

# 基于惯性传感器融合控制算法的聋哑手语识别



冉孟元<sup>1</sup> 刘礼<sup>1</sup> 李艳德<sup>2</sup> 王珊珊<sup>1</sup>

1 重庆大学大数据与软件学院 重庆 400044

2 兰州大学信息科学与工程 兰州 730000

(rmy@cqu.edu.cn)

**摘要** 聋哑人如何与外界进行有效沟通一直是一个备受关注的难点问题。文中提出了一种基于惯性传感器融合控制算法的手语识别方案,旨在实现高效准确的实时手语识别。该融合控制算法采用反馈控制思想,对两种传统的姿态信息计算方法进行融合,减少了环境对传感器的影响,可以准确获取被测对象在瞬时状态下的姿态信息。该算法通过对自采的聋哑手语数据进行数据融合、数据预处理和特征提取等处理,利用支持向量机、K-近邻法和前馈神经网络分类器自适应模型集成的分类方法进行分类。结果显示,所提传感器融合控制算法有效地得出了实时姿态,该手语识别方案对30种聋哑拼音手语的识别准确率达到96.5%。所提方案将为聋哑人手语识别打下坚实的基础,并为传感器融合控制的相关研究提供参考。

**关键词:** 传感器融合;姿态航向参考系统;聋哑手语识别;惯性传感器;动作捕捉

中图分类号 TP399

## Deaf Sign Language Recognition Based on Inertial Sensor Fusion Control Algorithm

RAN Meng-yuan<sup>1</sup>, LIU Li<sup>1</sup>, LI Yan-de<sup>2</sup> and WANG Shan-shan<sup>1</sup>

1 School of Big Data & Software Engineering, Chongqing University, Chongqing 400044, China

2 School of Information Science & Engineering, Lanzhou University, Lanzhou 730000, China

**Abstract** How to effectively communicate with the outside world for deaf people has always been a difficult issue that has attracted much attention. This paper proposes a sign language recognition scheme based on inertial sensor fusion control algorithm, which aims to achieve efficient and accurate real-time sign language recognition. This fusion control algorithm uses feedback control ideas to fuse two traditional attitude information calculation methods, reduces the impact of the environment on sensors, and can accurately obtain the attitude information of the measured object in the transient state. The algorithm performs data fusion, data preprocessing and feature extraction on the collected deaf-mute sign language data, and uses an adaptive model integration method composed of support vector machines(SVM), K-nearest neighbors(KNN) and feedforward neural networks(FNN) for classification. The results show that the proposed sensor fusion control algorithm effectively obtains real-time poses. The sign language recognition scheme achieves accurate recognition of 30 deaf-mute pinyin sign languages with a recognition accuracy of 96.5%. The work of this paper will lay a solid foundation for sign language recognition of deaf and dumb people, and provide references for related research on sensor fusion control.

**Keywords** Sensor fusion, Attitude and heading reference system, Deaf sign language recognition, Inertial sensor, Motion capture

## 1 引言

据世界卫生组织最新的统计数据显示,目前全球有超过5亿听力受损的残疾人,大约占世界人口的5%,已经成为一个不可忽视的群体。根据北京听力协会2017年的预估数据,我国听力受损残障人士已经超过7200万,其中有2700万听障者需用手语方式与人沟通交流。大多数语言和听力受损的人不能通过普通语言进行交流,手语是聋哑人与他人交流的主要途径。手语可以在无声环境下,结合手势、动作和面部表情表达意思<sup>[1]</sup>。这种语言具有手部动作多、词汇量有限、学

习困难等特点,使得聋哑人与正常人的交流存在巨大的困难。这种沟通障碍对聋哑人的生活和社会关系产生了不利影响<sup>[2]</sup>。因此,聋哑人需要借助翻译设备来与正常人进行交流。

传统的基于计算机视觉的手语识别算法可以有效地跟踪和识别手势,且无需与用户进行任何接触<sup>[3-4]</sup>。然而,基于视觉的手语识别可能会受光照条件的影响,这将限制其应用场景。随着传感器技术的飞速发展,惯性传感器越来越多地集成到消费电子产品中。基于惯性传感器的装置的一个显著优点是,它可以在没有任何外部参考或限制的环境下运行<sup>[5]</sup>。由于不同的人的动作速度和幅度不同,手语识别相对复杂。

因此,一些研究人员试图将惯性的数据与肌电图传感器<sup>[6-7]</sup>或视觉传感器<sup>[8-9]</sup>的数据结合起来,以提高系统的性能和鲁棒性。然而,多传感器融合增加了额外的成本和计算负担。

通过集成传感器的数据手套可以解决上述问题。数据手套不需要复杂的数据处理,不受运动类型的限制<sup>[10]</sup>。同时,基于数据手套的手部识别不受背景条件的影响<sup>[11-12]</sup>。Fu等制作了一款轻便的数据手套,可以进行手语识别<sup>[13-14]</sup>。Sriram等研制出一款识别系统,可以进行手语的学习<sup>[3]</sup>。此外,手势识别系统目前已经应用于多个领域,如手语识别、社会辅助机器人、沉浸式游戏、虚拟对象、远程控制和医疗保健等<sup>[15-20]</sup>。

在识别方法上,隐马尔可夫模型和动态时间规整作为两种重要的识别方法,被广泛应用于手语识别领域<sup>[6,21-25]</sup>。其他方法包括概率神经网络<sup>[5]</sup>、符号序列匹配<sup>[26]</sup>和随机线性形式文法<sup>[27]</sup>。通常,分类器训练的数据集与系统的最终用户是不一致的。为了开发一种与用户无关的算法,Akl等<sup>[27]</sup>提出一种基于惯性传感器的手语识别系统,该系统采用动态时间规整结合相似性传播的算法,在训练阶段为每个手势创建范例,并创建了一个由3780个轨迹组成、包含18种手势的数据库。该系统对18种手势在用户依赖和用户独立识别中分别达到了99.81%和94.60%的准确率。

本文提出一种基于惯性传感器融合控制算法的聋哑手语识别算法。手部动作产生的惯性传感器数据和弯曲传感器数据通过无线通信模块传输到个人终端。系统首先对数据进行惯性传感器数据融合,计算姿态信息;再对处理后的数据进行聋哑手语识别,包括:数据采集与信号预处理、时序数据分割、特征提取、分类器构建和模型融合。

本文的主要贡献包括:

1)针对九轴惯性传感器各自的优缺点,采用反馈控制的思想将传统的姿态计算方法融合起来,以减少环境对传感数据的影响,提高实时姿态获取的效率;

2)对于计算后的姿态信息,利用自适应模型集成方法对支持向量机、K-近邻法和前馈神经网络分类器进行模型集成,以优化最终结果。

本文第2节介绍基于惯性传感器和弯曲传感器的数据手套;第3节介绍通过惯性传感器融合控制算法,计算姿态信息;第4节介绍聋哑手语识别算法的具体流程;第5节分析实验结果,讨论方法的有效性;最后总结全文。

## 2 数据手套描述

本实验采用的数据手套由主控模块(Arduino)、九轴惯性传感器(MPU9250)和弯曲传感器组成,通过无线通信模块(HC-08)与个人终端进行信息传输。MPU9250是一款9轴运动跟踪装置,额定工作电压为2.4~3.6V,额定电流为3.5mA。它集成了两部分,一部分是3轴加速度计和3轴陀螺仪,另一部分是AKM公司的3轴磁力计(AK8963)。主控模块通过I2C总线可以直接输出9轴的全部数据,在40℃~

85℃时的精度为1%时钟漂移。数据手套的具体结构如图1所示。

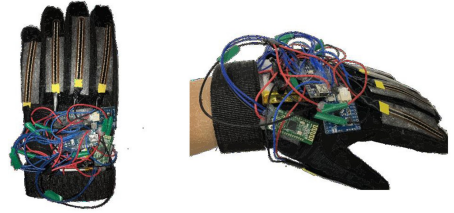


图1 数据手套  
Fig. 1 Data glove

陀螺仪测量用户手部运动产生的角速度模拟信号,通过内部A/D转换器对信号进行数模转换,输出用户可编程程(±250,±500,±1000/(°/s))的三轴(X,Y,Z)16位数字输出,其工作电流为3.2mA,休眠模式电流为8μA;加速度计则用于对手部运动产生的加速度进行测量,直接得到的模拟信号通过内部A/D转换器转化为用户可编程程(±2g,±4g,±8g,±16g)的三轴(X,Y,Z)16位数字信号,其工作电流为450μA,休眠模式电流为8μA;磁力计为3轴单片霍尔传感器,输出为14位(0.6μT/LSB)和16位(15μT/LSB)的分辨率输出。

本文采用的弯曲传感器为薄膜式压力传感器,其电阻会随弯曲角度进行变化,由于其价格低廉,使用简单(与电位器相似),目前被广泛用于竞技游戏、医疗设备和计算机外设等领域。

HC-08蓝牙串口通信模块是的基于Bluetooth Specification V4.0 BLE蓝牙协议的新一代数传模块。无线工作频段为2.4GHz ISM,采用GFSK调制方式,最大发射功率为4dBm,接收灵敏度为一93dBm。

## 3 惯性传感器融合控制算法

通过惯性传感器计算姿态信息有两种方法:一种是利用加速度计和磁力计测量重力上的分量和地磁投影分量;另一种是利用角速度积分来更新姿态信息。但许多外界因素(如温度等)会对角速度产生影响,同时外加磁场的变化也会影响磁力计的精度。为了弥补两种方法各自的不足,本文提出一种简单的惯性传感器融合控制方法,结合两种方法得到的姿态信息,相互修正,计算出更准确的结果,算法流程如图2所示。

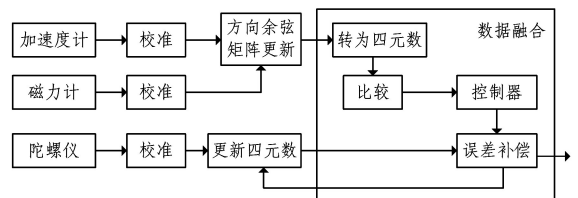


图2 惯性传感器融合控制算法

Fig. 2 Inertial sensor fusion control algorithm

### 3.1 聋哑手语介绍

本文研究的聋哑手语是基本的30个汉语拼音,根据中国

残联提供的手语标准规范,具体内容如图3所示。

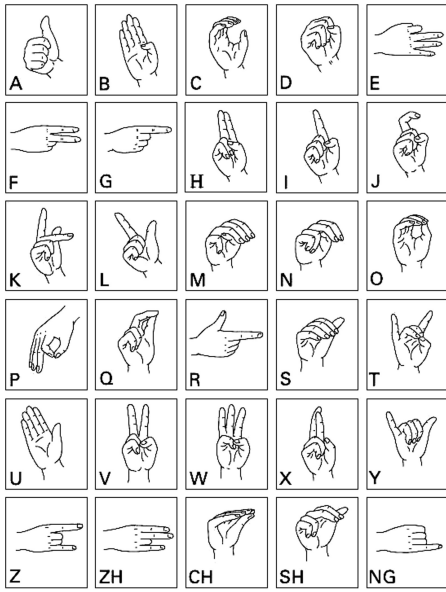


图3 汉语拼音聋哑手语图

Fig. 3 Chinese phonetic transcription deaf sign language

### 3.2 信号校准与处理

由于灵敏度和偏移误差,传感器的原始数据必须经过校准<sup>[29]</sup>。本文使用一种简单的校准方法:记录加速度计和磁力计的最大值与最小值,计算出加速度与磁场强度的均值,再将其分别与重力场和地磁场进行比较,校准偏移误差和灵敏度;同时,利用陀螺仪静止状态下的测量值来校正陀螺偏移误差。

校准后,需要对信号进行处理,以获得更可靠的信号。加速度计用于测量当前角加速度,由于测量原理,加速度计对高频信号敏感,振动环境中的高频信号对其影响较大。陀螺仪用于测量角速度,是一个间接测量角度的器件。它需要对角速度进行积分才能得到角度,噪声误差在积分作用下不断累积,最终会导致陀螺仪的低频干扰和漂移。磁力计用于测量载体与地磁场之间的夹角,低频特性较好,但易受周围磁场的影响。根据手部运动的信号特征,采用Mahony互补滤波器对传感器数据进行处理,对陀螺仪采用高通滤波,对加速度计和磁力计采用低通滤波。

### 3.3 方向余弦矩阵获取

方向余弦矩阵表示参考坐标系与局部坐标系之间的旋转关系。方向余弦矩阵的数学意义是将局部坐标系中的3个基向量用参考坐标系表示。如图4所示,方向余弦矩阵可以将每个单位向量从一个坐标转移到另一个坐标。

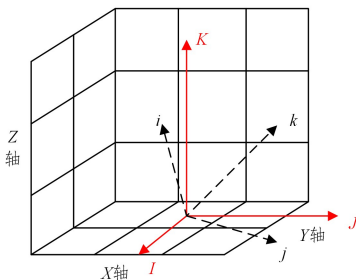


图4 局部坐标系与参考坐标系的关系

Fig. 4 Relationship between local coordinate system and reference coordinate system

方向余弦矩阵可以分为3列,每列中的3个元素分别表示局部坐标中的基向量和参考坐标中3个基向量的内积。方向余弦矩阵如下所示:

$$DCM_r^l = \begin{bmatrix} \vec{I}_r \cdot \vec{l}_0 & \vec{I}_r \cdot \vec{j}_0 & \vec{I}_r \cdot \vec{k}_0 \\ \vec{J}_r \cdot \vec{l}_0 & \vec{J}_r \cdot \vec{j}_0 & \vec{K}_r \cdot \vec{k}_0 \\ \vec{K}_r \cdot \vec{l}_0 & \vec{K}_r \cdot \vec{j}_0 & \vec{K}_r \cdot \vec{k}_0 \end{bmatrix} \quad (1)$$

方向余弦矩阵 $DCM_r^l$ 可以将局部坐标中的基准向量 $\vec{V}_0$ 转化为参考坐标系的基准向量 $\vec{V}_r$ ,其关系表达式如下:

$$\vec{V}_r = DCM_r^l \vec{V}_0 \quad (2)$$

根据 $DCM_r^l$ 的数学意义,它可以由惯性传感器中的加速度计和磁力计获得。设地磁北极方向为参考坐标系的X轴正方向,则方向余弦矩阵的第一列的值为磁力计的读数;再令重力方向为参考坐标系的Z轴正方向,则方向余弦矩阵第三列的值可以由加速度计获得。由于重力方向和地磁方向是正交的,并且在旋转之后仍然存在正交性,可以导出方向余弦矩阵的第二列, $DCM_r^l$ 的表达式如式(3)所示:

$$DCM_r^l = [\vec{m}_c \quad \vec{g} \times \vec{m} \quad \vec{g}] \quad (3)$$

其中, $\vec{m}$ 和 $\vec{g}$ 分别为磁力计和加速度计的读数, $\vec{m}_c$ 为磁力计校准后的数据。

### 3.4 陀螺仪姿态更新

另一种测量姿态的方法是对陀螺仪获得的角速度进行积分,不断更新姿态信息。本文采用四元数来表示陀螺仪更新后的姿态。四元数包括4个元素,第一个是标量,其余3个是组成向量。单位四元数的表示形式如式(4)所示:

$$q = s + \vec{v} = \cos \frac{\theta}{2} + \sin \frac{\theta}{2} \vec{k} \quad (4)$$

其中, $\vec{k}$ 表示坐标系方向的单位向量, $\theta$ 是旋转的角度。哈达马乘积 $\otimes$ 表示两个基于四元数的姿态信息之间的转化关系,其结果化简如式(5)所示。

$$q_1 \otimes q_2 = (s_1 s_2 - \vec{v}_1 \cdot \vec{v}_2) + (s_1 \vec{v}_2 + s_2 \vec{v}_1 + \vec{v}_1 \times \vec{v}_2) \quad (5)$$

利用陀螺仪,采用龙格库塔法对四元数进行更新,更新过程如下:

$$\dot{q} = \frac{1}{2} q \otimes t(\vec{\omega}) \quad (6)$$

$$q_n = q_{n-1} + \dot{q} \Delta t \quad (7)$$

其中, $\vec{\omega}$ 为直接从陀螺仪中获得的角速度。由于系统采样时间不可能无穷小,因此必须对式(6)中更新的四元数进行归一化,使其成为单位四元数。

### 3.5 误差补偿

由于温度、芯片封装等原因,陀螺仪会存在偏移误差,对角速度进行积分会积累误差。即使陀螺仪保持静止,这种现象也会引起数据的改变。为了解决这一问题,需要借助加速度计和磁力计更新姿态信息。但加速度计和磁力计也存在一些问题。首先,通过加速度计和磁力计获取方向余弦矩阵比陀螺仪输出四元数速度更慢;其次,方向余弦矩阵容易受线性加速度和铁磁材料引起的磁场畸变的干扰。

考虑到两种方法具有互补性,本文设计了一种控制器,使陀螺仪积分出的姿态信息接近通过加速度计和磁力计获得的方向余弦矩阵。

第1步 将方向余弦矩阵转化为四元数形式。

$$q_{AM} = a + bi + cj + dk$$

$$a = \frac{\sqrt{m_{11} + m_{22} + m_{33} + 1}}{2}$$

$$b = \frac{(m_{32} - m_{23})}{4a}, c = \frac{(m_{13} - m_{31})}{4a}, d = \frac{(m_{21} - m_{12})}{4a}$$

$$DCM = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \quad (8)$$

第2步 计算两个四元数 $q_{AM}$ 与 $q_g$ 之间的差。单位四元数可以表示三维变换,其进行逆运算表示逆变换,故两个四元数的差可通过 $q_{AM}$ 与 $q_g^{-1}$ 进行哈达马乘积获得。

$$\vec{q}_\Delta = \vec{q}_g^{-1} \otimes \vec{q}_{AM} \quad (9)$$

第3步 式(3)通过三角函数和反三角函数的计算,可以从四元数中获得坐标系方向的单位向量和旋转角度。

$$\vec{\theta}_\Delta = 2 * \cos^{-1}[\text{re}(\vec{q}_\Delta)] \quad (10)$$

$$\vec{V}_\Delta = \text{im}(\vec{q}_\Delta) / \sin(\vec{\theta}_\Delta / 2) \quad (11)$$

第4步 将 $\vec{\theta}_\Delta$ 进行伸缩,再将其乘以向量 $\vec{V}_\Delta$ 构成一个类似于虚拟角速度的矢量。

$$\vec{\omega}_\Delta = S(\vec{\theta}_\Delta) * \vec{V}_\Delta \quad (12)$$

第5步 通过式(6)、式(7)和式(12)对四元数进行更新,这时四元数表示被测对象在瞬时状态下的姿态信息。

## 4 聋哑手语识别

聋哑手语识别算法由数据采集与信号预处理、时序数据分割、特征提取、分类器构建和模型融合等组成。对数据手套传输过来的数据先进行数据清洗,补全缺失数据,去掉冗余数据;再对信号进行处理,以获得更可靠的信号,利用中值滤波将数据进行平滑处理,去除异常点。由于原始数据为时序数据,在对手语进行识别前还需要将数据分割成原子序列区间。在对时序数据进行分割后,需要先对数据进行特征提取,然后再对数据进行分类识别。本文提出的聋哑手语识别算法框图如图5所示。

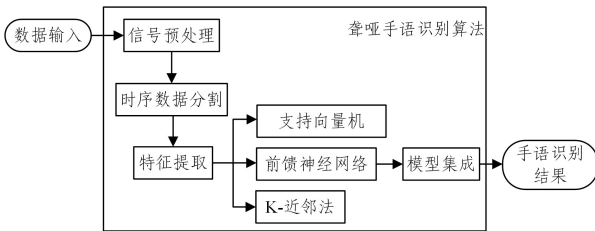


图5 聋哑手语识别框架

Fig. 5 Sign language recognition framework

### 4.1 数据采集与信号预处理

通过惯性传感器融合控制算法,手部运动的传感数据由姿态角和5个位于手指的弯曲传感器电压读数组成。为了降低训练集数据标注开销,采用按钮触发标注,即在执行手语识别时按下按钮,并在完成手语动作后释放。由于无线信号传输的可靠性不如有线传输,即便我们采取了一系列措施(如减少数据传输冗余、降低传输频率等),但还是有极少的数据出现了丢失。考虑到聋哑手语的应用场景对时间敏感度较小以及数据采集频率较大,本文采用克隆上条对应数据的方法。

惯性传感器数据融合算法针对惯性传感器数据特征对数据进行Mahony互补滤波,但这并没有保证数据的平滑性,同时忽略了弯曲传感器数据的噪声。为了解决这个问题,中值滤波可以有效提高数据的平滑性,同时其运算简单,不会影响手语识别的实时性,其表达式如下:

$$\text{pitch}[n] = \frac{1}{2M+1} \sum_{m=-M}^M \text{pitch}_r[n+m] \quad (13)$$

$$\text{yaw}[n] = \frac{1}{2M+1} \sum_{m=-M}^M \text{yaw}_r[n+m] \quad (14)$$

$$\text{roll}[n] = \frac{1}{2M+1} \sum_{m=-M}^M \text{roll}_r[n+m] \quad (15)$$

$$f_r[n] = \frac{1}{2M+1} \sum_{m=-M}^M f_r[n+m] \quad (16)$$

其中, $\text{pitch}_r[n]$ 、 $\text{yaw}_r[n]$ 、 $\text{roll}_r[n]$ 和 $f_r[n]$ 分别为俯仰角、偏航角、翻滚角和弯曲传感器的原始数据, $\text{pitch}[n]$ 、 $\text{yaw}[n]$ 、 $\text{roll}[n]$ 和 $f_r[n]$ 为滤波后的信号数据。本文将 $M$ 设为5,将滤波窗口设为11。

### 4.2 时序数据分割

时序数据分割的目的是识别出每个手语动作的开始和结束,连续的传感器数据可以分割成一个个单独的手语序列区间。在文献[17,28]中,通过按下和释放按钮可以直接对数据进行分割,但在实际使用中,并不会有人愿意在动作开始时按下按钮,结束后释放按钮。于是,本文提出一种时序数据的分割方案。图6所示为聋哑拼音手语“ $A$ ”的姿态信息,在未作出手语动作时,姿态数据是相对稳定的;相反,当作出手语动作时,姿态数据会产生较大的变化。设预处理后的姿态数据序列 $P = \{p[1], p[2], \dots, p[L]\}$ ,其中 $p[n] = (\text{pitch}[n], \text{yaw}[n], \text{roll}[n])$ 是一个三维向量, $d[n]$ 为 $p[n]$ 与 $p[n-1]$ 之间的欧氏距离。手语动作区间的 $d[n]$ 远大于无动作的 $d[n]$ ,可用于时序数据分割。

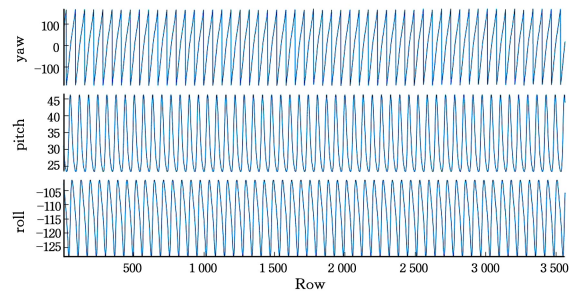


图6 聋哑拼音手语“ $A$ ”姿态信息

Fig. 6 Sign language phonetic transcription “ $A$ ”

### 4.3 特征提取

时序数据被分割成原子序列区间后,提取区间中的有效特征对于聋哑手语识别至关重要。考虑到惯性传感器的原始数据已经进行了融合,为了保证手语识别的实时性,本文从分割后的手部运动区间中提取了5个特征,包括均值、均方值、方差、标准差和角度特征。考虑到运动具有方向性,特征的计算都是基于传感器数据的绝对值进行。

1) 均值

$$E(\mathbf{X}) = \frac{1}{W} \sum_i^W |\mathbf{X}[i]| \quad (17)$$

其中, $W$ 为手语动作区间长度。

2)均方值

$$E(\mathbf{X}^2) = \frac{1}{W} \sum_i^W (\mathbf{X}[i])^2 \quad (18)$$

3)方差

$$D(\mathbf{X}) = E\{[\mathbf{X} - E(\mathbf{X})]^2\} \quad (19)$$

4)标准差

$$S = \sqrt{\frac{\sum_{i=1}^W (|\mathbf{X}| - E(\mathbf{X}))^2}{W}} \quad (20)$$

5)角度特征

本文构造了一个运动的特征向量,由不同时间段内的总的角度变化构成。对于同一个动作,无论运动快慢,其运动的角度变化都是一致的,故该特征与运动速度无关。图7为该特征向量的表现形式,图中的 $\Delta\theta_i$ 元素计算如下。

$$\Delta\theta_i = \sqrt{\left(\sum_{j=1}^i \omega_{jx} * \Delta t\right)^2 + \left(\sum_{j=1}^i \omega_{jy} * \Delta t\right)^2 + \left(\sum_{j=1}^i \omega_{jz} * \Delta t\right)^2} \quad (21)$$

其中, $\omega_{jx}$ 为第 $j$ 组陀螺仪的 $x$ 轴数据; $\omega_{jy}$ 为第 $j$ 组陀螺仪的 $y$ 轴数据; $\omega_{jz}$ 为第 $j$ 组陀螺仪的 $z$ 轴数据; $\Delta t$ 是两个样本之间的间隔时间。 $\Delta\theta_1$ 是第一个 $\Delta t$ s期间的总角度变化, $\Delta\theta_2$ 是第二个 $\Delta t$ s期间的总角度变化, $\Delta\theta_n$ 是前 $n$ 个 $\Delta t$ s期间的总角度变化。

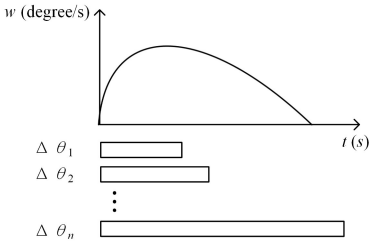


图7 角度变化

Fig. 7 Total angular change

## 4.4 分类器构建

### 4.4.1 支持向量机

支持向量机的本质是寻找一个分离超平面,实现间隔最大化。对于非线性问题,需要将训练样本从原始空间映射到一个高维空间,使得样本在高维空间中线性可分。令 $\phi(x)$ 表示将 $x$ 映射后得到的特征向量,则分离超平面所对应的模型表示如式(22)所示:

$$f(X) = w^T \phi(X) + b \quad (22)$$

其中, $w$ 为分离超平面法向量,则规划函数如下所示:

$$\min_{a,b} \frac{1}{2} \|w\|^2, y_i (w^T \phi(x) + b) \geq 1, i = 1, 2, \dots, m \quad (23)$$

### 4.4.2 K-近邻

K-近邻法的核心思想是:在特征空间中,有 $K$ 个与某个样本最相似的样本大多都属于同一个类别,则这个样本也属于该类。K-近邻法的前提是所选择的邻近点都已经完成分类。样本点之间的相似性是通过计算对象间距离实现的,通常距离使用欧氏距离或哈曼顿距离,表达式如下:

$$\text{欧氏距离: } d(x, y) = \sqrt{\sum_{k=1}^n (x_k - y_k)^2} \quad (24)$$

$$\text{曼哈顿距离: } d(x, y) = \sqrt{\sum_{k=1}^n |x_k - y_k|} \quad (25)$$

其中, $n$ 表示特征向量的维数。

### 4.4.3 前馈神经网络

前馈神经网络由多层有向图构成,每层与下一层之间的节点都采用全连接。本文的前馈神经网络由3层构成,如图8所示,分别是输入层、隐藏层和输出层。

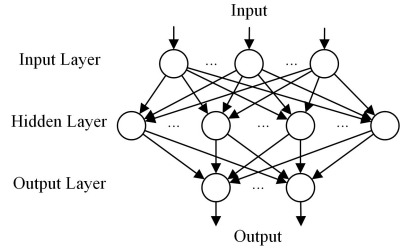


图8 三层前馈神经网络拓扑结构

Fig. 8 Feed forward neural network structure

分别假设 $N_i, N_h$ 和 $N_o$ 是输入层、隐藏层和输出层的节点个数,输出层的第 $k$ 个节点的表达式如下:

$$y_k = f\left(\theta_k + \sum_{h=1}^{N_h} \omega_{hk} f\left(\lambda_h + \sum_{i=1}^{N_i} v_{hi} x_i\right)\right) \quad (26)$$

其中, $x_i$ 为输入层的第 $i$ 个节点, $v_{hi}$ 为输入层和隐藏层节点之间的权重, $\omega_{hk}$ 为隐藏层和输出层节点之间的权重, $\theta_k$ 和 $\lambda_h$ 为偏差, $k=1, \dots, N_o$ <sup>[30]</sup>。除了输入层节点,每个节点都是具有非线性激活函数的神经元,在本文为logistic函数,其表达式如下:

$$f(z) = \frac{1}{1 + e^{-z}} \quad (27)$$

其中, $z$ 是输入突触的加权和。

### 4.4.4 模型集成

对于3个分类算法识别出的结果,本文采取Boosting模型进行集成。分类模型权值根据上一次错误率自适应地变化,为避免过拟合以及考虑到手语识别的实时性,本文将迭代次数设置为50。

## 5 实验设计与结果

本节通过实验验证文中所提手语识别算法的有效性。实验采用的训练数据来自3名志愿者(2名男性和1名女性),每个志愿者分别被要求重复每个手语动作50次,共计4500个手语动作,并将所有轨迹数据作为训练集。

在评估过程中,随机选取了另外两名志愿者,每个志愿者重复每个手语动作100次,聋哑手语识别算法对于各个聋哑拼音手语识别的准确率总体达到96.5%。由于聋哑拼音手语数量多,本文不一一列举,仅对两组在识别时容易混淆的聋哑拼音手语进行讨论。

拼音“C”“O”“CH”的聋哑手语如图9所示。



图9 “C”“O”“CH”聋哑手语

Fig. 9 Sign language phonetic transcription “C” “O” “CH”

其实验结果的混淆矩阵如图10所示。

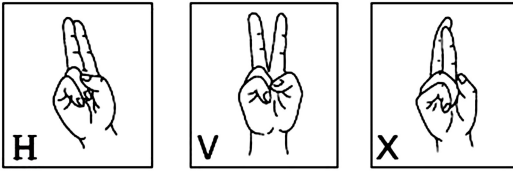


图 10 “C”“O”“CH”聋哑手语识别结果

Fig. 10 Results of sign language phonetic transcription “C” “O” “CH”

拼音“C”“O”“CH”聋哑手语动作姿态十分相似,只是在手指形态上有细微差别。由于弯曲传感器的阻值随弯曲角度的变化并非完全线性,导致了这组手语动作容易混淆。

同样,拼音“H”“V”“X”的聋哑手语动作的差别也大多在手指,如图 11 所示。

|     |     |     |     |     |
|-----|-----|-----|-----|-----|
|     |     | 实际值 |     |     |
|     |     | “H” | “V” | “X” |
| 识别值 | “H” | 172 | 15  | 12  |
|     | “V” | 20  | 179 | 5   |
|     | “X” | 8   | 6   | 183 |

图 11 “H”“V”“X”聋哑手语

Fig. 11 Sign language phonetic transcription “H” “V” “X”

其实验结果的混淆矩阵如图 12 所示。

|     |      |     |     |      |
|-----|------|-----|-----|------|
|     |      | 实际值 |     |      |
|     |      | “C” | “O” | “CH” |
| 识别值 | “C”  | 156 | 25  | 14   |
|     | “O”  | 35  | 163 | 10   |
|     | “CH” | 9   | 12  | 176  |

图 12 “C”“O”“CH”聋哑手语识别结果

Fig. 12 Results of sign language phonetic transcription “C” “O” “CH”

为了验证惯性传感器融合算法的有效性,本实验利用基于原始传感数据的方法与基于融合后数据的方法进行聋哑手语识别,并对结果进行对比,如表 1 所列。

表 1 基于原始数据与基于融合后数据的拼音聋哑手语识别方法准确率

Table 1 Recognition accuracy of sign language methods based on original and fused data

| Mthod                  | Accuracy/% |
|------------------------|------------|
| Raw data               | 94.5       |
| Data after Calculation | 96.5       |

结果表明,对惯性传感器数据进行融合可以提高拼音聋哑手语识别的准确率。表面上,融合计算会消耗部分计算资

源,但基于边缘计算技术将传感器数据融合计算设置在主控模块可以解决此问题。同时,由于融合后数据冗余减少、特征明显,反而提高了手语识别的效率。

实验中,惯性传感器数据融合后,将聋哑手语识别算法与传统的支持向量机、K-近邻法和前馈神经网络进行了比较,结果如图 13 所示。

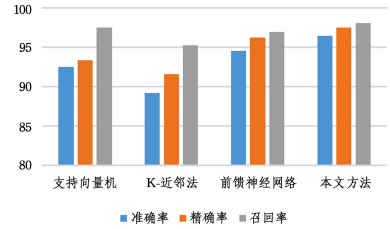


图 13 传感数据融合后各算法结果比较

Fig. 13 Results of different methods after data fusion

结果表明,本文提出的聋哑手语识别算法的识别效果相比传统的支持向量机、K-近邻法和前馈神经网络有所提升。原因在于本文提出的方法基于优化思想,采用 Boosting 模型对传统识别模型进行了集成。

**结束语** 本文提出了一种基于惯性传感器融合控制算法的聋哑手语识别算法,采用反馈控制思想,对两种传统姿态信息计算方法进行融合,减少了环境对传感器的影响。聋哑手语识别算法由数据采集与信号预处理、时序数据分割、特征提取、分类器构建和模型融合等组成。该算法对姿态融合后的数据进行预处理,并将数据分割成原子时间区间,通过自适应模型集成方法集成支持向量机、K-近邻法和前馈神经网络等分类算法,优化了最终结果。实验证明,针对 30 个汉语拼音聋哑手语,该方法的准确率相比传统的支持向量机、K-近邻法和前馈神经网络提升了 2%。

参 考 文 献

- [1] SHARMA V, KUMAR V, MASAGUPPI S C, et al. Virtual talk for deaf, mute, blind and normal humans[C]// 2013 Texas Instruments India Educators' Conference. IEEE, 2013: 316-320.
- [2] BHATNAGAR V S, MAGON R, SRIVASTAVA R, et al. A cost effective Sign Language to voice emulation system[C]// 2015 Eighth International Conference on Contemporary Computing (IC3). IEEE, 2015: 521-525.
- [3] SRIRAM N, NITHIYANANDHAM M. A hand gesture recognition based communication system for silent speakers[C]// 2013 International Conference on Human Computer Interactions (ICHCI). IEEE, 2013: 1-5.
- [4] PHI L T, NGUYEN H D, BUI T T Q, et al. A glove-based gesture recognition system for Vietnamese sign language[C]// 2015 15th International Conference on Control, Automation and Systems (ICCAS). IEEE, 2015: 1555-1559.
- [5] HOQUE M T, RIFAT-UT-TAUWAB M, KABIR M F, et al. Automated Bangla sign language translation system: Prospects, limitations and applications[C]// 2016 5th International Conference on Informatics, Electronics and Vision (ICIEV). IEEE, 2016: 856-862.
- [6] ELHAYEK H, NACOUZI J, KASSEM A, et al. Sign to letter

- translator system using a hand glove[C]// The Third International Conference on e-Technologies and Networks for Development (ICeND2014). IEEE,2014:146-150.
- [7] BEDREGAL B R C,DIMURO G P. Interval fuzzy rule-based hand gesture recognition[C]// 12th GAMM-IMACS International Symposium on Scientific Computing, Computer Arithmetic and Validated Numerics (SCAN 2006). IEEE,2006:12.
- [8] BORGHETTI M,SARDINI E,SERPELLONI M. Sensorized glove for measuring hand finger flexion for rehabilitation purposes[J]. IEEE Transactions on Instrumentation and Measurement,2013,62(12):3308-3314.
- [9] PŁAWIAK P,SONICKI T,NIEDWIECKI M,et al. Hand body language gesture recognition based on signals from specialized glove and machine learning algorithms[J]. IEEE Transactions on Industrial Informatics,2016,12(3):1104-1113.
- [10] TANYAWIWAT N,THIEMJARUS S. Design of an assistive communication glove using combined sensory channels[C]// 2012 Ninth International Conference on Wearable and Implantable Body Sensor Networks. IEEE,2012:34-39.
- [11] VUTINUNTAKASAME S,JAJONGRAK V,THIEMJARUS S. An assistive body sensor network glove for speech-and hearing-impaired disabilities[C]// 2011 International Conference on Body Sensor Networks. IEEE,2011:7-12.
- [12] KANWAL K,ABDULLAH S,AHMED Y B,et al. Assistive Glove for Pakistani Sign Language Translation[C]// 17th IEEE International Multi Topic Conference 2014. IEEE,2014:173-176.
- [13] FU Y F,HO C S. Static finger language recognition for handicapped aphasiacs[C]// Second International Conference on Innovative Computing, Information and Control (ICICIC 2007). IEEE,2007:299.
- [14] ARIF A,RIZVI S T H,JAWAID I,et al. Techno-Talk: An American Sign Language (ASL) Translator[C]// 2016 International Conference on Control,Decision and Information Technologies(CoDIT). IEEE,2016:665-670.
- [15] WU X,MAO X,CHEN L,et al. Trajectory-based view-invariant hand gesture recognition by fusing shape and orientation[J]. IET Computer Vision,2015,9(6):797-805.
- [16] WANG J S,CHUANG F C. An accelerometer-based digital pen with a trajectory recognition algorithm for handwritten digit and gesture recognition[J]. IEEE Transactions on Industrial Electronics,2011,59(7):2998-3007.
- [17] ZHANG X,CHEN X,LI Y,et al. A framework for hand gesture recognition based on accelerometer and EMG sensors[J]. IEEE Transactions on Systems,Man, and Cybernetics-Part A:Systems and Humans,2011,41(6):1064-1076.
- [18] 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,2014,44(2):293-299.
- [19] ZHOU S,FEI F,ZHANG G,et al. 2D human gesture tracking and recognition by the fusion of MEMS inertial and vision sensors[J]. IEEE Sensors Journal,2013,14(4):1160-1170.
- [20] LIU K,CHEN C,JAFARI R,et al. Fusion of inertial and depth sensor data for robust hand gesture recognition[J]. IEEE Sensors Journal,2014,14(6):1898-1903.
- [21] LU T. A motion control method of intelligent wheelchair based on hand gesture recognition[C]// 2013 IEEE 8th Conference on Industrial Electronics and Applications (ICIEA). IEEE,2013:957-962.
- [22] ZHU C,SHENG W. Wearable sensor-based hand gesture and daily activity recognition for robot-assisted living[J]. IEEE Transactions on Systems,Man, and Cybernetics-Part A:Systems and Humans,2011,41(3):569-573.
- [23] 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,2014,15(1):154-163.
- [24] HUSSAIN S M A,RASHID A B M H. User independent hand gesture recognition by accelerated dtw[C]// 2012 International Conference on Informatics, Electronics & Vision (ICIEV). IEEE,2012:1033-1037.
- [25] XU R,ZHOU S,LI W J. MEMS accelerometer based nonspecific-user hand gesture recognition[J]. IEEE Sensors Journal,2011,12(5):1166-1173.
- [26] ABID M R,PETRIU E M,AMJADIAN E. Dynamic sign language recognition for smart home interactive application using stochastic linear formal grammar[J]. IEEE Transactions on Instrumentation and Measurement,2014,64(3):596-605.
- [27] AKL A,FENG C,VALAEE S. A novel accelerometer-based gesture recognition system[J]. IEEE Transactions on Signal Processing,2011,59(12):6197-6205.
- [28] ZHANG Z Q. Two-step calibration methods for miniature inertial and magnetic sensor units[J]. IEEE Transactions on Industrial Electronics,2014,62(6):3714-3723.
- [29] CHEN C H,WU J C,CHEN J H. Prediction of flutter derivatives by artificial neural networks[J]. Journal of wind engineering and industrial aerodynamics,2008,96(10/11):1925-1937.
- [30] ZHANG L,WANG H,LIANG J,et al. Decision support in cancer base on fuzzy adaptive PSO for feedforward neural network training[C]// 2008 International Symposium on Computer Science and Computational Technology. IEEE,2008:220-223.



**RAN Meng-yuan**, born in 1995, postgraduate, is a student member of China Computer Federation. His main research interests include artificial intelligence and pattern recognition.



**LIU Li**, born in 1981, Ph.D, associate professor, is a member of China Computer Federation. His main research interests include big data analysis, artificial intelligence and smart wearable technology.