

文章编号:1005-0523(2005)05-0086-02

基于BP神经网络的静态手势识别方法

邓志国

(华东交通大学 基础科学学院,江西 南昌 330013)

摘要:手势识别正成为人机交互技术研究中的一种重要模式.本文针对运用摄像机和计算机视觉技术捕获来的静态手势,提出了一种手势特征提取和基于神经网络分类的手势识别方法.实验表明该方法的正确识别率可以达到98%左右,是一种非常有效的静态手势识别方法.

关键词:边缘跟踪;BP神经网络;手势识别

中图分类号:TP391

文献标识码:A

0 引言

随着多媒体技术的发展,计算机已经具备了处理语音、图形、图象和文字等多种通信媒体的能力,从计算机到用户的通信带宽得到了进一步的提高.为了提高计算机的使用效率,克服现有交互技术的种种不足,必须研究新的输入技术^[1]如语音输入技术、在线手写体输入技术、手势输入技术,以扩大从用户到计算机的通信带宽,使用户更加方便自然地和计算机系统交互.

广义地讲,手势是指人有意识地作出的手的运动(包括手指的弯曲、伸展,手腕的转动和手在空间的移动).不论是操纵工具或物体执行某种任务,还是进行交流,手势都表达出手势者的某种意图.由于长期的学习和人与人之间的相互影响,人们作出的手势表达的含义基本上是确定的,即在确定的场合下表达确定的含义.如我们握拳伸出大拇指朝上,表示对某人或某事的赞赏.因此,手势包括两方面的含义:手的运动及其表达出的手势者的意图或者说手势所表达的概念.在使用“手势”一词时,一般不作区别,可根据上下文判别.手势的识别就是根据用户的手势识别手势的含义.

手势识别是目标识别中的一种,是机器视觉系统中的一个关键部分.在绝大多数情况下,需要识别系统具有不变性识别能力(即平移、旋转及比例不变性).一般来说,目标识别的精度主要取决于目标特征的选择和所使用的分类器.与传统的识别方法相比,神经网络技术能够自动地从原始数据中获取良好的表达特征,从而避免了特征优化过程及所需的巨大运算量;神经网络结构本身具有鲁棒性,允许分类目标在一定的动态范围变化.

本文提出了一种基于BP神经网络的手势识别方法.实验表明该方法的正确识别率可以达到98%,是一种非常有效的静态手势识别方法.本文结构如下安排:第二部分介绍了手势图像的预处理,第三部分介绍了如何提取手势的特征向量,第四部分介绍了如何利用BP神经网络进行手势识别;第五部分是实验结果.

1 手势图像的预处理

对于一幅运用摄像机和计算机视觉技术捕获来的静态手势,我们首先要将手势图像利用阈值分割法把手势从背景中分割出来,然后再将分割出来的图像进行二值化获得黑白图像.为了后续的特征

收稿日期:2005-06-20

作者简介:邓志国(1978-),男,湖南长沙人,讲师.

提取,我们要将手势的轮廓跟踪出来.基于二值图像的轮廓跟踪的基本方法是:先根据“探测准则”找出目标轮廓上的像素,再根据这些像素找出目标物体上的其它像素.这个探测准则:从第一个边界点开始,定义初始的搜索方向为沿左上方;如果左上方的点是黑点,则为边界点,否则搜索方向顺时针旋转 45 度.这样一直到找到第一个黑点为止.然后把这个黑点作为新的边界点,在当前搜索方向的基础上逆时针旋转 90 度,继续用同样的方法搜索下一个黑点,直到返回最初的边界点为止.图 1 为利用该轮廓跟踪算法对手势跟踪的结果.

在图 1 中,(a)经过轮廓跟踪后得到(b)所示的情景.一般我们可以认为,对一幅手势图像进行轮廓跟踪后得到轮廓数据 $(x[m], y[m])$,其中 $0 \leq m \leq l-1$,为轮廓的像素数目.



图 1 手势图像的轮廓跟踪

2 手势特征提取

对每幅手势图像经过轮廓跟踪后就会得到轮廓数据 $(x[m], y[m])$,我们对该数据分别进行离散傅立叶变换,即

$$a[k] = \frac{1}{l} \sum_{m=0}^{l-1} x[m] e^{-jk(2\pi/l)m}$$

$$b[k] = \frac{1}{l} \sum_{m=0}^{l-1} y[m] e^{-jk(2\pi/l)m}$$

其中 $0 \leq k \leq l-1$.除去 $a[0], b[0]$,因为他们代表图像的中心位置信息.为了获得图像的不变性表示,令

$$r[k] = \sqrt{a[k]^2 + b[k]^2}$$

$$s[k] = r[k]/r[0], \text{其中 } 1 \leq k \leq l-1.$$

由上和傅立叶变换的性质容易看出 $r[k]$ 对平移和旋转具有不变性,进一步 $s[k]$ 是关于比例不变的.

由于傅立叶系数的低频部分主要由图像的整体轮廓所决定,其高频部分仅刻画了外形的细节,高频部分对目标识别的作用不大,并且随着系数 n 的增大, $s[n]$ 快速地衰减.所以我们只需要选取前

面几个比较大 $s[n]$ 的作为特征向量即可.

3 BP 神经网络分类器

在分类和识别的设计中,该文选用了 BP 神经网络^[3].BP 神经网络是一种前向传播的多层网络,网络除了输入节点以外还有一层或多层隐含节点,其训练方法被称为误差反传播算法,它利用实际输出和期望输出之差对网络的多层连接权由后向前逐层进行校正.

应用神经网络分类器的一个主要优点就是它具有并行分布计算能力,从而不必像传统的分类器那样从大量的数据中进行特征比较.BP 网络是应用最广泛的网络模型之一,它能满足一般的模式识别的需要^[2],本文与文献 2 不同的是,将 3 层 BP 网络结构应用于静态手势识别,该 BP 网络输入层的节点数由选取的特征向量 $s[n]$ 维数确定,神经元函数采用 S 型函数,且每个神经元有一阈值 θ .基本 BP 网络调整权值的选择方向是在梯度下降最速方向,在这个方向上,网络的目标值下降最快.

4 实验结果

本实验所使用的硬件设备为:P III 733 CPU, 256 MB RAM 个人电脑.实验对象为随机选取的 8 幅标准手势图像,如图 2 所示.

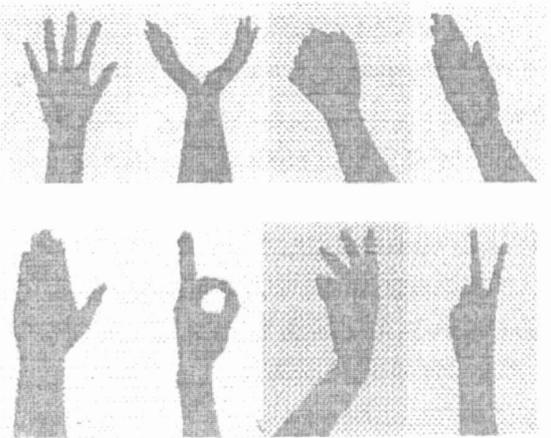


图 2 八幅标准的手势图像

我们选取的特征向量 $s[n]$ 的维数是 10 维,即取 $s[k]$ 中前面的 10 个值,从而 BP 网络的输入节点数是 10 个,输出节点数是 8 个.通过不断调整网络参数,当选取的中间节点数是 20 个,网络的学习率为 0.3,惯性系数为 0.5,经过 2 000 次学习周期后,误差可以减少到 0.03 以下.实际进(下转第 106 页)

The Application of Three-phase V/V Connected Traction Transformer in Engineering Design

QIU Chang-wen

(Guangzhou Subway Corporation, Guangzhou 515700, China)

Abstract: According to the structure principle and characteristics of three-phase V/V connected traction transformer, this paper introduces its application in engineering design, and proposes some suggestions.

Key words: three-phase V/V connected traction transformer; design; application

(上接第 87 页)行识别时,当每个输出节点的输出均小于 0.7 时,我们认为该目标无法识别。

将八幅标准手势图像进行缩放,旋转,移位得到 80 幅测试图像,并且通过多次重复实验,测试结果表明平均正确识别率为 98%。在实验中,我们对于手工绘制的另类图片也进行了测试,也得到了较好的识别效果。例如图 3(a)是一幅手工绘制的手势图像,通过我们的手势识别系统得出它与第一幅标准图像图 3(b)最相似。

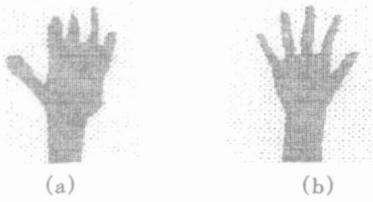


图 3 另类手势图像的识别

本文针对运用摄象机和计算机视觉技术捕获来的静态手势,提出了一种手势特征提取和基于神经网络分类的手势识别方法。实验表明该方法的正确识别率可以达到 98%,是一种非常有效的静态手势识别方法。

参考文献:

- [1] Kenneth Tolman. Multimodal communication. Proceedings of SPIE² The International Society for Optical Engineering, Soc. for Optical Engineering, WA, USA, v1785 1993, p 298—306.
- [2] Hong Bong. Object recognition of one of tools by a BP neural network. IEEE Trans Neural Networks, 1995, 6 (2): 484—487.
- [3] 殷勤业, 杨宗凯. 模式识别与神经网络[M]. 北京:机械工业出版社, 1992.

5 结论

A Hand Gestures Recognition Based on Neural Networks

DENG Zhi-guo

(School of Natural Sciences, East China Jiaotong University, Nanchang 330013, China)

Abstract: Hand gestures play a natural and intuitive communication mode for all human dialogs. This paper presents a hand gestures recognition algorithm that is based on neural networks. Experiment results show that recognition rate can reach to 98%.

Key words: contour-tracing; BP neural networks; hand gestures.