



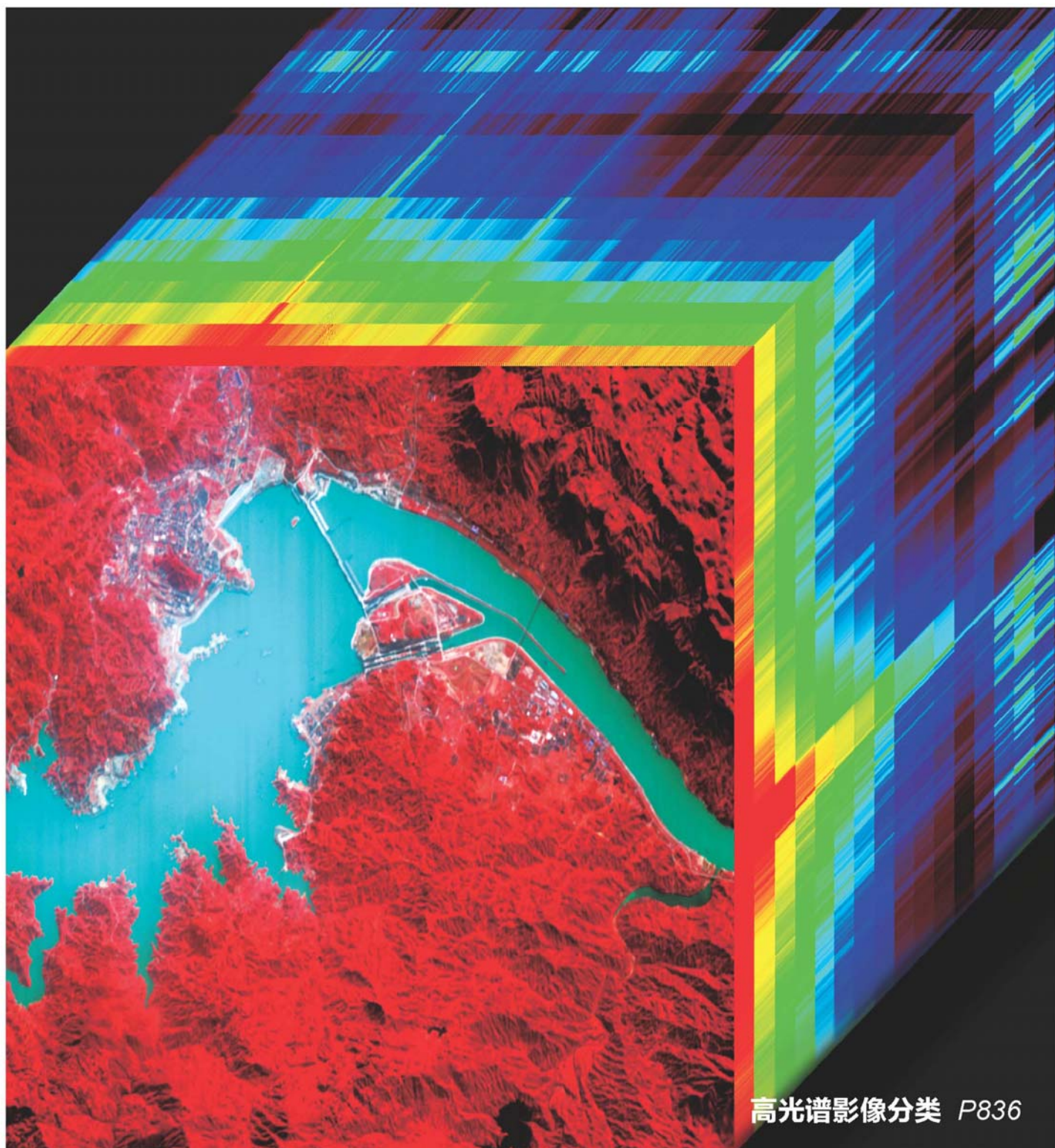
JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

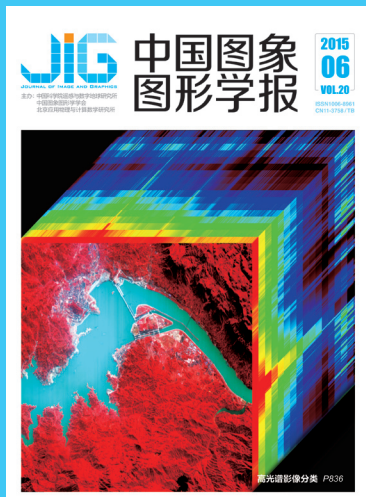
中国图象图形学报

2015
06
VOL.20

ISSN1006-8961
CN11-3758/TB



高光谱影像分类 P836



第20卷第6期（总第230期）

2015年6月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院

主办单位 中国科学院遥感与数字地球研究所

中国图象图形学学会

北京应用物理与计算数学研究所

主 编 李小文

编辑出版《中国图象图形学报》编辑出版委员会

邮政信箱 北京9718信箱

邮 编 100101

电子信箱 jig@radi.ac.cn

电 话 010-64807995

网 址 www.cjig.cn

广告经营许可证 京朝工商广字第0361号

总 发 行 北京报刊发行局

订 购 全国各地邮局

海外发行 中国国际图书贸易集团有限公司

（邮政信箱：北京399信箱 邮编：100048）

印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian

Monthly, Started in 1996

Superintended by Chinese Academy of Sciences

Sponsored by Institute of Remote Sensing and Digital Earth, CAS

China Society of Image and Graphics

Institute of Applied Physics and Computational Mathematics

Editor-in-Chief LI Xiaowen

Editor, Publisher Editorial and Publishing Board of Journal of

Image and Graphics

P.O.Box 9718, Beijing, P.R.China

Zip code 100101

E-mail jig@radi.ac.cn

Telephone 010-64807995

Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals

Domestic All Local Post Offices in China

Overseas China International Book Trading Corporation

(P.O.Box 399, Beijing 100048, P.R.China))

Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB

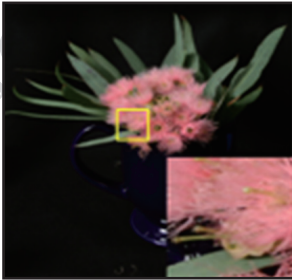
ISSN 1006-8961

CODEN ZTTXFZ

国外发行代号 M1406

国内邮发代号 82-831

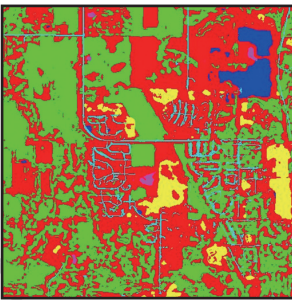
国内定价 50.00元



边缘结构保持的加权BDTV
全光场图像超分辨率重建
(第0733页)



基于超像素的多主体图像交
互分割(第0764页)



谐波分析光谱角制图高光谱
影像分类(第0836页)

图像处理和编码

边缘结构保持的加权BDTV全光场图像超分辨率重建

张旭东, 李梦娜, 张骏, 胡良梅, 王一 0733

基于Arnold映射的分块双层自适应扩散图像加密算法

徐亚, 张绍武 0740

结合梯度信息的特征相似性图像质量评估

苗莹, 易三莉, 贺建峰, 邵党国 0749

图像分析和识别

复杂场景下自适应背景减除算法

邵奇可, 周宇, 李路, 陈庆章 0756

基于超像素的多主体图像交互分割

伯彭波, 袁野, 王宽全 0764

采用金字塔纹理和边缘特征的图像烟雾检测

李红娣, 袁非牛 0772

结合手指检测和HOG特征的分层静态手势识别

刘淑萍, 刘羽, 於俊, 汪增福 0781

形状角计算改进及其在矩形目标识别中的应用

孟伟荣, 李辉, 方黎勇, 白金平 0789

融合图像感知哈希技术的运动目标跟踪

李子印, 朱明凌, 陈柱 0795

利用快速傅里叶变换的目标尺度自适应回归跟踪

张浪, 侯志强, 余旺盛, 许婉君 0805

基于点云属性信息的平面标靶特征提取

刘燕萍, 贾东峰 0815

融合局部思想和协作表达的鲁棒分类

楼宋江, 赵小明, 于海涛, 张石清 0822

计算机图形学

草绘3维人体建模的模板形变方法

李毅, 刘兴川, 孙亭 0828

遥感图像处理

谐波分析光谱角制图高光谱影像分类

杨可明, 刘飞, 孙阳阳, 魏华锋, 史钢强 0836

倾斜航空影像的城区DSM生成

吴军, 程门门, 姚泽鑫, 彭智勇, 李俊, 马峻 0845

封面图片由中国科学院遥感与数字地球研究所高光谱遥感应用技术研究室张霞研究员提供

CONTENTS

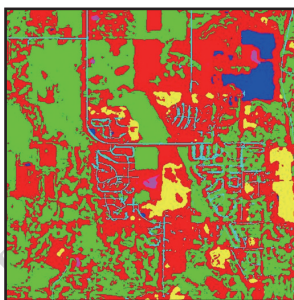
JOURNAL OF IMAGE AND GRAPHICS



Edge preserved light field image super-resolution based on weighted BDTV model (P0733)



Interactive multiphase image segmentation based on superpixels (P0764)



Classification algorithm of hyperspectral imagery by harmonic analysis and spectral angle mapping (P0836)

Image Processing and Coding

Edge preserved light field image super-resolution method based on weighted BDTV model

Zhang Xudong, Li Mengna, Zhang Jun, Hu Liangmei, Wang Yi 0733

Encryption algorithm of image blocking and double adaptive diffusion with Arnold mapping

Xu Ya, Zhang Shaowu 0740

Image quality assessment of feature similarity combined with gradient information

Miao Ying, Yi Sanli, He Jianfeng, Shao Dangguo 0749

Image Analysis and Recognition

Adaptive background subtraction approach of Gaussian mixture model

Shao Qike, Zhou Yu, Li Lu, Chen Qingzhang 0756

Interactive multiphase image segmentation based on superpixels

Bo Pengbo, Yuan Ye, Wang Kuanquan 0764

Image based smoke detection using pyramid texture and edge features

Li Hongdi, Yuan Feiniu 0772

Hierarchical static hand gesture recognition by combining finger detection and HOG features

Liu Shuping, Liu Yu, Yu Jun, Wang Zengfu 0781

Improved shape angle computation and its application in detection of rectangular

Meng Weirong, Li Hui, Fang Liyong, Bai Jinping 0789

Object tracking algorithm based on perception hash technology

Li Ziyin, Zhu Mingling, Chen Zhu 0795

Scale adaptive regression tracking method based on fast Fourier transform

Zhang Lang, Hou Zhiqiang, Yu Wangsheng, Xu Wanjuan 0805

Feature extraction of flat target based on point cloud attribute information

Liu Yanping, Jia Dongfeng 0815

Locality constrained collaborative representation classification for robust classification

Lou Songjiang, Zhao Xiaoming, Yu Haitao, Zhang Shiqing 0822

Computer Graphics

Sketch-based method for 3D human modeling using template deformation

Li Yi, Liu Xingchuan, Sun Ting 0828

Remote Sensing Image Processing

Classification algorithm of hyperspectral imagery by harmonic analysis and spectral angle mapping

Yang Keming, Liu Fei, Sun Yangyang, Wei Huafeng, Shi Gangqiang 0836

Automatic generation of high-quality urban DSM with airborne oblique images

Wu Jun, Cheng Menmen, Yao Zexin, Peng Zhiyong, Li Jun, Ma Jun 0845

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2015)06-0781-08

论文引用格式: Liu S P, Liu Y, Yu J, Wang Z F. Hierarchical static hand gesture recognition by combining finger detection and HOG features[J]. Journal of Image and Graphics, 2015, 20(6): 0781-0788. [刘淑萍, 刘羽, 於俊, 汪增福. 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): 0781-0788.] [DOI:10.11834/jig.20150607]

结合手指检测和 HOG 特征的分层静态手势识别

刘淑萍¹, 刘羽¹, 於俊¹, 汪增福^{1,2}

1. 中国科学技术大学自动化系, 合肥 230027; 2. 中国科学院合肥智能机械研究所, 合肥 230031

摘要: **目的** 基于手势的交互方式在人机交互中发挥着越来越重要的作用, 手势识别是大多数手势交互系统的核心技术。当手势种类较多时, 目前已有的大多数手势识别方法往往无法获得足够高的识别率。为此, 提出了一种结合手指检测和梯度方向直方图(HOG)特征的分层静态手势识别方法。**方法** 提出一种基于形态学操作的手指检测算法作为手势识别方法的基础。首先由肤色模型从输入图像中提取出手部区域, 然后利用手指检测算法识别出手势包含的手指个数, 并根据手指个数从事先训练好的支持向量机分类器集中选取一个, 最后提取手部区域的 HOG 特征, 并利用选择好的分类器完成识别任务。**结果** 对 25 种常用手势进行了识别实验, 将本文方法与单独使用 HOG 特征的方法进行对比。本文方法可以将传统 HOG 方法的识别率提高 20% 左右。**结论** 基于手指个数的分层识别策略可以有效地解决传统单层识别方法在手势种类较多时识别率不高的问题。在手部区域能被成功检测的情况下, 提出的结合手指检测和 HOG 特征的方法可以取得较理想的手势识别结果, 且能达到实时性要求。

关键词: 人机交互; 静态手势识别; 分层策略; HOG 特征; 支持向量机

Hierarchical static hand gesture recognition by combining finger detection and HOG features

Liu Shuping¹, Liu Yu¹, Yu Jun¹, Wang Zengfu^{1,2}

1. Department of Automation, University of Science and Technology of China, Hefei 230027, China;

2. Institute of Intelligent Machines, Chinese Academy of Sciences, Hefei 230031, China

Abstract: **Objective** Hand gestures have recently been regarded as an important communication modality in human-computer interaction. These gestures are considerably advantageous over traditional interfaces such as the keyboard and mouse given their convenience and naturalness. Hand gesture interfaces have been successfully employed in many areas over the past 20 years, including robotics, sign language communication, clinical medicine, and the entertainment industry. Recognition is the key technique in most interaction systems based on hand gestures. Such recognition can be grouped into two categories: dynamic and static hand gesture recognition. Most existing methods for static hand gesture recognition usually perform poorly given the large number of hand gesture categories. In this study, we propose a hierarchical method of static hand gesture recognition that combines finger detection and histogram of oriented gradient (HOG) features to overcome this limitation. **Method** A hierarchical strategy is designed based on the number of outstretched fingers to solve the “large-cate-

收稿日期: 2015-02-10; 修回日期: 2015-03-11

基金项目: 国家自然科学基金项目(61472393, 61303150); 国家科技计划项目(2012GB102007); 安徽省科技攻关项目(13Z02008)

第一作者简介: 刘淑萍(1990—), 女, 中国科学技术大学自动化系模式识别与智能系统专业在读硕士研究生, 主要研究方向为计算机视觉、手势识别、人机交互。E-mail: fengya@mail.ustc.edu.cn

Supported by: National Natural Science Foundation of China (61472393, 61303150)

gory” problem described above. The basic idea is as follows: Entire categories of gestures are divided into several subsets according to the number of outstretched fingers. This number is first recognized by a certain finger detection method as an input gesture. Then, we derive the appropriate gesture from the subset that contains the gestures featuring the given number of fingers via a certain feature extraction approach. Thus, the recognition problem is solved in two stages. Following the first finger-based stage, the number of categories that must be addressed by the second layer decreases significantly. To achieve this target, we propose a new finger detection method based on morphological operations. The proposed method consists of the following steps. First, the hand region is detected with a skin color model. Then, the number of outstretched fingers is identified by the finger detection method. Accordingly, a support vector machine classifier is selected from a set of pre-classified classifiers. Finally, we calculate the HOG feature of the input gesture, which is then inputted into the selected classifier to obtain a result. **Result** In our experiments, 25 popular static gestures are tested to verify the effectiveness of the proposed recognition method. The one-stage HOG-based recognition method is mainly used for comparison. Experimental results demonstrate that the proposed method can improve the recognition accuracy of the conventional HOG-based method by approximately 20%. **Conclusion** This study makes three main contributions to literature. First, we present a new finger detection method that can accurately determine finger position and identify the number of outstretched fingers. Second, we design a finger-based hierarchical recognition strategy that can effectively overcome the “large-category” limitation for conventional one-stage recognition methods. Third, we propose a two-stage method of static hand gesture recognition that combines finger detection and HOG features. When hand regions can be detected effectively, the proposed method can perform satisfactorily with real-time efficiency. In the future, we will design a practical system for hand gesture recognition that employs a robust hand detection approach. Furthermore, the presented finger detection method can be used to normalize rotations as necessary.

Key words: human-computer interaction; static hand gesture recognition; hierarchical strategy; HOG features; support vector machine

0 引言

近年来,手势逐渐发展为人机交互的一个重要方式。相对于鼠标、键盘等传统交互方式,基于手势的交互方式更加自然,在方便性和实用性上都具有明显的优势。在过去的二十年里,手势交互在诸多领域得到了广泛的应用,包括服务型机器人^[1]、手语翻译^[2]、医疗^[3]、娱乐^[4]等。目前,手势交互系统有两种常用的信息捕获方式:基于数据手套采集手部的运动数据和基于相机拍摄手部的图像视频。后者通过图像处理和模式识别技术对输入图像视频进行处理,实现手势交互的目的,具有对设备要求低、便于操作、交互自然等优点。

手势识别是绝大多数手势交互系统的核心技术,如何快速准确地判断一个手势的含义是需要解决的关键问题。一般来说,基于视觉的手势识别可以分为动态手势识别和静态手势识别两大类。动态手势是指一段时间内手部的运动过程,识别系统的任务是理解整个运动过程代表的意义;静态手势则关注某个时刻手部的一个特定形状,识别系统的任

务是判断该手部形状的含义。这两类手势识别系统在实际中有着不同的应用环境,研究方式也有很大区别。本文的研究对象是基于视觉的静态手势识别。图1所示是一个典型的静态手势识别系统流程图。对于一幅输入图像,首先进行手部区域的检测,再根据需要对检测结果进行旋转和尺度等归一化操作,然后采用某种方式提取特征来定量表示该手势的特性,最后采用某种分类器对手势进行分类判决。

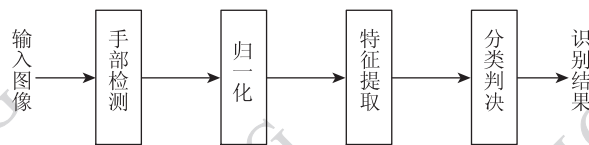


图1 典型的静态手势识别系统流程图

Fig. 1 Typical flowchart of static hand gesture recognition systems

在图1所示的各个模块中,特征提取步骤对手势识别系统的性能有着至关重要的作用,同时影响着识别的准确率和效率。用于描述手势的特征主要有两种形式:基于轮廓的特征和基于区域的特征。常用的轮廓描述方法^[5]包括曲率、链码、傅里叶描

述子等。此外, Hausdorff 距离等直接基于轮廓点空间位置的形状匹配方法也在手势识别中得到了应用^[6]。然而, 这类方法普遍对轮廓变形、视角变化和噪声比较敏感, 手部运动的高度灵活性往往限制了这类方法在手势识别中的应用。基于区域的特征提取方法利用整个目标区域的像素提取特征, 对噪声和形变更加鲁棒。在这类方法中, 图像的几何矩和多项式矩常被用来描述手势特征。文献^[7]利用 Hu 不变矩^[8]提取手势区域的特征。Hu 矩具有旋转和尺度不变性, 但仍对手部形变与视角变化比较敏感。最近, 有学者^[9]提出基于 2 维 Krawtchouk 矩^[10]的手势特征提取算法, 对空间形变与噪声鲁棒, 取得了不错的效果。同时, Krawtchouk 多项式是离散正交的, 可以保证较低的特征维度即能有效地表示出目标特性, 提高了分类效率。但是, Krawtchouk 矩不具有旋转不变性, 使得该方法在特征提取前必须进行旋转归一化。为了充分地利用手部区域的灰度信息, 有学者^[11-12]提出了一些基于图像局部特征的手势识别方法, 例如视觉特征词袋(BoF)^[13]和梯度方向直方图(HOG)^[14]等。这类特征提取方法对光照变化和遮挡具有较好的鲁棒性, 可以一定程度克服手部检测结果不准确对识别率的影响。基于 BoF 的方法具有旋转不变性, 但需要足够多的特征点来保证识别的准确性, 而手部区域的特征点往往较少。基于 HOG 的方法可以描述整个手部区域的灰度变化情况, 虽不具有旋转不变性, 但位置和方向空间的量化可以一定程度上抑制旋转带来的影响。

值得注意的是, 受限于当前常用特征的区分能力, 目前已有的大多数手势识别算法往往只能在手势类别数量较少时获得不错的性能。文献^[9, 11-12]中的方法是近年来手势识别领域具有代表性的几种方法, 它们处理的手势类别数分别 10、7 和 9 (更早的一些方法^[6-7]虽然提到能够处理较多类别的手势识别问题, 但它们对输入数据有严格的规范性要求, 旋转、形变和光照等因素都会对识别性能产生明显的影响)。当需要处理的手势种类增多时, 不同手势的类间距离迅速变小, 使得这些方法的识别率难以保证。特别地, 当类别数量较多时(如 20 种以上), 容易存在仅以旋转区分的手势, 这使得无法使用具有旋转不变性的特征, 也不能对手部区域进行旋转归一化。这样相当于增加了每一类的类内距离, 给识别任务带来了更大的困难, 一个微小的旋转

就可能导致识别错误。

为了解决上述问题, 设计了一种基于手指个数(指一个手势伸出来的手指数量)的分层识别策略, 其基本思想是: 将所有手势按照手指个数的不同分为若干个子集, 对于一个输入手势, 首先使用某种方法识别出其包含的手指个数, 然后在具有相同手指个数的手势子集中通过提取某种特征识别出该手势种类。这样, 经过第 1 层的手指个数识别, 第 2 层识别问题所面对的手势种类将显著减少。本文贡献主要包括以下 3 个方面:

- 1) 提出了一个基于形态学的手指检测方法, 可以准确地定位手指区域, 进而识别出手指数;
- 2) 提出了一种基于手指个数的分层识别策略, 可以有效解决传统方法在手势种类较多时识别率不高的问题;
- 3) 提出了一个结合手指检测和 HOG 特征的静态手势识别新方法。

1 本文方法

图 2 所示是本文静态手势识别方法的流程图, 包括以下几个主要步骤: 对输入图像进行手部区域检测, 得到二值图像及其对应的灰度图像; 对二值图像利用手指检测方法, 识别出手指数, 从事先训练的 SVM 分类器集合中选择相应的分类器; 对灰度图像, 计算其 HOG 特征; 将计算得到的特征输入已选择的 SVM 分类器进行分类判决, 得到识别结果。下面对以上几个步骤分别进行详细描述。

1.1 手部检测

本文主要研究的是分层识别策略在处理类别数量较多的手势识别问题时相对于传统方法的优点, 而手部区域的检测不是本文的关注重点。因此, 这里对手部检测问题做出简化处理, 假设输入图像中手部区域是所有肤色区域中最大的, 且前臂部分不裸露出来。在此基础上, 采用文献^[15]中的肤色模型提取手部区域。事实上, 复杂情况下的手部区域检测问题目前已有相关的研究成果。例如, 文献^[11]提出了一种去除脸部区域的方法, 文献^[9]提出了一种去除前臂区域的方法, 文献^[16]提出了一种基于深度相机的手掌检测方法。这些方法都可以自然地结合到图 2 所示的方法中, 使其可以处理输入更加复杂的情况。

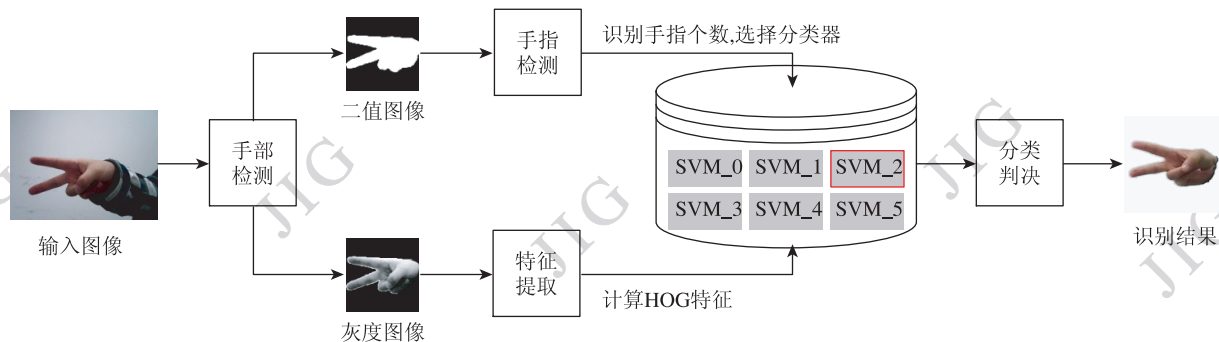


图2 本文静态手势识别方法的流程图

Fig. 2 Flowchart of the proposed static hand gesture recognition method

文献[15]通过结合 YUV 和 YIQ 彩色空间的信息来检测肤色像素。在 YUV 空间中,利用 U - V 分量计算色调因子

$$\theta = \tan^{-1}(V/U) \quad (1)$$

在 YIQ 空间中,同相彩色分量 I 的计算公式为

$$I = 0.596R - 0.274G - 0.322B \quad (2)$$

式中, R 、 G 和 B 表示 RGB 彩色空间的像素值。根据文献[15],亚欧人的肤色区域像素应满足

$$105^\circ \leq \theta \leq 150^\circ, 30 \leq I \leq 100 \quad (3)$$

利用式(3)得到的结果往往包含一些噪点,也会漏掉一些属于手部的像素。在此基础上,保留面积最大的区域并填充该区域内部的小洞,即可得到较理想的二值图像。为了后续处理的方便,取出手部区域的外接正方形区域,即先找出外接矩形,再将矩形的中心作为正方形中心、矩形的长边作为正方形边长得到该正方形。在获得手部区域的二值图像后,再将此二值图像和输入图像的对应区域进行点乘运算,即可获得手部区域的灰度图像,用于后续 HOG 特征提取。

1.2 手指检测

手指检测算法在检测得到的二值图像上进行,算法流程图如图3所示,包含以下步骤:

- 1) 对检测得到的二值图像进行尺度归一化,采样到 100×100 像素大小;
- 2) 对归一化后的二值图像进行开运算操作,其中结构元素是一个半径为 12 像素的圆;
- 3) 用步骤 1) 的结果减去步骤 2) 的结果,得到包含手指区域的二值图像;
- 4) 对步骤 3) 的结果进行后处理,先腐蚀以消除细小的连接部分,再去面积很小的噪声区域;
- 5) 统计步骤 4) 结果中连通区域的个数,即为手指个数。

指个数。

上述步骤中,步骤 2) 的目的是消除手指这样的细长区域,所以结构元素的直径要大于手指的宽度,但过大会导致开运算的腐蚀操作消除过多区域,影响最终结果。实验发现,当二值图像大小为 100×100 像素时,结构元素半径的理想取值范围是 10 ~ 15 像素,实验中该值固定为 12。

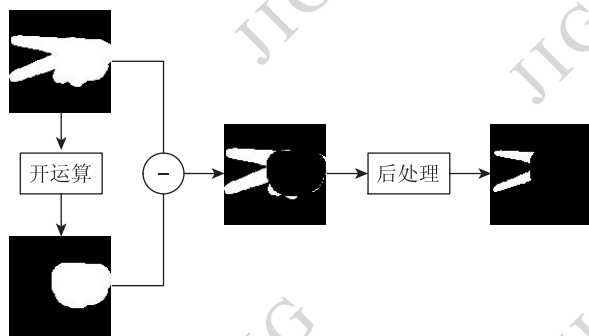


图3 手指检测算法流程图

Fig. 3 Flowchart of finger detection algorithm

前已提及,在手势种类较多时,容易存在仅以旋转区分的手势。因此,本文方法在检测完成后不进行旋转归一化的操作。事实上,上述手指检测算法不仅可以识别出手指数,还可以求得指尖的位置。因此,当处理的问题不包含仅以旋转区分的手势时,还可以利用该手指检测方法得到的指尖位置进行旋转归一化。

1.3 特征提取

为了减少 HOG 特征的计算开销,将手势灰度图像归一化到 64×64 像素,特征计算的步骤如下:

- 1) 计算归一化后图像的梯度信息,包括梯度幅值和方向;
- 2) 将图像划分为 8×8 个 Cell, 每个 Cell 包含

8×8 个像素, 相邻 2×2 个 Cell 构成 1 个 Block, Block 之间重叠 1 个 Cell 的宽度, 故共有 7×7=49 个 Block, 如图 4 所示;

3) 对每个 Block, 按照文献[14]的方法进行梯度权值投影, 形成一个 2×2×9=36 维的梯度方向直方图, 并进行 L₂ 范数归一化以减少光照影响;

4) 将所有 Block 的梯度描述子按顺序连接起来, 形成一个 49×36=1 764 维的描述子, 作为该手势图像的特征向量。

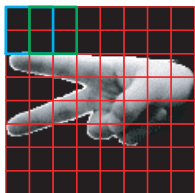


图 4 HOG 特征提取

Fig. 4 HOG feature extraction

1.4 分类判决

在训练阶段, 将具有相同手指个数的手势归为一个训练集, 每个训练集训练得到一个 SVM 分类器。一般来说, 不同手势的手指个数从 0 变化到 5, 但是有些手势的手指个数检测结果往往不具有唯一性, 此时可以在所有可能的类别中都加入该手势的训练样本。

在测试阶段, 根据手指检测算法识别出手指数, 选择相应的分类器对该手势的 HOG 特征向量进行分类判决, 得到最终的识别结果。

2 实验与分析

使用图 5 所示的 25 种手势(包含了数字、字母等生活中常用的手势)来验证本文方法的有效性。为便于叙述, 从 1 到 25 对这 25 种手势进行了编号, 这里编号不代表手势的实际含义。其中编号为 1~19 的手势来源于右手, 编号为 20~25 的手势来源于左手, 这里假设左右手相同的动作代表不同的含义。由图 5 可以看出, 这 25 种手势存在两组具有严格旋转关系的手势, 即手势 9 与手势 19、手势 11 与手势 14。此外, 还存在若干组近似存在旋转关系的手势, 如手势 3 与手势 22、手势 4 与手势 23、手势 5 与手势 24 等等。表 1 列出了不同手指个数下的手势编号, 可以看出有些手势出现了不止一次。实验

中, 对每种手势选取大量的样本进行手指个数识别测试, 通过统计所有可能的结果得到表 1。在这种情况下, 可以保证手指个数的识别正确率基本达到 100% (只要不违背表 1 的内容即认为识别正确)。



图 5 实验中用到的 25 种手势

Fig. 5 Twenty-five hand gestures used in the experiments

表 1 不同手指个数下的手势

Table 1 Hand gestures with different numbers of fingers

手指个数	手势序号
0	10, 11, 12, 14, 15
1	1, 2, 8, 10, 13, 17, 18, 20, 21
2	1, 3, 7, 9, 19, 20, 22
3	4, 16, 23
4	5, 16, 24
5	6, 25

实验中, 采集了 10 个人的手势图像序列, 手势的旋转角度均尽量控制为 0, (按图 5 所示的方向, 称为标准样本)。对每一种手势, 从每个人的数据中选取 70 幅图像构成标准样本库。为了测试手势旋转对识别效果的影响, 将这 70 个手势随机等分为 7 份, 每份 10 个, 将其中 1 份保持 0 度不变, 其余 6 份分别旋转 -30°、-20°、-10°、+10°、+20° 和 +30°, 得到旋转后的样本库。在此基础上, 构造了两种不同的训练集, 训练集 1 从标准样本库中选取训练样本, 训练集 2 从旋转后的样本库中选取训练样本。对每种训练集, 都选取相同 8 个人的手势数据, 这样每种手势共包含 8×70=560 个训练样本。从旋转后的样本库中选取其余 2 个人的数据作为测试集, 这样每种手势共包含 2×70=140 个测试样本。

使用传统的单层 HOG 手势识别方法^[12]与本文方法进行比较。两种方法均采用上述的训练集与测试集,且用于提取 HOG 特征的参数完全一致。实验中,所有的分类器均使用 OpenCV 库所带的 SVM 函数训练得到,其中核函数选用径向基函数,并采用自动求取最优参数的模式。

表 2 给出了 HOG 方法与本文方法对每一种手势的识别率。在训练集 1 中,只包含标准样本,由于

表 2 HOG 方法与本文方法的识别率比较

Table 2 Comparison between the HOG-based method and the proposed method on recognition accuracy

手势	训练集 1		训练集 2	
	HOG 方法	本文方法	HOG 方法	本文方法
	/%			
1	70.7	78.6	84.3	100
2	41.4	45.7	62.1	98.6
3	54.3	63.6	70.0	100
4	53.6	57.9	56.4	100
5	59.3	78.6	78.6	100
6	72.1	100	79.3	100
7	75.7	92.9	87.9	100
8	67.9	72.9	71.4	100
9	58.6	65.7	71.4	100
10	60.0	69.3	65.0	100
11	42.9	52.9	65.7	100
12	100	100	100	100
13	64.3	70.7	73.6	100
14	71.4	80.0	86.4	100
15	79.3	92.1	85.7	99.3
16	62.1	91.4	81.4	100
17	50.0	72.9	70.7	100
18	100	100	100	100
19	65.0	75.7	75.0	100
20	85.7	100	92.9	100
21	72.1	85.0	72.9	100
22	49.3	60.0	63.6	100
23	35.7	69.3	60.7	100
24	42.9	97.1	56.4	100
25	75.7	100	100	100
平均	64.4	78.9	76.5	99.9

手势种类较多且测试样本存在旋转,HOG 方法的平均识别率只有 64.4%,好几种手势的识别率不足 50%。本文方法在这种训练集下可以取得 78.9%的平均识别率,且每种手势的识别率较 HOG 方法都有所提高,但由于旋转样本的影响,识别率仍难以令人满意。在训练集 2 中,训练集和测试集的覆盖范围一致,HOG 方法的识别率有了明显的提高,但由于手势种类较多且存在旋转,类间距离很小,单独依靠 HOG 特征进行分类仍显得力不从心,平均识别率只提高到 76.5%,甚至略低于本文方法在训练集 1 中的识别率。本文方法在训练集 2 中可以得到近乎 100% 的识别率,只有两种手势在 140 个测试样本中出现了 1 到 2 个的识别错误。此结果表明,通过手指个数进行第 1 层分类,每个子问题的类别数量显著减少,使得 HOG 特征具有足够的能力完成第 2 层的分类任务。

为了更好地体现分层识别策略的优势,图 6 给出了两种训练集下,HOG 方法与本文方法的识别率与旋转角度关系曲线。这里的识别率是指一种旋转角度下所有 25 种手势的平均识别率。此时,对每种旋转角度,共包含 $25 \times 2 \times 10 = 500$ 个测试样本(测试集和上面是相同的,25 种手势,两个人,每个人每种手势共 10 个该角度下的测试样本)。由图 6(a)可以看出,在训练集 1 下,HOG 方法可以对旋转角度为 0° 的测试样本基本做到正确分类(正确数量是 499),但当旋转角度增加到 10° 以后,识别率迅速降低。本文方法在该训练集下的曲线变化趋势和 HOG 方法相同,但识别率降低的速率要明显低于 HOG 方法。在训练集 2 下,如图 6(b)所示,HOG 方法在 0° 和 10° 下的识别率基本达到 100%,但其他角度的识别率差异很大,这主要是由手势集包含的具体手势决定的。本文所用的 25 种手势中可能大部分手势逆向旋转更容易与其他手势产生混淆,导致 HOG 方法对正向旋转手势的识别率明显高于逆向旋转手势的。出现这种现象的根本原因在于在手势种类较多且存在旋转的情况下,HOG 特征的区分能力不足,可能对某些旋转角度下的某些手势做到准确识别,但很难做到全面地准确。和表 2 体现出的结果一致,本文方法在训练集 2 下可以得到理想的识别效果,只是在旋转角度为正 30° 时错分了 500 个样本中的 3 个,其他 6 种旋转角度下,识别率均为 100%。

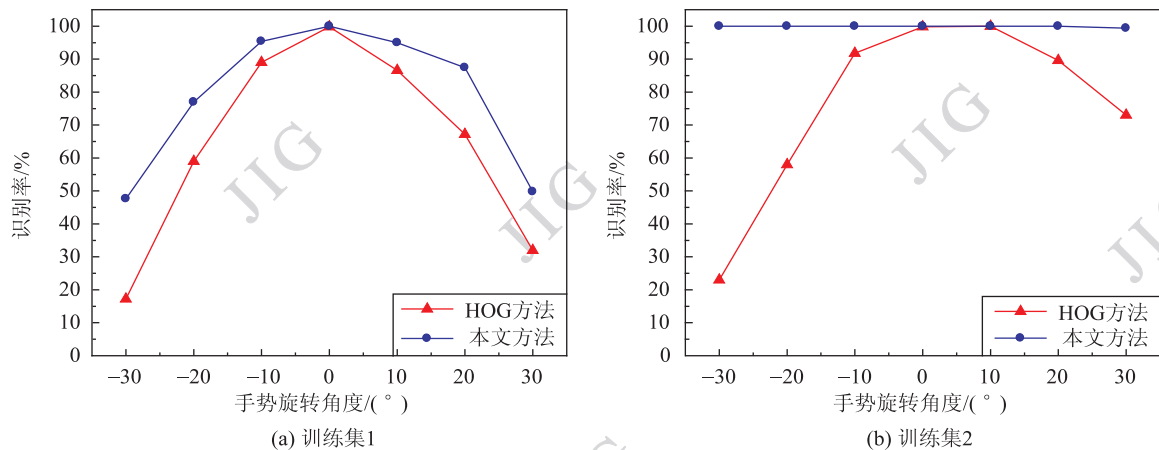


图6 两种方法的识别率与旋转角度关系曲线

Fig. 6 The relationship between recognition accuracy and rotation angle of two different methods

最后,对上述两种方法的效率进行测试。由于两种训练集包含样本数量一致,区别仅在于训练样本是否带有旋转性,因此只对训练集2(更加实用)下的情况进行测试。两种方法的硬件环境均为3.0 GHz的CPU和4 GB的内存。表3列出了两种方法处理测试集中所有图像的平均运行时间。可以看出,两种方法在效率上相差无几,每秒可以处理23帧左右,基本能达到实时性要求。

表3 HOG方法与本文方法的运行效率比较

Table 3 Comparison between the HOG-based method and the proposed method on computational efficiency

	HOG 方法	本文方法
每帧时间/ms	43.05	43.36
每秒帧数	23.23	23.06

3 结 论

近些年来,手势识别技术在人机交互领域受到越来越多的关注。然而,受限于特征的区分能力,传统的单层识别方法在手势种类数目较多时往往不能取得令人满意的结果。为解决此问题,设计了一个基于形态学手指检测算法的分层识别策略,并结合HOG特征提出了一个分层的手势识别方法。对25种常用手势的实验结果表明,基于手指个数的分层识别策略可以有效地解决传统单层识别方法在手势种类较多时识别率不高的问题。此外,在手部区域能被成功检测的情况下,提出的结合手指检测和

HOG特征的方法可以取得较理想的手势识别结果,且能达到实时性要求。

对手部区域检测算法做了简化处理,未来工作在于完善这一部分。通过结合相关的后处理方法或使用深度相机,使手部检测算法可以在复杂环境下得到理想的结果,以提高整个手势识别方法的实用性。

参考文献 (References)

- [1] Chen Y T, Tseng K T. Developing a multiple-angle hand gesture recognition system for human machine interactions [C]//Proceedings of the 33rd Annual Conference of the IEEE Industrial Electronics Society. Taipei, China: IEEE, 2007: 489-492. [DOI: 10.1109/IECON.2007.4460049]
- [2] Ong S C W, Ranganath S. Automatic sign language analysis: a survey and the future beyond lexical meaning [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(6): 873-891. [DOI: 10.1109/TPAMI.2005.112]
- [3] Wachs J P, Stern H J, Edan Y, et al. A gesture-based tool for sterile browsing of radiology images [J]. Journal of the American Medical Informatics Association, 2008, 15(3): 321-323. [DOI: 10.1197/jamia.M241]
- [4] Leyvand T, Meekhof C, Wei Y C, et al. Kinect identity: technology and experience [J]. Computer, 2011, 44(4): 94-96. [DOI: 10.1109/MC.2011.114]
- [5] Zhang D, Lu G. Review of shape representation and description techniques [J]. Pattern Recognition, 2004, 37(1): 1-19. [DOI: 10.1016/j.patcog.2003.07.008]
- [6] Zhang L G, Wu J Q, Gao W, et al. Hand gesture recognition based on hausdorff distance [J]. Journal of Image and Graphics, 2002, 7(11): 1144-1150. [张良国, 吴江琴, 高文, 等. 基于

- Hausdorff 距离的手势识别 [J]. 中国图象图形学报, 2002, 7(11): 1144-1150. [DOI: 10.11834/jig.2002011341]
- [7] Chalechale A, Safaei F, Naghdy G, et al. Hand posture analysis for visual based human machine interface [C]//Proceedings of APRS Workshop on Digital Image Computing. St Lucia, Australia: The University of Queensland, 2005: 91-96.
- [8] Hu M K. Visual pattern recognition by moment invariants [J]. IRE Transaction on Information Theory, 1962, 8(2): 179-187. [DOI: 10.1109/TIT.1962.1057692]
- [9] Priyal S P, Bora P K. A robust static hand gesture recognition system using geometry based normalizations and Krawtchouk moments [J]. Pattern Recognition, 2013, 46(8): 2202-2219. [DOI: 10.1016/j.patcog.2013.01.033]
- [10] Yap P T, Paramesran R, Ong S H. Image analysis by Krawtchouk moments [J]. IEEE Transactions on Image Processing, 2003, 12(11): 1367-1376. [DOI: 10.1109/TIP.2003.818019]
- [11] Triesch J, Von der Malsburg C. Real-time hand gesture detection and recognition using bag-of-features and support vector machine techniques [J]. IEEE Transactions on Instrumentation and Measurement, 2011, 60(11): 3592-3607. [DOI: 10.1109/TIM.2011.2161140]
- [12] Re Y, Gu C C. Hand gesture recognition based on HOG characters and SVM [J]. Bulletin of Science and Technology, 2011, 27(2): 211-214. [任彧, 顾成成. 基于 HOG 特征和 SVM 的手势识别 [J]. 科技通报, 2011, 27(2): 211-214.]
- [13] Lazebnik S, Schmid C, Ponce J. Beyond bags of features: spatial pyramid matching for recognizing natural scene categories [C]//Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition. New York, USA: IEEE, 2006: 2166-2178. [DOI: 10.1109/CVPR.2006.68]
- [14] Dalal N, Triggs B. Histograms of oriented gradients for human detection [C]//Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition. San Diego, USA: IEEE, 2005: 886-893. [DOI: 10.1109/CVPR.2005.177]
- [15] Teng X, Wu B, Yu W, et al. A hand gesture recognition based on local linear embedding [J]. Journal of Visual Languages and Computing, 2005, 16(5): 442-454. [DOI: 10.1016/j.jvlc.2005.04.003]
- [16] Yao Z W, Pan Z G, Teng G D. The wrist recognition and the center of palm estimation based on depth camera [J]. Journal of Image and Graphics, 2014, 19(3): 463-470. [姚争为, 潘志庚, 滕国栋. 基于深度相机的手腕识别与掌心估测 [J]. 中国图象图形学报, 2014, 19(3): 463-470.] [DOI: 10.11834/jig.20140317]