

基于陀螺传感器的三维手势识别方案

黄启友^{1,2}, 戴 永¹, 胡明清², 刘军发², 陈益强²

(1. 湘潭大学信息工程学院, 湖南 湘潭 411105; 2. 中国科学院计算技术研究所普适计算中心, 北京 100190)

摘 要: 传统输入设备无法满足人们的随意性输入需求。为此, 提出一种基于陀螺仪传感器的三维手势识别方案。硬件架构由陀螺传感器信息采集模块、单片机信息处理模块以及射频无线传输模块组成。利用多功能滤波器进行数据预处理, 设计一种基于角度的特征提取算法, 提取三维手势特征。实验结果表明, 该方案的平均识别率达到 99.3%, 能较好地实现 3D 空间的鼠标输入功能和键盘输入功能。

关键词: 3D 输入; 手势识别; 陀螺传感器; 人机交互; 支持向量机; 多类分类器

3D Gesture Recognition Scheme Based on Gyro Sensor

HUANG Qi-you^{1,2}, DAI Yong¹, HU Ming-qing², LIU Jun-fa², CHEN Yi-qiang²

(1. College of Information Engineering, Xiangtan University, Xiangtan 411105, China;

2. Research Center for Pervasive Computing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China)

【Abstract】 Traditional input devices are increasingly unable to meet people's needs of random input. Aiming at the defect, a 3D gesture recognition scheme based on gyroscope sensor is proposed. The hardware architecture of scheme includes a high accuracy data acquisition hardware system based on gyroscope and a MCU based on atmega16 and a wireless transmission module based on nrf24API. The gesture recognition algorithm includes data preprocessing module based on multisection filter. A feature extraction algorithm is based on angular movement to extract the 3D gesture feature. Experimental results show that the average recognition rate of this scheme is 99.3% which can well realize 3D space mouse input function and the keyboard function.

【Key words】 3D input; gesture recognition; gyro sensor; human-computer interaction; Support Vector Machine(SVM); multi-class classifier

DOI: 10.3969/j.issn.1000-3428.2011.22.050

1 概述

随着信息技术的发展, 在很多人机交互场合, 3D 输入设备越来越重要。而传统的输入设备存在许多不足, 如触摸屏、普通键盘它们的交互模式还停留在二维阶段, 不能脱离特定的平面。基于传感器的 3D 输入设备不仅为使用者解除空间和线缆的束缚, 在自由的空间完成相应的人机交互, 并且能够支持输入各种形状图案和文字数字信息。这种新型的输入设备越来越受到研究领域和工业界的重视。

在研究领域, 目前已经有许多相关的工作。对于手势识别一般分为 2 类, 基于视觉信息的手势识别^[1]和基于传感器信息的手势识别。基于视觉信息的识别轨迹跟踪算法复杂度普遍较高^[2], 此外, 基于视觉的技术普遍对周围环境的光线和用户所处的位置和方向有强烈的依赖性。文献[3]利用当前广泛存在的加速度传感器捕捉人的体感动作, 但文献[4]指出加速度传感器在低速和匀速操作中灵敏度不高。在结合视觉信息和运动传感器信息方面, 文献[5]采用摄像头和数据手套结合的方法, 利用人眼和指尖的连线确定用户指向屏幕位置进行手势识别达到简单的鼠标输入功能, 但仍然无法克服视觉信息手势识别的缺陷如光照影响和复杂的图像计算处理。本文提出一种使用低速高精度的角速度陀螺仪进行手势识别的方案。

2 硬件架构

手势识别的硬件平台框架如图 1 所示, IDG300 角速度陀螺仪的感应度为 2.0 mV/°/s, 能够感知人手势动作如向上、向下、向左、向右等。把检测到的模拟信号通过自身装备的 A/D 转换器转换成数字信号。Atmega16 是整个手势识别平台

的中心单元, 负责陀螺传感器的初始化, 从其数据接口读取传感器数据, 并把数据打包发送给无线收发器 nrf24API。nrf24API 是手势识别平台的无线通信模块, 工作在 2.4 GHz 的共同频段。为了能够实现多传感器数据的有效处理, 还设计了上位机硬件设备。该设备采用 USB 接口, 能够有效实现与上位机的连接。上位机设备如图 1 所示, 主要包括 nrf24API 和 Atmega16 单片机等。Atmega16 系列单片机带有 USB 接口, 因此, 可以无缝地跟主机连接。



图 1 手势识别的硬件平台框架

3 软件架构

除了上述的手势识别硬件平台, 还设计了上位机软件。上位机主要负责实现鼠标功能、陀螺传感器数据的预处理、手势识别。鼠标的功能主要通过计算手势动作在水平和垂直方向的相对偏移, 然后把相对移动角度转化成鼠标在屏幕上的实际位移。对陀螺传感器数据的预处理和识别, 主要由上位机完成。手势识别的软件平台框架见图 2, 包括手势动作

基金项目: 湖南省高校创新平台开放基金资助项目(09K040)

作者简介: 黄启友(1984—), 男, 硕士研究生, 主研方向: 智能人机交互, 模式识别; 戴 永, 教授; 胡明清, 助理研究员、博士后; 刘军发, 博士; 陈益强, 研究员、博士

收稿日期: 2011-05-23 **E-mail:** qiyouhuang@163.com

分割、数据预处理、数据序列归一化、特征提取、支持向量机(Support Vector Machine, SVM)多类分类等处理模块组成。

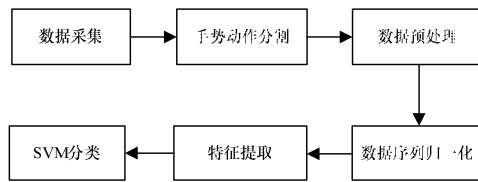


图2 手势识别的软件平台框架

4 手势识别算法

手势识别算法主要包括手势动作分割、数据预处理、数据序列归一化、特征提取和支持向量机多类分类器。

4.1 动作分割

手势识别硬件平台专门设计了一个按钮。通过按键的按下和松开区分手势信息和垃圾信息。

4.2 数据预处理

本文采用3个滤波器,分别是空闲状态滤波器、异常数据滤波器以及平滑滤波器。空闲滤波器主要过滤掉陀螺仪传感器静止时的数据,一般在手势识别动作的开始和结束阶段出现。因为静止时的数据对识别效果的贡献小,且增加了后续识别的处理难度。异常数据滤波器过滤由于动作分割阶段按键按下和松开时手的抖动以及电压不稳造成的异常数据。平滑滤波器针对原始数据的曲线不平滑,这里采取滑动窗口结合中值滤波算法,滑动窗口大小设置为20。

4.3 数据序列归一化

由于人们在输入同一个动作时的速度和幅度不同,因此导致所采集到的数据序列呈现不同的长度。即使对于同一个人而言,输入相同一个动作,所花费的时间和动作的幅度也不尽相同,因此需要消除这方面的依赖性。本文采用亚采样和插值方法把不等长的原始数据序列归一化成等长的数据序列。一方面消除了对动作幅度、持续时间的依赖性;另一方面便于后续的数据处理。

4.4 特征提取

手特征提取是影响手势识别性能最关键的一个环节。由于人们在输入手势动作时,不同的人动作幅度和速度不一样,使用统计方法很难找到合适的具有较强鲁棒性的特征。此外,从用户角度来看,用户输入手势动作后期望在短时间能得到响应(给出识别结果),因此,要求特征提取具有较低的算法复杂度。两轴陀螺传感器的输出为如图3所示的2个方向的瞬时角速度。通过大量的统计分析,本文发现如图3所示的瞬时运动角度信息可以作为一个可靠的特征。这样的特征可以表征人们的手势。因此,在假设X轴水平放置的情况下,可以选取运动方向与水平的夹角作为特征描述。

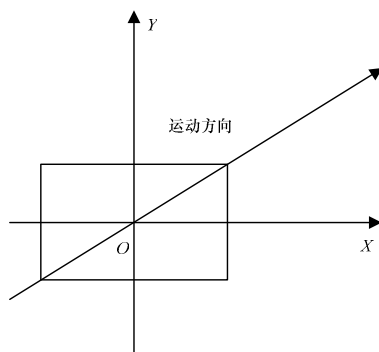


图3 手势运动方向的角度分解示意图

具体地,本文选取如下方法表征角度特征,即:

$$F_x(i) = \frac{x_i}{\sqrt{x_i^2 + y_i^2}} (i=1, 2, L, K)$$

$$F_y(i) = \frac{y_i}{\sqrt{x_i^2 + y_i^2}} (i=1, 2, L, K)$$

其中, x_i 、 y_i 是所有预处理操作完成后 i 时刻陀螺传感器输出的水平和垂直方向的瞬时角速度; K 是预处理完成后实际数据序列长度,这里取 250。上述特征反映了 i 时刻输入动作的角度信息。

4.5 支持向量机多类分类

首先基于大量的数据做离线训练,把特征数据序列作为训练样本,最终输出支持向量模型对应的向量空间,拉格朗日乘子和偏移量参数。在线实时预测阶段,对预处理好的特征向量与训练好的支持向量机模型求决策值,最大决策值对应的下标是其所在的类别。

5 实验

5.1 数据集的设计与采集

本文采用手势识别平台实现对 0~9 的数字识别,以此验证上述预处理方法和特征提取方法的有效性。首先,需要对识别的数字进行定义,这样定义一方面需要考虑输入方法要满足人们的书写习惯,还要兼顾较高的区分度。本文设计如图4所示的数字书写风格,其中,圆圈是起始点;箭头代表手势的方向。由图4可以看到,每个数字都能一笔写成,书写方便,且输入符合大多数人的书写习惯。图5是相对应的特征曲线图(限于篇幅,选取部分图片罗列),表示数据序列归一化后时刻 t 的角度特征值 y 。从图5可以看出,识别模式曲线区分度大。

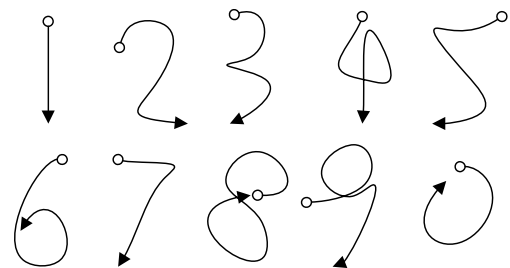
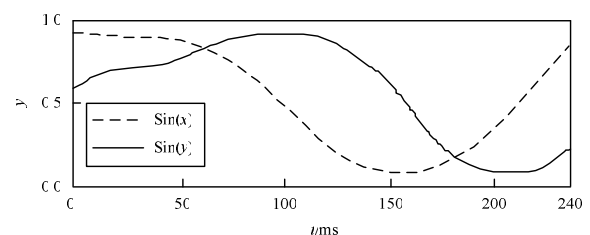
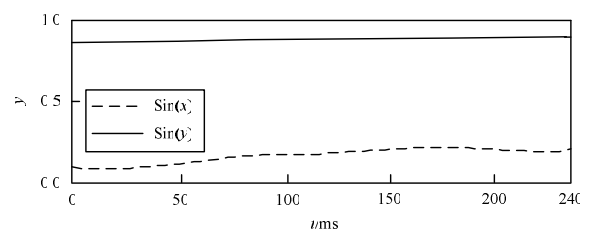


图4 0~9数据集设计



(a)数字0特征曲线



(b)数字1特征曲线

图5 数字特征曲线

在数据采集方面,采集 22 名志愿者,其中,14 位男性,8 位女性,年龄在 29 岁~45 岁之间,对 0~9 共 10 个要识别的数字每人分别采集 200 个样本。总共采集 44 000 个样本。这些样本经过预处理和特征提取后作为训练支持向量机模型的输入数据。

5.2 结果分析

对每个数字做 100 次测试,实验结果如图 6 所示。

输入	输出									
	0	1	2	3	4	5	6	7	8	9
0	100									
1		99						1		
2			100							
3				100						
4					96					4
5						100				
6							100			
7								100		
8									100	
9					2					98

图 6 用户独立模式下的数字识别矩阵

在图 6 中,输入是用户实际手势动作输入;输出表示实际识别结果。平均识别率达到 99.3%。从上述结果看,识别效果理想,完全符合该系统实际应用的要求。

6 结束语

本文采用陀螺仪传感器感知人性化自然的手势输入数

据,结合 RF 无线通信、USB 接口、手势识别等技术,并提出一种使识别率效果显著的鲁棒特征提取算法。该系统采用的硬件设备具有低成本、非依赖性特点,很好地实现 3D 空间的鼠标输入功能和键盘输入功能。该系统在实现数字调台以及上网冲浪时的文本输入有较好的表现,并且在三网融合领域具有较好的应用前景。

参考文献

[1] 任海兵,祝远新,徐光佑,等.基于视觉手势识别的研究综述[J].电子学报,2000,28(2):118-121.

[2] Abe K, Saito H, Ozawa S. 3D Drawing System via Hand Motion Recognition from Two Cameras[C]//Proc. of International Conference on Systems, Man and Cybernetics. Nashville, Tennessee, USA: IEEE Press, 2000: 840-845.

[3] Zhang Shiqi, Yuan Chun, Zhang Yan. Handwritten Character Recognition Using Orientation Quantization Based on 3D Accelerometer[C]//Proc. of the 5th Annual International Conference on Mobile and Ubiquitous Systems. Dublin, Ireland: [s. n.], 2008.

[4] 曹 青,欧阳红林,胡 彪,等.基于角速率陀螺仪的无线鼠标的设计与实现[J].微计算机信息,2008,24(35):230-232.

[5] 肖志勇,秦华标.基于视线跟踪和手势识别的人机交互[J].计算机工程,2009,35(15):198-200.

编辑 陆燕菲

(上接第 152 页)

识别同一感知对象的这 2 种特征的神经元自身发放脉冲的强度就是相同的。所以,就要看识别灰度和形状特征的神经元的发放脉冲强度的关系是怎样的。

在实验中,识别形状和颜色的神经元坐标都可以通过像素坐标来表示。对于每一种特征,任取 4 组不同的神经元坐标,计算出自身发放脉冲的强度,其计算公式如式(1)~式(6)所示。在式(1)、式(2)中,连接矩阵 W 、 M 均为 3×3 的矩阵,其连接强度计算如下式所示:

$$W,M=1/\sqrt{(i-\bar{i})^2+(j-\bar{j})^2}$$
 (8)

其中, $i,j\in(1,2,3)$; \bar{i},\bar{j} 为 3×3 矩阵中心元素的平均横、纵距离,取定值为 1.5。

重复计算 50 次,即每种特征取 200 个不同的神经元坐标,每次得到的同一物体中不同特征对应的神经元发放脉冲强度都一样,计算结果如表 1 所示。

表 1 不同特征神经元自身的发放脉冲强度

特征名称	神经元坐标(i,j)	发放强度	感知对象
长方形	$i\in(88,268),j\in(18\sim251)$	7	图形 1
圆形	$i\in(92\sim187),j\in(280\sim379)$	26	图形 2
深色	$i\in(100\sim250),j\in(30\sim240)$	7	图形 1
浅色	$i\in(90\sim260),j\in(380\sim470)$	26	图形 2

在表 1 中,通过计算每种特征对应神经元发放脉冲的强度,可以得到对应于深色的长方形的神经元的脉冲发放强度均为 7,而浅色的圆形的脉冲发放强度均为 26。那么,可以得到深色与长方形捆绑在一起作为感知对象 1,浅色和圆形捆绑在一起作为感知对象 2。相同的脉冲发放强度对应同一

感知对象,从而实现了特征捆绑的任务。

5 结束语

本文在 PCNN 的基础上提出了一种基于强度的 PCNN 模型,通过仿真实验,得到了如下结论:神经元的邻域内神经元脉冲发放总强度将不同的特征分离开来,而神经元自身的脉冲发放强度又将属于同一感知对象的不同特征捆绑起来。该模型较好地解决了特征的分离与捆绑问题,为今后特征捆绑问题的进一步研究提供了一定的参考价值。

参考文献

[1] Livingstone M, Hubel D. Segregation of Form, Color, Movement and Depth: Anatomy, Physiology, and Perception[J]. Science, 1988, 240(4853): 740-749.

[2] Treisman A, Gelede G. A Feature-integration Theory of Attention[J]. Cognitive Psychology, 1980, 12(1): 97-136.

[3] Watanabe W, Nakanishi K, Aihara K. Solving the Binding Problem of the Brain with Bi-directional Functional Connectivity[J]. Neural Networks, 2001, 14(4/5): 395-406.

[4] 石志伟,史忠植,刘 曦,等.特征捆绑的计算模型[J].中国科学: C 辑,2008,38(5):485-493.

[5] Eckhorn R, Reitboeck H J, Arndt M, et al. Feature Linking via Synchronization Among Distributed Assemblies: Simulation of Results from Cat Cortex[J]. Neural Computation, 1990, 2(3): 293-307.

[6] 罗美淑,刘世勇,石 磊.基于微分进化的 PCNN 图像分割方法[J].计算机工程,2010,36(21):225-227.

编辑 任吉慧

