

融合图像轮廓矩和傅立叶描叙子的手势识别

雷海龙, 杨 胜, 刘红新

(湖南大学 计算机与通信学院, 湖南 长沙 410082)

摘 要: 针对手势识别中人手是复杂变形体, 手部特征描叙容易受到环境因素影响的特点, 提出了一种基于混合轮廓特征的手势识别的新方法. 首先根据手的颜色信息将手从复杂的背景中分离出来, 然后提取出手轮廓的矩不变量和 Fourier 描叙子特征, 然后将 2 种特征融合构成混合特征, 最后将新特征输入 BP 神经网络进行识别. 实验表明: 此方法具有较好的鲁棒性和较高的识别率.

关键词: 手势识别; 矩不变量; 傅立叶描叙子; BP 神经网络

中图分类号: TP391

文献标识码: A

文章编号: 1000—7180(2011)04—0103—04

Fusion of Image Contour Moments and Fourier Descriptors for the Hand Gesture Recognition

LEI Hai-long, YANG Sheng, LIU Hong-xin

(Institute of Computer and Communications, Hunan University, Changsha 410082, China)

Abstract: Hand is a highly variable organ and hand features are easily affected by environmental factors. Considering the characteristics of hand gesture, a novel hand gesture recognition algorithm based on fusion contour moments is presented. First, According to the color cue the hand is available to extract from the complicated background, then the contour moment invariant and Fourier Descriptor are extracted and fused into a mixed feature, finally the mixed feature are put into BP network to identity. The experimental results show that the method proposed has better robustness and higher recognition rate.

Key words: hand gesture recognition; moment invariant; Fourier descriptors; BP neural network

1 引言

手势为一种自然而直观的信息表达方式, 已经成为一种重要的人机交互方式. 以人手直接作为计算机的输入设备, 利用手势识别技术对机器进行控制^[1], 使人机间的通讯不再需要中间媒体. 这方面的研究成果主要有基于数据手套的手势识别和基于视觉的手势识别技术^[2], 由于基于数据手套的识别需要使用者戴上笨重的识别设备, 限制了交互的自然性和方便性, 并不实用, 因此当前主要研究基于计算机视觉的手势识别技术, 而基于计算机视觉的手势识别主要由手势分割, 手势特征提取, 手势识别等阶段组成.

在复杂背景下进行手势分割, 由于背景的复杂性和环境光照变化的不稳定性, 人手的颜色容易受到前景和复杂背景中的类肤色的干扰, 出现误检, 因此成功的手势识别面临着诸多的困难. 不少的研究者采用对手势图像加上种种限制的方法, 例如采用特殊颜色的背景或者识别者带有特殊颜色的手套来简化分割的难度, 但这人为的增加了诸多的限制, 不利于自然的人机交互.

手势的特征选择与轮廓提取也是手势识别的难点, 由于人手是非刚体物体, 很容易产生形变. 因此人的手势具有多样性, 多义性, 时间和空间上的差异性. 对于这种形状较为复杂且易多变的物体, 可以采用轮廓矩和 Fourier 描叙子的方法来提取其轮廓,

收稿日期: 2010—05—04; 修回日期: 2010—07—08

基金项目: 国家自然科学基金项目(60873184)

轮廓矩和 Fourier 描叙子这二种方法除了对目标轮廓描叙准确,充分以外,而且还具有平移,缩放以及旋转不变性,因此在基于轮廓的目标识别中得到了广泛的应用.何海清^[2]等采用几何矩来识别手势,但其拒识率较高,刘寅^[3]采用傅立叶描叙子进行手势识别,但对产生形变的同一类手势的识别效果不理想.轮廓矩和 Fourier 描叙子描叙物体的轮廓,二者各有优势,但轮廓矩对局部变化不如 Fourier 描叙子的敏感,对具有相同轮廓的物体区分能力差,而 Fourier 描叙子对细节特征非常敏感,对属于同类但局部差异较大的物体有很大误识.

文中提出一种基于轮廓矩和 Fourier 描叙子的混合特征构建方法,用于手势的识别.首先,在颜色空间中利用人手的颜色信息把手从复杂背景中提取出来,然后利用轮廓矩和 Fourier 描叙子的混合特征来进行识别,克服了单独识别时候的缺点,提高了特征的平移、缩放、旋转不变性,取得了较好的识别率.

2 手势识别系统的设计

手势识别系统由四部分构成,首先手势分割阶段把手的形状从复杂的背景中分离出来,然后进行手势的特征提取,再进行特征的融合,然后进行训练,再用来进行手势的识别.如图 1 所示.

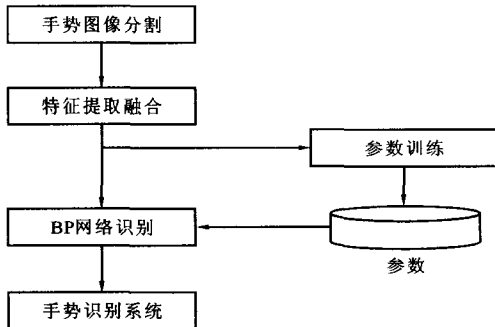


图 1 手势识别系统的流程图

2.1 手势分割

从复杂的彩色图像中快速划分出手势,是手势识别的关键步骤,它处理的好坏直接影响到手势的识别效果.常用的彩色表示方法是基于三基色的 RGB 颜色空间,但如果直接在 RGB 颜色空间里面进行肤色检验的话,那么检测的结果将会包含很多的和肤色相近的区域,例如黄色,桔黄色等等.因此,在复杂背景情况下,特别是存在近似肤色的区域的话,将无法分割出手势区域.文中在分析现有方法的基础上,采用一种双模型的基于复杂背景下的手

势分割新方法.首先,在 RGB 模式空间里面利用杨盈昀^[4]提出的人肤色检测函数 Skin Test;输入的图片为只含有人体肤色的小块区域 $M * N$,分别输出肤色特征的 RGB 的平均值:

$$R = \frac{\sum_1^{M * N} R}{M * N}, G = \frac{\sum_1^{M * N} G}{M * N}, B = \frac{\sum_1^{M * N} B}{M * N}$$

经过大量的实验发现,尽管因为光线或其他的外界因素的影响,人的肤色特征 RGB 值变化范围较大,但是由于人体肤色特征的色调的稳定性, R/G 的范围保持在 $(1.0, 1.3)$ 之间.首先建立第一个肤色模型,进行粗略的分割,然后将其转换到 YCbCr 模型里面,再对其进行了精确的分割.在 YCbCr 色彩空间中, Y 表示像素的亮度, Cb 表示蓝色分量, Cr 表示红色分量,通常对于 RGB 格式的图像,可以通过以下的公式将图片从 RGB 格式空间转换到 YCbCr 空间

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.2990 & 0.5870 & 0.1140 \\ -0.1687 & -0.3313 & 0.5000 \\ 0.5000 & -0.4187 & -0.0813 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

通过限制 Cb, Cr 的值可以确定人体的肤色区域,实验证明,采用 $Y \in [130, 165]$, $Cb \in [108, 129]$, $Cr \in [127, 146]$, 可以有效的提取出图像中的人体肤色区域.图 2 为原始手势图.图 3 为分割后的手势图.



图 2 原始的手势图



图 3 分割后的手势图

2.2 轮廓矩

不变矩理论^[5-6]是模式识别和计算机视觉中的一个重要内容. 常用的基于区域的矩不变理论最早由 Hu 提出, 因为尺度规范化矩具有平移, 缩放以及旋转不变性, 使得矩特征在模式识别中获得了不断的发展与广泛的应用. 因为 Hu 矩不变量是基于区域的, 其计算量非常的大, 为使得矩特征适合轮廓特征的提取, 陈文静^[6]对基于区域的矩进行了修改, 用对曲线的积分代替对区域的积分, 则:

$$m_{pq} = \int_C x^p y^q ds \quad p, q = 0, 1, 2, \dots$$

其中 $\int_C$ 代表沿曲线 C 的线性积分: $ds = \sqrt{(dx)^2 + (dy)^2}$, 由此可以看出, 零阶矩 $m_{0,0}$ 是曲线的弧长.

$(p+q)$ 阶中心矩可以定义为:

$$\mu_{pq} = \iint_C (x - \bar{x})^p (y - \bar{y})^q ds$$

其中, $\bar{x} = m_{10}/m_{00}$; $\bar{y} = m_{01}/m_{00}$.

对于数字图像

$$\mu_{pq} = \sum_{(x,y) \in C} (x - \bar{x})^p (y - \bar{y})^q$$

非常明显, 轮廓中心矩是对平移不变的. 为了得到尺度不变性, 同区域矩一样, 也要引进尺度规范化矩的定义, 即 $\eta'_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma}$, 其中 $\gamma = p+q+1$ ($p+q = 2, 3, \dots$).

经过规范化后的轮廓矩具有平移和尺度不变性. 同基于区域的矩一样, 也可以构造出如下7个标准矩, 并且7个标准矩同时具有平移, 尺度和旋转不变性. 这些矩不变量可以作为轮廓特征, 用于目标的识别.

$$\begin{aligned} \varphi_1 &= \mu_{2,0} + \mu_{0,2} \\ \varphi_2 &= (\mu_{2,0} + \mu_{0,2})^2 + 4\mu_{1,1}^2 \\ \varphi_3 &= (\mu_{3,0} - 3\mu_{1,2})^2 + (3\mu_{2,1} + \mu_{0,3})^2 \\ \varphi_4 &= (\mu_{3,0} + \mu_{1,2})^2 + (\mu_{2,1} + \mu_{0,3})^2 \\ \varphi_5 &= (\mu_{3,0} - 3\mu_{1,2})(\mu_{3,0} + \mu_{1,2}) \times [(\mu_{3,0} + \mu_{1,2})^2 - 3(\mu_{2,1} + \mu_{0,3})] + (3\mu_{2,1} - \mu_{0,3})(\mu_{2,0} + \mu_{0,3}) \times [3(\mu_{3,0} + \mu_{1,2})^2 - (\mu_{2,1} + \mu_{0,3})^2] \\ \varphi_6 &= (\mu_{2,0} - \mu_{0,2})[(\mu_{3,0} + \mu_{1,2})^2 - (\mu_{2,1} + \mu_{0,3})^2] + 4\mu_{1,1}(\mu_{3,0} + \mu_{1,2})(\mu_{2,1} + \mu_{0,3}) \\ \varphi_7 &= (3\mu_{2,1} - \mu_{0,3})(\mu_{3,0} + \mu_{1,2}) \times [(\mu_{3,0} + \mu_{1,2})^2 - 3(\mu_{2,1} + \mu_{0,3})^2] - (\mu_{3,0} - 3\mu_{1,2})(\mu_{2,1} + \mu_{0,3}) \times [3(\mu_{3,0} + \mu_{1,2})^2 - (\mu_{2,1} + \mu_{0,3})^2] \end{aligned}$$

2.3 Fourier 描叙子

由于区域的边界是一条封闭的曲线, 因此对边界上某一固定的起始点 b_0 , 沿边界曲线上的一个动

点 b 的坐标变化是一个周期函数. 通过规范化后, 周期函数可以展开成 Fourier 级数. Fourier 级数中的一系列系数直接与边界曲线的形状有关, 可作为形状的描述, 称为傅里叶描叙子^[7]. 对于二维平面中的封闭边界, 共有 K 个边界点. 将边界点 (x_0, y_0) , $(x_1, y_1), \dots, (x_{K-1}, y_{K-1})$ 采用复数形式来表示, 即 $s(k) = x(k) + jy(k)$, $k = 0, 1, 2, \dots, K-1$. 这样就将二维数列变成一维数列.

计算 $s(k)$ 的 Fourier 变换可以得到:

$$\alpha(u) = \frac{1}{K} \sum_{k=0}^{K-1} s(k) e^{-j2\pi uk/K} \quad (1)$$

$\alpha(u)$ 即称为边界的 Fourier 描叙子. 对其进行逆变化, 可以恢复出原来的 $s(k)$, 有:

$$s(k) = \frac{1}{K} \sum_{k=0}^{K-1} \alpha(\mu) e^{j2\pi \mu k/K} \quad (2)$$

这样通过采用部分 Fourier 系数来代替所有的 Fourier 系数, 从而实现轮廓的重建, 例如: 只采用前面 P 个系数对图像的轮廓进行重建, 此时

$$\bar{s}(k) = \frac{1}{K} \sum_{p=0}^{P-1} s(k) e^{-j2\pi \mu k/K} \quad (3)$$

因为 Fourier 变化的高频分量对应轮廓的细节分量, 而低频分量对应着轮廓的基本形状, 因此可以采用少量的低频 Fourier 系数来实现图像轮廓的重建.

2.4 特征融合

混合特征的构建过程实际上就是二种特征融合的过程. 信息融合一般分为像素级融合、特征级融合和决策级融合^[7]. 其中特征级融合在信息融合中占有重要的地位, 其优点是可以保留参与融合的多特征的有效鉴别信息, 并在一定程度上实现信息的压缩. 融合能够提高分类性能的前提是不同的特征有一定的独立性和互补性. 因为轮廓矩对局部变化不太敏感, 对具有相近轮廓的手势区分能力差, 而 Fourier 描叙子则对细节特征非常敏感, 对同种手势但局部差异较大的手势识别效果差. 所以, 针对手势的特点, 如果能将二种特征结合起来就可以使总体的识别率提高. 而因为随着分类问题的不同, 模式的特征以及分类器等都有很大的不同. 故对于特征融合没有固定的格式, 文中采用的基本思想为: 设某类模式样本空间为 Ω , 其中 A 和 B 分别是二组特征. 对于任一样本 $\xi \in \Omega$. 其对应的 2 个特征向量分别为 $\alpha \in A$ 和 $\beta \in B$, 则其合并后的特征为 $\gamma = \begin{Bmatrix} \alpha \\ \beta \end{Bmatrix}$.

3 BP 神经网络的手势识别

采用神经网络作为分类器, 神经网络的各个输

人对于分类具有同等的影响. 文中采用以下方法^[9]来实现混合特征的构建, 设计如下:

$$F = [k |s(1)|, k |s(2)|, \dots, k |s(p)|, \varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5, \varphi_6, \varphi_7]$$

其中, $|s(1)| \cdots |s(p)|$ 为前 p 个低频的 Fourier 系数的幅值, k 为权值, $\varphi_1 \cdots \varphi_7$ 是 7 个轮廓矩不变量.

确定了混合特征的构建形式后, 接下来是确定 k . 不同的 k 值决定了 2 种特征对混合特征的影响程度, 如何获得最佳的 k 值是值得研究的问题, 它不仅和分类的具体问题有关、原始特征的形式有关, 而且和后续的分类器也有关系. 通过研究发现, 对于各种手势, 在达到较高的识别率的情况下, 其 k 值基本上满足

$$k \sum_{i=1}^m |s^i(1)| = \sum_{i=1}^m (|\varphi_1^i| + |\varphi_2^i| + |\varphi_3^i| + |\varphi_4^i| + |\varphi_5^i| + |\varphi_6^i| + |\varphi_7^i|)$$

其中, k 为需要确定的权值, $|s^i(1)|$ 为第 i 个训练样本的第一个 Fourier 描叙子的幅值, m 为训练样本的个数, $|\varphi_1^i| \cdots |\varphi_7^i|$ 为第 i 个样本的 7 个轮廓矩值. 因此, 就可以求的 k 的值.

4 实验结果

本实验系统中, 预先定义了 10 中手势(对应于人用手势表示的数字 0 到 9), 每种手势取 300 个样本, 其中 150 个作为训练集合, 150 个作为测试集合. 实验的结果如表 1~表 3 所示. 其中单独使用轮廓矩时的平均正确率为 92.4%, 单独采用 Fourier 描叙子时的识别正确率是 91.5%, 而混合特征的平均正确率 96.4%. 可以看出, 采用混合特征大幅度的提高了识别正确率.

表 1 单独采用轮廓矩时的识别率

手势	0	1	2	3	4
正确率	90%	94.3%	97.3%	95.3%	93.3%
手势	5	6	7	8	9
正确率	91%	95.7%	90.3%	92.7%	89.7%

表 2 采用 Fourier 描叙子时候的识别率

手势	0	1	2	3	4
正确率	87.3%	97.3%	94.7%	94%	91.7%
手势	5	6	7	8	9
正确率	88.7%	86.7%	91.3%	92.3%	91%

表 3 采用混合特征时候的识别率

手势	0	1	2	3	4
正确率	92.7%	98%	98.3%	96.7%	96.7%
手势	5	6	7	8	9
正确率	95.3%	96%	95.7%	97.3%	94.6%

5 结束语

文中利用双肤色模型的方法, 在一定的程度上解决了纷乱背景下的手势分割问题, 结合轮廓矩和 Fourier 描叙子, 以 BP 网络作为分类器, 利用其较好的学习能力, 解决了手势在空间上不一致的问题, 达到了较高的识别率. 但在存在光照变化和阴影的情况下, 手势分割的结果仍不是很理想, 仍然需要我们做进一步的工作.

参考文献:

[1] 任海兵, 祝远新, 徐光祐, 等. 复杂背景下的手势分割与识别[J]. 电子学报, 2000, 28(2): 256—261.
[2] 何阳清, 葛元, 王林泉. 应用几何矩和边缘检测的手势识别算法[J]. 计算机工程, 2005, 31(15): 165—174.
[3] 刘寅, 腾晓龙, 刘重庆. 复杂背景下基于傅立叶描叙子的手势识别[J]. 计算机仿真, 2005, 22(15): 158—161.
[4] 杨盈昀, 谢婷婷, 施美楠. 基于肤色的人脸检测算法研究[J]. 北京广播学院学报: 自然科学版, 2002, 9(4): 11—20.
[5] 张蕾, 普杰信, 范庆辉. 一种基于改进 BP 神经网络的物体识别方法[J]. 微电子学与计算机, 2008, 25(4): 152—155.
[6] 陈文静, 陈峰, 苏显渝, 等. CCD 抽样过程对傅立叶变换轮廓术测量的影响[J]. 光电子激光, 2005, 16(9): 1074—1079.
[7] 李学勇, 路长厚, 李建美. 融合轮廓矩和 Fourier 描叙子特征的压印字符识别[J]. 光电子激光, 2007, 18(10): 1244—1248.

作者简介:

雷海龙 男, (1985—), 硕士研究生. 研究方向为图像处理、智能视频.
杨 胜 男, (1977—), 副教授. 研究方向为图像处理、智能视频、数据挖掘.
刘红新 男, (1974—), 硕士研究生. 研究方向为图像处理、智能视频.