

# 一种复杂背景下的手势提取方法

李瑞峰 贾建军

(哈尔滨工业大学 机器人技术与系统国家重点实验室, 黑龙江 哈尔滨 150001)

**摘要:** 介绍了一种在一般复杂背景条件下将手势与背景分离的新方法. 该方法首先在一般背景下对获得的手势图片进行卷积滤波处理, 验证了肤色的聚类特性, 然后综合利用手势的肤色特征和其特有的几何特征, 成功地将手势与背景分离, 最后通过不同背景、不同手势提取实验验证了该方法的有效性. 将该手势分割方法用于 8 种手形手势识别实验, 静态手势识别率能达到 99 %.

**关键词:** 图像处理; 手势; 识别; 分类; 特征; 复杂背景

**中图分类号:** TP242.2 **文献标识码:** A **文章编号:** 1671-4512(2008) S1-0080-03

## A method of extracting gesture from complex background

Li Ruifeng Jia Jianjun

(State Key Laboratory of Robotics and System, Harbin Institute of Technology, Harbin 150001, China)

**Abstract:** A new method which can segment gesture from general and complex backgrounds is introduced. First, this method deals the gesture picture which is get in the general background by convolute filter, and verifies clustering character of complexion, then by the complexion and geometry characters of the gestures, this method segments gestures from the backgrounds successfully. At last, by experiments of different backgrounds and eight different gestures, the validity of the method is verified, static state identify rate is 99 %.

**Key words:** image processing; gesture; recognition; classification; character; complex background

人机交互是目前机器人重点发展和研究的领域, 人机交互可以通过遥控装置、人自身的声音、肢体语言等来实现, 而基于手势识别的人机交互越来越成为研究的热点<sup>[1]</sup>.

基于计算机视觉的手势识别方法使用摄像机直接拍摄手势的运动过程, 从手势图像序列中分割出人手, 根据图像序列前后帧之间的关系计算出运动参数. 如美国 MIT 媒体实验室的 Starner, 通过提取左右质心的运动轨迹、手形的形状等特征参数, 结合语法规则识别 40 个美国手语, 正确率达到 97 % (没有语法规则的正确率为 90.7 %). 另外, Microsoft Korea 的 Hyeon-KyuLee, 利用基于 HMM 的阈值模型 (HMM banded threshold model) 从单手的运动轨迹识别出 9 种手势命令, 平均识别率达到 98.19 %<sup>[2]</sup>. 在基于单目或

者双目摄像头的手势识别中, 把图像中的人手与背景进行分离始终是一个难点. 实际情况中, 背景往往是不可预测的, 并且是多变的. 以往的实验常常采取佩戴带标记的手套, 或者复杂背景简单化等手段, 这样的方法可以获得高的识别率, 但这并不能满足现实条件下人机交互的需要. 本研究提出一种在实验室一般复杂背景条件下, 有效提取手势的方法.

## 1 图片的获取及处理工作

### 1.1 实验环境

由于有强烈光照<sup>[3]</sup>或者特别黑暗的环境会使图片的颜色产生失真, 从而对机器人的识别产生严重干扰, 因此采取一般光束避免过强或过弱光

收稿日期: 2008-07-15.

作者简介: 李瑞峰(1965-), 男, 教授, E-mail: lrf100@hit.edu.cn.

基金项目: 国家高技术研究发展计划资助项目 (2007AA04Z221); 长江学者团队计划项目资助 (IRT0423).

照导致的失或干扰. 图 1 为实验人员在实验室背景下,与摄像头距离大约为 1.5 ~ 1.8 m,向机器人发出手势命令,手势命令表示左手数字 3. 获取到的图片格式为 24 bit ,bmp 格式,图片大小为 768 ×576.

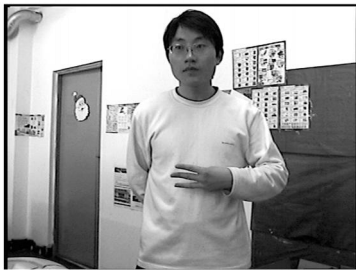


图 1 表示左手数字 3 的手势

1.2 图片的平滑处理

手势图片往往由于摄像头本身的系统误差,或多或少的在手势的边缘处存在锯齿,这对后期各种手势特征的提取工作有很大的影响. 因此,有必要在提取手势之前对图片进行平滑处理. 平滑操作本质上是对图片进行低通滤波,常用的处理方法有平均领域法、中值滤波,高斯平滑、卷积滤波等. 由于图片格式为 24 bit ,因此采取卷积滤波<sup>[4]</sup>的方法. 卷积操作是以像素的灰度值为基础的,应分别用同一卷积对各像素的  $R, G, B$  颜色分量分别卷积,并将 3 个分量的卷积结果作为像素心  $R, G, B$  值的 3 个分量. 其中卷积核矩阵为

$$\{1/9, 1/9, 1/9, 1/9, 1/9, 1/9, 1/9, 1/9, 1/9\}.$$

经过平滑操作后的图片如图 2 所示. 由于卷积系数是浮点数,若直接用于计算,则运算量会很大,影响系统实时性. 因此,本文用一个简单的技巧来解决这个问题. 先对卷积核中的卷积系数乘以其最小公倍数,得到一个所有的卷积系数都为整数的新卷积核. 用新卷积核进行卷积操作,然后将结果除以原卷积核各卷积系数的最小公倍数,就可以求得真正的卷积结果,而且运算效率有很大的提高.



图 2 平滑处理后的效果图

实验证明对图片进行 2 或 3 次的平滑操作对后期特征提取的效果是最好的.

2 手势与背景的分割

2.1 通过色度理论进行分割

本文中手势的提取是基于 24 bit 彩色图片的色度空间理论,根据手的肤色与背景的不同来进行分割. 首先通过手的颜色来获取到手颜色的聚类空间<sup>[5]</sup>,判断是否可以通过颜色来进行分割. 在此将处于  $RGB$  空间的图片转换到  $HSV$  空间. 其中:  $H$  为色度;  $S$  为饱和度;  $V$  为亮度.

图 3 是用来检验色度空间  $H$  值聚类的图片. 图 4 是  $H$  值的直方图分布.

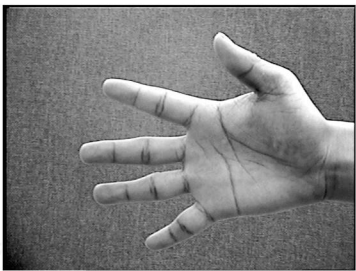


图 3 用来检测  $H$  值聚类特性的图片

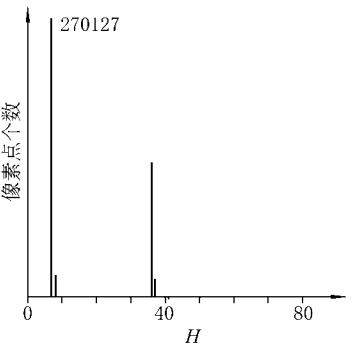


图 4  $H$  值直方图

由图 4 可以看出,左边的分布为背景蓝色,右边的分布为手的肤色,  $H \in [30, 45]$ , 因此手势的肤色具有很好的聚类特性. 同理也可得到  $S \in [35, 200]$ ,  $V \in [20, 255]$ . 运用上述  $H, S$  和  $V$  阈值对图 3 进行阈值分割,将处在阈值范围内的像素点转变为黑色,即 0; 否则转变为白色,即 255.

由此可以得到如图 5 所示的效果图. 由图 5 可见,运用颜色空间理论对手势图片进行分割,可以过滤掉大约 80 % 的无用信息. 但仍有许多无用信息,如面部、背景颜色接近肤色部分. 下面运用手势的几何特性对手势图片进行再次分割.

2.2 通过几何特性<sup>[6]</sup>进行分割

2.2.1 通过面积阈值进行分割

显然,图 5 中存在大量面积不等的连通区域,因此先对图 5 进行二值图像标记处理. 采用 8 连通判别算法<sup>[7]</sup>,从左到右,从上到下逐个像素进行



图 5 经过颜色分割后的效果图

扫描,若发现某像素点像素值为 0,则依次检测该点的右上、正上、左上、及左前 4 个点的像素值,然后根据 8 连通判别原则对像素进行标号,最后达到对整个连通区域进行标号.标记结束后,对各个连通区域进行像素点总和统计,即面积统计.最后通过面积阈值  $[S_{\min}, S_{\max}]$  对图像进行分割.通过面积阈值能过滤掉剩下的 10 % 的无用信息.

### 2.2.2 通过矩形长宽阈值进行分割

手势的长和宽具有一定的阈值,在上一步面积分割后,可以得到只含有少数连通区域的手势图片,那么可以在上一步程序中嵌入判断长宽阈值的程序.如果剩下的连通区域为标号为  $I$ ,那么计算  $I$  区域最大外接矩形的长和宽.通过长和宽的阈值  $[w_{\min}, w_{\max}]$ ,  $[h_{\min}, h_{\max}]$  对图像进行分割.

### 2.2.3 通过矩形 1/4 处的交点间隔进行分割

很显然,手指的宽度可以做为一个阈值来分割.图 6 为此分割方法的原理图.

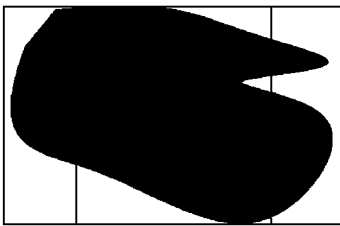


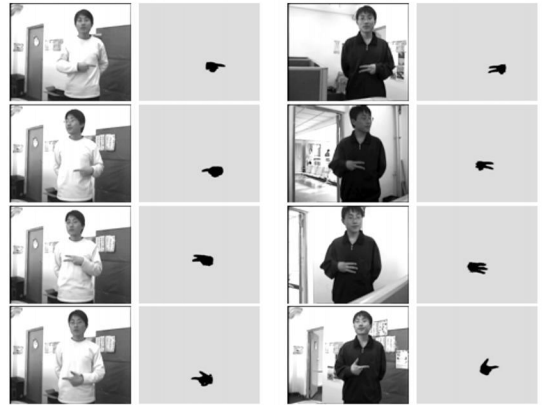
图 6 矩形分割原理图

在通过面积和长宽阈值分割后,剩下的连通区域  $I$  中,找出其外接矩形,然后用矩形左右 1/4 处的竖线进行切割,并标记竖线上颜色变化处的点.标记完成后,计算黑色区域的交点间隔宽度选定手指宽度阈值为  $[l_{\min}, l_{\max}]$ ,若存在一个交点间隔在阈值范围允许范围内,则认为该阈值为手势区域,经过上述步骤,可以得到最终的手势图片.

## 3 实验与评价

在实验室环境下,通过摄像头得到若干个不同实验背景下的不同手势图片.通过本算法对其

进行手势与背景分割实验,分割前后图片对比如图 8.



(a) 相同背景

(b) 不同背景

图 7 两种背景手势分割后的效果图

从上面的图片可以看出,本算法对一般复杂环境下手势与背景的分割有很好的效果,所提取出来的手势轮廓分明而且比较完整,对于后期各种手势特征的提取工作做了很好的准备工作.

将该手势分割方法用于手势识别实验,手形采用图 7 中的 8 种手形,静态手势识别率能达到 99 %.用摄像头采集若干帧连续帧图像,提取手形质心运动的轨迹,并将轨迹分为上下、左右,不确定,这样形成 24 种轨迹,对于该动态手势的识别率达到 97 %,和其他识别方法相比,本识别方法不用高亮标记的视觉手套,算法简单实用,与其他手势识别方法具有相当的识别率.

本文所描述的提取手势的算法综合利用了手势的肤色特征、手势的几何形状特征,很好地解决了在一般复杂背景下的手势提取的难点.

## 参 考 文 献

- [1] 任海兵,祝远新,徐光祐,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] 任海兵,祝远新,徐光祐,等.复杂背景下的手势分割与识别[J].自动化学报,2002,28(2):256-261.
- [3] 刘丽华.人脸识别方法综述[J].云南师范大学学报,2005,25(6):13-18.
- [4] 周长发.精通 Visual C++ 图像处理编程[M].3 版.北京:电子工业出版社,2007.
- [5] 吴飞.模式识别——原理、方法及应用[M].北京:清华大学出版社,2003.
- [6] 周航,阮秋琦.基于表现特征的手势识别[J].广西师范大学学报,2003,21(1):25-30.
- [7] 杨淑莹.VC++ 图像处理程序设计[M].2 版.北京:清华大学出版社,2006.