

学校编码: 10384

分类号_____密级_____

学号: 24320111152276

UDC_____

廈門大學

硕士学位论文

基于单目摄像头的手势识别方法研究

Research on Gesture Recognition Method Based on
Single Camera

高雅萍

指导教师姓名: 段 鸿 副 教 授

专业名称: 计算机软件与理论

论文提交日期: 2014 年 04 月

论文答辩日期: 2014 年 05 月

学位授予日期: 2014 年 06 月

指导教师: _____

答辩委员会主席: _____

2014年5月

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下，独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果，均在文中以适当方式明确标明，并符合法律规范和《厦门大学研究生学术活动规范（试行）》。

另外，该学位论文为（）课题（组）的研究成果，获得（）课题（组）经费或实验室的资助，在（）实验室完成。（请在以上括号内填写课题或课题组负责人或实验室名称，未有此项声明内容的，可以不作特别声明。）

声明人（签名）：

年 月 日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

1.经厦门大学保密委员会审查核定的保密学位论文，于 年 月 日解密，解密后适用上述授权。

2.不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月 日

摘 要

随着计算机和网络的迅速发展,人机交互已经是人们日常生活中的一个重要组成部分。旧有的键盘、鼠标等利用物理媒介进行人机交互的方式已经不能完全满足用户的需要。一些新型的人机交互方式逐渐成为研究的热点,如手势识别、语音识别、人脸识别等。将手势作为人机交互手段,相比其他交互方式更加自然,也更加直接,在智能系统、机器人控制、多媒体教学、娱乐游戏等领域拥有广泛的应用前景。

本文研究的是单目视觉下利用手势识别方法进行虚拟输入。涉及到的领域包括图像分割、指尖检测、手势跟踪以及识别。所做的研究主要包括以下几个方面:

1、在手势分割方面,研究了目前常见的肤色空间模型,分析了 YCbCr 颜色空间的优缺点,提出了结合肤色和运动信息的基于相邻帧的自适应高斯建模方法进行肤色检测,得到了较好的分割效果。

2、在指尖检测方面,研究了基于曲率的指尖检测方法,利用 K-余弦算法以及 K-medoids 聚类方法对指尖点进行筛选。对 K-余弦算法提出了自适应计算 K 值的方法,提高了指尖候选点的选取精度,从而提高检测的准确性。

3、在手势跟踪阶段,采用了经典的 Camshift 跟踪算法。由于 CamShift 是半自动的算法,需要初始化目标窗口。本文以张开的手掌作为起始动作,并将此手掌的区域作为初始化窗口传递给 CamShift 算法。

4、在手势识别阶段,本文通过选取几个明显的特征值作为特征向量,利用模板匹配的方式进行分类识别,对于动态手势则采用简单的 DTW 方法进行识别。

最后,本文对手势识别做了交互系统的实现,利用动态手势来控制 Picasa 中图片的浏览和操作,获得了较好的效果。

关键词: 手势识别; 图像分割; 指尖检测; CamShift; 模板匹配

Abstract

With the computer and network developing rapidly, human-computer interaction has become an important part in people's daily life. Some traditional way of human-computer interaction, such as a mouse, keyboard, cannot fully meet the needs of users. Some new human-computer interaction has become a hot topic, such as gesture recognition, speech recognition, face recognition and so on. Compared to other interaction, the gesture, as a means of human-computer interaction, is more natural and direct. It has wide application prospect in intelligent system, robot control, multimedia, entertainment etc...

This study is the virtual input method based on hand gesture recognition under monocular vision, involved in the field to the hand gesture segmentation, fingertip detection, hand tracking and gesture recognition. The research mainly includes the following aspects:

1. In the part of gesture segmentation, it studied the present common color space model, analyzed the advantages and disadvantages of YCbCr color space. Combined of skin color and motion information, it proposed an adaptive Gauss modeling method based on adjacent frame for skin color detection, and got better effect of segmentation.
2. In the fingertip detection, it studied the fingertip detection method based on curvature, screened the fingertips by K-cosine algorithm and K-medoids clustering algorithm. To improve the accuracy of the selected candidate points, it presented a method to adaptive computation of K on K-cosine algorithm. Through the experiment, it improved the detection accuracy.
3. In the hand gesture tracking stage, it used the classic Camshift algorithm. Because it was semi-automatic, we need to initialize the target window. In this dissertation, we made the open hands as the starting action, and passed the palm area to the CamShift algorithm as the initial window.
4. In the step of gesture recognition, this dissertation selected several obvious

characteristic values to form the characteristic vector, and using template matching method for classification and recognition. For the dynamic hand gesture, it used a simple identification, DTW algorithm.

Finally, this dissertation made an interactive system for gesture recognition, using dynamic hand gestures to control picture browsing and operation in Picasa, good results were obtained.

Keywords: Gesture Recognition; Image Segmentation; Fingertip Detection; CamShift; Template Matching

厦门大学博硕士学位论文摘要库

目 录

| | |
|------------------------------|-----------|
| 第一章 绪论 | 1 |
| 1.1 研究背景及意义 | 1 |
| 1.2 国内外研究现状 | 2 |
| 1.3 主要研究内容与特色 | 3 |
| 1.4 本文结构安排 | 3 |
| 第二章 相关技术综述 | 5 |
| 2.1 基于视觉的手势识别技术 | 5 |
| 2.1.1 基于单目摄像头的手势识别..... | 5 |
| 2.1.2 基于多目摄像头的手势识别..... | 5 |
| 2.2 关键技术介绍 | 8 |
| 2.2.1 手势分割..... | 8 |
| 2.2.2 特征提取..... | 10 |
| 2.2.3 手势跟踪..... | 11 |
| 2.2.4 手势识别..... | 13 |
| 2.3 本章小结 | 14 |
| 第三章 手势分割方法研究 | 15 |
| 3.1 基于肤色的手势分割 | 15 |
| 3.1.1 常用的颜色空间模型..... | 16 |
| 3.1.2 基于 YCbCr 空间的肤色检测 | 19 |
| 3.2 基于运动的手势分割 | 20 |
| 3.2.1 时间差分法..... | 21 |
| 3.2.2 背景差分法..... | 21 |
| 3.3 基于肤色和运动信息的手势分割方法 | 22 |
| 3.3.1 基于相邻帧的自适应肤色检测..... | 22 |
| 3.3.2 背景去噪..... | 25 |
| 3.4 实验结果与分析 | 26 |

| | |
|-------------------------------|-----------|
| 3.5 本章小结 | 27 |
| 第四章 指尖检测方法研究 | 28 |
| 4.1 基于曲率的指尖检测算法 | 28 |
| 4.1.1 轮廓曲率检测..... | 28 |
| 4.1.2 峰值点聚类..... | 30 |
| 4.1.3 指尖信息获取..... | 30 |
| 4.2 改进的指尖检测算法 | 31 |
| 4.2.1 基于距离变换获取掌心掌宽..... | 32 |
| 4.2.2 K-余弦算法中自适应的 K 值计算 | 34 |
| 4.2.3 排除干扰信息..... | 35 |
| 4.3 实验结果与分析 | 36 |
| 4.4 本章小结 | 38 |
| 第五章 手势识别方法研究与实现 | 39 |
| 5.1 手势跟踪技术 | 39 |
| 5.2 静态手势识别 | 40 |
| 5.2.1 特征提取..... | 40 |
| 5.2.2 模板匹配..... | 43 |
| 5.2.3 实验结果与分析..... | 43 |
| 5.3 动态手势识别 | 45 |
| 5.3.1 DTW 算法原理 | 45 |
| 5.3.2 轨迹特征选择..... | 47 |
| 5.3.3 实验结果与分析..... | 48 |
| 5.4 手势识别交互系统的实现 | 49 |
| 5.4.1 系统简介..... | 49 |
| 5.4.2 交互应用..... | 50 |
| 5.5 本章小结 | 51 |
| 第六章 总结与展望 | 53 |
| 6.1 总结 | 53 |

| | |
|-----------------|----|
| 6.2 展望 | 53 |
| 参考文献 | 55 |
| 硕士期间发表的论文 | 59 |
| 致 谢 | 60 |

厦门大学博硕士论文摘要库

Contents

| | |
|---|-----------|
| Chapter 1 Introduction..... | 1 |
| 1.1 The Research Background and Significance | 1 |
| 1.2 The Research Status at Home and Abroad..... | 2 |
| 1.3 The Main Research Contents and Characteristics | 3 |
| 1.4 The Structure Arrangement..... | 3 |
| Chapter 2 Related Technology Review | 5 |
| 2.1 Gesture Recognition Technology Based on Vision | 5 |
| 2.1.1 Gesture recognition technology based on single camera | 5 |
| 2.1.2 Gesture recognition technology based on multi camera | 5 |
| 2.2 Introduction of Key Technologies..... | 8 |
| 2.2.1 Gesture segmentation..... | 8 |
| 2.2.2 Feature extraction..... | 10 |
| 2.2.3 Gesture tracking | 11 |
| 2.2.4 Gesture recognition | 13 |
| 2.3 Chapter Summary..... | 14 |
| Chapter 3 Research on Gesture Segmentation | 15 |
| 3.1 The Gesture Segmentation Based on Skin Color | 15 |
| 3.1.1 The common color space models..... | 16 |
| 3.1.2 Skin color detection based on YCbCr space | 19 |
| 3.2 The Gesture Segmentation Based on Motion | 20 |
| 3.2.1 Temporal difference method | 21 |
| 3.2.2 Background subtraction method | 21 |
| 3.3 Gesture Segmentation Based on Skin Color and Motion..... | 22 |
| 3.3.1 Adaptive skin color detection based on adjacent frame | 22 |
| 3.3.2 Background denoising | 25 |
| 3.4 Experimental Results and Analysis | 26 |

| | |
|---|-----------|
| 3.5 Chapter Summary | 27 |
| Chapter 4 Study on the Fingertip Detection Method | 28 |
| 4.1 The Fingertip Detection Algorithm Based on the Curvature | 28 |
| 4.1.1 Contour curvature detection..... | 28 |
| 4.1.2 Peak clustering | 30 |
| 4.1.3 Information acquisition of fingertips | 30 |
| 4.2 The Improved Fingertip Detection Algorithm | 31 |
| 4.2.1 Obtain the palm based on the distance transform | 32 |
| 4.2.2 Adaptive calculation of K value in the K- cosine algorithm..... | 34 |
| 4.2.3 Remove interference information | 35 |
| 4.3 Experimental Results and Analysis | 36 |
| 4.4 Chapter Summary | 38 |
| Chapter 5 Research and Implementation of Gesture Recognition | 39 |
| 5.1 Hand Tracking | 39 |
| 5.2 Static Gesture Recognition | 40 |
| 5.2.1 Feature extraction..... | 40 |
| 5.2.2 Template matching | 43 |
| 5.2.3 Experimental results and analysis | 43 |
| 5.3 Dynamic Gesture Recognition | 45 |
| 5.3.1 The principle of DTW algorithm | 45 |
| 5.3.2 Feature selection of trajectory..... | 47 |
| 5.3.3 Experimental results and analysis | 48 |
| 5.4 Implementation of Gesture Recognition Interaction System | 49 |
| 5.4.1 System introduction | 49 |
| 5.4.2 Interactive application..... | 50 |
| 5.5 Chapter Summary | 51 |
| Chapter 6 Conclusion and Prospect | 53 |
| 6.1 Conclusion | 53 |
| 6.2 Prospect | 53 |

| | |
|--|-----------|
| References | 55 |
| Publication as the Degree Candidate | 59 |
| Acknowledgement | 60 |

厦门大学博硕士论文摘要库

第一章 绪论

1.1 研究背景及意义

人机交互（HCI）是研究人与计算机之间通过相互理解的交流与通信，在最大程度上为人们完成信息管理、服务和处理等功能，使计算机真正成为人们工作和学习的助手的一门技术科学，它是伴着计算机的诞生就发展起来的^[1]。早先，人机交互需要借助一定的物理设备，如键盘、鼠标、遥控器等来操作计算机，并不符合人们的自然交流习惯。随着计算机视觉、模式识别、人工智能等技术的不断发展，人们开始追求自然和谐的人机交互模式，当前人机交互的研究热点是如何利用人自身的生物特性，使人与计算机可以直接进行自然交互。其中，语音识别、人脸表情识别、触屏识别、手势识别、虹膜识别等等的研究最多。总的来说，人机交互技术已经从以计算机为中心逐步转移到以人为中心，是多种媒体、多种模式的交互技术^[2]。手是人体最为灵活的部位，相较其他交互方式而言，将手势作为人机交互的手段显得更加自然，也更加直接，因此手势识别技术是人机交互的一大研究热点。

从输入方式来看，当前手势识别的研究可分为基于数据手套、基于触摸屏，以及基于视觉等等。相对来说，基于视觉的手势识别无需借助键盘、鼠标等其他输入设备，而是利用摄像头，通过定义合适的手势来控制计算机，交互更加自然。计算机视觉根据输入视频的摄像机数目不同，可分为单目视觉和多目视觉，两者最大的差异在于多目视觉可以多得到一个深度信息，因此多目视觉的建模难度更高，计算量更大，表达的意义也更丰富准确。随着硬件技术的不断发展，摄像头在个人计算机、手机、MP4 等数码产品和家用电器中更加普及。同时价格也不高，作为视频输入设备用到手势识别系统中十分合适，因此基于视觉的手势识别技术被广泛地应用到各个产业中。目前，该技术在智能系统、机器人控制、多媒体教学、娱乐游戏等领域都有很好的应用前景。不过，由于手势本身具有多样性、多义性以及时间和空间上的差异性等特点，加之人手是复杂变形体以及视觉本身的不适定性，所以手势识别成为多学科交叉的研究课题^[3]。

1.2 国内外研究现状

目前,对基于视觉的手势识别的研究已经进行了几十年之久,并取得了许多成果。国外从事该领域研究的机构有麻省理工大学、美国 MIT 媒体实验室、密歇根州立大学计算机系等,而国内的有浙江大学 CAD&CG 国家重点实验室,清华大学人机交互与媒体集成研究所等等。

国外对该领域的研究相对国内更早。Hans Rijkema 等人于 1991 年研究了用于机器人的基于知识的手动作,它以经验为基础观察人类活动,利用获得的知识判断手的动作,并将手的动作应用于人工智能技术^[4]。K Grobel 等人于 1997 年将图像处理技术用到手势识别中,利用隐马尔可夫模型(HMM)对获得的手势特征进行识别,识别率可达 94%^[5]。1999 年,微软韩国的 HK Lee 等人,利用隐马尔可夫的阈值模型,识别出 9 种动态手势命令,平均识别率高达 93.14%^[6]。目前,基于视觉的手势识别研究已经有了许多新的突破,该技术也越来越多的被广泛应用在我们日常生活中。日本东芝公司在 Qosimo 产品线中加入了图片识别和手势识别技术,成功制造出可以进行手势交互的电视机和笔记本^[7]。微软在 2010 年推出的 XBOX 360 体感周边外设 Kinect 是一种 3D 体感摄影机,由于导入了及时动态捕捉、影像识别等功能,因此不需要任何控制器,只需通过摄像头捕捉三维空间中玩家的运动信息,就能对其进行识别并实现控制。该产品的出现使人机互动游戏进入了一个新的纪元^[8]。

在我国,各大高校和研究院对手势识别的研究十分热衷。清华大学的祝远新等人通过对帧间图像进行运动分割估计出运动信息,并将其与颜色、形状等多种信息结合在一起进行手势识别,这是一种新的基于表观的手势识别技术,其识别率超过 90%^[9-11]。王凯等人通过单目摄像头获取手势视频片段,并利用 Adaboost 算法和光流匹配法进行实时手势识别,可对手势做出较为准确的识别^[12]。随着技术的成熟,手势识别产品开始出现在国内市场上。2011 年,腾讯公司推出“QQ 手势达人”,利用手势识别控制 PPT 播放^[13];2012 年康佳公司推出国内首款手势交互式电视机,不久之后,长虹、TCL 等电视厂家,也都相继推出带有手势交互功能的电视机^[14]。

1.3 主要研究内容与特色

本文主要研究复杂环境下裸手的手势识别，并在单目视觉下进行实现。在总结国内外已有研究方案的基础上，本文以单目视觉下的单手手势为研究对象，从图像分割和指尖提取两方面加以改进，通过与鼠标操作的结合来控制 Google 图片管理软件 Picasa，并验证该方法的可行性。

系统采用的输入设备是普通的 USB 摄像头，可以获取图像的颜色信息。利用颜色信息，通过 CamShift 跟踪方法，对人手区域进行快速定位跟踪。结用肤色和运动信息准确地将人手从背景中分割出来。在对比分析多种颜色空间的手势分割效果以后，在 YCbCr 颜色空间中，肤色的聚集程度更高，且将亮度与色度分量分离开来，因此选择 YCbCr 颜色空间进行手势分割。在特征提取与识别方面，由于指尖是人手的重要特征，采用基于曲率的方法进行指尖检测，并采用 K-medoids 聚类算法定位指尖，获取指尖个数及方向信息，计算量小。最后，本系统使用 Visual Studio 2010 及 OpenCV2.4.3 工具来进行系统实现。

本文的创新之处主要有以下两点：

- (1) 提出了一种自适应的手势分割方法，该方法结合了肤色和运动信息，通过利用相邻两帧快速构建肤色模型，减少不同人种人手颜色不同以及光照的影响。利用背景差分方法进行运动检测，提高手势分割的准确性。
- (2) 在基于曲率的指尖检测方法中，对 K-余弦算法提出改进，利用手势轮廓中的弧长、弦长以及掌宽信息，来对 K 值进行自适应计算，从而减小误判漏判率，提高算法准确性。

1.4 本文结构安排

本文的章节安排如下：

第一章介绍了基于手势识别的虚拟输入方法的研究背景及意义，说明其发展现状，指出现有手势识别系统存在的问题，并简要说明本文的研究工作、特色以及结构安排。

第二章对手势识别技术进行综述，介绍了单目与多目视觉下的手势识别方法，并从手势分割、特征提取、手势跟踪、分类识别四个方面进行技术阐述。

第三章是对手势分割方法的研究，介绍了常用肤色模型，提出了一种融合肤色信息和运动信息的手势分割方法，实现了手部区域的完整分割。

第四章是对指尖检测方法的研究，介绍了基于曲率的指尖检测算法，并对K-余弦算法提出改进，提高了多指尖检测的准确率。

第五章是对手势识别技术的研究，介绍了手势跟踪方法，静态手势识别以及动态手势识别，并实现了一个动态手势交互系统。

第六章总结了全文的主要工作，提出了存在的问题，并对未来的发展进行展望。

厦门大学博硕士论文摘要库

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士学位论文摘要库