

Study of Hand Gesture Detection and Identification Algorithm Based on OpenVC

Zhousheng Ma

Henan Economy and Trade Vocational College, Zhengzhou Henan
Email: 1635085706@qq.com

Received: Sep. 22nd, 2017; accepted: Oct. 7th, 2017; published: Oct. 13th, 2017

Abstract

In view of the problem of lack of real-time recognition and low efficiency of dynamic gesture recognition, a dynamic gesture tracking and digital recognition algorithm based on OpenCV is proposed. The method uses the skin color feature to extract the gesture profile, and HMM is used to track gestures in the form of predictions and relocations. The identification of gestures is done by first positioning the fingertip profile. After the experiment, the framework can track the palm area stably, and the recognition accuracy is high, and there is no error and can achieve the accuracy of the target tracking based on the recognition rate.

Keywords

Skin Color Feature, Gesture Tracking, OpenCV, Gesture Recognition, HMM

基于OpenCV动态手势检测识别方法的研究

马州生

河南省经贸职业学院, 河南 郑州
Email: 1635085706@qq.com

收稿日期: 2017年9月22日; 录用日期: 2017年10月7日; 发布日期: 2017年10月13日

摘要

针对动态手势识别实时性不足和识别效率低的问题, 本文提出一种基于OpenCV的动态手势检测和数字识别方法。该方法采用肤色特征提取手势轮廓, 以预测和重定位的形式采用HMM来对手势进行跟踪, 手势的识别采用先对指尖轮廓进行定位的方式。经过实验, 本文方法能够稳定地跟踪手掌区域, 识别准确度高, 并能够在保证识别率的基础上实现对手指数目的准确识别跟踪。

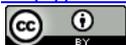
关键词

肤色特征, 手势跟踪, OpenCV, 手势识别, HMM

Copyright © 2017 by author and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

基于生物特征的身份认证方法和系统的研究一向都是图像识别技术[1]和模式识别领域研究的热点, 手势图像识别在人机交互中的应用成为一大热点, 出现了各种各样的识别方法。手势作为人机交互 (Human Computer Interaction, HCI) 接口的一种, 具有以人为中心, 简单、高效、人性化的特点, 是目前智能识别领域人机交互的重要手段之一。

针对手掌的提取主要是为了将手掌从庞大的背景中提取出来, 因为背景复杂, 采用灰度图像二值化的方法无法取得好的效果。采用颜色手套的解决方案能取得较好效果, 但会有其他问题。而在本框架中, 采取肤色分割提取手掌轮廓的方式, 通过测试, 在大多情况下可以取得理想效果。

对于手势的跟踪, 经典手势跟踪算法[2] [3]都存在不同的问题。均值漂移 (Meanshift) 算法不能很好地适应目标尺寸变化和目标的形变; 连续自适应均值漂移 (Camshift) 算法能很好地利用颜色信息进行跟踪, 但是容易受到相同或相近颜色的干扰; 卡尔曼滤波 (Kalman Filter) 算法的假设条件有可能不成立; 文献[4]借助 TLD 算法能有效地实现对手势稳定地跟踪, 且改进的 DTW 识别算法在保证识别率的基础上, 有效地提高了识别效率但为了保证部分手势的识别率, 增加了保留模板的数目, 但是这在一定程度上限制了其他手势识别效率的提升; 文献[5]用似然函数来降低粒子滤波算法中重要性权值的方差, 提高了抽样效率, 跟踪准确性也有所提高, 但算法复杂度较高; 文献[6]提出了 TLD (Tracking Learning Detection) 算法, 利用在线学习机制解决了长时间跟踪的稳定性问题, 但需要通过手动框选的方式来指定目标的初始化位置, 框选位置的不同对跟踪的结果有较大的影响, 会导致跟踪不稳定。

静态手势识别算法一种是采取手指指尖定位, 以手指指尖数目为静态特征, 此方法识别率高且失误率低; 还有一种是通过 SVM [7]对设定好的手势先分类, 获取特征向量。经测试比对, 该方案能辨识的手势图像多, 达到二十多种, 可是对测试环境要求苛刻。我们测试的静态手形简单, 因此采取以手指数目作为静态特征的方法有更好的表现。

本文框架对于动态手势识别, 无需太过复杂的识别效果, 只要实现对手掌质心的跟踪, 相邻帧图中的有用手掌轮廓质心位置做差可得到相邻帧图像的动态特征向量。

为了实现定位手势轮廓的高效准确的跟踪, 本文采取预测和重定位的形式, 采用基于卡拉曼滤波的粒子滤波算法来对手势进行跟踪。

本文提出的基于 OpenCV 的动态手势跟踪识别框架如图 1, 首先根据肤色特征对手势轮廓进行准确完整的提取和分离并对采集的图片进行形态学处理获得手势区域, 然后再使用粒子滤波算法对定位到的手势轮廓进行高效准确的跟踪和识别。本文后续会从 OpenCV 计算机视觉库介绍, 肤色分割模型的构建, 图像的预处理, 静态手势特征提取方法的构建, 对定位的手势轮廓进行高效准确的跟踪, 手势识别及实验结果分析等方面论述原理及效果。

2. 预处理与肤色分割

2.1. 图像预处理

在提取图像中肤色前，为了实验效果能更明显，必须先对图像进行必要的处理，去除噪声干扰，让颜色相近的区域更连贯。并且需要清晰的图像边缘，以达到提高识别率和保证算法稳定性的效果。

本文采用高斯滤波来对图像进行预处理，作用是对信号做平滑处理，数字图像后期的使用，其噪声干扰成为最大的问题，因为误差会累计，Gauss 滤波器可以用来获取高信噪比的图像(真实信号)。图像预处理过程如图 2 所示。

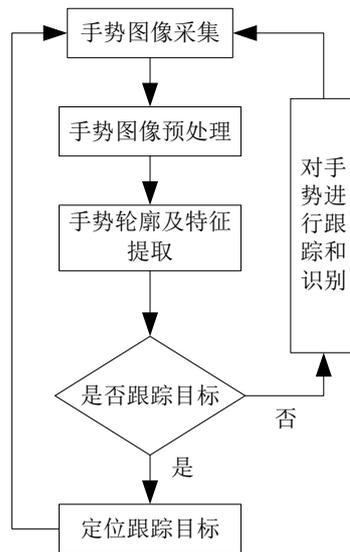


Figure 1. Framework of hand gesture identification
图 1. 基于 OpenCV 的动态手势跟踪识别框架

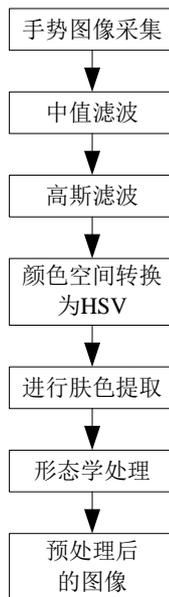


Figure 2. Image preprocessing
图 2. 图像预处理

为了得到良好的图像边缘, 经过 Gauss-Laplace 变换, 需要对图像做 Gauss 平滑滤波, 去除干扰噪声。

2.2. 肤色检测手部区域

利用肤色来检测手部是手部检测最直接的方法。皮肤颜色稳定, 不轻易受到缩放、平移和旋转的影响, 且对图像的尺寸、拍摄方向和角度的依赖性较小。

颜色空间主要有两种, 分别是 YCbCr 颜色空间和 HSV 颜色空间, 根据 Enamin D. Zarit 等人对肤色在这些彩色空间分布的研究, 以及在检测中性能的分析, 本文通过对手部皮肤颜色进行特征分析(选取黄种人的肤色), 选用 HSV 颜色空间进行手部区域检测。其转换公式为:

$$H = \begin{cases} H_1/360 & \text{if } B \leq G \\ (360 - H_1)/360 & \text{else} \end{cases} \quad (1)$$

其中,

$$H_1 = \cos^{-1} \left\{ \frac{0.5[(R-G) + (R-B)]}{\sqrt{(R-G)^2 + (R-B)(G-B)}} \right\} \quad (2)$$

$$S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)} \quad (3)$$

$$V = \frac{\max(R, G, B)}{225} \quad (4)$$

通过对大量肤色和非肤色在 H 和 S 空间分布的分析, 建立的肤色信息分割[8]公式如下:

$$F = \begin{cases} 0 & \text{if } Hd < H < Hu \parallel S > Su \\ 1 & \text{otherwise} \end{cases} \quad (5)$$

其中 F 为分割后得到的图像通过训练样本获得的在 H 和 S 空间肤色分布的参数, 再通过(5)式分割彩色图像获取灰度二值图像。

3. 手指跟踪与手势识别

3.1. 静态手势特征提取

常见的静态手势特征有质心、周长、手指分布等, 手的轮廓可以用不间断的点来表示, 在手掌轮廓不易提取时, 或者为了提高计算效率, 也可以通过计算手掌轮廓特征作为手势特征。

本文主要通过计算手指的个数, 来识别手势特征并进行跟踪[9]。首先要提取手掌轮廓, 计算轮廓形状特征有: 轮廓的质心、轮廓的最短最长径长、轮廓的外接圆(圆心和半径)、轮廓的周长和面积、轮廓在图像中的矩形框位置、轮廓的外包络点集合、轮廓的点集合、轮廓的各阶矩、轮廓的有效特征向量的提取、手指指尖的定位。手的位置特征是指手掌的质心位置, 针对二维图像, 质心位置是可以通过计算零阶距和 X 、 Y 的一阶距得到的。假设二值化之后的图像为 $I(X, Y)$, 质心 (x_c, y_c) 计算公式如下:

$$M_{00} = \sum_x \sum_y I(x, y) \quad (6)$$

$$M_{10} = \sum_x \sum_y xI(x, y) \quad (7)$$

$$M_{01} = \sum_x \sum_y yI(x, y) \quad (8)$$

$$x_c = \frac{M_{10}}{M_{00}}, \quad y_c = \frac{M_{01}}{M_{00}} \quad (9)$$

手势轮廓特征提取[10]的关键在于手指特征的提取，具体是利用手指指尖距手掌中心最远的这个规律，把手的轮廓点的坐标转换成以手掌中心位置为原点的极坐标系，然后分析轮廓曲线中的局部最大值，如图3所示，通过映射规则把点的位置映射为手指。

本系统中测试人员将手置于交互空间，手掌心正对摄像头[11]，指尖向上。本框架设定如图4的6种不同手形。测试人员可以移动不同的位置做出这些手形来人机交互。

3.2. 动态手势识别

动态手势具有时间和空间上的变化，在不需要实现识别过于复杂手势的前提下，只需获得当前手势轮廓的质心，通过对比相邻帧间轮廓的质心位置变化便可以获得动态特征向量。把得到的动态特征向量[12]转换为角度进行编码，用HMM进行动态手势识别。

隐马尔可夫模型是从马尔可夫(Markov)链的基础上发展来的，马尔可夫链是随机过程，HMM是双重随机过程，马尔可夫链中观察值与状态一一对应，HMM中状态转移与观察值不对应，不能直接观察状态转移，所以才称为隐马尔可夫模型[13]。隐马尔可夫模型是一个三元组名 $= (A, B, c)$:

1) 状态集合 $S = \{S_1, S_2, \dots, S_N\}$ ，其中 N 为状态数，表示 t 时刻的状态。状态之间的转移是根据转移概率矩阵 A 决定的。

2) 观察符号的集合 $O = \{V_1, V_2, \dots, V_M\}$ ，其中 M 是集合中的符号数量。每个状态输出 O 是根据观察概率矩阵 B 决定的。

3) 状态转移概率分布矩阵 $A = \{a_{ij}\}$ ，其中 $a_{ij} = P(q_{t+1} = S_j | q_t = S_i)$ ，表示从 S_i 转移到 S_j 的概率。

4) 观察概率矩阵 $B = \{b_i(j)\}$ ， $b_i(j) = P(O_t = V_j | q_t = S_i)$ ，表示状态 S_i 观察到 V_j 的概率。

5) 初始状态概率分布 $c = \{c_i\}$ ， $c_i = p(q_1 = S_i)$ ， c_i 表示初始状态为 S_i 的概率。

HMM 从一个初始状态开始，每个状态根据观察概率矩阵，输出一个观察符号，然后根据状态转移矩阵，转移到一个新的状态。一个具有 N 个状态数的隐马尔可夫模型在一共有 N^T 种不同的长度为 T 状态转移序列，具有很大的状态空间。隐马尔可夫模型具有很强的学习能力，对序列的长短变化适应性比较好，适用于对动态时间序列数据的识别且识别稳定性有显著提高。识别过程如图5所示。

4. 实验结果与分析

本文实验是在 VS2013 集成开发环境下，基于 OpenCV 视觉库用 C++ 开发，测试视频来源为普通数字视频摄像头拍摄，使用的计算机为普通桌面 Pc 机，采用 Intel Pentium 双核 2.7GHz CPU，2G RAM，微软 Windows10 32 位操作系统，摄像头采用普通数字视频摄像头，为验证本框架的稳定性和高效性，我们采用 6 种手势，如图4分别代表 6 种不同的手指状态。

实验分别进行静态识别和动态识别。在静态识别中，选取 8 名实验人员，分别用右手采集 6 段图像，长度为 500 帧，每段图像只使用一个手形，共需要 3000 帧图像数据，之后逐帧将采集的图像输入，来检验本文方法的正确性。在动态识别中，让系统在实验室通常环境下即时测试，实验对象在摄像头前做出既定手形，可以实现对既定的 6 种手势状态实时稳定的跟踪识别，能够使本文系统有更好的识别效果[14]，测试了系统的稳定性和鲁棒性。测试过程实时手指指尖确定如图6所示。

通过实验可以看出在识别准确性上，本框架与基于 TLD 和 DTW 的动态手势跟踪识别算法相比识别精度相近，都达到了 95.8% 以上，具有较高的识别率。测试结果如表1所示。并且与 Zhou 等人的实验相

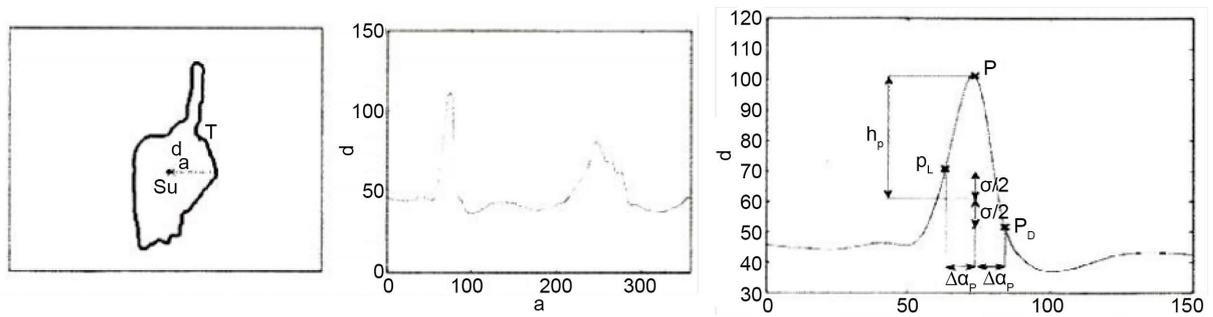


Figure 3. Local minima of finger location
图 3. 手指定位局部极值法



Figure 4. Six gesture digits extraction
图 4. 检测提取 6 种手势表示的数字

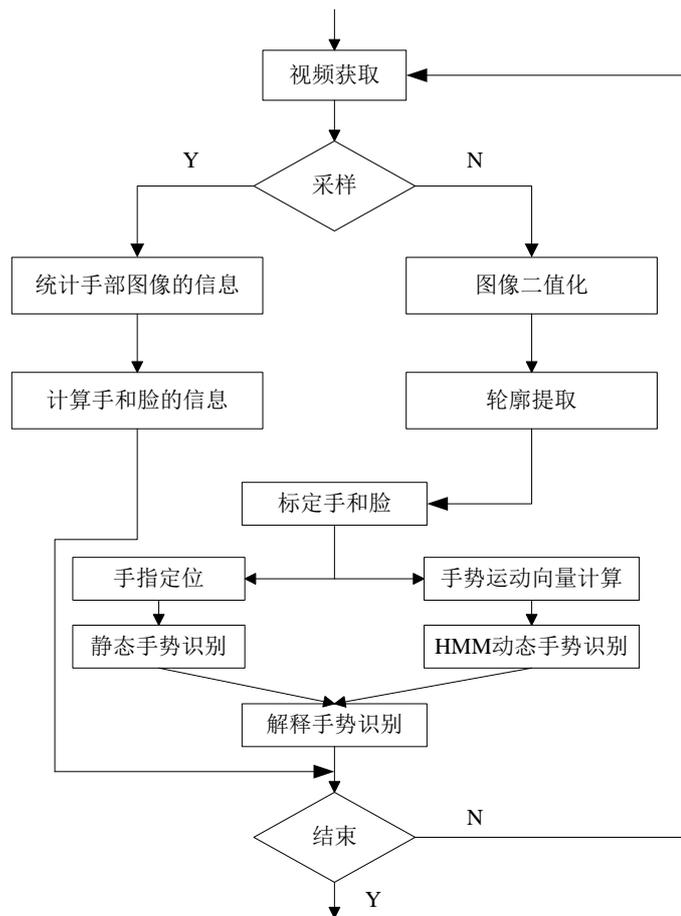


Figure 5. Global procedure of gesture recognition
图 5. 手势识别整体流程



Figure 6. Gesture recognition and mark of positioning the fingertips

图 6. 实时手指指尖标记图

Table 1. Result of gesture recognition testing

表 1. 手势识别测试结果

手势数字状态	0	1	2	3	4	5	6
识别率/%	96.2	98.3	95	97.5	96	98	100

比, 本文框架的处理速度分别是其两种方法的 5 倍和 45 倍, 实验人员按上述手形进行即时实验, 每帧图像处理时间约为 82.71 ms, 速率范围 11~14 帧/s, 体现了很好的实时性且系统稳定性良好, 设备获取便捷, 测试对环境的要求低, 场景布置简单, 适用于工作间、实验室等场景的人机交互。

5. 结论

本文主要针对手势跟踪识别实时性和识别效率的问题, 提出了一种基于 OpenCV 的动态手势跟踪识别框架。实验结果表明, 本文方法根据肤色特征分割能有效的实现手势轮廓的检测, 且通过手指指尖定位与 HMM 相结合在保证高识别率的基础上, 系统的实时性与稳定性显著提高。文中在动态识别阶段, 为增强对既定手势的快速跟踪识别[15], 保证实时性, 没有设定其他更复杂的手势, 在一定程度上限制了其他种类手势的识别效率, 如何快速自适应的实时识别跟踪多种复杂手势, 是本文下一步要研究的方向。

基金项目

河南省自然科学基金(2015A071008)资助项目。

参考文献 (References)

- [1] 王淼, 杜毅, 张忠瑞. 无人机辅助巡视及绝缘子缺陷图像识别研究[J]. 电子测量与仪器学报, 2015, 29(12): 2-3.
- [2] 彭艳芳. 视频运动目标检测与跟踪算法研究[D]: [硕士学位论文]. 武汉: 武汉理工大学, 2010.
- [3] 高文, 朱明, 贺柏根, 吴笑天. 目标跟踪技术综述[J]. 中国光学, 2014, 7(3): 365-375.
- [4] 余超, 关胜晓. 基于 TLD 和 DTW 的动态手势跟踪识别[J]. 计算机系统应用, 2015, 24(10): 2-4.
- [5] 何广. 基于粒子滤波的目标跟踪算法研究[D]: [硕士学位论文]. 哈尔滨: 哈尔滨工业大学, 2006.
- [6] Kalal, Z., Mikolajczyk, K. and Matas, J. (2012) Tracking-Learning-Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **34**, 1409-1422.
- [7] 张佳佳, 王贺, 刘宇, 蒋涛, 武春爱. SVD 与 SVM 叠加算法在玻璃温度表图像识别中的应用[J]. 电子测量技术, 2015, 38(1): 1-3.
- [8] 肖秦琨, 赵艳, 高嵩. 基于 RGB 彩色和深度信息的人体关节点定位[J]. 国外电子测量技术, 2015, 34(2): 4-5.
- [9] 刘鑫辰, 傅慧源, 马华东. 基于 RGB-D 摄像头的实时手指跟踪与手势识别[J]. 计算机科学, 2014, 41(10): 1-5.
- [10] 路凯, 李小坚, 周金祥. 基于肤色和边缘轮廓检测的手势识别[J]. 北方工业大学学报, 2006, 18(3): 2-3.
- [11] 江超, 艾娇燕. 基于 OpenCV 的摄像头动态手势轨迹识别及其应用[J]. 计算机应用, 2012, 32(S1): 1-2.
- [12] 王志龙. 基于运动跟踪的动态手势识别算法研究[D]: [硕士学位论文]. 武汉: 华中科技大学, 2013.

-
- [13] 王西颖, 戴国忠, 张习文, 张凤军. 基于 HMM-FNN 的复杂动态手势识别[J]. 软件学报, 2008, 19(9): 6-11.
- [14] 余永维, 殷国富, 殷鹰, 杜柳青. 基于深度学习网络的射线图像缺陷识别方法[J]. 仪器仪表学报, 2014, 35(9): 2-7.
- [15] 朱建伟. 基于多摄像头的手势识别研究与实现[D]: [硕士学位论文]. 广州: 中山大学, 2010.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2329-1273, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: jisp@hanspub.org