

基于几何特征的桌面静态手势识别

张 辉^{1,2}, 邓继周¹, 周经纬¹, 朱成顺^{1,2}, 刘李明¹, 张胜文^{1,2}

(1. 江苏科技大学 机械工程学院, 江苏 镇江 212003;

2. 江苏科技大学 江苏省船海机械装备先进制造重点实验室, 江苏 镇江 212003)

摘 要: 为满足桌面环境下的手势交互需求, 研究一种将手指数目、指间距及指间夹角特征相结合的桌面静态手势识别方法。以提高手势识别率、降低识别难度为目标, 根据指间距及指间角度特征进行手势建模。利用分层策略将手势识别分为两层, 将多种类手势识别转化到每层下较少种类的识别, 根据每层的识别结果作为判别依据共同完成手势识别任务。实验结果表明, 在常规桌面背景下, 该方法比普通的模板匹配方法具有更高的识别准确率。

关键词: 人机交互; 手势识别; 静态手势; 几何特征; 手势建模; 分层策略

中图分类号: TP391.4 **文献标识码:** A **文章编号:** 1000-7024 (2020) 10-2977-05

doi: 10.16208/j.issn1000-7024.2020.10.045

Static gesture recognition of desktop based on geometric features

ZHANG Hui^{1,2}, DENG Ji-zhou¹, ZHOU Jing-wei¹, ZHU Cheng-shun^{1,2},

LIU Li-ming¹, ZHANG Sheng-wen^{1,2}

(1. School of Mechanical Engineering, Jiangsu University of Science and Technology, Zhenjiang 212003, China;

2. Jiangsu Provincial Key Laboratory of Advanced Manufacturing for Marine Mechanical Equipment,
Jiangsu University of Science and Technology, Zhenjiang 212003, China)

Abstract: To satisfy the needs of gesture interaction in desktop environment, a desktop static gesture recognition method combining the number of fingers, the distance and the angle between fingers was studied. A gesture model was established according to the finger spacing and angle characteristics to improve the recognition rate of gestures, and to reduce the difficulty of recognition. The process of gesture recognition was divided into two layers by using hierarchical strategy. Multiple types of gesture recognition were converted to less kinds of recognition in each layer, and the gesture recognition task was completed according to the recognition result of each layer. Experimental results show that the proposed method can achieve higher recognition accuracy than the template matching method under the conventional desktop background.

Key words: human-computer interaction; gesture recognition; static gesture; geometric characteristics; gesture modeling; hierarchical strategy

0 引 言

近年来, 基于视觉的手势识别由于具有操作简单、自然的优势而成为人机交互领域研究热点^[1]。相比于基于模板匹配^[2]、基于机器学习^[3,4]等手势识别方法, 基于几何特征的识别方法^[5]在简单背景的桌面应用条件下具有计算简单、识别速度快的优点^[6]。该方法主要利用手势中的手指

数目、角度、相对位置等特征进行手势的分类与判别。但是由于个体差异等原因, 上述几何特征即使在相同的手势中依然存在较大范围的变化, 很难根据单一的几何特征实现手势的分类。为了提高该方法的识别率, 许多学者从多角度考虑手势几何特征, 将不同的几何特征进行组合来识别手势。文献 [7, 8] 采用提取手势指尖个数结合 Hu 特征的方法, 构建加权融合的特征距离, 将待识别手势与模

收稿日期: 2019-01-28; **修订日期:** 2020-07-31

作者简介: 张辉 (1981-), 男, 江苏镇江人, 博士, 副教授, CCF 会员, 研究方向为 CAD/CAM、增强现实、视觉测量检测; 邓继周 (1994-), 男, 江苏南通人, 硕士研究生, 研究方向为视觉测量检测; 周经纬 (1993-), 男, 江苏宿迁人, 硕士研究生, 研究方向为计算机视觉与图像处理; 朱成顺 (1975-), 男, 江苏扬州人, 硕士, 副教授, 研究方向为 CAD/CAM; 刘李明 (1980-), 男, 安徽安庆人, 硕士, 讲师, 研究方向为工业设计、室内设计、虚拟现实; 张胜文 (1963-), 男, 安徽阜阳人, 博士, 教授, 研究方向为智能制造。
E-mail: zimmerman@126.com

板进行相似度匹配来识别手势。使用多种特征结合的手势识别方法虽然可以获得较高的手势识别率,但会带来特征数据变多,计算复杂的缺点。对此本文从多方面考虑手势几何特征,研究了一种分层的手势识别算法,利用手势几何特征进行手势建模并分类,避免特征数据融合带来计算复杂的问题,提高了手势识别率。

1 算法流程

本文基于几何特征的手势识别方法主要根据手指数、指间距以及指间夹角,具体的手势识别算法流程如图 1 所示。首先利用肤色信息将手势区域与其它干扰区域进行分割,获得手势图像。其次对手势图像进行形态学处理^[9],统计出手指数。根据手指数的不同,将手势识别拆分为两层。层 1 对于某些特定的手势,其仅依靠手指数目即可完全判别,此时直接输出识别结果。层 2 是在层 1 的基础上,提取指间距以及指间夹角作为特征参数,与标准模板进行匹配,实现对手势的进一步识别。

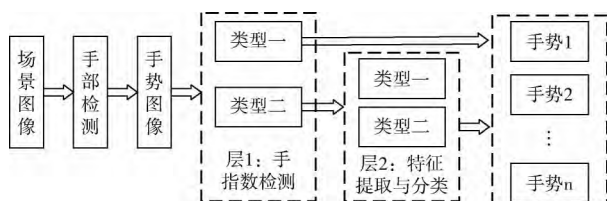


图 1 手势识别算法流程

2 手势区域检测

作为手势识别的关键步骤,手势区域检测就是将场景图像进行分割,获得手势图像。与其它干扰目标相比,人手最明显的特征就是肤色,因此可以利用肤色信息进行手势分割。本文在相对简单的桌面应用背景下获取场景图像,选用 HSV (hue-saturation-value) 色彩空间将场景图像亮度信息分离,利用人体肤色信息,对场景进行分割,排除非肤色物体的干扰,实现手势区域的检测。在 HSV 色彩空间中,用色度 H 、饱和度 S 和亮度 V 对色彩进行量化表示,符合人眼的直观特性^[10]。将图像从 RGB 彩色空间转换到 HSV 彩色空间采用如下公式

$$H = \begin{cases} \frac{(G-B)}{\max(R,G,B)-\min(R,G,B)} \times \frac{\pi}{3}, & \text{当 } R = \max(R,G,B) \\ \frac{(B-R)}{\max(R,G,B)-\min(R,G,B)} \times \frac{\pi}{3}, & \text{当 } G = \max(R,G,B) \\ \frac{(R-G)}{\max(R,G,B)-\min(R,G,B)} \times \frac{\pi}{3}, & \text{当 } B = \max(R,G,B) \end{cases}$$

$$V = \max(R,G,B)$$

$$S = 1 - \min(R,G,B)/V \quad (1)$$

根据文献 [11], 在 HSV 颜色空间中检测肤色时, 人

类的肤色区域像素应满足

$$2 \leq H \leq 28, 50 \leq S \leq 200 \quad (2)$$

利用上述肤色分割获得的场景二值图像往往会存在噪声、孔洞及其它无关区域,对此需要去除无关区域的干扰。由于在肤色分割过后手部区域成为场景二值图中的主要部分,通过边缘检测算法对手势图像进行轮廓检测,将检测到的最大轮廓作为手势图像绘制输出,其过程如图 2 所示。

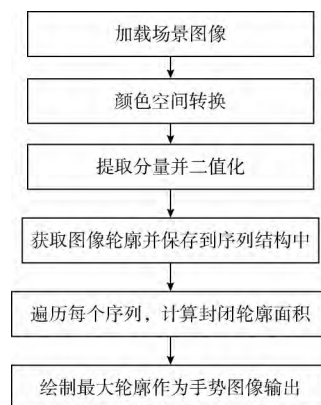


图 2 手势图像检测流程

3 手势识别

手势识别方法在上述获得的手势图像中进行,包括手势几何特征的提取、建模与分类,进而实现手势的识别。其中本文几何特征主要包括手势图像中手指数目以及利用指间距和指间夹角参数,同时利用分层识别的思想实现手势结果最终判定。其中,对于可以根据手指数目进行唯一判定的手势图像可以直接输出结果。对于相同手指数目的不同手势图像来说,可根据指间距以及指间夹角的不同进行二次判定。

3.1 手指数检测

本文基于手势几何特征,实现手势图像的识别,因此对手指数的检测是实现分层策略的前提^[12]。手指数的检测方法包含以下步骤:

(1) 对上述步骤获得的手势图像进行开运算操作,获得手掌区域图像。其中结构元素为半径为 15 像素的圆;

(2) 将手势图像与步骤 (1) 的手掌图像进行差运算,得到手指区域图像;

(3) 对于步骤 (2) 中得到的手指区域图像进行腐蚀操作,去除小连通区域以及其它噪声干扰,此时统计出连通区域的个数即为手指数。

3.2 手势建模与识别

通过手势区域检测以及手指数的统计,对于仅依靠手指数作为唯一特征的手势图像可直接输出识别结果。为了进一步扩大手势识别的种类,本文基于手势几何特征采用分层策略实现相同手指数目不同手势图像的识别。常见的

手势几何特征包括手指数、指间距、指间夹角以及手形的长宽比等, 由于个体的差异, 采用单一的几何特征对手势图像进行手势识别中存在局限性。例如, 对于同一操作者, 不同的手势其指间距存在较大的区别, 可作为判断依据。但是对于不同的操作者, 手掌大小异同, 此时不同的手势之间提取到的手势存在较大误差, 造成手势识别准确率的下降。因此, 在手势识别过程中, 往往需要融合多个手势几何特征, 实现手势的判断。针对上述情况, 本文在完成手指数统计的基础上, 提出以指间距及指间夹角作为手势参数, 对手势图像进行进一步判别。一般情况下, 在手势图像中手指呈长条状, 为了提高手势识别的准确性以及实时性, 通过构造手指的最小外接矩形近似代替手指。其中以各个矩形中心作为手指的质心, 通过计算对应矩形质心之间的距离, 选取最大的质心距离作为该手势的指间距。如图3所示, d_1 、 d_2 、 d_3 为任意两个矩形之间的距离, 以外接矩形最外侧两个边所在的直线构建向量 $\overrightarrow{OP_1}$ 、 $\overrightarrow{OP_2}$, 计算其夹角 θ , 作为指间夹角, 用于进一步加强对手势几何特征的描述, 通过设定的模板参数, 实现手势的最终识别。具体的手势模型构建方法如下:

(1) 根据统计手势图像中连通域的个数作为手指数 m , 当 $m \in \{2, 3, 4\}$ 时, 此时单凭手指数无法直接输出识别结果, 仍需提取手势指间距以及指间夹角作为判定依据;

(2) 构建手指区域的最小矩形包围盒, 其中包围盒中心为该手指的质心位置, 计算任意两个包围盒的质心距离 d , 以最大的质心距离作为该手势图像的指间距;

(3) 以手势图像中最外侧的两个矩形包围盒最外侧的两个边所在的直线构建向量, 计算两个向量之间的夹角, 作为手势图像的指间夹角 θ ;

(4) 定义手势模型计算参数 $K_{i-j} = d_{i-j}^{\max} \times \theta$, 其中 $i-j$ 为手势图像的编号, 计算手势图像中手势模型计算参数, 并与标准的手势模型参数进行比较, 从而输出最终的识别结果。

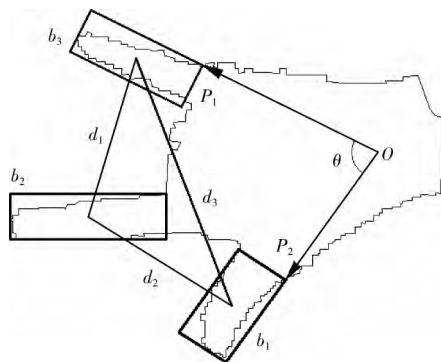


图3 特征提取

所示的9种静态手势作为识别对象。为了便于区分, 将图4中的手势进行如下编号: 图(a)、图(b)、图(i)这3种类型手势分别用0、1、5表示, 图(c)~图(h)这6种类型手势分别用2-1、2-2、3-1、3-2、4-1、4-2表示。编号首位代表手指数, 而第二位代表相应手指数下的手势类型号。实验中, 10位实验对象在相同环境下做出9种手势图像, 其中每种手势均以不同角度呈现于桌面背景中, 每类手势图像采集5次, 总共450幅图像。

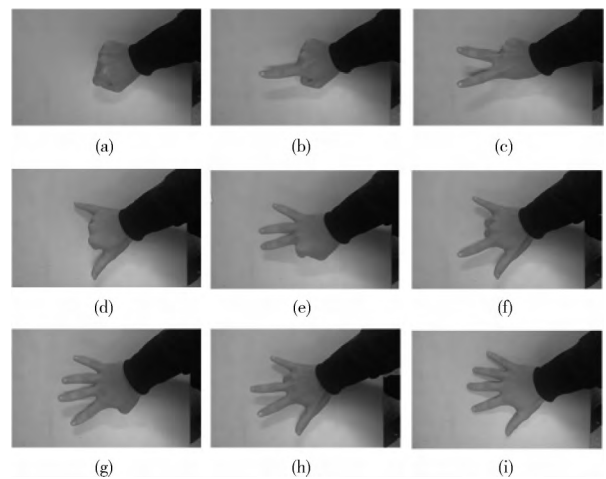


图4 实验手势

本文借助 OpenCV (open source computer vision library) 视觉库进行相关的图像处理操作。OpenCV 封装了多种通用算法用于实现计算机图像处理, 成为视觉领域中最有力的研究工具之一。

在实验中, 根据分层策略实现手势图像的识别。首先统计出手势图像中的手指数目, 根据手指数目对需要进一步判断的图像使用上述提出的手势模型参数, 进行最终判定。根据上节提出的手势建模方法, 对需要进一步判别的手势图像进行特征提取, 构建手指包围盒, 计算指间距以及指间夹角, 得到手势特征值。本文以图4所示的手势图作为标准手势图像, 采用相同方法计算标准手势特征值 K_{i-j} , 同时考虑到个体的差异对于手势特征值提取带来的误差, 对实际获得的手势特征 K'_{i-j} 按照表1中设置的阈值对手势进行判定。

表1 手势特征范围划分

手势编号	特征值范围
$i-1$	$K'_{i-1} \leq \frac{1}{2}(K_{i-1} + K_{i-2})$
$i-2$	$K'_{i-2} \geq \frac{1}{2}(K_{i-1} + K_{i-2})$

4 实验与分析

为了验证本文提出识别方法的有效性, 本文采用如图4所示的9种静态手势作为识别对象。图5给出了实验中每种手势下的特征数据分布情况, 将特征值进行了归一化处理, 图中纵坐标 $I'_{i-j} =$

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

$\frac{2K_{i-1}^n}{K_{i-1} + K_{i-2}}$ ，为归一化的特征值，此时的手势分割阈值为 1，横坐标 n 为拍摄次数，其中图 (a)、图 (b)、图 (c) 分别对应首编号为 2、3、4 下的两种手势特征值。

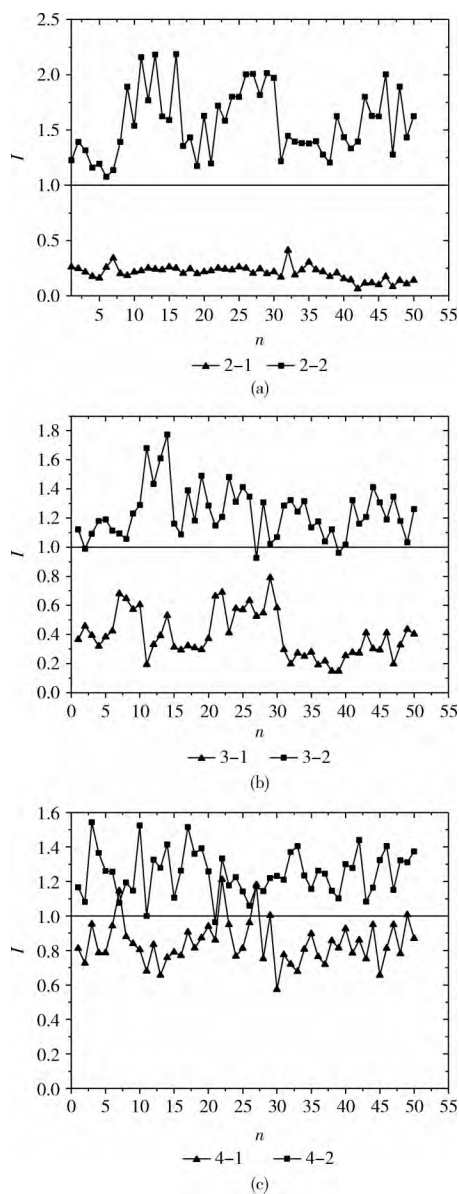


图 5 手势特征值

首先按照本文的手指检测算法可以成功获得所有手指个数，对编号为 0、1、5 的手势识别率可以达到 100%。然后对编号为 2、3、4 手势进行进一步识别。如图 5 所示，在第二层中，利用本文提出的分层手势识别方法将手指数目为 2、3、4 的 6 种手势分为 3 个小类，在每个小类下只有两种手势的存在，成功将多目标分类利用分层策略变成两项分类，降低了基于几何特征的手势识别难度。

为了对比实验结果，使用传统的基于 Hu 矩特征的模板匹配法^[13]与本文方法进行比较。同样，模板匹配法也使

用上述的手势图像组，并且取本文采集标准手势特征所用的手势图像作为标准模板。计算手势轮廓 Hu 矩值，取其中抗噪声干扰较好的 4 个值作为手势特征值，采用欧式距离作为模板匹配时的相似度评价，以最小值对应的模板手势作为识别结果，将此模板匹配法作为方法 I。此外利用本文的手指数检测方法先进行手势的初步分类，对于需进一步分类的手势使用模板匹配法进行识别，将此方法作为方法 II，进行比较。表 2 给出了 3 种手势识别方法对于 9 种手势的最终识别结果。

表 2 手势识别率

序号	方法 I	方法 II	本文方法
0	50	100	100
1	44	100	100
2-1	92	96	100
2-2	36	58	100
3-1	56	90	100
3-2	48	82	92
4-1	44	72	90
4-2	58	74	96
5	72	100	100

从表 2 可以看出，在单独使用 Hu 特征进行模板匹配的方法 I 中，识别率随着手势的不同有着很大差距，总体处于一个相对较低水平。方法 II 利用分层的识别方法，在手指数识别的基础上进行模板匹配，只需与该数目下的两个模板进行相似度匹配，使得识别率相比较方法 I 得到了明显的提高。然而在同样进行分层识别的情况下，模板匹配法的识别率不如本文构建的手势模型识别率高，虽然随着手指数目的增多，每个小类下的两种手势的特征值差距会有所降低，但使用本文方法仍能够保持识别率在 90% 以上。

5 结束语

本文对基于几何特征的桌面静态手势识别方法进行了研究。在简单的桌面背景下，针对单一几何特征对手势识别准确率带来不利的影响，提出了分层的思想，将手势识别分成两个步骤。首先，利用肤色模型对手势图像进行分割，其次根据设置的阈值统计封闭轮廓面积作为手指数。对于某些特定的手势仅根据手指的数目多少就可以做出唯一判据，直接输出识别结果。否则，在此基础上，构建手势模型，计算出手势指间距以及指间夹角，并进行模板匹配，从而得到最终的识别结果。本文设计的识别算法能够识别预定义的 9 种自然手势类型，简单有效，能够满足桌面静态手势识别的需求。

参考文献:

- [1] Hasan H, Abdul-Kareem S. Human-computer interaction using vision-based hand gesture recognition systems: A survey [J]. Neural Computing and Applications, 2014, 25 (2): 251-261.
- [2] Rodrigo I, Álvaro S, Teyseyre A, et al. Approximate string matching: A lightweight approach to recognize gestures with Kinect [J]. Pattern Recognition, 2017, 62 (2): 73-86.
- [3] Wahid M F, Tafreshi R, Al-Sowaidi M, et al. Subject-independent hand gesture recognition using normalization and machine learning algorithms [J]. Journal of Computational Science, 2018, 27 (4): 69-76.
- [4] LIU Xiaojian, ZHANG Yuan. Gesture recognition based on multi-feature and SVM classification [J]. Computer Engineering and Design, 2017, 38 (4): 953-958 (in Chinese). [刘小建, 张元. 基于多特征提取和 SVM 分类的手势识别 [J]. 计算机工程与设计, 2017, 38 (4): 953-958.]
- [5] LIN Shuiqiang, WU Yadong, CHEN Yonghui. Gesture recognition method based on geometric characteristics [J]. Computer Engineering and Design, 2014, 35 (2): 636-640 (in Chinese). [林水强, 吴亚东, 陈永辉. 基于几何特征的手势识别方法 [J]. 计算机工程与设计, 2014, 35 (2): 636-640.]
- [6] QI Jing, XU Kun, DING Xilun. Vision-based hand gesture recognition for human-robot interaction: A review [J]. ROBOT, 2017, 39 (4): 565-584 (in Chinese). [齐静, 徐坤, 丁希仑. 机器人视觉手势交互技术研究进展 [J]. 机器人, 2017, 39 (4): 565-584.]
- [7] WANG Yan, XU Shiyi, CHEN Haiyun. Gesture recognition based on weighted feature distance [J]. Computer Science, 2017, 44 (S1): 220-223 (in Chinese). [王艳, 徐诗艺, 湛海云. 基于特征距离加权的手势识别 [J]. 计算机科学, 2017, 44 (S1): 220-223.]
- [8] CAO Jie, ZHAO Xiulong, WANG Jinhua. Gesture recognition method based on improved finger tip and Hu moments [J]. Computer Engineering and Applications, 2017, 53 (21): 138-143 (in Chinese). [曹洁, 赵修龙, 王进花. 融合改进指尖点和 Hu 矩的手势识别 [J]. 计算机工程与应用, 2017, 53 (21): 138-143.]
- [9] LIU Shuping, LIU Yu, YU Jun, et al. Hierarchical static hand gesture recognition by combining finger detection and HOG features [J]. Journal of Image and Graphics, 2015, 20 (6): 781-788 (in Chinese). [刘淑萍, 刘羽, 於俊, 等. 结合手指检测和 HOG 特征的分层静态手势识别 [J]. 中国图象图形学报, 2015, 20 (6): 781-788.]
- [10] Wang M, Chen W Y, Li X D. Hand gesture recognition using valley circle feature and Hu's moments technique for robot movement control [J]. Measurement, 2016, 94 (12): 734-744.
- [11] ZHU Yue, LI Zhenwei, YANG Xiaoli, et al. Static gesture recognition system based on vision [J]. Computer Technology and Development, 2019, 29 (2): 69-72 (in Chinese). [朱越, 李振伟, 杨晓利, 等. 基于视觉的静态手势识别系统 [J]. 计算机技术与展, 2019, 29 (2): 69-72.]
- [12] ZHOU Jingwei. Development of auxiliary assembly platform based on information projection technology [D]. Zhenjiang: Jiangsu University of Science and Technology, 2019: 1-69 (in Chinese). [周经纬. 基于信息投影技术的辅助装配平台开发 [D]. 镇江: 江苏科技大学, 2019: 1-69.]
- [13] YI Jingguo, CHENG Jianghua, KU Xishu. Review of gesture recognition based on vision [J]. Computer Science, 2016, 43 (S1): 103-108 (in Chinese). [易靖国, 程江华, 库锡树. 视觉手势识别综述 [J]. 计算机科学, 2016, 43 (S1): 103-108.]