

文章编号: 1001-9081(2013)S2-0193-04

基于 OpenCV 的数字手势识别算法

赵 健*, 张冬泉

(北京交通大学 机械与电子控制工程学院, 北京 100044)

(* 通信作者电子邮箱 11121557@bjtu.edu.cn)

摘 要: 基于视觉的手势识别提供了一种更加友好的人机交互方式。提出一种基于 OpenCV 视觉库的快速实用的静态手势识别算法, 通过将运动信息与肤色相结合的方法, 去除与肤色相近的背景区域, 实现了 HSV 色彩空间中手势图像的准确分割; 提取图像的 Hu 不变矩作为手势的特征向量, 解决了手势的旋转、平移和缩放问题; 通过计算输入手势与模板手势图像的特征向量的距离实现静态手势识别。实验结果表明, 此算法的实时性好, 10 个数字手势的平均识别率可以达到 97.4%。

关键词: OpenCV; 手势识别; 平均背景法; 肤色分割; 数字手势; Hu 矩

中图分类号: TP391.41 **文献标志码:** A

Number gesture recognition method based on OpenCV

ZHAO Jian*, ZHANG Dongquan

(School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing 100044, China)

Abstract: The gesture recognition based on computer vision can provide a more friendly interactive way. A fast and practical static gesture recognition algorithm based on OpenCV computer vision library was proposed in this paper. The motion information was combined with the skin color to remove background regions with similar color to the skin. By this way gesture images were split out accurately in HSV color space. Extracting the Hu invariable moments of the gesture images as the feature vector to solve the problem of the rotation, scale and translation of gestures. Static gesture recognition was completed by comparing the Hu moments of the input gesture image and the template gesture image. Ten number gestures can be recognized using this method. Experimental results demonstrate that this algorithm is real-time with the average recognition rate of 97.4%.

Key words: OpenCV; gesture recognition; average background method; segmentation based on skin color; number gesture; Hu moment

0 引言

手势是人们交互的重要信息, 随着计算机视觉的发展, 基于计算机视觉的手势识别方法已成为近些年来人机交互领域的一大研究热点^[1-2]。将手势识别系统用于人机接口, 从而实现直接的、自然的人机交互方式, 是手势识别的发展趋势和基本目标。

根据手势的输入介质不同, 可以将手势识别系统分为两种: 一是利用数据手套, 二是基于计算机视觉。利用数据手套实现的手势识别技术使用不方便, 而且成本较高^[3-4]。

基于视觉的手势识别一般可分为手势图像的采集、手势图像的分割、特征提取及手势识别四个步骤。通过肤色分割获取手势的二值图像, 再进一步提取手势轮廓特征^[5-6]是应用最广泛的一种方法。但人手表面的颜色受人类个体差异、光源的颜色以及光照的角度不同所造成的高光和阴影等诸多因素的影响^[7]。近年来, 不少学者开始利用手势的内部灰度变化来实现手势的检测与识别^[8]。手势识别方法主要有模板匹配法(Template Matching)、神经网络法(Neural Network, NN)、动态时间规整法(Dynamic Time Warping, DTW)及隐马尔可夫模型法(Hidden Markov Model, HMM)等。其中, 模板匹配法多用于静态手势识别中^[9-10], 而基于隐马尔可夫模型

的识别和基于动态时间规整的识别则主要用于动态手势识别^[11-12]。

本文首先将平均背景法运动检测算法与 HSV 颜色空间肤色分割算法相结合, 分割出手势图像, 避免了将背景中接近肤色的区域误检测为手势区域的情况; 接着提取手势的轮廓特征, 并计算轮廓图像的 Hu 矩作为特征向量, 通过计算输入手势与模板手势的距离, 实现对 10 个数字手势的识别。本文的算法是基于 OpenCV 开源计算机视觉库(open source computer vision library)实现的, OpenCV 具有强大的图像和矩阵运算能力, 其代码简洁且执行效率高。

1 手势分割

1.1 运动检测

本文采用平均背景法进行运动检测, 其基本思路是计算每个像素的平均值和标准差作为它的背景模型^[13]。首先对前 35 幅图像进行累计学习, 用函数 cvAbsDiff() 计算每帧图像间的绝对差图像, 累积到 35 帧图像后, 用这个累计的图像除以输入图像累计的数目 35, 来计算平均原始图像和绝对差分图像; 然后选取平均原始图像(初始背景)与绝对差分图像的适当倍数作为分割的上、下限, 将输入图像像素点与设定的阈值范围进行比较, 阈值范围内的为背景, 否则为前景。如图

收稿日期: 2013-06-07; 修回日期: 2013-07-19。

作者简介: 赵健(1988-), 女, 天津人, 硕士研究生, 主要研究方向: 人体行为识别; 张冬泉(1967-), 男, 北京人, 副教授, 博士, 主要研究方向: 轨道交通控制与嵌入式系统。

1 为原始图像,图2 为平均背景法检测的图像。



图1 原始图像

图2 检测图像

1.2 肤色分割

对不同肤色在RGB、HSV颜色空间下的颜色直方图进行

比较,如图3和图4,图中横坐标分别代表不同人种肤色的R、G、B及H、S、V分量的值,纵坐标表示相对应的像素个数。

由图3和4可以看出,不同人种的肤色的R、G、B颜色分量分布分散,且差异较大,但在HSV颜色空间中,不同人种的肤色的色调(Hue)基本是一致的,不同的只是亮度和饱和度。观察不同人种的H值的直方图可以看出,H值大部分都集中在0.055~0.09范围内。根据查阅资料和笔者的实际研究,最终确定肤色的阈值为0.03~0.128^[14]。

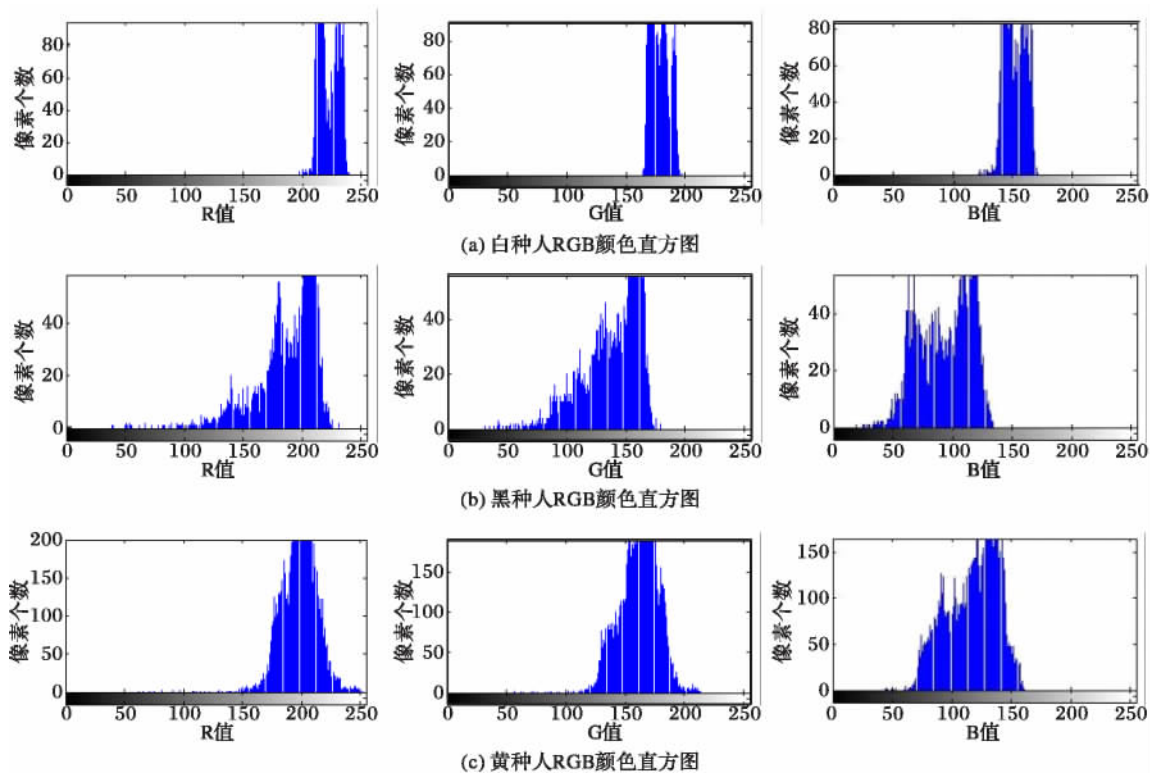


图3 不同人种 RGB 颜色直方图

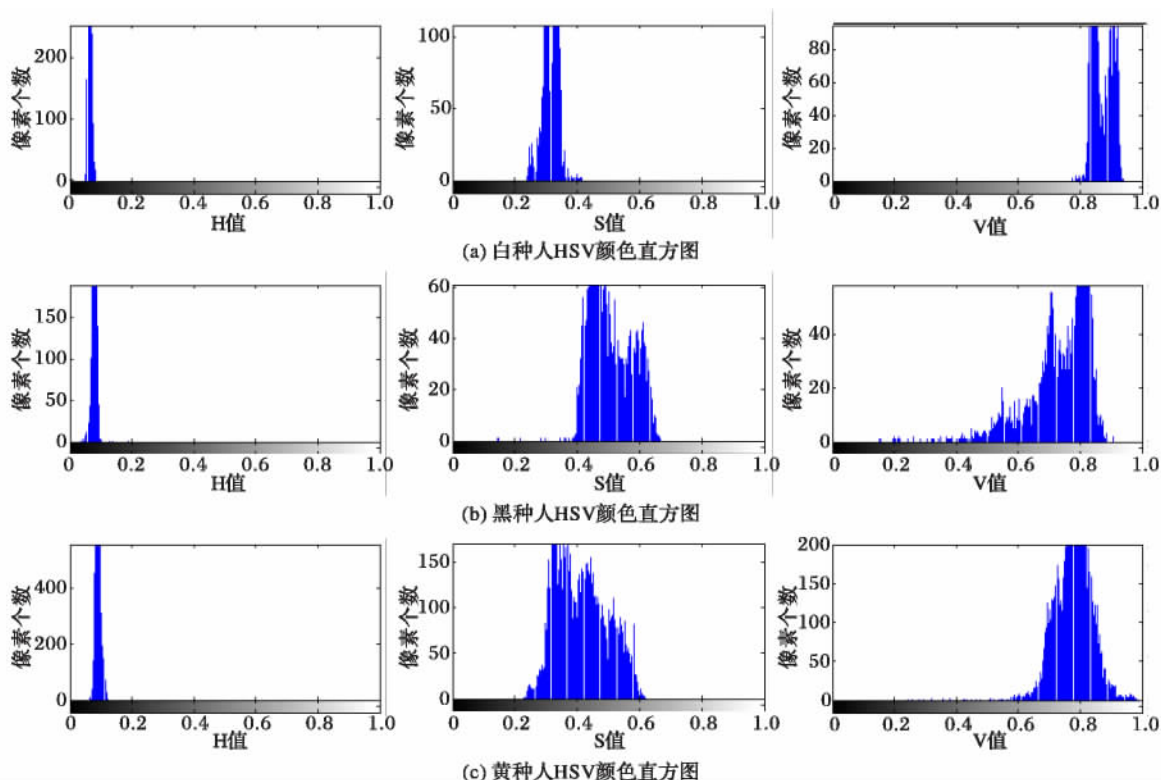


图4 不同人种 HSV 颜色直方图

经过阈值分割,将在阈值范围内的像素置为 1,其余像素置为 0,如式(1)所示。得到分割后的二值图像,如图 5 所示。

$$f(x, y) = \begin{cases} 1, & 0.03 < H < 0.128 \\ 0, & \text{其他} \end{cases} \quad (1)$$

对运动检测得到的二值图像与肤色分割得到的二值图像进行与操作,得到最终的手势图像,如图 6 所示。由图可以看出,通过这种方法,背景中误检测为肤色的区域和没有运动的人脸区域被去除。



图5 肤色分割图像



图6 手势图像

2 手势识别

2.1 特征提取

1962 年, Hu 首次提出图像识别的几何矩理论,并证明了所提出的 7 个矩对图像的平移、旋转和比例变化均保持不变。

本文即利用几何矩进行特征提取,并应用到手势识别中去,很好地解决了手势识别过程中手势的旋转、缩放、尺度变换所带来的问题,使得手势识别系统具有很好的稳定性。

对于二值图像 $f(x, y)$, 其 $p+q$ 阶几何矩按照式(2)定义:

$$\mu_{pq} = \sum_x \sum_y x^p y^q f(x, y) \quad (2)$$

一阶矩 (μ_{01}, μ_{10}) 用于确定图像质心 (\bar{x}, \bar{y}):

$$\begin{cases} \bar{x} = \mu_{10} / \mu_{00} \\ \bar{y} = \mu_{01} / \mu_{00} \end{cases} \quad (3)$$

若将坐标原点移至 \bar{x} 和 \bar{y} 处,就得到了对于图像位移不变的中心矩:

$$M_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y) \quad (4)$$

其中: (x, y) 表示图像像素的空间位置坐标, (\bar{x}, \bar{y}) 为图像质心。

根据 Hu 氏理论,通过各种不同级别的几何矩的数学组合,得到 7 个特征量。这 7 个特征量分别是:

$$\varphi_1 = \eta_{20} + \eta_{02} \quad (5)$$

$$\varphi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (6)$$

$$\varphi_3 = (\eta_{30} - 3\eta_{12})^2 + (\eta_{03} - 3\eta_{21})^2 \quad (7)$$

$$\varphi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{03} + \eta_{21})^2 \quad (8)$$

$$\begin{aligned} \varphi_5 = & (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) [(\eta_{30} + \eta_{12})^2 - \\ & 3(\eta_{03} + \eta_{21})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \\ & [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (9)$$

$$\begin{aligned} \varphi_6 = & (\eta_{20} - \eta_{02}) [(\eta_{30} + \eta_{12})^2 - \\ & (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{30} + \eta_{21}) \end{aligned} \quad (10)$$

$$\begin{aligned} \varphi_7 = & (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12}) [(\eta_{30} + \eta_{12})^2 - \\ & 3(\eta_{03} + \eta_{21})^2] - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03}) \\ & [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (11)$$

为了保持几何矩的尺度不变性,为归一化后的几何矩:

$$\eta_{pq} = M_{pq} / M_{00}^{1/2(p+q+2)} \quad (12)$$

利用上述公式计算图 7 中模板手势图像的几何矩,其结

果见表 1。为了减少计算量,此处提取了手势的轮廓图像进行计算,这 10 个手势(见图 7 中的图(a)~(j))分别表示数字 1~10。考虑到实际效果和计算复杂度,选取前四个矩作为特征向量,使用这些不变矩特征量组合形成特征空间 ($\varphi_1, \varphi_2, \varphi_3, \varphi_4$)。



图7 模板手势图像

表1 手势图像的几何不变矩

手势	$\varphi_1 (10^{-4})$	$\varphi_2 (10^{-10})$	$\varphi_3 (10^{-15})$	$\varphi_4 (10^{-16})$
1	7.042	346.239	14.423	123.097
2	7.918	1467.752	42.107	947.721
3	7.668	1000.596	133.246	1568.886
4	7.410	710.629	139.305	1485.391
5	6.980	26.926	7.789	475.615
6	6.914	120.213	180.786	9.785
7	7.008	314.813	39.221	2.517
8	6.833	43.443	4.698	72.292
9	6.912	54.441	2.029	214.444
10	6.823	8.769	12.831	14.763

2.2 模板匹配

式(13)定义了输入手势图像与 10 个模板手势图像中任一手势之间的距离 D_M :

$$D_M = \sum_{i=1}^7 \omega_i |\varphi_i - G_i| \quad (13)$$

其中: G_i 为输入手势的七个 Hu 矩特征分量, φ_i 为模板手势的 Hu 矩特征分量, ω_i 为各特征分量的权值,为了调节特征向量中各矩分量数量级的一致,在实际实验中取 ω_i 的值为 $10^4, 10^{10}, 10^{15}, 10^{16}$ 。 D_M 即为输入手势与模板手势的特征矩距离, D_M 越小表示越接近识别结果,即找到与输入手势特征距离最小的模板手势,那么这个手势即为识别结果。

3 实验与结果

具体实验中,请 10 个实验者在数字视频摄像头前打手势,每个实验者对 10 个手势每个手势累计共打 6 遍,这样采集到 600 个手势样本(每个手势对应 60 组样本),然后分别从每种手势的 60 个样本中,任取 10 个作为学习样本,按上述算法,计算各手势的几何矩,并取其平均值作为标准值,依此建立手势库;剩下的 50 套手势作为测试集。测试结果如表 2 所示。

表2 手势识别正确率

数字手势	1	2	3	4	5	6	7	8	9	10
识别率/%	98	100	94	98	100	98	92	98	96	100

4 结语

近些年手势识别成为许多学者的研究热点,对于一些需

要无声沟通的场合(如手术中),或对于一些没有掌握键盘或鼠标输入技能的人,手势识别提供了一种更加方便快捷的人机交互方式,具有广阔的应用前景。

本文是在 Windows 环境下,利用 OpenCV 提供的各种图像算法函数,实现手势识别。OpenCV 具有很高的移植性,可以很容易地将同样的代码在 Mac 和 Linux 系统下运行。本文提出的静态手势识别算法具有识别速度快、准确率高的特点,结合了运动检测的肤色分割算法在背景中出现接近肤色的区域时也可以将手势图像准确地分割出来,且 Hu 矩的不变性很好地解决了图像的旋转、平移和缩放问题。

参考文献:

- [1] 王西颖,戴国忠,张习文,等.基于 HMM-FNN 模型的复杂动态手势识别[J].软件学报,2008,29(9):2302-2312.
- [2] RAUTARAY S S, AGRAWAL A. Real time multiple hand gesture recognition system for human computer interaction[J]. International Journal of Intelligent Systems and Applications, 2012, 5(8): 56-64.
- [3] LEE C, XU Y. Online interactive learning of gestures for human/robot interfaces[C]// Proceedings of the 1996 IEEE International Conference on Robotics and Automation. Washington, DC: IEEE Computer Society Press, 1996: 2982-2987.
- [4] KADOUS M W. Machine recognition of Auslan signs using PowerGloves: Towards large-lexicon recognition of sign language[C]// Proceedings of the 1996 Workshop on the Integration of Gesture in Language and Speech. Wilmington, DE: IEEE Computer Society Press, 1996: 165-174.
- [5] ROOMI S M M, PRIYA R J, JAYALAKSHMI H. Hand gesture recognition for human-computer interaction[J]. Journal of Computer Science, 2010, 6(9): 994-999.
- [6] 陶霖密,彭振云,徐光祐.人体的肤色特征[J].软件学报,2001,12(7):1032-1041.
- [7] SURAL S, QIAN G, PRAMANIK S. Segmentation and histogram generation using the HSV color space for image retrieval[C]// Proceedings of the 2002 International Conference on Image Processing. Piscataway: IEEE, 2002, 2: 589-592.
- [8] BARCZAK A, DADGOSTAR F. Real-time hand tracking using a set of cooperative classifiers based on Haar-like features[R]. Palmerston North, New Zealand: Massey University, Institute of Information and Mathematical Sciences, 2005.
- [9] 郭兴伟,葛元,王林泉.基于形状特征的字母手势的分类及识别算法[J].人工智能及识别技术,2004,30(18):130-132.
- [10] CUI Y T, WENG J J. View-based hand segmentation and hand-sequence recognition with complex backgrounds[C]// Proceedings of the 13th International Conference on Pattern Recognition. Piscataway: IEEE, 1996, 3: 617-621.
- [11] LEE H-K, KIM J H. An HMM-based threshold model approach for gesture recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999, 21(10): 961-973.
- [12] LIU J Y, WANG Z, ZHONG L, et al. uWave: Accelerometer-based personalized gesture recognition and its application[C]// Proceedings of the Seventh Annual IEEE International Conference on Pervasive Computing and Communications. New York: IEEE Computer Society Press, 2009: 1-9.
- [13] BRADSKI G, KAEHLER A. 学习 OpenCV[M]. 于仕琪,刘瑞祯,译.北京:清华大学出版社,2010:299.
- [14] 田巍,庄镇泉.基于 HSV 色彩空间的自适应肤色检测[J].计算机工程与应用,2004,40(14):81-85.

(上接第 146 页)

值,例如 $\Gamma_s = 0.80$ 。系统按此 Γ_s 值运行一段时间无安全事故,再逐渐调低 Γ_s 值。选择适当的 Γ_s 值并依据反馈不断进行调整,最终可以在保障安全和减少认证数量之间找到一个动态的平衡点。在 $\Gamma_s = 0.80$, $\Delta = 0.85$ 时,由 Mamdani 推理方法可推算出 $\Phi = 0.95$ 。再由 $\Phi = 1 - \alpha^n$ (取 $\alpha = 0.9$),可得出 $n = 28.4$ 。所以,对于开放度 $\Delta = 0.85$ 的交易,信用度 $\Phi < 0.95$ 的用户,其安全度小于阈值 0.8,仍需进行认证;而信用度 $\Phi \geq 0.95$ (即已经成功交易过 28 次以上)的用户,其安全度超过了阈值 0.8,免于认证。

3 结语

电子政务和电子商务的每一项交易没有必要都先执行严格认证,在电子商务中的一种建议是随机验证小额交易^[2],这有一定的盲目性。本文在基于安全性量化评估的基础上,提出了一个信任模型:对于安全度高于安全度阈值 Γ_s 的交易不进行认证,在兼顾安全性的基础上减少了认证的数量。系统对安全度的评估开销远小于认证的开销。进一步的工作是将此信任模型运用于一个实际的系统中进行测试,并估算出节省的开销。

参考文献:

- [1] BRUSCHI D, CURTI A, ROSTI E. A quantitative study of public key infrastructures[J]. Computer and Security, 2003, 22(1): 15-21.
- [2] MANCHALA D W. E-commerce trust metrics and models[J]. IEEE Internet Computing, 2000, 4(2): 36-44.
- [3] POVEY D. ISO/IEC 15408, Developing electronic trust policies using a risk management model[S]. Geneva: ISO, 1999.
- [4] SSE-CMM Model Description Document Version 2.0 [EB/OL]. (1999) [2013-01-01]. <http://www.sse-cmm.org>.
- [5] ISO/IEC JTC 1/SC 27/WG 1. ISO/IEC TR 13335, Guidelines for the Management of IT Security (GMITS) [R]. Geneva: ISO, 1996.
- [6] BETH T, BORCHERDING M, KLEIN B. Valuation of trust in open network[C]// GOLLMANN D, ed. Proceedings of the 1994 European Symposium on Research in Security, LNCS 875. London: Springer-Verlag, 1994: 3-18.
- [7] JOSANG A, KNAPSKOG S J. A metric for trusted systems[C]// Global IT Security. Wien: Austrian Computer Society, 1998: 541-549.
- [8] ABDUL-RAHMAN A, HAILES S. A distributed trust model[C]// Proceedings of the 1997 Workshop on New Security Paradigms. New York: ACM Press, 1998: 48-60.
- [9] KOSKO B. Fuzzy engineering[M]. Upper Saddle River: Prentice Hall, 1997.