO YU P, FEI F, YANG D H, et al. Research on Data Glove for Hand Rehabilitation [J]. Chinese Journal of Scientific Instrument, 2017, 38(10): 2415-2421.
- [13] LIM J G, KIM S Y, KWON D S. Pattern recognition-based real-time end point detection specialized for accelerometer signal [C]. IEEE/ASME International Conference on Advanced Intelligent Mechatronics, 2009: 203-208.
- [14] HSU Y L, CHU C L, TSAI Y J, et al. An inertial pen with dynamic time warping recognizer for handwriting and gesture recognition [J]. IEEE Sensors Journal, 2015, 15(1): 154-163.
- [15] XIE R, CAO J. Accelerometer-based hand gesture recognition by neural network and similarity matching [J]. IEEE Sensors Journal, 2016, 16(11): 4537-4545.
- [16] CHO S J, OH J K, BANG W C, et al. Magic wand: A hand-drawn gesture input device in 3-D space with inertial sensors [C]. Ninth International Workshop on Frontiers in Handwriting Recognition, 2004.
- [17] ZHANG L, ZHANG D. Domain adaptation extreme learning machines for drift compensation in E-nose systems [J]. IEEE Transactions on Instrumentation & Measurement, 2015, 64(7): 1790-1801.

作者简介



叶素芬, 2017 年于南京航空航天大学获得学士学位, 现为南京航空航天大学硕士研究生, 主要研究方向为基于惯性传感器的手势识别。

Email: yesufen@nuaa.edu.cn

Ye Sufen received her B. Sc. degree from Nanjing University of Aeronautics and Astronautics in 2017. She is currently a M. Sc. candidate at Nanjing University of Aeronautics and Astronautics. Her main research interest is gesture recognition based on inertial sensors.



赖际舟 (通信作者), 分别在 2000 年和 2005 年于南京航空航天大学获得学士学位和博士学位, 现为南京航空航天大学教授、博士生导师, 主要研究方向为惯性导航及多信息可靠融合导航技术。

Email: laijz@nuaa.edu.cn

Lai Jizhou (Corresponding author) received his B. Sc. and Ph. D. degrees both from Nanjing University of Aeronautics and Astronautics in 2000 and 2005, respectively. He is currently a professor and Ph. D. supervisor at Nanjing University of Aeronautics and Astronautics. His main research interests include inertial navigation and integrated navigation system.