



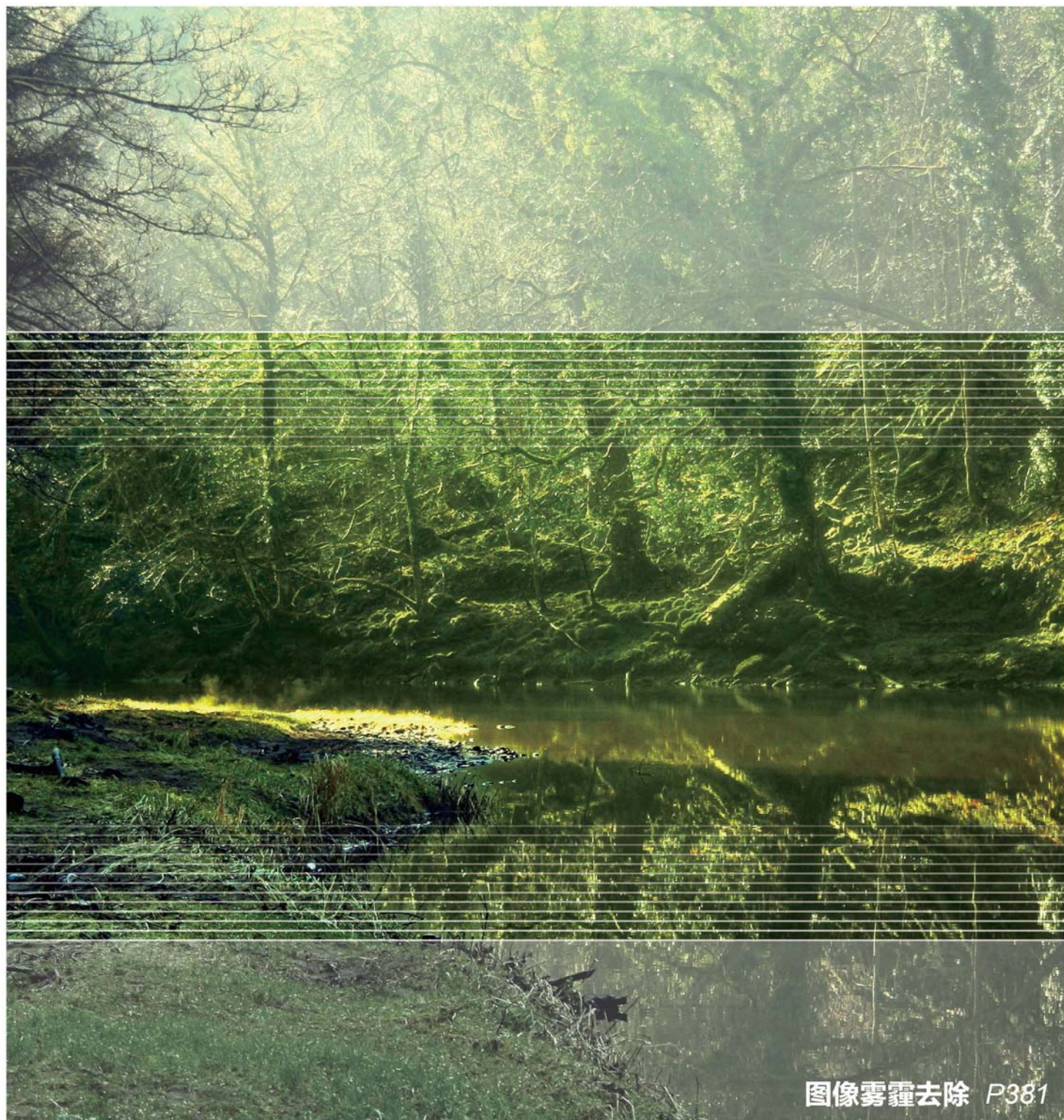
JOURNAL OF IMAGE AND GRAPHICS

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

2014
03
VOL.19

ISSN1006-8961
CN11-3758/TB



图像雾霾去除 P381



第19卷第3期（总第215期）

2014年3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

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@irsa.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@irsa.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

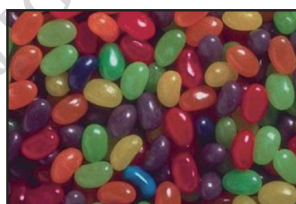
国外发行代号 M1406

国内邮发代号 82-831

国内定价 50.00元



控制关键帧选择的H.264熵编码加密算法(第358页)



显著性检测指导的高光区域修复(第393页)



由灭点进行径向畸变的自动校正(第407页)

综述

图像场景分类中视觉词包模型方法综述

赵理君, 唐娉, 霍连志, 郑柯 0333

面向对象的高分辨率SAR图像处理及应用

张红, 叶曦, 王超, 张波, 吴樊, 汤益先 0344

图像处理和编码

控制关键帧选择的H.264熵编码加密算法

张小红, 袁春经 0358

小波域视觉密码零水印算法

曲长波, 杨晓陶, 袁铎宁 0365

双重轮廓演化曲线的图像分割水平集模型

王相海, 李明 0373

暗原色先验单幅图像去雾改进算法

孙小明, 孙俊喜, 赵立荣, 曹永刚 0381

基于双边滤波的图像去雾

王一帆, 尹传历, 黄义明, 王洪玉 0386

显著性检测指导的高光区域修复

王祎璠, 姜志国, 史骏, 张浩鹏 0393

应用DCT块标准差的自适应隐写算法

凌轶华, 蔡晓霞, 陈红 0401

由灭点进行径向畸变的自动校正

刘丹, 刘学军, 王美珍 0407

图像分析和识别

黎曼流形上的保局投影在图像集匹配中的应用

曾青松 0414

结合全局和局部信息的“两阶段”活动轮廓模型

戚世乐, 王美清 0421

图像理解和计算机视觉

基于超像素的点追踪方法

罗会兰, 钟睿, 孔繁胜 0428

遥感图像处理

结合相位一致与全变差模型的高分辨率遥感图像边缘检测

黄秋燕, 肖鹏峰, 冯学智, 王珂 0439

第9届和谐人机环境联合学术会议专栏

人眼视觉感知特性的非下采样Contourlet变换域多聚焦图像融合

杨勇, 郑文娟, 黄淑英, 方志军, 袁非牛 0447

面向手绘军标图形的旋转自由识别方法

吴玲达, 张友根, 邓维, 宋汉辰 0456

基于深度相机的手腕识别与掌心估测

姚争为, 潘志庚, 滕国栋 0463

多信息融合的快速车牌定位

王永杰, 裴明涛, 贾云得 0471

H.264/AVC中基于决策树的P帧快速模式选择

王萍, 张晓丹, 张磊 0476

互信息域中的无参考图像质量评价

董宏平, 刘利雄 0484

CONTENTS

JOURNAL OF IMAGE AND GRAPHICS



H.264 video entropy coding encryption by controlling key frames(P358)



Algorithm for highlight area inpainting guided by saliency detection(P393)



Automatic approach of lens radial distortion correction based on vanishing points(P407)

Review

- Review of the bag-of-visual-words models in image scene classification
Zhao Lijun, Tang Ping, Huo Lianzhi, Zheng Ke 0333
- The processing and applications of high resolution SAR images with object-based image analysis
Zhang Hong, Ye Xi, Wang Chao, Zhang Bo, Wu Fan, Tang Yixian 0344

Image Processing and Coding

- H.264 video entropy coding encryption by controlling key frames
Zhang Xiaohong, Yuan Chunjing 0358
- Zero-watermarking visual cryptography algorithm in the wavelet domain
Qu Changbo, Yang Xiaotao, Yuan Duoning 0365
- Level set model for image segmentation based on dual contour evolutionary curve
Wang Xianghai, Li Ming 0373
- Improved algorithm for single image haze removing using dark channel prior
Sun Xiaoming, Sun Junxi, Zhao Lirong, Cao Yonggang 0381
- Image haze removal using a bilateral filter
Wang Yifan, Yin Chuanli, Huang Yiming, Wang Hongyu 0386
- Highlight area inpainting guided by saliency detection
Wang Yifan, Jiang Zhiguo, Shi Jun, Zhang Haopeng 0393
- Adaptive steganographic algorithm utilizing block standard deviation of DCT coefficients
Ling Yihua, Cai Xiaoxia, Chen Hong 0401
- Automatic approach of lens radial distortion correction based on vanishing points
Liu Dan, Liu Xuejun, Wang Meizhen 0407

Image Analysis and Recognition

- Locality preserving projection on Riemannian manifold for image set matching
Zeng Qingsong 0414
- "Two-stage" active contour model driven by local and global information
Qi Shile, Wang Meiqing 0421

Image Understanding and Computer Vision

- Method of point tracking based on superpixel
Luo Huilan, Zhong Rui, Kong Fansheng 0428

Remote Sensing Image Processing

- Edge detection from high resolution remote sensing image combining phase congruency with total variation model
Huang Qiuyan, Xiao Pengfeng, Feng Xuezhi, Wang Ke 0439

Special Issue on HHME 2013

- Multi-focus image fusion based on human visual perception characteristic in non-subsampled Contourlet transform domain
Yang Yong, Zheng Wenjuan, Huang Shuying, Fang Zhijun, Yuan Feiniu 0447
- Rotation free recognition of hand-drawn military marking symbols
Wu Lingda, Zhang Yougen, Deng Wei, Song Hanchen 0456
- The wrist recognition and the center of palm estimation based on depth camera
Yao Zhengwei, Pan Zhigeng, Teng Guodong 0463
- License plate detection based on multiple features
Wang Yongjie, Pei Mingtao, Jia Yunde 0471
- Fast intermode decision based on the decision tree for H.264/AVC P-frame encoding
Wang Ping, Zhang Xiaodan, Zhang Lei 0476
- No-reference image quality assessment in mutual information domain
Dong Hongping, Liu Lixiong 0484

中图法分类号: TP391.9 文献标识码: A 文章编号: 1006-8961(2014)03-0463-08

论文引用格式: 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.20140217]

基于深度相机的手腕识别与掌心估测

姚争为, 潘志庚, 滕国栋

杭州师范大学数字媒体与人机交互研究中心, 杭州 311121

摘要: **目的** 当手和手臂都进入深度相机所设定的有效深度范围时, 它们将被作为一个整体来提取, 若处理时也把它们作为一个整体, 这可能会影响手势交互的一些重要算法, 如掌心估测、手朝向估测、手的跟踪等。掌心是手势交互中较为稳定的点, 掌心与手簇中心的连线常被用来估测手的朝向。因此提高掌心估测算法的性能有助于提高手势交互的整体性能。**方法** 为了有效地分割手与手臂, 从分析手腕的运动特征和手的轮廓特点入手, 并利用内切矩形的几何特征, 提出手腕识别算法; 为了提高掌心估测的性能, 从手势交互的特点入手, 分析了锐角三角形和最大内切圆的几何特征, 提出新的掌心估测算法。**结果** 本文算法在空气多点触摸系统中进行了实验, 新的掌心估测算法较之原算法在性能上提高了近 7 倍, 且仍然能保持掌心坐标的稳定性, 坐标偏差不大于 3 个像素。同时手腕识别算法的引入也提高了掌心估测的准确性。**结论** 实验结果表明, 手腕识别算法能较好地分割出手与手臂, 新的掌心估测算法能很好地支持实时交互。

关键词: 手腕识别; 掌心估测; 深度相机; 手部分割; 手势交互

The wrist recognition and the center of palm estimation based on depth camera

Yao Zhengwei, Pan Zhigeng, Teng Guodong

Digital Media and HCI Research Center, Hangzhou Normal University, Hangzhou 311121, China

Abstract: **Objective** When the hand and the forearm enter the available depth range of the depth camera, the data of the hand and the forearm will be extracted together. Processing these data as a whole may affect some important algorithms such as the center of palm estimation, the orientation of hand estimation and hand tracing. The center of the palm is quite stable in the gesture interaction, so the line through the center of the palm and the center of the hand cluster is usual used as a hand orientation indicator. As a result, improving the performance of the center of palm estimation is favorable for increasing the overall performance. **Method** In order to correctly separate the hand from the forearm, the research begins with finding the motion features of the wrist and the contour features of the hand, and takes advantage of the geometric characteristics of an inscribed rectangle. At last a wrist recognition algorithm has been proposed. For improving the performance of the center of palm estimation, we start with analyzing the geometric characteristics of an acute triangle and an inscribed circle, and combine the features of the hand interaction. Finally a new algorithm of estimating the center of palm is proposed. **Result** The algorithms above are tested in an air multi-touch system. The proposed algorithm in this paper runs nearly 7 times quicker than the original algorithm, and still can keep the stability of the estimated coordinate of the center of palm, coordinate deviation not more than 3 pixels. Moreover, the success rate of the center of palm estimation is improved by using the wrist recognition algorithm. The accuracy of the center of palm estimation is improved due to using the wrist recognition algo-

收稿日期: 2013-09-09; 修回日期: 2013-11-03

基金项目: 国家自然科学基金重点项目(61332017); 浙江省教育厅年度科研项目(21120954)

第一作者简介: 姚争为(1977—), 男, 讲师, 2010 年于上海大学获计算机应用技术专业博士学位, 主要研究方向为增强现实和人机交互。

E-mail: yzwiron@hznu.edu.cn

rhythm. **Conclusion** Our experiments proved that the wrist recognition algorithm can separate the hand from the forearm well, and the new algorithm of the center of palm estimation can support real time interaction well.

Key words: wrist recognition; the center of the palm estimation; depth camera; hand segment; gesture interaction

0 引言

几十年来人机交互一直是众多学者关注的一个重要研究课题,这些年来交互方式层出不穷,但都在为实现自然交互,凭直觉的交互而努力,即只依靠人们在日常生活中已习得的知识经验,无需学习或只需少量学习就能操控系统^[1]。人手在交流与操控方面具有无可比拟的灵巧性,成为人类最高效、用途最广泛的交互工具之一,因此手势被广泛应用于人机交互,尤其在虚拟现实和增强现实领域^[2-3]。按照输入设备来分,手势交互系统可以分为4类:基于数据手套、基于标识^[4-5]、基于手部图像^[6-8]和基于手绘模型^[9-10]。本文系统属于第3类。在这一类中手部的分割与提取是成功实现手势交互的重要因素,被认为是整个手势交互处理链中最重要的一环^[11]。对于基于视频的手势交互系统而言,背景的复杂性以及光照会直接影响手部分割的效果^[12]。深度相机的出现可以解决视频相机遇到的一些问题,如低光照环境,环境中存在与肤色比较接近的其他对象等。较简单的手部分割是设置深度阈值将手从其他对象中分离出来^[13-14]。该方法认为在近距离阈值和远距离阈值之间的点是属于手的点,但这需要事先预测并要求用户在该范围内活动。一种关于设置深度阈值的改进方法是根据身体的其他部位的位置来决定手的深度,而不是假设手是场景中距离摄像

机最近的物体。Cerlinca^[15]和Van^[16]都是使用鲁棒的脸部检测算法找到头的位置,再来估测手的位置。Fujimura^[17]使用基于Karhunen-Loeve的分解算法来找到身体,然后再来对手进行估测。其实无论采用哪种方法,要想准确地将手和前臂分割(找到手腕)都是不容易的。Ren等人^[14]的方法需要用户戴上腕带以帮助准确地分割手与前臂。Lee等人^[18]所提的方法虽无需分割手和前臂就能提取出指尖和掌心。但是当前臂大面积进入有效深度范围内时,系统就无法稳定地获取掌心数据。而掌心是非常重要的信息,在手势交互中掌心与手簇中心的连线常被用来估测手的朝向,在空气鼠标,空气多点触摸系统中也常被用来代表手的位置。

本文系统是在CandescentNUI空气多点触摸系统的基础上进行开发的,提出了手腕识别算法,该算法能较准确地对手和手臂进行分割,这为之后的掌心估测打下了良好的基础。还提出了新的掌心估测算法,该算法引入锐角三角形与最大内切圆的几何特性,使得掌心估测更加高效。

1 手部分割与掌心估测

Candescent NUI系统采用OpenNI框架从Kinect传感器获取深度信息。图1为与本文相关的主要手势交互过程。

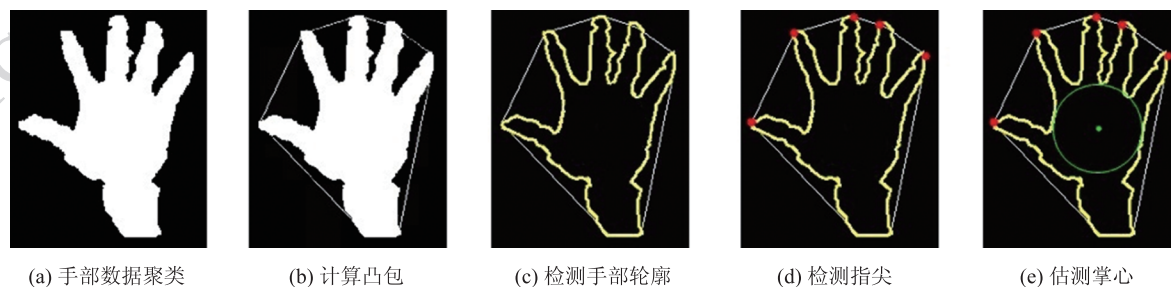


图1 手势交互部分过程

Fig. 1 The partial procedure of the gesture interaction system

1) 手的检测。通过人工设置深度阈值把手从背景中分离出来。系统将深度相机取到的手部数据映射成2维空间中的点,然后采用K均值聚类算法

将2维数据点分类,因此该系统可支持双手操作。

2) 手的凸包检测。采用Graham扫描算法来计算手簇的凸包。

3) 手的轮廓检测。聚类后系统采用摩尔邻居跟踪算法来得到手部轮廓。

4) 指尖检测。找到同时在凸包和轮廓上的点集 C 。对于其中的每个点 p_0 , 在点的两侧轮廓线上各取一个点 p_1, p_2 ; p_1, p_2 到 p_0 的距离为预先设定的值。计算 p_1, p_2 中点到 p_0 的距离, 只有这个距离大于特定阈值, 该点才识别为指尖点。

5) 掌心检测。取手轮廓内部的每个点, 计算这些点到轮廓上每一个点的距离, 取最小值。在这些最小距离值中找出最大值。这个最大值被认为是手轮廓内部最大圆的半径, 而这个圆心就认为是掌心。实际上为了减少执行时间, 系统每隔 8 个点取一个。

掌心是手势交互的重要数据, 因为手掌无论是张开, 握紧还是旋转, 掌心都是稳定不变的。但当用户手臂有较大部分进入有效深度范围内时, 原有算法将手臂也看成手的一部分, 这很容易错误地将掌心定位在手臂处。此外, 原掌心估测算法虽容易理解, 但是计算量较大。针对上述问题进行了研究, 提出手腕识别算法和新的掌心估测算法。

1.1 手腕识别算法

从手轮廓特征可知, 手腕点是凹陷点; 从手腕的运动特征可知, 手腕可进行较大幅度的转动。因此在交互时, 手腕的轮廓特征可能时刻变化着, 某一时刻在轮廓两侧可能只能找到一个手腕凹陷点, 如图 2 所示。



图2 手腕一侧的点

Fig. 2 One of the wrist points

本文方法是先找出一个明显的凹陷点, 然后再通过构建内切矩形来找出另一个手腕凹陷点。步骤如下:

1) 先找到凸包与轮廓的相交点集合, 根据这些点的顺序定义凸包线段 (每条凸包线段的端点由两个连续的点构成), 选取出符合条件的候选凸包线段, 这些候选凸包线段可能是手腕点所对应的凸包

线, 候选凸包线段的选取原则如下:

(1) 候选凸包线段的两个端点不能都是指尖点;

(2) 凸包线段的距离不能小于设定的阈值。

如图 3 所示, p_0p_1, p_5p_6, p_6p_7 为候选凸包线段。

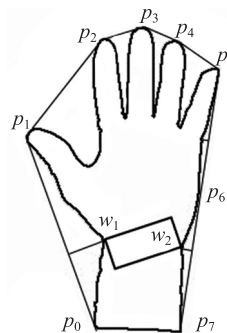


图3 手腕识别示意图

Fig. 3 The sketch of the wrist recognition

2) 找出与候选凸包线所对应的轮廓线。轮廓线的起点和终点就是候选凸包线两端的端点, 如果该轮廓线上还含有其他的凸包和轮廓的交点, 表明轮廓线方向取反了, 应换个方向搜索。

3) 计算轮廓线到候选凸包线的最大距离。如果有多条候选凸包线, 在所得到的最大距离中, 再取最大值, 该值所对应的点就为手腕点 (图 3 中的 w_1 点)。

4) 当一个手腕点确定后, 将该手腕点与轮廓上的每个点相连, 连线作为矩形的对角线, 判断矩形是否在轮廓线内部。若不在轮廓线内部, 该点就不是手腕点; 若在内部, 记录下该点和对角线长度。在所有记录的点中使对角线距离最短的点即为手腕另一侧的点 (图 3 中的 w_2 点)。

判断矩形是否在轮廓线内部, 采用如下方法: 依次判断轮廓线上的点是否在矩形内部, 若所有的点都不在矩形内部, 表明矩形在轮廓线内部, 反之亦然。判断点是否在矩形内部方法如下: 将矩形的 4 条边分成两组, 每一组是两条平行线, 如果这个点在两条平行线的同一侧, 表明这个点在矩形外面。如果这个点不在两条平行线的同一侧, 判断这个点和另一组平行线的关系。只有当这个点都不在这两组平行线的同一侧时, 这个点才是矩形内部的点。

找到了手腕两侧的点之后, 需要判断手腕哪一侧的轮廓属于手部。判断的依据是上一次掌心

的位置,与上一次掌心位置属同一侧的轮廓为手部轮廓。

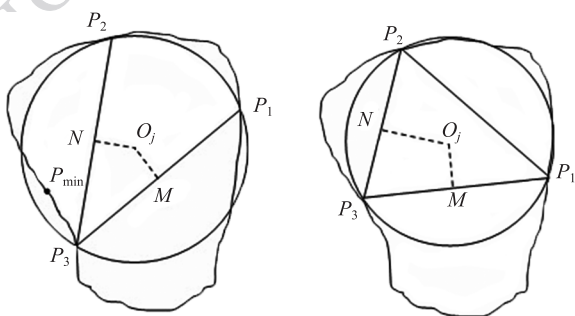
1.2 掌心估测算法

原有的算法(Stefan 提出的算法^[19])需取手部轮廓内的每个点,并计算每个点到轮廓的最短距离。在所有的最短距离中取最大值。这个最大值就是手轮廓内部最大圆的半径,而这个圆的圆心就是掌心。从算法可知,该算法的计算量比较大。因此本文参照 Sun 的思想^[20]来提高掌心估测算法的效率。但 Sun 的方法适用于轮廓边缘较为光滑,凹凸变化不明显的情况。而 Kinect 深度传感器本身就有一定的噪声,手的形状也较为复杂多变,因此该方法无法直接应用到系统中。

本文对 Sun 的终止条件进行了改进(Sun 的终止条件是轮廓上的所有点到圆心的距离不小于圆的半径)。考虑到一般的手势交互并不是高精度作业,因此掌心估测算法的终止条件未必一定要找到轮廓内最大圆,可找出接近最大圆的圆,即允许轮廓线上有极少量像素点出现在圆内。

同时为了去除由噪声造成的掌心错误估测,例如圆心错误地定位在了手指上等情况(此时找到的圆与手指轮廓相切,圆的直径就是手指的宽度),用最小圆半径进行限定。如果找到的圆的半径小于最小圆半径($R_{\min}$),需继续搜索掌心。 $R_{\min}$ 的值容易设定,只要 $R_{\min} \times 2$ 大于手指的宽度并小于手掌的宽度即可,由于手指宽度远远小于手掌宽度,因此有许多值可以选择,本文取 33。掌心估测算法步骤如下:

1) 首先在被测轮廓上选取 3 个初始点(如图 4 中的 3 点 P_1 、 P_2 、 P_3),并保证这 3 点构成一个锐角三角形。



(a) 初始状态下的圆

(b) 接近于轮廓最大内切图的圆

图 4 求掌心方法原理图

Fig. 4 The principal of estimating the center of the palm

2) 取这个锐角三角形的两条相交弦(P_1P_3 和 P_2P_3),并求它们中垂线的交点(O_j),该交点为候选的圆心点。候选圆心点为

$$X_j = \frac{\begin{bmatrix} [(Y_M - Y_N)(Y_3 - Y_1)(Y_3 - Y_2) + \\ X_M(X_3 - X_1)(Y_3 - Y_2) - \\ X_N(X_3 - X_2)(Y_3 - Y_1)] \end{bmatrix}}{\begin{bmatrix} (X_3 - X_1)(Y_3 - Y_2) - \\ (X_3 - X_2)(Y_3 - Y_1) \end{bmatrix}} \quad (1)$$

$$Y_j = \frac{\begin{bmatrix} (X_N - X_M)(X_3 - X_1)(X_3 - X_2) - \\ Y_M(Y_3 - Y_1)(X_3 - X_2) + \\ Y_N(Y_3 - Y_2)(X_3 - X_1) \end{bmatrix}}{\begin{bmatrix} (X_3 - X_1)(Y_3 - Y_2) - \\ (X_3 - X_2)(Y_3 - Y_1) \end{bmatrix}} \quad (2)$$

式中, X_j 、 Y_j 为点 O_j 的坐标值; X_1 、 Y_1 是点 P_1 的坐标值; X_2 、 Y_2 是点 P_2 的坐标值; X_3 、 Y_3 是点 P_3 的坐标值; X_M 、 Y_M 是弦 P_1P_3 的中点坐标值; X_N 、 Y_N 是弦 P_2P_3 的中点坐标值。

该圆的半径为之前选取的 3 点中任意一点到该圆心的距离为

$$R_j = \sqrt{(X_1 - X_j)^2 + (Y_1 - Y_j)^2} \quad (3)$$

3) 计算轮廓上各点到该圆心的距离序列

$$R_{ji} = \sqrt{(X_i - X_j)^2 + (Y_i - Y_j)^2} \quad (4)$$

式中, $i, j = 1, 2, \dots, n$, n 为轮廓上像素点的数量, X_i 、 Y_i 为轮廓上的点。

如果上述距离序列中值小于 R_j 的个数小于之前设定的阈值 T (轮廓线在圆内的像素个数),并且 R_j 大于之前设定的阈值 $R_{\min}$ (手掌的最小半径长度),表明该圆的圆心就是所要求取的掌心。否则进入步骤 4)。

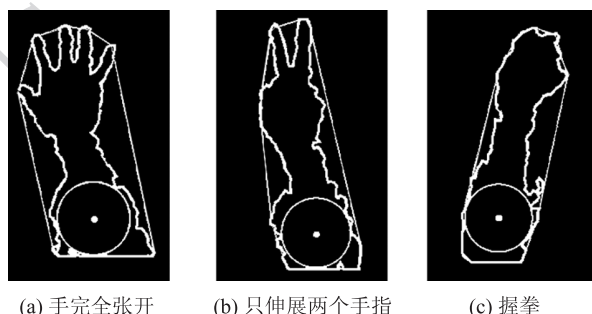
4) 用最短距离所对应的被测轮廓上的点 $P_{\min}$ 代替上述 3 点之一,并保证新的 3 点仍然形成一个锐角三角形,然后重复步骤 2)~4),直至条件满足。最后一次计算所得到的圆心就是被测手的掌心。

2 实验结果

采用 OpenNI 框架从 Kinect 传感器获取深度信息。CPU 为 AMD Athlon(tm) 64 × 2 Dual Core Processor 5200+, 主频 2.7 GHz, 内存 4 GB, 显卡为 NVIDIA GeForce 9600GT, 操作系统为 Window7 32 位。系统支持的手的个数设为 2, 有效深度区域为 0.5 ~ 0.8 m, 圆内允许的最大轮廓点个数 (T) 分别

取 25 和 50, 手掌的最小半径设为 33, 候选凸包线的最小距离设为 50。

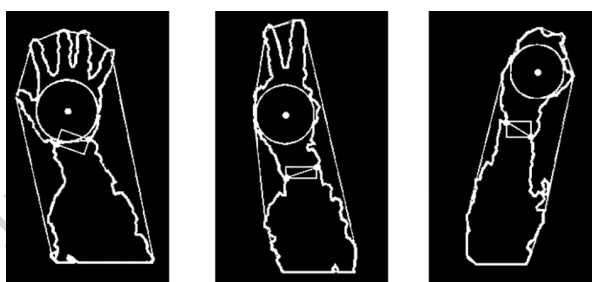
本文系统是基于 Candescent NUI 系统开发的, 因此本文手腕和掌心识别方法所用到的手的信息, 如手的凸包、轮廓、指尖等, 均由原系统获得。由于原系统在有大面积前臂进入有效区时, 容易错误估测掌心, 如图 5 所示。因此着重对有大面积前臂进入有效区域的情况进行了分析测试, 测试分 3 种情况: 手完全张开, 只伸展两手指和握拳。为了对新旧算法进行比较, 从真实的交互系统中截取了 3 幅图像进行测试。为了突出手腕识别效果, 该实验只在原有算法的基础上增加了手腕识别功能(掌心估测算法仍采用原算法), 实验结果如图 6 所示。从实验可以看出虽然手腕识别算法算不上精准, 但是达到了预期的效果, 即能将手与手臂分割开来, 防止了之后的掌心估测错误。



(a) 手完全张开 (b) 只伸展两个手指 (c) 握拳

图 5 未使用手腕识别算法的效果

Fig. 5 Estimate the center of the palm not using the wrist recognition algorithm



(a) 手完全张开 (b) 只伸展两个手指 (c) 握拳

图 6 加入手腕识别算法后的效果

Fig. 6 Estimate the center of the palm using the wrist recognition algorithm

掌心估测可以通过找出手掌中的最大内切圆来实现。考虑到手势交互并不是高精度的交互, 因此将终止条件修改为寻找接近最大内切圆的圆。但实验发现, 掌心容易被错误地定位在手指上, 如图 7 所示。

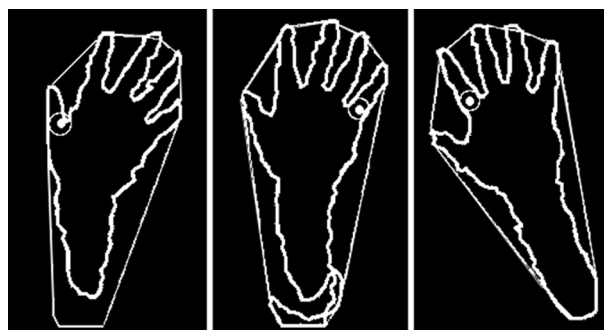


图 7 错误估测掌心

Fig. 7 The errors of estimating the center of palm

这是因为“接近最大内切圆的圆”在算法上表述为: 手轮廓线上有少量像素点出现在圆内。这样的表述不够全面, 没有对圆的最小半径进行限定。

图 8 为采用本文最终算法进行掌心估测的结果。为了对新老掌心估测算法进行比较, 我们根据手轮廓内像素点的数量, 取了 3 组数据。第 1 组数据见图 8(a)~(c), 手轮廓线上的像素点个数分别为 625、470、505, 轮廓线内像素点个数分别为 7 085、5 919、7 335; 第 2 组数据见图 8(d)~(f), 手轮廓线上的像素点个数分别为 1 074、1 356、1 155, 轮廓线内像素点个数分别为 11 205、12 362、13 047; 第 3 组数据如图 8(g)~(i), 手轮廓线上的像素点个数分别为 1 407、1 478、1 024, 轮廓线内像素点个数分别为 16 716、17 961、18 990。

由于原掌心估测算法(Stefan 算法)是求最大内切圆, 因此圆心具有唯一性。本文算法是求接近最大内切圆的圆, 这样的圆有多个, 这就需要通过实验来验证是否得到的掌心坐标具有稳定性, 以及本文算法的执行效率。实验对每幅图像测试 10 次, 然后取平均。表 1 为采用新方法的测试结果。

从表 1 可以看出本文算法有很高的执行效率, 能支持实时操作。尤其在图 8(a)~(i)等情况下, 本文算法能快速找到掌心。这说明本文算法的执行时间与轮廓线内像素的个数并无太大关联, 主要取决于轮廓线本身: 轮廓线越光滑, 变化越平缓, 执行效率越高。因此当用户握拳时, 该算法具有最高的执行效率。

从表 1 还可以看出, 适当地放宽限制条件(轮廓线在圆内的像素个数), 算法的执行时间就能迅速缩短, 而且从坐标的标准差可以看出掌心坐标仍然能保持较好的稳定性。因此该算法能根据系统的实际需要平衡速度和精度。

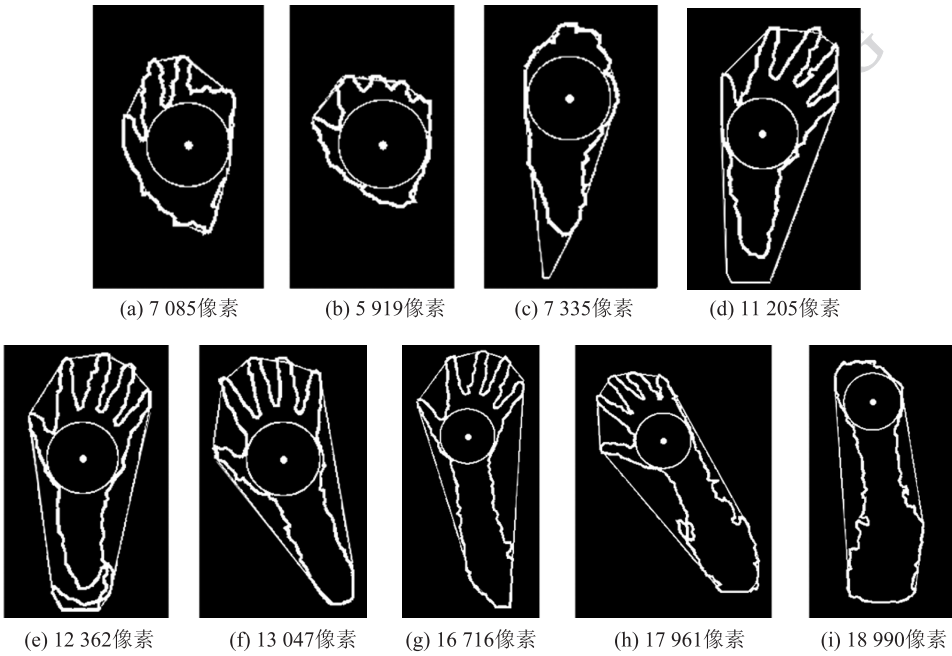


图 8 按轮廓线内像素点数量分成的 3 组数据

Fig. 8 Divide into three groups based on the number of the points inside the hand contour

表 1 采用新的掌心估测算法结果

Table 1 The results of using the new algorithm of estimating the center of the palm

图 8	阈值 T	平均耗时/s	轮廓线在圆内的平均个数	掌心 X 坐标标准差/像素	掌心 Y 坐标标准差/像素
(a)	25	0.016 354	14.8	2.002 776	6.963 396
	50	0.009 256	35.9	2.806 738	8.327 331
(b)	25	0.023 305	17.7	0.875 595	1.751 19
	50	0.009 176	32.9	1.286 684	2.674 987
(c)	25	0.020 285	16.2	0.471 405	0.948 683
	50	0.010 28	32.2	1.549 193	5.850 926
(d)	25	0.069 835	18.1	5.349 974	7.501 111
	50	0.027 444	32.1	5.719 363	7.633 988
(e)	25	0.152 618	17.7	1.766 981	3.018 462
	50	0.040 003	34.4	2.780 887	3.938 415
(f)	25	0.069 672	14.9	1.135 292	6.733 828
	50	0.028 626	38.2	1.159 502	6.835 366
(g)	25	0.085 116	11.4	2.699 794	7.211 103
	50	0.038 471	33.9	2.875 181	7.367 345
(h)	25	0.085 419	18.1	2.616 189	8.643 559
	50	0.025 23	36.7	3.620 927	9.254 54
(i)	25	0.024 086	16.7	2.503 331	3.502 38
	50	0.010 914	36.5	2.945 807	4.473 378

表 2 为 Stefan 算法、改进的 Stefan 算法, 本文算法之间的耗时对比。其中改进的 Stefan 算法是指: 由于原有的算法并不适用于实时操作, 因此改进算法不选取轮廓线上的所有点, 而是每隔 8 个点取一

个。轮廓线内点的选取也采用相同方式。此处的本文算法与之前的新方法也略有不同, 对于轮廓线上的点也是每隔 8 个点取一个。从结果可以看出本文算法在速度上具有较明显的优势。

表 2 掌心估测算法耗时比较

Table 2 Time-consuming comparison among algorithms

	图 8(a)	图 8(b)	图 8(c)	图 8(d)	图 8(e)	图 8(f)	图 8(g)	图 8(h)	图 8(i)
Stefan 算法	0.432 5	0.270 9	0.358 5	1.149 2	1.609 9	1.447 8	2.259 4	2.575 9	1.889 1
改进的 Stefan 算法	0.006 8	0.004 1	0.005 8	0.021 2	0.026 6	0.022 8	0.034 3	0.039 1	0.028 6
本文算法	0.000 2	0.000 8	0.002 8	0.003 8	0.002 4	0.004 7	0.008 5	0.004 5	0.002 1

/s

3 结 论

本文手腕识别算法对手和前臂进行分割较为准确, 新的掌心估测算法提高了掌心的查找效率。当手臂大面积进入有效深度范围内时, 本文算法能识别出手腕, 将手与手臂进行准确的分割。这为之后稳定、可靠地获取掌心做了充分且必要的准备, 而且这也降低了对 Kinect 传感器放置位置的要求。同时本文提出的掌心估测算法, 较之原有算法在运行效率上有了较大幅度的提升。这两个算法虽应用于空气多点触摸系统, 但它们也可作为手势识别的前期步骤。由于本文没有采用 Kinect 的骨架结构, 因此用户可以坐在电脑面前近距离操控电脑, 进行手指级别的相关操作。

本文系统采用 OpenNI 框架, 但当改用 Kinect SDK 时, 系统的性能会有明显的下降, 无法达到实时交互。同时本文系统要求用户在规定的深度范围内 (0.5 ~ 0.8 m) 交互, 这对交互的自然性会有一定的影响。因此今后的工作将关注在 Kinect SDK 环境下如何提高系统的性能以及如何减小交互深度范围的限制。

参考文献 (References)

- [1] Wachs J P, Kölsch M, Stern H, et al. Vision-based hand-gesture applications [J]. Communications of the ACM, 2011, 54(2): 60-71. [DOI: 10.1145/1897816.1897838]
- [2] Lee B, Chun J. Manipulation of virtual objects in marker-less AR system by fingertip tracking and hand gesture recognition [C]// Proceedings of the 2nd International Conference on Interaction Science. Seoul, Republic of Korea: ACM, 2009: 1110-1115. [DOI: 10.1145/1655925.1656127]
- [3] Sun C, Zhang M M, Li Y. Hand based natural interaction in augmented reality [J]. Journal of Computer-Aided Design & Computer Graphics, 2011, 23(4): 697-704.
- [4] Paul G K, Adeline P, Adrien B, et al. Hand navigator: hands-on interaction for desktop virtual reality [C]//The 15th ACM Symposium on Virtual Reality Software and Technology. Bordeaux, France: ACM, 2008: 53-60. [DOI: 10.1145/1450579.1450591]
- [5] Wang R Y, Popovic J. Real-time hand-tracking with a color glove [C]//Proceedings of the ACM SIGGRAPH. New Orleans, USA: ACM, 2009:63 [DOI: 10.1145/1531326.1531369]
- [6] Ravikiran J, Kavi M, Suhas M, et al. Finger detection for sign language recognition [C]//Proceedings of the International Multiconference of Engineers and Computer Scientists. Hong Kong, China: IAENG, 2009:489-493.
- [7] Youssef K, Woo P Y. Difference of the absolute differences-a new method for motion detection [J]. International Journal of Intelligent Systems and Applications, 2012, 4(9): 1-14. [DOI: 10.5815/ijisa.2012.09.01]
- [8] Pan Z G, Li Y, Zhang M M, et al. A real-time multi-cue hand tracking algorithm based on computer vision [C]//Proceeding of the VR 2010. Waltham, MA: IEEE 2010: 219-222. [DOI: 10.1109/VR.2010.5444787]
- [9] Lee W S, Kara L B, Stahovich T F. An efficient graph-based recognizer for hand-drawn symbols [J]. Elsevier Computers and Graphics, 2007, 31(4): 554-567. [DOI: 10.1016/j.cag.2007.04.007]
- [10] Kara L B, Stahovich T F. An image-based, trainable symbol recognizer for hand-drawn sketches [J]. Elsevier Computers and Graphics, 2005, 29(4): 501-517. [DOI: 10.1016/j.cag.2005.05.004]
- [11] An H, Kim D. Hand gesture recognition using 3D depth data [R]. Korea: Department of Computer Science and Engineering, Pohang University of Science and Technology, 2009.

- [12] Erol A, Bebis G, Nicolescu M, et al. Vision-based hand pose estimation: a review [J]. Computer Vision and Image Understanding, 2007, 108: 52-73. [DOI: 10.1016/j.cviu.2006.10.012]
- [13] Uebbersax D, Gall J, Van den Bergh M, et al. Real-time sign language letter and word recognition from depth data [C]//Proceedings of the International Conference on Computer Vision Workshops. Barcelona: IEEE, 2011: 383-390. [DOI: 10.1109/ICCVW.2011.6130267]
- [14] Ren Z, Yuan J, Zhang Z. Robust hand gesture recognition based on finger-earth mover's distance with a commodity depth camera [C]//Proceedings of the ACM International Conference on Multimedia. Scottsdale, Arizona: ACM, 2011: 1093-1096.
- [15] Cerlinca T I, Pentiu S P. Robust 3D hand detection for gestures recognition [C]//The 5th International Symposium on Intelligent Distributed Computing. Delft, Netherlands: Springer, 2011: 259-264.
- [16] Van den Bergh M, Van Gool L. Combining RGB and ToF cameras for real-time 3D hand gesture interaction [C]//Proceedings of the Workshop on Applications of Computer Vision. Kona, Hawaii: IEEE, 2011: 66-72. [DOI: 10.1109/WACV.2011.5711485]
- [17] Fujimura K, Xia L. Sign recognition using depth image streams [C]//Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition. Southampton, UK: IEEE, 2006: 381-386. [DOI: 10.1109/FGR.2006.101]
- [18] Lee U, Tanaka J. Hand controller: image manipulation interface using fingertips and palm tracking with Kinect depth data [C]//Proceedings of the 10th Asia Pacific Conference on Computer Human Interaction. Matsue, Japan: ACM, 2012: 705-706
- [19] Stefan S. Candescent NUI [CP/OL]. 2012-01-04 [2012-10-24]. <http://candescentnui.codeplex.com/releases/view/518198>.
- [20] Sun Y Q, Che R S. Novel method for solving maximum inscribed circle [J]. Optics and Precision Engineering, 2003, 11(2): 78-84. [孙玉芹, 车仁生. 求解最大内切圆的一种新方法[J]. 光学精密工程, 2003, 11(2): 78-84.]