

文章编号: 2095-2163(2020)01-0271-04

中图分类号: TP391

文献标志码: A

基于 Leap Motion 的三维动态手势识别研究

严雨灵, 陈闵叶, 吕亚辉

(上海工程技术大学 航空运输学院, 上海 201620)

摘要: 手势识别作为人机交互的重要手段,其识别率的提高是当今重要的研究方向。通过分析当今三维手势识别技术的发展状况,采用 Leap Motion 设备针对三维动态手势进行数据采集和处理,使用 SVM 与 PNN 模式识别算法,分别进行了两类手势区分以及多类手势区分的仿真实验。结果表明,经过优化的 PNN 对动态手势均较 SVM 能达到更高的识别率,且经过 PCA 处理后运算效率较高。

关键词: 模式识别; Leap Motion; SVM; PNN

Research on 3D dynamic gesture recognition based on Leap Motion

YAN Yuling, CHEN Minye, LÜ Yahui

(School of Air Transportation, Shanghai University of Engineering Science, Shanghai 201620, China)

[Abstract] Gesture recognition is an essential means of human-computer interaction. The improvement of recognition rate is an important research direction today. By analyzing the development of today's 3D gesture recognition technology, Leap Motion equipment is used for data acquisition and processing of 3D dynamic gestures. Using SVM and PNN pattern recognition algorithms, two kinds of gesture differentiation and multi-class gesture differentiation simulation experiments are carried out. The results show that the optimized PNN can achieve higher recognition rate than dynamic SVM, and the computational efficiency is higher after PCA processing.

[Key words] pattern recognition; Leap Motion; SVM; PNN

0 引言

随着计算机技术的不断发展,三维手势作为最近兴起的人机交互手段,已然受到学界的广泛关注。三维手势需要的输入包含深度的信息,目前主要分为基于数据手套的手势识别方法与基于视觉的手势识别方法^[1]。基于数据手套的主要有 Wiseglove、5DT Glove、Cybertouch 等。该方法可以获得精确的手部动作,但其设备价格较高且会产生束缚感,影响使用者的沉浸感^[2]。麻省理工学院著名科学家 Dertouzos 指出“未来的技术应该适应人的需求,而不是与之相反”^[3],手势识别技术也应当主动适应客户的需求,而不是要求用户做出太多的变化去适应设备。

1 Leap Motion 设备

2010年,微软推出了 Kinect for Windows,能够捕捉人体形状、动作,识别和完成面部的跟踪,首次引起了手势识别技术的研究热潮,但其识别精度有限^[4]。2013年,Leap Motion 公司发布了 Leap Motion Controller 小型体感器设备,能够捕捉 150°范

围视场以及约 8 平方英尺的交互式空间内的多个物体,其针对手部区域进行识别,感应区间精确到 0.01 mm^[5],远远超过 Kinect 的 4 mm。本文研究的虚拟维修过程针对手部动作,因此采用 Leap motion 设备进行数据的采集^[6]。

2 手势数据采集

本次研究中,实验过程设计采集 5 个人的 4 组动态手势,分别为握拳转动、手指张开的转动、抓取以及推压,用以保证算法的可靠性。Leap Motion 采集到的数据主要包括掌心数据 handdirection、handpalmposition (用 C 表示),手指数据 fingerposition (用 F_i 表示,根据设备设定 $i = 0, 1, 2, 3, 4$)。

由于每个人的手型各不相同,因此本文采用指尖至掌心的距离作为特征值,并进行归一化处理以增强适应性和鲁棒性。研究拟采用指标简述如下。

(1) 归一化因子。其数学公式为:

$$G = \frac{\sum_{i=1}^5 \|F_i - C\|}{5}, \quad (1)$$

作者简介: 严雨灵(1994-),男,硕士研究生,主要研究方向:模式识别;陈闵叶(1962-),男,教授,主要研究方向:计算机视觉、虚拟现实;吕亚辉(1995-),男,硕士研究生,主要研究方向:机器人智能控制。

通讯作者: 严雨灵 Email:13788979506@163.com

收稿日期: 2019-09-19

(2) 距离 D 。其数学公式为:

$$D = \frac{\sum_{i=1}^5 \|F_i - C\|}{G} \quad (2)$$

组合各个特征值, 得到特征值向量集 $V = [D_1, D_2, D_3, D_4, D_5, F_1, F_2, F_3, F_4, F_5]$ 。单个手势将每一帧采集到数据的特征向量集合并, 得到新的特征向量集进行后续操作。采集过程如图 1 所示。

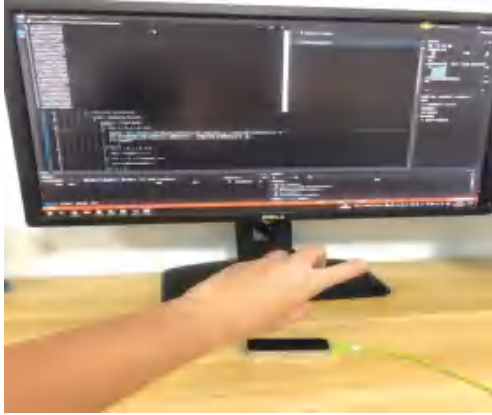


图 1 动态手势识别试验

Fig. 1 Dynamic gesture recognition test

3 手势识别方法

3.1 支持向量机

1995 年, Vapnik 利用统计学习理论作为基础, 提出了支持向量机。最初该理论的提出是针对模式识别中一对一的分类问题, 在此基础上发展成多分类问题^[7]。

设样本集为 (X_i, Y_i) , $i = 1, \dots, n$, $x \in R$, $y \in \{+1, -1\}$ 。

对于线性可分情况, 样本集由一个超平面区分开, 把这个超平面记为 $w \cdot x + b = 0$ 。而最优的超平面要求不但要将两类数据正确地分开, 还要使这两类的间隔达到最大。构造最优超平面的问题可以转变为在式(4)的约束下求式(3)的最小值的问题。此时用到的公式为:

$$\phi(w) = \frac{1}{2} \|w\|^2, \quad (3)$$

$$\text{s.t. } y_i(w \cdot x_i + b) \geq 1, i = 1, 2, \dots, N. \quad (4)$$

由于一部分样本点会偏离较远, 导致线性不可分。因此, 引入了松弛变量(非负)来允许部分样本点在一定范围内的偏离。惩罚系数 C 用来控制松弛变量大小, 从而得到更准确的分类平面^[8]。

针对非线性分类问题, SVM 将问题转化为高维特征空间的线性问题, 从而求取最优线性分类平面。

由核函数代替最优分类超平面的点积, 从而免去了复杂的高维空间运算。常用的核函数主要有线性核函数、多项式核函数、径向基核函数等。

本文采用台湾大学林志仁教授开发设计的一个快速有效的 SVM 软件包进行手势数据的分类识别。其中提供了编译好的执行文件, 通过 Mex 编译器即可在 Matlab 平台中使用, 代码开源, 而且也提供了针对具体问题优化的空间。

3.2 概率神经网络

概率神经网络由 Specht 首先提出, 并以径向基函数网络发展而来, 就是一种前馈型神经网络(Probabilistic Neural Networks, PNN)。通过采用多变量 Parzen 窗估计不同类的概率密度函数。特点有训练时间短、结构固定, 能产生贝叶斯后验概率输出, 具有强大的非线性识别能力^[9]。

径向基网络和竞争神经网络共同组成的概率神经网络总共有 3 层, 分别是: 输入层、径向基层和竞争层(输出层)。竞争层的输出结果就是整个网络的最终分类结果, 输入的样本向量的个数与径向基层中神经元的个数相同, 训练样本数据的分类数等于输出层的神经元个数, 径向基层的神经元具有阈值。利用输入向量以及对应的期望类别, 构建 PNN 网络对输入向量进行正确分类。数据以及相应的类别通过向量设计网络, 再将数据回代得到结果。

4 动态手势识别实验

4.1 实验准备

本文采用支持向量机和概率神经网络两种算法对动态手势进行识别, 在 Matlab 平台中, 采用分类算法进行手势识别的具体流程如图 2 所示。

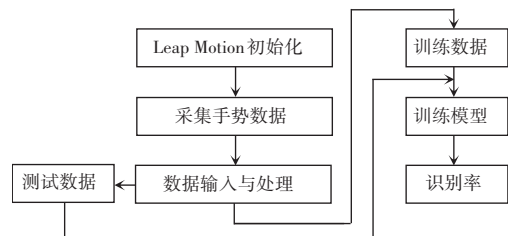


图 2 手势识别流程图

Fig. 2 Gesture recognition flowchart

在进行训练前, 须归一化处理数据, 从而为后续处理做好准备并保证程序运行时收敛更快。在过程中需用到如下公式:

$$X_i = \frac{x_i - \min(x)}{\max(x) - \min(x)}, i = 1, \dots, n. \quad (5)$$

其中, X_i 表示特征 X 的某个值; $\max(x)$ 表示该

特征的最大特征值; $\min(x)$ 表示该特征中的最小特征值。

4.2 结果分析

采用 Libsvm 和 PNN 进行分类试验, 两两区分, 标签设为 1 和 2。单个手势样本总量为 1 000 组, 按 8:2 的比例确定训练集和测试集样本, 即每一种手势中的 800 组用作训练集样本, 200 组用作测试集样本。

本文采取隔 50 ms 取一次数据的形式记录手势数据, 再合并成为样本特征, 从而提高数据的有效性并使运算量可控。一对一手势分类结果见表 1, 一对多手势分类结果见表 2。

表 1 一对一手势识别效果

Tab. 1 One-to-one gesture recognition effect

	性质参数	识别率/%	运行时间/s
SVM	线性核	89.5	0.050
	多项式核	89.0	0.139
	RBF 核	88.0	0.053
	Sigmoid 核	50.0	0.039
PNN	$S = 0.125$	95.0	0.881

一对一的实验结果表明, 线性核函数可达到 89.5% 的识别率, 而多项式核函数在达到 89.0% 的优良识别率下运行时间最短。同时分析指出, PNN 神经网络经过高斯参数的优化后最高可达到 95.0% 的识别率, 明显优于支持向量机算法。

表 2 一对多手势识别效果

Tab. 2 One-to-many gesture recognition effect

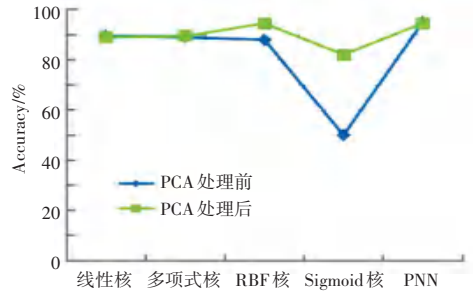
	性质参数	识别率/%	运行时间/s
SVM	线性核	70.25	0.113
	多项式核	75.50	0.123
	RBF 核	67.25	0.309
	Sigmoid 核	25.00	0.351
PNN	$S = 0.44$	84.00	0.878

一对多的实验结果表明, 多分类的识别率和两类分类相比较低, 但总体规律不变。SVM 算法中采用多项式核函数识别率较高, 而 PNN 神经网络经过高斯参数的优化后识别率仍明显高于支持向量机算法。

4.3 PCA 算法处理

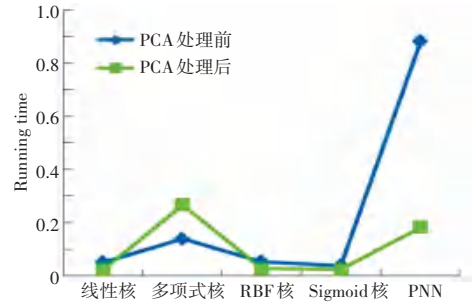
为了进一步提高识别效果与降低 PNN 神经网络的运行时间, 采用了主成分分析 (PCA) 对特征值进行降维操作, 旨在排除噪声数据的干扰从而提高准确率。实验结果表明, SVM 算法的准确率有所提高, PNN 算法在保证准确率的情况下由于 PCA 的降

维作用使运行时间大幅减少。一对一使用后结果对比如图 3 所示。一对多使用后结果对比如图 4 所示。



(a) 准确率

(a) Accuracy

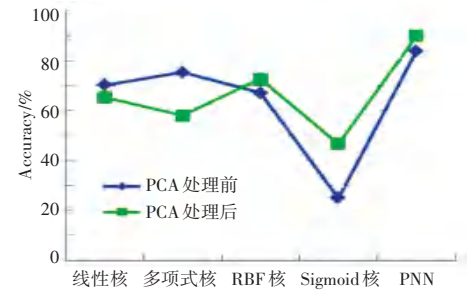


(b) 运行时间

(b) Running time

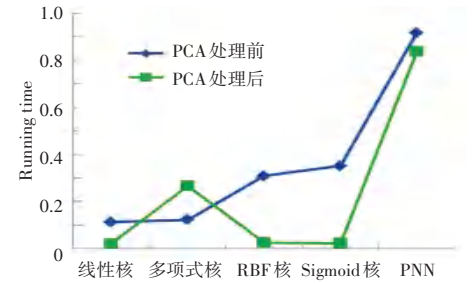
图 3 一对一使用 PCA 处理后识别结果

Fig. 3 One-to-one recognition results using PCA



(a) 准确率

(a) Accuracy



(b) 运行时间

(b) Running time

图 4 一对多使用 PCA 处理后识别结果

Fig. 4 One-to-many recognition results using PCA