|  |  |
| --- | --- |
|  |  |
|  |  |

Diagram

Description automatically generated

本论文标题

English Subject

作者姓名

（软件工程师）

公司：安波福电子（苏州）有限公司

部门：AS&UX CAM&TEL TEAM

邮箱：xxx@aptiv.com

导师：导师姓名

日期：2022/10/30

**安波福电子**

**论文原创性声明**

本人郑重声明：所提交的论文是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不含其他个人或集体已经发表或撰写过的研究成果。对本文的研究作出重要贡献的个人和集体，均已在文中以明确方式标明。

论文作者签名： 年 月 日

本论文标题

中文摘要

目前基于大屏幕的人机互动技术受到人们的广泛关注。大屏幕互动领域的人机互动技术大致可以分成两种：基于语音识别的人机互动技术和基于计算机视觉与图像识别的人机互动技术。基于语音识别的人机互动技术经常会受到复杂嘈杂环境的限制，应用范围较小，大多应用在小型的民用人机互动领域。基于计算机视觉实现的人机互动系统主要是借助传感器等设备采集图像或者信号，然后通过人机互动软件对传感器所采集到的图像信号或者是数字信号完成信号预处理操作，再将所需跟踪的目标从背景中分割出来，经由目标跟踪，动作识别等一系列的操作，完成人机互动。与此同时，基于计算机视觉和图像识别技术所设计的人机互动成本相对较低，互动效果良好。鉴于计算机视觉实现的人机互动技术的优良性能，该技术被广泛的应用在大屏幕人机交互领域。

现有的基于大屏幕的人机互动操作系统在设计时主要存在三个问题。一是怎样得到投影仪和摄像头之间精确的映射关系；二是存在复杂区域时，如何准确的分割出动态手势区域，并进行准确的动态手势的定位与跟踪；三是互动传感器采集到的深度图像里的干扰噪声如何有效且快速的滤除。

**关键词：**大屏幕人机互动系统；Kinect传感器；联合滤波算法；手势跟踪；深度图像滤波

**English Subject**

**Abstract**

At present, human-computer interaction technology based on the big screen has drawn more and more people's attention. Large-screen human-computer interaction technologies nowadays are mainly divided into two types: a voice-based human-computer interaction technology and a machine vision-based human-computer interaction technology. Voice-based human-computer interaction technology is often limited by complex and noisy environments, so the applications are relatively small and mostly applied in the field of small-scale human-computer interaction. The human-computer interaction system based on machine vision is mainly based on the sensor and other equipment to capture images or signals, and then the interactive software is employed to capture the image or signal processing, and then the target from the background is segmented. Then a series of operations as target tracking, gesture recognition and so on are applied to complete the corresponding human-computer interaction. Therefore, the cost of human-computer interaction based on machine vision is relatively low, and the interaction effect is well obtained. Therefore, human-computer interaction technology based on machine vision is widely used in large-screen human-computer interface.

There are three major issues about the large screen human-computer interaction system currently. The first is how to get the exact mapping relationship between the projector and the camera. The second is how to accurately segment the dynamic gesture area when there exists complex areas, and locate and track dynamic gestures accurately. The third is how to effectively filter the interference in the depth image collected by the interactive sensor.

**Keywords:** Large-screen human-computer interaction system; Kinect sensor; Federated filtering algorithm; Gesture tracking; Depth image filtering

目 录

[第一章 绪 论 1](#_Toc515267458)

[1.1 人机互动技术介绍 1](#_Toc515267459)

[1.1.1 人机互动技术课题研究背景 1](#_Toc515267460)

[1.1.2 人机互动系统研究意义 1](#_Toc515267461)

[1.2 课题国内外研究现状分析 2](#_Toc515267462)

[1.2.1 大屏幕投影校准技术的研究现状 3](#_Toc515267463)

[1.2.2 手势跟踪算法的研究现状 3](#_Toc515267464)

[1.2.3 图像滤波去噪分析的研究现状 6](#_Toc515267465)

[1.3 课题研究目标及内容 7](#_Toc515267466)

[1.4 本文结构安排 8](#_Toc515267467)

[第二章 基于3D传感器的人机互动系统介绍 10](#_Toc515267468)

[2.1 互动系统框架组成 10](#_Toc515267469)

[2.2 硬件选型与介绍 11](#_Toc515267470)

[2.2.1 硬件选型 11](#_Toc515267471)

[2.2.2 硬件传感器简介 12](#_Toc515267472)

[2.3 本章小结 13](#_Toc515267473)

[第三章 基于几何标定的大屏幕投影校准 14](#_Toc515267474)

[3.1 投影仪摄像头系统的几何映射 14](#_Toc515267475)

[3.2 基于棋盘格校准的Kinect摄像头的标定 15](#_Toc515267476)

[3.2.1 摄像头内参标定 19](#_Toc515267477)

[3.2.2 摄像头外参标定 20](#_Toc515267478)

[3.3 基于射影变换的投影仪摄像头系统校正 22](#_Toc515267479)

[3.4 本章小结 24](#_Toc515267480)

[第四章 基于动态手势跟踪定位的互动算法设计 25](#_Toc515267481)

[4.1 基于canny检测算子的手势轮廓分割 25](#_Toc515267482)

[4.2 基于区域生长填充手势轮廓 27](#_Toc515267483)

[4.3 椭圆手势模型建立 28](#_Toc515267484)

[4.4 基于粒子滤波与均值漂移相结合的动态手势跟踪算法 29](#_Toc515267485)

[4.4.1 粒子滤波与均值漂移(Mean shift)算法介绍 29](#_Toc515267486)

[4.4.2 基于粒子滤波和均值漂移结合的动态手势滤波算法设计 33](#_Toc515267487)

[4.5 动态手势跟踪定位实验验证 35](#_Toc515267488)

[4.5.1 手势轮廓分割实验研究 35](#_Toc515267489)

[4.5.2 基于软件平台的手势位置定位实验 37](#_Toc515267490)

[4.5.3 手势跟踪效果对比 39](#_Toc515267491)

[4.6 本章小结 45](#_Toc515267492)

[第五章 基于联合滤波算法的红外图像去噪 46](#_Toc515267493)

[5.1 红外图像的联合滤波算法设计 46](#_Toc515267494)

[5.2 基于八连通算法的红外笔散斑标记算法设计 49](#_Toc515267495)

[5.3 基于红外图像的抗干扰性能实验研究 50](#_Toc515267496)

[5.4 本章小结 52](#_Toc515267497)

[第六章 任意墙面触控互动系统功能演示及其应用 53](#_Toc515267498)

[6.1 系统初始化过程 53](#_Toc515267499)

[6.2 手势定位控制功能演示 55](#_Toc515267500)

[6.3 红外笔点击控制功能演示 56](#_Toc515267501)

[6.4 应用实例介绍 58](#_Toc515267502)

[6.5 本章小结 60](#_Toc515267503)

[第七章 总结与展望 61](#_Toc515267504)

[7.1 总结 61](#_Toc515267505)

[7.2 展望 62](#_Toc515267506)

[参考文献 64](#_Toc515267507)

[攻读学位期间本人出版或公开发表的论著、论文 70](#_Toc515267508)

[致谢 71](#_Toc515267509)

# 第一章 绪 论

## 1.1 人机互动技术介绍

### [1.1.1 人机互动技术课题研究背景](#_Toc28209)

人机交互技术(Human-Computer Interaction Techniques)是通过相应的输入和输出设备，并且基于软件程序，将计算机和人的操作动作有机的结合起来，从而实现互动交流的相关技术[1]。人机互动技术自从出现以后，就一直作为相对独立、重要的领域而备受关注。随着计算机行业的发展，与人机互动相关的硬件、软件技术也在飞速发展，这些技术的发展也在推动着人机互动技术的进步。与此同时，计算机的操作者也开始从相关领域的专家逐渐扩展到普通的使用者，这样更加扩展了人机互动技术的应用范围，极大了刺激了该领域的迅速发展[2]。

目前，大屏幕人机互动领域主要是基于语音识别的互动技术或者是计算机视觉与图像识别的互动技术实现的。然而在嘈杂的环境下，基于语音识别的互动技术有着诸多限制，因此语音识别的互动技术应用范围较小，且大多应用在小型的民用人机互动领域。基于计算机视觉实现的人机互动技术主要是借助传感器等设备采集图像或者信号，然后通过一系列的图像识别处理操作，完成相应的人机互动。因此基于计算机视觉的人机互动技术受互动环境的限时较小，且互动成本相对较低，互动效果良好[3-4]。鉴于计算机视觉实现的机互动技术的优良性能，该技术成为大屏幕人机交互的相关领域互动技术的首选。

### 1.1.2 人机互动系统研究意义

目前，人机互动技术处于高速发展阶段，人们的日常生活中出现了越来越多的人机互动系统。这些互动技术的输入已经不满足于传统的键盘或者是手柄输入，更多的已经变成手指的动作，声波的振动甚至是使用者的眨眼操作，因此为了实现多种输入方式的智能人机互动操作，人机互动技术的研究迫在眉睫[5]。

然而目前国内人机互动技术的研究还大大落后于国外的同类型研究。其劣势主要在于国内的重点是对互动的界面和互动硬件的研究，而国外已经将研究的重点转移到对互动系统本身原理的研究，并准备颠覆目前传统的互动技术框架，形成一种新型的人机互动模式。本课题研究的主体是基于投影画面的大屏幕人机互动技术。现有的投影系统大屏幕互动技术在实现人机交互时，由于操作者的走动或者是周围环境的影响。往往会出现以下问题：

**I**：如何快速建立投影仪和摄像头间的精确映射关系？

**II**：当存在复杂区域时，如何准确的分割出动态手势区域，并进行精准的动态手势的定位与跟踪？

**III**：传感器采集到的深度图像中的干扰如何有效的滤除？

### [1.2.2 手势跟踪](#_Toc7193)算法的研究现状

手势跟踪算法主要包括3大类方法：基于优化匹配搜索手部的跟踪方法、基于滤波预测的手部跟踪方法、基于轮廓和特征匹配目标的跟踪算法。主要分类如图1.1所示。



图1.1 基于 Kinect手势跟踪的算法分类

当前运用比较普遍的手势跟踪方法是基于轮廓和特征匹配的手部跟踪算法，该方法通过计算动态手势区域轮廓上各个点与其附近邻接点构成的向量的方向及其夹角来判断指尖选取操作点，但是该种方法需要对动态手势区域进行精确的分割，而分割的精确度严重制约着跟踪的精度。文献[13]里提及一种基于 Kinect 传感器进行手势识别的工作系统，该工作系统是由手工位置检测、手指位置与轮廓识别三个部分组成。在该手势识别工作系统里识别步骤所采用手势的轮廓跟踪算法是：a) 扫描互动屏幕上的所有像素点，并设置好手势运动出发点; b) 对被设定为起点的当前像素进行邻域检查，并计算邻域相关参数; c) 顺时针检测像素点，进行循环反复检测，直到检测的像素点数量超出设置的最大值。但是基于轮廓和特征匹配的手部跟踪算法主要有两个缺点：(1) 当所要跟踪的目标与摄像头之间存在遮挡或者跟踪环境背景光线出现变化时，都会对跟踪的效果产生极大的影响。(2) 由于基于特征匹配的算法是将目标模板同许多待匹配的特征点进行匹配的，匹配的时长就限制了跟踪的速度，因此基于匹配的跟踪算法一般实时性都不是很强[13-18]。

基于滤波预测的手部跟踪算法，其实是对运动手势在下一帧图像中可能出现的位置进行预估，然后在动态手势运动区域寻找全局最优点。Tang对于3D 手势的跟踪提出了相应的滤波跟踪算法，该预测算法具有局部最优、预测稳定的优点，同时该方法还兼具了计算量小、计算实时性强，准确把握预测目标的速度与位置等优点[19-20]，然而该种基于3D手势的预测算法仅仅适用于呈高斯分布的线性系统。针对上述情况，Alamsyah提出的粒子滤波系统可以解决该问题[21]。粒子滤波是针对于非线性系统提出的一种滤波算法，从理论上来说，粒子滤波可以达到全局最优的效果。粒子滤波也存在着一定的缺点，例如粒子滤波的跟踪模型使得该滤波算法实时性不高，不能很好的满足工程应用的要求。其次，在粒子滤波中随着迭代次数的增加，还会存在粒子退化的问题。这些都是亟待解决的问题。

基于优化匹配搜索的手势跟踪方法，文献[22]提出利用Kinect传感器的深度摄像头来采集深度图像，通过Mean shift算法来跟踪动态手势对象，然后生成对应的手部轮廓图像，再对这些轮廓图像做好部分轮廓标签，使得每个骨骼关节都会出现在轮廓标记的中心位置[23]; 接着需要手动配置人体骨架的相关参数并形成相应的大型采集数据集，然后训练相关的随机决策树; 再采用Mean shift算法来预估作为操作点动态手势的关节中心，利用上述步骤所产生的框架来实现实时性强的手势位置预估。作为相应的实验验证，美国手语位数识别便是通过该种手部跟踪算法来实现的[23]。相较于其他基于优化匹配搜索的跟踪方法，基于Mean shift滤波算法的优势在于该方法不需要预先知晓特征空间的特点，唯一需要的就是样本数据，然后根据这些样本点进行滤波。因此该种算法计算量较小，实现难度不大，实时性强，比较适合对实时性要求高的系统。但是该算法不能用来处理体积小的目标或者是快速运动中的目标[22]。

# 参考文献

[1] Crespo R G, Gonzalez C. Special issue on human computer interaction[J]. IEEE Latin America Transactions, 2015, 13(2): 399-400.

[2] 杨柳青. 语音人机交互及其在智能调度中的应用[D]：[硕士学位论文]. 济南：山东大学, 2013.

[3] 王修晖, 华炜, 鲍虎军. 面向多投影显示墙的手势交互系统设计与实现[J]. 计算机辅助设计与图形学学报, 2007, 19(03): 318-322.

[4] 梁卓锐. 计算机视觉手势交互的交互映射研究[D]：[博士学位论文]. 广州：华南理工大学, 2015.

[5] Zeng Z, Pantic M, Roisman G I, et al. A survey of affect recognition methods: audio, visual, and spontaneous expressions[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2009, 31(1):39-58.

[6] 许庶. 投影仪画面的自适应几何矫正[D]：[硕士学位论文]. 杭州：浙江大学, 2013.

[7] Zollmann S, Langlotz T, Bimber O. Passive-active geometric calibration for view-dependent projections onto arbitrary surfaces[J]. Journal of Virtual Reality and Broadcasting, 2007, 4(6): 1-11.

[8] Horn B K P. Shape from shading: A method for obtaining the shape of a smooth opaque object from one view[D]. Cambridge, 1970.

[9] Chen C L, Tai C L, Lio Y F. Virtual binocular vision systems to solid model reconstruction[J]. The International Journal of Advanced Manufacturing Technology, 2007, 35(3): 379-384.

[10] Tomasi C, Kanade T. Shape and motion from image streams under orthography: a factorization method[J]. International Journal of Computer Vision, 1992, 9(2): 137-154.

[11] Faugeras O D. What can be seen in three dimensions with an uncalibrated stereo rig[C]. Proc of IEEE Conference on Computer Vision and Pattern Recognition. 1992: 563-578.

[12] Michael K, Wolfgang O, Werner J. Active industrial surface with the inverse projected-fringe-technique[J]. International Society for Optical Engineering, 2001, 4596:37-47.

[13] Li Y. Hand gesture recognition using Kinect[C]. Proc of the 3rd International Conference on Software Engineering ＆ Service Science. 2012, 197-198,

# 致谢

转眼研究生的三年学业生活就快要结束了，本篇毕业论文也接近完成。在此，我要向研究生生涯中给予过我帮助和鼓励的同窗们，老师们，朋友们致以深深的敬意和由衷的感谢。

。。。。