# 鱼眼成像全景漫游系统的研究

# Research of Panoramic Roaming System Based on Fisheye Images

学科专业：控制科学与工程

研 究 生：王志舟

指导教师：张 军 副教授

天津大学电气与自动化工程学院

二零一五年十一月

独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作和取得的研究成果，除了文中特别加以标注和致谢之处外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得 **天津大学** 或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

学位论文作者签名： 签字日期： 年 月 日

学位论文版权使用授权书

本学位论文作者完全了解  **天津大学** 有关保留、使用学位论文的规定。特授权 **天津大学** 可以将学位论文的全部或部分内容编入有关数据库进行检索，并采用影印、缩印或扫描等复制手段保存、汇编以供查阅和借阅。同意学校向国家有关部门或机构送交论文的复印件和磁盘。

（保密的学位论文在解密后适用本授权说明）

学位论文作者签名： 导师签名：

签字日期： 年 月 日 签字日期： 年 月 日

**摘 要**

**ABSTRACT**

**目 录**

[第一章 绪论 1](#_Toc435388227)

[**1.1** 研究背景 1](#_Toc435388228)

[**1.2** 虚拟现实的应用领域 1](#_Toc435388229)

[**1.3** 场景漫游的两种实现方式 2](#_Toc435388230)

[**1.4** 国外研究现状 3](#_Toc435388231)

[**1.5** 国内研究现状 3](#_Toc435388232)

[**1.6** 论文的结构安排 4](#_Toc435388233)

[第二章 基于鱼眼图像的全景漫游相关技术介绍 5](#_Toc435388234)

[**2.1** 鱼眼镜头的成像原理 5](#_Toc435388235)

[**2.2** 鱼眼图片的种类 6](#_Toc435388236)

[**2.3** 鱼眼图像的畸变 6](#_Toc435388237)

[**2.4** 畸变校正的方法 7](#_Toc435388238)

[**2.5** 摄像机镜头模型 8](#_Toc435388239)

[**2.6** 校正图像的全景拼接 10](#_Toc435388240)

[2.6.1 图像的配准方法 11](#_Toc435388241)

[2.6.2 图像的融合方法 12](#_Toc435388242)

[2.6.2.1 加权系数法 12](#_Toc435388243)

[2.6.2.2 Toet算法 12](#_Toc435388244)

[2.6.2.3 多分辨率样条法 13](#_Toc435388245)

[2.6.2.4 颜色空间变换融合算法 13](#_Toc435388246)

[**2.7** 全景图跟传统图像的区别 13](#_Toc435388247)

[**2.8** 全景浏览模型 14](#_Toc435388248)

[**2.9** 本文的实验环境 15](#_Toc435388249)

[第三章 鱼眼图像有效区域的提取 16](#_Toc435388250)

[**3.1** 各种提取方法的比较 16](#_Toc435388251)

[3.1.1 面积统计法 16](#_Toc435388252)

[3.1.2 霍夫圆变换法 16](#_Toc435388253)

[3.1.3 线扫描法 17](#_Toc435388254)

[3.1.4 三种方法的处理结果对比 17](#_Toc435388255)

[**3.2** 改进的线扫描法——变角度线扫描法 17](#_Toc435388256)

[第四章 鱼眼图像的校正 22](#_Toc435388257)

[**4.1** 畸变图像到视球面的转化 22](#_Toc435388258)

[**4.2** 视球面向平面的转化——纵向压缩柱面投影校正 24](#_Toc435388259)

[第五章 多幅鱼眼图像的全景拼接 27](#_Toc435388260)

[**5.1** 图像拼接的基本流程 27](#_Toc435388261)

[**5.2** 图像的拼接 28](#_Toc435388262)

[5.2.1 图像配准 28](#_Toc435388263)

[5.2.1.1 相似度阈值筛选 29](#_Toc435388264)

[5.2.1.2 特征点的对称规则筛选 29](#_Toc435388265)

[**5.3** 图像的融合 31](#_Toc435388266)

[**5.4** 生成全景图像 32](#_Toc435388267)

[第六章 全景漫游的实现 33](#_Toc435388268)

[**6.1** 全景图像的存储方式 33](#_Toc435388269)

[**6.2** 单位球面到视平面的重投影 34](#_Toc435388270)

[**6.3** 视平面场景的运动变化调整 36](#_Toc435388271)

[**6.4** 算法性能的讨论 37](#_Toc435388272)

[**6.5** 全景漫游的实验结果 37](#_Toc435388273)

[第七章 总结与展望 39](#_Toc435388274)

[**7.1** 总结 39](#_Toc435388275)

[**7.2** 进一步的工作 39](#_Toc435388276)

[参考文献 40](#_Toc435388277)

[发表论文和参加科研情况说明 41](#_Toc435388278)

[致谢 42](#_Toc435388279)

## 绪论

本章首先阐述课题的研究背景，以及虚拟现实的应用领域，说明虚拟现实发展的巨大潜力。同时对比国内外的研究现状，导出场景漫游的具体研究方向，并介绍了实现场景漫游的两种主要方式。最后介绍本文的研究内容以及文章的组织结构。

### 研究背景

本课题来源于项目：智能交通监控——黄色网格线内违章停车检查。得到天津市创新基金(13ZXCXGX40400, 13ZXCXGX40500)、天津市小巨人创新基金(2012-XJR21017)资助。

虚拟现实技术涉及的领域比较宽广，包括人工智能、传感器技术、计算机图形学、多媒体技术以及网络技术等众多方面，是这些技术领域的集成。它通过人工仿造或创造媒体空间，产生虚拟场景，能够通过多种媒体传感交互设备给人以身临其境的感觉。

当前，使用图像绘制的虚拟全景技术能以较低的制作成本获得一定程度的虚拟现实效果，并且在立体物体展示、虚拟场景构造、产品广告、旅游景区、房产销售广告、三维游戏等方向都得到了应用。值得一提的是，百度地图的全景浏览功能就是一个很好的例子，它不仅能够查到某个地理位置，还能够进一步的查看该地点的真实环境，取得了很好的虚拟现实效果。

要想在虚拟场景中获得身临其境的感觉，虚拟现实系统的实时性和画面质量是衡量虚拟场景优劣的两个重要指标。在有限的系统计算能力下，这两个指标往往是一对矛盾体，需要在两者之间进行权衡来满足一定的性能要求。

可以肯定的是，虚拟现实的潜在市场非常大，一旦技术上取得相应突破，将会对人类的交互方式产生深远的影响。

### 虚拟现实的应用领域

虚拟现实对医学、娱乐、军事航天、室内设计、房产开发、工业仿真等领域都有十分重要的现实意义。

医学上，通过虚拟环境可以建立虚拟人体模型，借助于跟踪球、头戴式可视设备(HMD)、感觉手套，学生可以很容易的去学习和了解人体内部的各种器官结构，也可以进行模拟尸体解剖和各种手术练习，从而不受标本、场地等条件的限制。它还可以帮助医生寻找最佳手术方案，对于新药物的研制等方面也能够发挥十分重要的作用。

室内设计方面，虚拟现实可以成为设计工具，以视觉形式反映设计者的思想，形象直观，避免复杂的设计图纸只有内行人才能看懂，设计者完全可以按照自己的构思创建虚拟房间并直接展现出来，既节约了时间，又节省了模型制作成本，还提高了效率。

房产开发中，传统的展示手段如平面图、表现图、沙盘、样板房等已经远远满足不了消费者的需要。国内已经出现了利用虚拟实景技术进行房产销售的尝试。

军事航天中，模拟训练一直是军事与航天中的重要一环，这为虚拟现实提供了广阔的应用前景。利用虚拟现实技术可以模拟零重力环境，替代非标准的水下训练宇航员的方法。

虚拟现实技术与网络技术结合，可以将文物展示、旅游业带向一个崭新的发展阶段。它可以将文物实体的影像数据或旅游景区的场景数据采集并建立起三维模型数据库，通过网络以一种全面、生动、逼真的形式展示给终端用户，摆脱了地域限制，实现更为广泛的资源共享。

三维游戏也是虚拟现实技术的重要应用方向，并且在该领域也一直非常活跃，因为其实时性、交互性与场景真实性深深的震撼着玩家的心灵，随着软硬件技术的不断发展，不远的将来，真正意义的虚拟现实游戏必将为人类娱乐、教育和经济发展做出新的贡献。

### 场景漫游的两种实现方式

目前的场景漫游技术主要分为：基于几何模型的漫游技术和基于图像绘制的漫游技术。

基于几何模型的漫游技术，首先要建立场景的三维模型，使用曲线、多边形、曲面等几何元素构建精确、完整的场景模型，显示时还要进行如纹理映射、着色以及光影明暗处理等复杂操作。所以使用几何模型的方法对系统的计算能力有要求，如果场景复杂程度较高的话，实时性的要求必然会使系统很有压力，这时特殊的硬件设备就不得不用来解决这个问题了，所以系统的成本也相应的增大。除了这些负面因素，几何模型的方法对比图像绘制的方法有着显著的优势，它可以构造预想中的场景，并且能够随意控制浏览的路线和速度。一些大型3D游戏如：刺客信条、使命召唤，已经在游戏场景运用这种几何模型的方法，自由度非常高。

而基于图像绘制的技术则对成本控制非常有利，它技术实现相对简单，也是近十年来兴起的热门技术，与几何模型的方法不同，它从一开始就是从现有的场景图像出发来构建不同视点的图像，优点是建模容易、真实感强、绘制快，成本显著降低，适用于复杂场景的虚拟浏览，它不需要大量的计算来渲染场景的每一部分，但缺点是数据的完整性、精确性不及几何模型的方法。

基于图像的漫游方法在绘制速度和真实感上的明显优势，在计算机领域引起了浓厚兴趣，吸引了一大批的专家学者对它进行研究，并取得了一批有价值的成果。

### 国外研究现状

美国是虚拟现实技术的发源地，目前美国在该领域的基础研究主要集中在感知、用户界面、后台软件和硬件四个方面，美国宇航局的AMEN实验室完善了HMD（头戴式可视设备），使其成为可用性较高的产品。NASA研究的重点放在对空间站操纵的实时仿真上。

麻省理工学院建立了一个虚拟环境下的对象运动跟踪动态系统。

苹果公司推出的全景视频产品，第一次使人们领略了具有照片质量的虚拟现实环境，它利用软件把相同视点的若干张边缘稍有重叠的照片缝合在一起，组成一张360°全景图像，通过热点将不同视频的全景图进行连接，通过在全景图上开窗，实时生成用户所要观察的场景，经不同全景图的切换来实现虚拟环境中的漫游。

斯坦福研究所建立了“视觉感知计划”，研究现有虚拟现实技术的进一步发展，还利用遥感技术进行外科手术仿真。

日本对虚拟现实技术的研究居于世界领先地位，在虚拟现实游戏方面研究颇深，目前主要致力于建立大规模虚拟现实知识库。

富士通实验室正在研究虚拟生物与虚拟现实环境的相互作用。

### 国内研究现状

我国虚拟现实技术起步较晚，但也有了相当不错的发展，得到了国家的支持和重视。

北京航空航天大学计算机系是国内最早进行虚拟现实技术研究的，主要研究了虚拟环境中物体物理特性的表示与处理，在虚拟现实中的视觉接口方面开发出了部分硬件，实现了分布式虚拟环境网络设计。

浙江大学 CAD＆CG国家重点实验室开发出了一套桌面型虚拟建筑实时漫游系统以及相应的漫游算法。

哈工大计算机系已经成功地虚拟出人的高级行为中特定人脸图像的合成，在表情的合成和唇动合成上取得成果，现在研究人说话时头势和手势动作。

西安交通大学信息工程研究所对虚拟现实中的关键技术——立体显示技术进行了研究。

北方工业大学VAD研究中心完成了体视动画的自动生成部分算法与合成处理软件。

国防科技大学的虚拟实景空间系统是目前国内最完整的一个系统。它还实现了同心球拼图的室内三维漫游系统。

中国科学院计算机研究所实现了一个基于球面全景图固定视点的室内漫游系统。

西北工业大学CAD＆CAM研究中心，上海交通大学图像处理及模式识别研究所，华东船舶工业学院计算机系，安徽大学电子工程与信息科学系也在虚拟现实技术方面进行了一些相关的研究工作和尝试。

国内全景漫游方面在不断的发展和逐步成熟，但由于一些技术原因和人的观念以及互联网环境的惯性，在推广方面还要作一些努力。

### 论文的结构安排

本文研究了基于鱼眼图像校正并绘制的虚拟场景漫游相关技术，文章组织结构如下：

第一章引入课题研究的背景。虚拟现实技术在各个领域的突出应用、场景漫游的两种不同实现方式。同时还介绍了研究课题在国内和国外的不同研究现状等。

第二章对鱼眼图像全景漫游的相关技术进行介绍。包括鱼眼镜头的成像原理、鱼眼图片的种类和畸变类型、鱼眼图片的校正方法以及摄像机镜头的相关数学模型，并对图像的拼接与融合技术作了简单介绍，最后介绍了浏览全景图的几种方式。

第三章具体分析了鱼眼镜头的原理以及鱼眼图像畸变产生的原因，在此基础之上通过对各种区域提取方法的比较（如面积统计法、霍夫圆变换法、传统线扫描法），提出了一种改进的鱼眼畸变图像圆形区域提取方法——变角度线扫描法，它可以更灵活、更准确的提取出畸变图像的圆形有效区域。

第四章通过实验阐述了鱼眼畸变图像的校正方法，经过对经纬映射校正法以及柱面校正法的分析与比较，在柱面映射校正法的基础上提出了纵向压缩的柱面投影校正方法，该方法参数可调节，有一定的灵活性。并对两种方法的处理结果进行了实验对比。

第五章重点阐述了多幅校正后的鱼眼图像进行拼接融合的原理与实验结果，生成了鱼眼图像的球面全景图。

第六章说明了对鱼眼球面全景图进行浏览的方法和思路，并实际的制作了全景图浏览器，达到了对全景图进行浏览的目的。

第七章对文章的整体内容作了梳理和总结，并表达了对鱼眼全景浏览的未来的看法。

## 基于鱼眼图像的全景漫游相关技术介绍

### 鱼眼镜头的成像原理

鱼眼镜头是一种短焦距广角镜头，它的视角接近或等于180°，所能观察的范围远大于人眼的视角范围，因其结构和工作原理与鱼眼相似而得名。鱼眼镜头之所以能够拍摄较大范围的场景，是因为它的镜头采用了类球形结构，这有别于一般的镜头。近年来，鱼眼镜头的应用也渐渐多了起来，它在视频会议、全景图像拼接、虚拟漫游、机器人导航以及大范围公共场所的安防监控等领域都扮演着重要角色，通常一个鱼眼镜头可以代替多个安装角度不同的普通镜头，并且也不需要单镜头全景监控相机的那一套复杂的机械旋转结构，理论上两个鱼眼镜头就可以实现全景成像。但是由于鱼眼镜头的大视角是以牺牲图像直观性为代价的，所以用鱼眼镜头拍摄的图片都存在着很大程度的畸变，会使人眼观察感到一些不适应，最显著的就是“桶形畸变”。虽然这样的畸变给摄影爱好者们带来了惊喜，使他们创造出一些别致的作品，但对于鱼眼镜头在其它方面（如监控）的应用，这样的畸变或多或少也带来了一些麻烦。因此，鱼眼图像的校正技术就应运而生了。

通常我们使用球面透视成像模型来描述鱼眼镜头的成像过程。如图2.1所示，鱼眼镜头的表面可以看成是一个球面，外界景物的光线投射到球面上，形成一幅外界景物球面映射图，景物光线通过光心O之后，伴随光的折射，以不同于入射光线直射的方向投射于相机的感光部件（CCD图像传感器阵列），从而形成一幅鱼眼图像。

鱼眼成像原理图

图2.1 鱼眼镜头成像原理以及图像坐标系说明

### 鱼眼图片的种类

由于CCD图像传感器的感光部分一般设计为矩形，而鱼眼镜头成像区域为圆形，所以随着镜头焦距的不同，成像区域大小与传感器大小的不同配合就会形成如下三种图像形式，分别是：圆形鱼眼图、全帧鱼眼图和鼓形鱼眼图。对于不同类型的鱼眼图像，生成全景图所需要的拍摄技术复杂度和图片张数有所不同。对于圆形鱼眼图像来说，由于它在各个方向上都能达到视域180º，所以仅需要水平拍摄两到三张就可用于全景图生成，并且拍摄技术简单。全帧鱼眼图需要水平拍摄一圈四张外加朝天朝地两张共六张图片才能生成全景图，操作很复杂。至于鼓形鱼眼图，由于它在垂直方向上可以达到视域180º，所以需要水平一圈拍摄四张图像才能生成一张全景图。本文主要以圆形鱼眼图像作为研究的主要内容。

|  |  |  |
| --- | --- | --- |
|  |  |  |
| 圆形鱼眼图 | 全帧鱼眼图 | 鼓形鱼眼图 |

图2.2 常见的三种不同形式的鱼眼图像

### 鱼眼图像的畸变

鱼眼镜头所产生的图像畸变类型大概可以分为三种：

径向畸变(Radial Distortion)，即从圆心开始各个放射方向上，在图像点到圆心的距离上所产生的偏差。

偏心畸变(Decentering Distortion)，图像的中心不在鱼眼镜头的光轴上所引发的图像畸变。

切向畸变(Tangential Distortion)，从圆心到图像点的向量在切线方向发生的偏离，也就是向量的角度偏差引起的畸变。

在以上的三类畸变中，径向畸变是最主要的因素，这种畸变是以圆心对称分布在图像上的，在广角镜头中普遍存在，图像上的点沿半径方向偏离理想位置，如果与理想位置作比较，畸变点如果沿半径方向偏向圆心，这种畸变称为枕形畸变；如果畸变点沿半径方向偏离圆心，这种畸变称为桶形畸变。鱼眼或广角镜头会产生桶形畸变，一般情况下的偏心畸变和切向畸变比较轻微，可以忽略不计。

|  |  |  |
| --- | --- | --- |
|  |  |  |
| 标准图 | 枕形畸变 | 桶形畸变 |

图2.3 径向畸变的两种情形

### 畸变校正的方法

针对存在畸变的鱼眼图片，为了消除这些畸变，许多学者都对其作过不同程度的研究，并取得了一些令人满意的效果，不过这些校正方法都有一定的应用场景，并不能适用于所有的畸变情况，有一部分还需要视图片内容、拍摄方法以及图片场景的限制。

Zorin和Barr提出了一种局部映射校正方法，这种方法可以保证场景中的直线和物体的形状在校正后的图像中依然保持直线和本来的形状，他们使用的单参数映射函数簇可以通过插补法在线性映射法和立体映射法之间对直线和物体形状特性的保持进行权衡。最近，Zelnik-Manor等人通过拼接的方法把多个线性映射场景组合成为一个多平面融合的整体，具体思路是通过把鱼眼图像分割成不同的块，对这些小块进行局部图像校正，然后再把校正结果拼接回一张平面图，拼接图片的边缘是通过人工调整的方法使过渡尽量平滑，例如图片中两面墙之间的接合处的处理。Kopf等人在Zelnik-Manor等人工作的基础之上作了一些简化操作，他们不是校正所有区域，而是对那些畸变比较严重的区域进行校正，这样做需要在图像的重构方面做一些很复杂的处理工作。Robert Carroll等人则使用了最小平方优化的思想，试图保持图像的局部共形映射，它们的方法并不保证空间直线在图像上一定是直线，需要涉及大量的用户干涉操作。Shah提出了一种鱼眼镜头摄像机标定方法，它简单实用，但这种方法标定的摄像机内部参数个数较少，不够精确。Xiong提出了另一种鱼眼镜头的标定方法，只考虑了径向畸变，没有考虑切向畸变的因素，所以得到的标定参数精度较低。英向华和胡占义提出的基于球面透视投影约束的标定校正方法，以空间直线在鱼眼球面投影为大圆作为约束条件，建立了鱼眼镜头的多参数校正模型，将空间直线在鱼眼投影曲线上的点映射为球面点，通过球面到大圆的球面距离最小来进行模型参数的优化，从而利用优化后的模型对鱼眼图像进行校正，这些空间直线投影在鱼眼图像上的点需要人为的选择出来，才能进行标定相关的操作。

从上面学者的研究内容来看，鱼眼图像的校正方法大致分为两种方式：通过与实体摄像机密切相关的优化参数标定方法进行校正以及通过已知的鱼眼映射设计模型进行相关的校正工作。标定的方法可以相对精确的还原畸变的鱼眼图像，一般需要实体摄像机的配合，因为涉及到参数优化过程，且鱼眼镜头的模型相对比较复杂，所以优化的过程耗时较多。而使用已知鱼眼镜头模型的方法优点是可以较方便快捷的进行校正处理，通用性更强一些，在对精确度要求不太高的情况下是一种较好的选择，同时，鱼眼镜头在生产的过程中就是按着一定的映射关系进行设计的，所以模型比较标准，缺点是还原的精确度与生产制造的准确程度相关。本文中使用已知的鱼眼图像映射模型进行相关研究。

### 摄像机镜头模型

数字图像是数码相机把通过镜头的外界光线用图像传感器捕获后形成的。把三维场景映射成为二维平面图像这本身就说明场景和图像之间存在的联系。我们使用模型的方式来从数学上描述这种对应关系。



图2.4 镜头成像模型

根据透镜成像原理，我们知道有下面的公式是成立的：



在计算机视觉中，相机模型可以有多种形式的简化处理，首先我们可以忽略镜头的形状，只把它看成一个极微小的孔，只考虑那些通过镜头光心的光线在理论上不会对所成的像有任何影响。因为大多数情况下物距都远远大于像距，即*do*>>*di*，我们可以假设像平面就处于镜头的焦距处( *di=f* )。从几何的角度来看，物体的像与它本身是颠倒的，我们可以把它的像上下翻转一下，并且放在镜头前方的焦距处，从而简化成像模型。很明显，这种简化方式在物理上是说不通的，但从数学的角度来讲，它们是完全等价的。这就是我们常说的小孔成像模型，如图2.5(a)所示，此时有关系：



|  |  |
| --- | --- |
|  |  |
| (a)小孔成像模型 | (b)三维空间成像模型 |

图2.5 二维和三维成像模型

有了摄像机的针孔成像模型，我们就可以进一步来证明三维点坐标(*X, Y, Z*)与二维图像坐标(*x, y*)在指定像素坐标系下的关系了。图2.5(b)中三维坐标系*Y*轴向下是为了和图像坐标系一致，由公式知，三维坐标一点将会被映射为图像平面中的点(*fX/Z*, *fY/Z*)，如果我们想把三维坐标转化为以像素为单位的二维图像坐标，就需要把坐标除以现实中像素的宽度(*px*)和高度(*py*)，我们发现，如果用焦距分别除以像素的宽度和高度就获得了焦距分别用水平像素单位和垂直像素单位的表达，即*fx=f/px*, *fy=f/py*，这样就可以把完整的映射关系表达如下：



上式可以被重写为矩阵的形式。通过使用齐次坐标，可以把二维坐标表示为三维，三维坐标表示为四维，令*s*为缩放因子，上式可以表示为：



上式第二个矩阵是一个简单的映射矩阵，第一个矩阵包括了一个摄像机的所有参数，称为内部参数矩阵，并且内部参数矩阵对于同一个摄像机来说总是不变的。更加一般化的情况是，图2.5(b)中的参考平面中心不在三维坐标的*Z*轴上时，我们需要考虑它们之间的旋转和平移的表达，这种描述旋转和平移的元素组成了一个矩阵，被叫作外部参数矩阵，外部参数矩阵通常随观察物体相对于相机位置的不同而变化。考虑到这些，摄像机的一般化模型可以表示为：



许多使用摄像机标定技术的图像校正方法都是基于式来进行的，他们使用事先做好的标定板来获取一些三维坐标与二维图像坐标的已知对应关系，然后利用优化理论对摄像机模型的参数进行求解，从而获得摄像机的内部和外部参数，以用于之后的校正过程当中。

上述是摄像机的一般模型，在镜头生产中往往按照事先设计好的模型进行制造。常见的这类模型有：透视投影模型(perspective projection)、立体投影模型(stereographic projection)、等距投影(equidistant projection)、正弦投影(sine-law projection)以及等立体角投影(equi-solid angle projection)。假设视球面是一个单位球面，中心是原点，向量*C*=(0,0,1)表征镜头凝视的方向，即*Z*轴正方向，那么视球面上一点*P*(*X, Y, Z*)就可以用两个角度确定下来，就是我们常用的球面参数化坐标。如0(a)所示，, , 。像机会把点*P*映射到二维平面距*x*轴正方向角度，距离原点的平面点上。长度和角度的不同关系形成了不同的投影模型，对应于之前所说的五种，其中*k*是比例因子：



|  |  |
| --- | --- |
|  |  |
| (a)球面参数化坐标 | (b) *xy*平面的像素映射关系(*z*轴方向的俯视图) |

图2.6 球面参数化坐标与相关映射关系

理论上窄视角镜头可以用透视投影模型来描述，一些广角或超广角镜头可以用等距投影、正弦投影以及等立体角投影模型来描述。通过这些模型的学习可以为畸变图像的校正提供可靠的依据。

### 校正图像的全景拼接

对于不同的应用情况，所需要生成的全景图是不同的，从而需要使用不同的全景图生成方法，相应的拼接过程也各不相同。但总体上来说，全景图的拼接过程都可以按照四个步骤来进行：获得图像帧序列、建立投影模型、对图像帧序列进行配准以及对处理结果进行图像融合处理使拼接处尽量过渡自然、平滑。

#### 图像的配准方法

图像拼接中的关键是图像的配准技术，相邻的两幅待处理图像总是存在一定程度的重叠，这是图像拼接的基本要求，配准处理所做的工作就是要将这些重叠区域尽可能识别出来，并根据这些重叠信息估计出两幅图片间的转换关系，从而把两幅图像映射到同一模型中，实现图像的拼接。

图像的配准技术大致包括：基于图像区域的配准、基于图像特征的配准以及基于相位的配准。

基于图像区域的配准方法也叫作模板匹配法，应用于存在简单变换关系的图像拼接中，例如简单的平移。模板的选取可以是图像中的一列、一块指定大小的窗口或者整个图像区域。不同的选择需要在配准精度和计算速度上作权衡。它在拼接技术发展的早期被广泛应用，实现比较简单，但是存在一定的局限：如果待拼接的图像上除了水平或垂直的的运动之外，还包括一些不规则形状的变形时，两幅图像就很难被配准，因为所选择的模板通常是矩形的，它在另一幅图中的形状很可能就不是一个矩形区域了，此时进行的配准误差就会很大。这些缺点影响了这个方法的应用广度。



图2.7 区域配准示意图

基于图像特征的配准方法需要先从待配准图像上获取特征信息，然后通过对获得的特征进行筛选，选出那些在两幅图像重叠区域中对应的特征信息，从而实现对图像的配准。图像的特征可以是点、线（边）或者对象轮廓，相应的有很多已有的特征检测方法可以使用。特征点的检测方法有：Harris角点检测、Shi-Tomasi角点检测、SIFT特征点检测以及SURF特征点检测。Sobel、Robert、Prewitt、Laplacian、Canny等算子可用于特征线（边）的检测。霍夫变换法可以用于规则对象轮廓检测，边界链码以及一些边界检测算法也可用于不规则对象的轮廓提取。在图像的拼接中，使用最多的是利用特征点来进行图像的配准，因为通过它可以建立任何转换关系之间的图像拼接，适用范围比较广，计算量也可以较好的控制，是拼接技术发展中的主要方向，但它对特征不明显的图像拼接效果不理想，例如天空和海洋的场景。

基于相位的配准方法多用于航空照片和卫星遥感图像的配准处理。最早由Kuglin和Hines提出，该方法把待拼接的图像通过傅立叶变换在频域中计算互功率谱，从而得出图像间的平移量。对于图像间存在旋转的情况，它先把图像坐标变换成极坐标，这样就可以把旋转量转换成为位移量来进行相同的处理。此方法的优点是简单，精确度高，缺点是配准图像需要有较大比例的重叠区域才能不至于造成太大的偏差，对于图像之间存在仿射变换和投影变换的情况，处理效果不理想。

#### 图像的融合方法

图像在拼接之后，一般会在边界区域出现过渡不自然的可见缝合线，这可能是由于拍摄时的场景光照明暗不同、对比度和灰度值不一致造成的。这时就需要对拼接图像的重叠部分进行融合处理。融合的方法有很多种，也相对比较成熟。常用的方法主要分为：基于灰度的融合、基于颜色空间变换的融合以及基于变换域的融合算法。

##### 加权系数法

一种最简单的图像融合算法。两幅图像对应像素点分别以不同的权重相加后得到融合图像。设加权系数为，那么融合图像的过程可以用下式表示：



如果令加权系数随着图像坐标不同按一定的规律变化，例如从1逐渐变化为0，那么融合图像相当于从图像*I*1逐渐过渡到图像*I*2，从而实现了融合图像的平滑过渡，消除了拼接痕迹，该方法能有效的去除边界缝隙，实质上是一种平滑效果，有可能使拼合处出现模糊现象。

基于兴趣区域（ROI）的图像融合法是将加权系数法应用到拼合图像的重叠区域中而不是对整幅图像处理，减少了计算量的同时也能很好的消除边界缝隙。

##### Toet算法

需要先获取拼接图像的共同成份，这可以通过计算灰度较小值来获得：



然后从拼接的两幅灰度图像中分别减去共同成份，得到两幅图像的特征成份：



从图像*I*1和*I*2中分别去掉另一幅图的特征成份：



使用两幅图像的不同部分作为背景成份：



融合处理表达为下式：



其中，加权系数满足关系：。

##### 多分辨率样条法

该方法是比较好的融合方法，用它拼合的图像具有清晰无缝衔接的特点，避免了重影问题，一定程度上还可以处理轻微运行物体图像的拼接工作，运算量较大。这是由Burt Adelson提出的，利用输入图像构建图像金字塔，将原始图像在频域上分成不同频段的带通层，分别对这些带通层进行拼接处理，最终将拼接好的各带通层还原为拼接好的融合图像。这样的分解组合过程是一个无损可逆过程，所以可以有效的完成图像融合的任务。

##### 颜色空间变换融合算法

将颜色空间RGB模型与其它颜色模型（YIQ、ISH、Lab）进行转换，在转换后的颜色模型空间中执行图像的融合操作，然后反变换回RGB颜色空间完成融合过程。这种方法是借助于不同颜色模型空间的显示优势，突出重叠区域的像素特征，从而间接的提高融合的效果。

基于变换域的图像融合算法其实包括了颜色空间变换的方法，它是更广义的名称。不同的变换会对融合过程有不同的简化能力，同时在不同变换空间中，融合所遵循的各种规则的复杂程度不同，选择一种好的变换可以相当程度的降低融合的难度。一些学者使用过wavelet、Ridgelet、Curvelet、Contourlet等等不同的变换进行图像融合的研究，取得了一些令人振奋的研究成果，它们都在各自独特的领域中被广泛的应用着。

### 全景图跟传统图像的区别

全景技术本质上是基于图像的虚拟现实技术，它具有3D效果，可以让用户在浏览全景时具有身临其境的感觉；全景图像通过特殊的全景播放器来播放浏览，可以上下左右放大缩小来任意观看，交互性强。

全景跟二维图像的比较：全景图能捕捉到更多的图像信息，能自由观看任意角度，交互性强，更具真实感和沉浸感。

全景跟视频比较：文件占存储空间更小，更方便于网络应用，生成容易，制作成本低，交互性更强。

全景与flash做出来的环视相比：交互性强，操控方便。球型全景上下视角不受限制，通过播放器的透视处理，使图像沉浸感更强烈，更真实，flash环视没有经过透视处理，沉浸感和真实感差。全景图的生成简单，制作成本更低，对制作人员技术要求更低。

### 全景浏览模型

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

图2.8 三种不同形式的全景图及浏览方式

球形全景图是与人眼模型最接近的一种全景描述，人们在观察周围的事物时是通过视网膜获取图像信息并通过透视变换投影到眼球部分的，以视点为中心的单位视球面就是最自然的全景模型。为了便于计算机存储和表示，球形全景图将球面的经度和纬度坐标直接映射为水平和垂直坐标的一点，所形成图像的高度大约是宽度的两倍。因此从赤道到两极，横向拉伸不断加剧，南北两个极点被拉伸成了扁平状分布在整个图像的上下部边缘。球形全景图可以实现整个水平和竖直的360º全景浏览。

柱形全景图类似球形全景图，只是随着目标接近南北两极，纵向也会拉伸，两极会发生无限的纵向拉伸（因此它的顶部和底部没有水平线）。由于这个原因，柱面投影不太适合具有非常大的垂直视角的图像。柱面投影是传统摆动镜头全景胶片相机所提供的标准投影方式。它对于目标尺寸的保持比直线投影更准确，然而这样就将平行于观测者视线的直线渲染成了曲线。柱面可以方便的展开为一个矩形平面，像素按列均匀分布，这种模型可以简化全景浏览器的设计，数据的采集和生成全景图的过程相对其它两种较容易，适用于模拟摄像机水平绕轴旋转的全景浏览方式，垂直方向上的浏览范围受到限制。

立方体全景图是将全景图分成了前后左右上下六个面，浏览的时候将六个面结合成一个密闭空间来现实整个水平和竖直的360º全景。立方体全景图主要用于场景的全方位展示，它的获取过程比较复杂，不仅需要具有复杂的摄像技术，在浏览器的设计过程中，立方体的角点处的视线也需要特殊的处理。它的数据存储比较方便，屏幕显示时对应的重采样区域边界多为规则多边形，有利于显示，主要应用于虚拟展示。

### 本文的实验环境

本文研究过程中使用的台式机处理器为Intel® Core™ i3-2120 3.30GHz主频，安装内存为4.00GB，安装了Windows10\_x64操作系统。使用Microsoft Visual Studio 2015作为开发环境IDE，使用C++语言及OpenCV2.4.8图像处理库作为处理工具，编译了OpenCV在Qt环境下的GUI工具库，后文中的全部实验内容包括全景浏览器的设计实现都是在这个GUI工具库下完成的，使用了Git版本控制工具，并且将源代码发布到了GitHub中，仓库地址为：

https://github.com/wangzhizhou2014GitHub/Fish-eye-Image-Correction-Code。

## 鱼眼图像有效区域的提取

想要校正一幅鱼眼图像，首先要做的工作便是对圆形有效成像区域的提取（这里主要是针对圆心坐标和半径的参数求取），有效区域提取的准确与否直接关系到图像校正结果的好坏以及之后的图像拼接能否成功，常用到的方法有：面积统计法、霍夫圆变换法、线扫描法。

### 各种提取方法的比较

#### 面积统计法

利用求均匀平面图形的形心（重心）公式来完成圆形区域参数的提取。原始图像经过灰度变换后，再选定一个阈值进行灰度图像的二值化，从而形成了圆形区域的平面图形，在其上应用求形心的公式，就可计算出圆形有效区域的圆心坐标和半径。

计算过程如下：

 ，其中 

上面形心计算公式要应用在数字图像处理中就要做离散化变形，如下：



上式中是二值图中所有白色像素点的总个数，和分别是二值图中每一个白色像素点的横纵坐标，这里的坐标系使用图像坐标系，坐标单位为像素，原点在图像的左上角，向右和向下分别是横纵轴的正方向，以图像坐标表示，即为圆心，半径为：。

#### 霍夫圆变换法

霍夫变换法是图像处理中识别几何形状的基本方法之一，应用范围很广，也存在很多改进算法和变体。最常使用的有霍夫直线检测和霍夫圆检测，这也是图像处理基础理论知识。这里使用霍夫圆变换的思路是：将原始图像经灰度变换后，利用Canny边界检测算子进行图像二值化并将圆形有效区域的轮廓提取出来，之后利用霍夫圆变换得出圆形有效区域的圆心和半径参数。

#### 线扫描法

原理很直观，它的思路是将原始图像灰度化，然后在灰度图上，分别使用水平和垂直两对直线，从上下左右四个方向夹逼圆形有效区域，获得四个切点的坐标，之后利用这两对坐标，计算出圆心坐标和半径大小。在扫描线移动的过程中，只要同一扫描线上的像素最大亮度值与最小亮度值之差大于设定的阈值，便认为该扫描线已经与圆形有效区域相切，最大亮度值的像素点即为切点。

假设两对切点的图像坐标值分别为：

、、、

则圆心和半径的图像坐标值可以这样得到：



#### 三种方法的处理结果对比

实验中，我们发现以上三种方法对于圆形区域内部和边界暗点较少的图片处理结果较好，但对于圆形有效区域边界较暗的鱼眼图像存在不同程度的提取误差。其中，霍夫圆变换法对圆形区域轮廓不明显的图像给出的结果较差（如图3.2样例1第3列），线扫描法对圆形区域边界存在噪声点的鱼眼图像存在提取误差（如图3.2样例2第四列，由于图像顶部阳光太强，导致圆形区域的顶部附近存在亮噪声点，从而使线扫描法不能很好的与圆边界相切，造成误差），面积统计法对圆形区域内部暗点较多且暗点区域分布不对称的图像提取效果很差（如图3.2样例2第二列）。

由图3.2的对比结果可以看出，三种方法所能处理的图片类型都有一定的局限性。通过实际观察这三种方法的处理结果并仔细分析后，本文在线扫描法的基础上进行了适当的改进，经过对实验结果的观察比较，发现改进的线扫描法对图3.2中两类具有代表性的图片以及大部分正常拍摄质量的圆形鱼眼图片均具有较高的提取准确度见（图3.4），并且兼顾了线扫描法的直观性，从而使线扫描这一方法有了更广的适用范围。下面介绍改进的线扫描法。

### 改进的线扫描法——变角度线扫描法

考虑到传统的线扫描法仅仅从上下左右四个方向逼近圆形区域，如果圆形区域刚好在这四个顶点附近存在较多噪声点，必然会导致该方法产生误差，针对这一情况，本文使用成对的不同倾斜角的直线，按照与传统线扫描法相同的判断方法：

**在扫描线移动的过程中，只要同一扫描线上的像素最大亮度值与最小亮度值之差大于设定的阈值，便认为该扫描线已经与圆形有效区域相切，最大亮度值的像素点即为切点。**

从各个方向对圆形区域夹逼，求取切点坐标，并对这些切点进行筛选，剔除一些无效切点，利用剩余的切点坐标进行圆拟合，进而计算出圆心坐标和半径，以供下一步的校正使用。

以下使用图3.2中样例二的源图片为例，具体说明整个算法流程：

首先，令直线对的倾斜角从0开始，以递增（为便于说明原理，这里取，可以增大的值来提高准确性），分别从圆形区域两侧逼近（如图3.3(a)所示），从而找到各对切点的坐标值（这里的坐标是在图像坐标系下取得，以像素为单位，图像坐标系以图片左上角为原点，向右为轴正方向，向下为轴正方向，如图3.3(c)所示）。

之后，对这些切点坐标数据进行分析，如表3.1所示，从中发现一个规律，使用无效切点对计算出的半径明显大于有效切点对（这也可以从图3.3(b)中看出），例如表3.1所示的第一对和最后一对切点。所以，剔除无效切点的办法是：

对所有切点对计算出的半径取平均值，然后把那些半径大于均值的切点对坐标剔除。

表3.1中切点对序号为2~5的切点坐标为剔除后的有效数据，使用它们进行圆拟合。关于圆拟合有三种有效的方法可供使用，分别是：Taubin method、Pratt method和Kasa method。前两种方法属于强健和精确的圆拟合，即使数据点只是在很小的圆弧范围内，也能得到令人满意的结果。我们这里采用Kasa圆拟合法，因为它的速度很快，计算式简单，但稳定性不如前两种方法。它的计算过程如下：

假设是用于圆拟合的切点坐标数据，对其进行列填充扩展，构造下面的两个矩阵：



令，按下式计算（其中是*A*的伪逆矩阵）：



则拟合圆的圆心坐标和半径*R*为：



至此，圆形有效区域的提取工作完成，实验结果见图3.4。该方法对一些特殊的鱼眼图像圆形区域的提取也有一定的效果，如图3.5所示。下一步进行图像校正。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  | 面积统计法 |
|  |  | 霍夫圆方法 |
|  |  | 线扫描法 |
| 原始图像 | 灰度化处理 | 圆域提取过程 | 圆域提取区域 |  |

图3.1 三种圆域提取方法的处理过程示意

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 样例1 | org1 | area2 | Hough2 | scan1 |
| 样例2 | org3 | area3 | hough3 | scan3 |
|  | 原始图片 | 面积统计法 | 霍夫圆变换法 | 线扫描法 |

图3.2 三种提取方法的处理结果比较

|  |  |  |
| --- | --- | --- |
| 变角线扫描原理 | 观察无效点并剔除 | 图像坐标系 |
| (a) 扫描线逼近圆域 | (b) 连接各对切点 | (c) 图像坐标系 |

图3.3 变角度线扫描法原理

表3.1 剔除无效切点对的数据分析表

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 扫描线倾斜角 | 切点对序号 | 切点坐标 | 对应半径 | 与半径平均值(419.30)比较 |
| 0° | 1 | (446, 6 ), (430, 881) | 437.57 | > |
| 30° | 2 | (229, 109), (625, 834) | 413.04 | < |
| 60° | 3 | ( 72 , 270), (780, 692) | 412.11 | < |
| 90° | 4 | ( 21 , 469), (846, 467) | 412.50 | < |
| 120° | 5 | ( 69 , 665), (788, 258) | 413.10 | < |
| 150° | 6 | (531, 25), (224, 823) | 427.50 | > |

|  |  |  |  |
| --- | --- | --- | --- |
| 1 | 3 | 4 | 5 |
| *n*=2，*t=*10 | *n*=5，*t*=20 | *n*=6，*t*=40 | *n*=6，*t*=30 |
| 7 | 6 | 8 | 9 |
| *n*=5，*t*=20 | *n*=8，*t*=16 | *n=*7，*t*=32 | *n=*5，*t*=40 |

图3.4 变角度线扫描法提取多种类型鱼眼图片圆形区域的处理结果(n是切点对数，t是判定切线成立所设定的阈值)

|  |
| --- |
| 2 |
| *n*=9*, t*=149 |

图3.5 特殊鱼眼图像区域提取

## 鱼眼图像的校正

鱼眼图像的校正过程可以用图4.1表示：

图像校正流程

图4.1 图像校正流程

整个校正过程分为以下两个阶段：

1. 利用鱼眼镜头的成像模型将畸变的图像还原到单位视球面上；
2. 将单位视球面通过特定的映射关系，变换成为我们通常所见的二维图像。

通常，鱼眼镜头在设计时是按照以下的几种成像模型进行的：

体视投影：



等距投影：



等立体角投影：



正交投影：



其中，是镜头光轴与入射光线的夹角，是成像点到图像主点的距离，*f*是镜头光心到图像主点的距离（如图4.2(b)所示，其中），这里假定图像的主点与畸变图像圆形有效区域的中心重合，忽略相机生产时可能存在的微小误差。考虑到市场上应用等距投影模型设计的鱼眼镜头占大多数，所以我们采用它作为成像模型，将鱼眼畸变图像还原到单位视球面上。

### 畸变图像到视球面的转化

下面我们建立单位视球面上以经纬度为坐标的点与鱼眼畸变图像上以图像坐标为坐标的点之间的对应关系。

图像平面到视球面的转换

(a) (b)

图4.2 鱼眼图像向视球面的转换

由图4.2(a)可知，图像坐标系下，圆形区域内的一点，满足关系式：，转换为平面直角坐标有：



再转换为图4.2(b)中所示的极坐标有：



考虑到鱼眼镜头的等距映射模型及图片圆域半径*R*，有，于是我们可以确定出参数，其中是鱼眼镜头的半视场角(假设为已知量)，从而：



这样我们就得到了单位半球面上任一点的球面坐标参数和，于是，球面坐标为：



再转换为经纬度坐标，如图4.2(b)右半部分的半球面所示，，实际上，由这个经纬度坐标可以确定另一个球面坐标，它与式得到的球面坐标存在如下关系：



至此经纬度坐标与鱼眼图像坐标建立了对应关系。下一步便是将半球面向通常的图像平面转化。

### 视球面向平面的转化——纵向压缩柱面投影校正

根据微分几何的知识，球面不是可展曲面，不能展开成平面或平面的子集。所以，通常我们采用投影的方式来成像。实际上我们一般的相机采用的是透视投影模型，只是当镜头的视场角较小时，可以忽略它所产生的畸变。标准镜头视场角在30度左右，广角镜头的视场角都大于90度，而鱼眼镜头的视场角更是接近或超过180度。那么，要校正鱼眼图像，就必须把鱼眼图像重新投影回视球面，然后再投影为通常见到的图像。但是由于鱼眼图像的视场范围大于通常拍摄的图像，所以要想在不损失图片信息的情况下进行这种转换，就必须在对待畸变的问题上有一定程度的取舍。本文使用柱面投影模型并对其进行了改进，以适应鱼眼图像校正的需要。

我们可以把鱼眼镜头看做半个视球面，然后将视球面映射到柱面上去，如图4.3(a)所示。

|  |  |
| --- | --- |
| 视球面到平面的转换 | 图像校正原理图 |
| (a)视球面到平面的转化 | (b)纵向压缩的柱面投影原理 |

图4.3

半视球面上每一条经线都对应着柱面纵向的一条色带，经线上每点按照其所在纬度以一定关系进行映射，如果我们按照透视映射进行的话，会发现当纬度或时，需要的柱面高度将会是无穷大（如图4.3(b)所示，带箭头虚线是透视映射的轨迹），为了解决这个问题，我们采用将透视映射线人为的弯折一定角度之后再投射到柱面上去（图4.3(b)带箭头实线所示），即视球面上一点到同经度*Z*轴方向最高点的连线与投影线的夹角固定为（实验时我们取），这样就可以把本来投影到无穷远点的部分压缩到柱面映射生成的图像上。

在图4.3(b)中，对应用正弦定理有：



其中.

对于直角三角形、，由勾股定理知：



又因为：



所以得：



又因为：



同时考虑到当时，,为保证是锐角，我们对取绝对值，从而有：



将式(3-13)和式(3-14)带入式(3-10)并整理得下式：

, 其中

根据正弦定理，有如下的关系：

，其中

从而，半视球面上的一点与柱面展开结果图上的对应点的坐标关系如下，为指定常量：



其中。

实际上，计算机内部是处理离散数字信息的，数字图像的处理也是一样，所以校正后的柱面图像上的点和原来的畸变图像上的点不是满射关系。如果采用正向映射校正，柱面上将因为部分像素与原图没有对应关系而空缺（如图4.4第二列所示），所以在这里我们采用反向映射的校正方式，即遍历柱面图上的每一个像素的位置，根据之前推导的对应关系逆向从原始图片中取出对应的像素，至于对应不上的那一部分像素，则采用双线性插值计算的方法来填补空缺。

|  |  |  |  |
| --- | --- | --- | --- |
| org1 | forwardMap1 | backwardMap1 | long_lati_Map1 |
| org2 | forwardMap2 | backwardMap2 | long_lati_Map2 |
| 原图 | 正向映射校正  （本文的方法） | 反向插值映射校正  （本文的方法） | 经纬映射法  （反向插值校正） |

图4.4 鱼眼图像的校正结果比较

对两幅不同的图片分别用纵向压缩校正法和经纬映射校正法进行处理，并比较处理结果。从第一个校正示例的处理结果来看，使用经纬映射的校正方法没有完全把弯曲的楼栋校正过来，而改进的校正方法却很好的弥补了这一点，但改进的方法在水平方向上的校正不理想，因为图中的车辆被明显拉长了。从第二个校正示例的处理结果来看，使用经纬映射的校正方法在图片的边缘（尤其是左上角和左下角）产生了很明显的拉伸和扭曲且造成一定程度的模糊，使处理结果显示的不很自然，而改进的校正方法弥补了这一点。

## 多幅鱼眼图像的全景拼接

前面所说的鱼眼图像的最大视场为180度，理论上说只需要两张一前一后的鱼眼图像即可完成对全景图像的构建。实际上，由于拍摄图像时不可能精确的实现让两张鱼眼照片分别容纳前后180的场景内容，并且在图像校正的过程中也不能保证对校正图像的边缘进行完美的处理，所以，一般情况下会在空间一个视点上，利用三角架，水平等角度的旋转镜头，从而获得三到四张鱼眼照片，覆盖视点上360º内的场景内容。这几张照之间会有一部分场景内容是重叠的，正是它们之间这些重叠区域的存在为全景图的拼接提供了可能性。

### 图像拼接的基本流程

图像拼接技术就是将数张有重叠部分的图像拼成一幅无缝衔接的大图的过程。其中，图像的配准和图像融合是图像拼接的两个关键技术。

图像配准是图像融合的基础，图像配准算法一般计算量非常大，因此，图像拼接技术要发展，很大程度上要依赖图像配准技术的进一步发展。图像拼接的方法有很多种，不同的算法的步骤有一定差异，一般来说，都包括以下五个步骤：

图5.1 图像拼接基本流程

1. 图像预处理：包括对输入图像的去噪、边缘提取、直方图处理、建立图像的匹配模板以及对图像进行某种变换(如傅里叶变换、小波变换等)操作。
2. 图像配准：就是采用一定的匹配策略,找出待拼接图像中的特征点在参考图像中对应的位置,进而确定两幅图像之间的变换关系。
3. 建立变换模型：根据图像特征之间的对应关系,计算出数学模型中的各参数值,从而建立两幅图像的数学变换模型。
4. 统一坐标变换：根据建立的数学变换模型,将待拼接图像转换到参考图像的坐标系中,完成统一坐标变换。
5. 融合重构：将待拼接图像的重合区域进行融合得到平滑无缝全景图像。

图像融合部分，一般耗时不太大，现有的几种主要方法效果差别不大，总体来说，图像融合在算法上比较成熟。

### 图像的拼接

图像拼接依赖于拍摄到的图像中相邻两幅图像之间重叠区域的相似性来实现，所使用的算法必须能够对相似区域之间的相似程度做出合理判断，以确保相似区域的数据在拼接后能够相互覆盖在一起。有时图像在拼接后，边界区域会有明显的明暗区别或稍有一点失准，在允许的范围内可以通过图像融合技术对边界处进行处理，消除明显差异，达到无缝衔接。

#### 图像配准

图像的配准一般有两种方式：基于区域相似度的匹配和基于特征点相似度的匹配。

基于区域相似度的匹配是将图像分成较小的区域，通过计算这些小区域的相似度来确定图像配准的位置文献中提出的基于差值图像极值点搜寻的算法就是一种基于区域的匹配方法。但是这类方法的缺点是需要拼接的两幅图像大小相同，并且由于这类匹配算法要对图像的大面积区域像素点进行计算，对于尺寸较大、分辨率较高的图像进行处理时就会产生非常大的计算压力，处理效率不很理想。

基于特征点的图像拼接算法不是直接利用图像的像素值，而是通过提取特征点、特征线或边来进行配准，这些特征被检测到后进一步提取为向量描述符，从而应用向量间的各种距离定义相似程度，被称为特征描述符。

图像处理中特征点的检测方法有很多种：有基于高斯拉普拉斯算子检测的方法(LOG)，利用像素点Hessian矩阵（二阶微分）及其行列式值的方法(DOH)。2004年Lowe提出的高效的尺度不变特征算法(SIFT)，使用该方法提取出的特征点对图像尺度和旋转保持不变，可以用于一个物体或场景不同视角下的可靠匹配，同时对光线变化、噪声、仿射变化都具有鲁棒性，SIFT方法能够从图像中提取出大量的特征点，这些特征点有很高的独特性并且均匀分布于图像中，在物体识别方面应用广泛。2006年，Bay和Ess等人基于SIFT算法的思路，提出了加速鲁棒特征的算法(SURF)，使用了近似Harr小波方法来提取特征点，通过在不同尺度上利用积分图像，有效的计算出近似Harr小波值，简化了二阶微分模板的构建，提高了尺度空间的特征检测效率。其它的还有Harris角点检测以及FAST特征点检测算法。各种检测算法适用于不同的应用情形，因为不同的取景条件或者镜头不同的空间敏感度可能会使相应的物理意义上的特征点相异，一个不正确的特征点可能会导致图像配准的失败，同时检测方法需要具有一定的稳定性，以保证在出现一些特征点的变化或者噪声的情况下不会受到影响。

本文采用基于特征点相似度的图像配准方法进行处理，先从两幅相邻的有重叠区域的图像中分别提取一定数量的特征点并计算出描述符，然后对这些特征点描述符进行相似度计算，找出匹配点。首先基于SURF特征检测算法从图片中检测出一定数量的待选特征，在提取出特征描述符后，需要加入对特征点的筛选环节，以增加图像配准的精确程度。

##### 相似度阈值筛选

对于两幅图中的每一个特征点，我们都使用*k*-近邻法找到另一幅图中与它最相似的两个特征点，这样每个特征点都在另一幅图像中有两个侯选的可匹配点。如果特征点与它最近邻候选匹配点的相似度明显大于它与次近邻候选匹配点的相似度，我们就可以选最近邻候选匹配点为最佳匹配点，因为这两个候选点足够不同，选择最近邻的第一个候选点作为最佳匹配点出错的机率很小。如果特征点与它两个候选匹配点的相似度都相近的话，那么选择其中之一为最佳匹配就会有一定错误机率，因为候选特征点之间也非常相似，这时我们应该排除这个特征点和它的两个候选匹配点。这样，一大部分具有二义性的特征点可以被剔除出去，从而减少配准时的计算量。特征点与它两个候选特征点的相似度比值可以作为一个阈值进行调整，控制特征点被筛掉的数量。

##### 特征点的对称规则筛选

通过上一步的过滤，现在我们有两个特征点配对集，一个是从第一幅图到第二幅图的，另一个是从第二幅图到第一幅图的，我们还可以进一步检测两个方向的特征点配对是不是对称的，从而再一次滤除一些无效特征点对应关系，进一步减少配准计算量。

|  |  |
| --- | --- |
|  |  |
| 原图1 | 原图2 |
|  | |

图5.2 经过相似度阈值和对称规则筛选后的特征点对应关系

经过以上两个步骤的处理，特征点的筛选结果如图5.2所示。对应的特征点筛除数据如表5.1所示，从表中可以看出相似度阈值筛选后，特征点的个数有了明显的减少，91%左右的无效特征被剔除了，原图中所剩余的特征点个数分别为：130和147。在此基础上再应用对称规则筛选，进一步剔除了40%左右的特征点，此时两幅图中的有效特征点个数相同，均为79个，可以极大的提高配准的准确程度和效率。

表5.1 特征点筛选过程的相关数据

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 处理步骤 | 原图1特征点数 | 原图2特征点数 | 图1减少 | 图2 减少 |
| SURF特征检测 | 1570 | 1668 |  | |
| 相似度阈值筛选 | 130 | 147 | 91.72% | 91.19% |
| 对称规则筛选 | 79 | 79 | 39.23% | 46.26% |

有了待配准图像的一些较正确的特征点对应关系后，我们就可以使用一些多视场投影几何的理论来具体的计算配准图像之间的变换关系。

通过对R.Hartley和A.Zisserman所著《Multiple View Geometry in Computer Vision》一书的学习，了解到可以用单应矩阵(homography matrix)来描述从固定点只通过单纯旋转镜头所拍摄的两个不同场景间的变换关系。

从公式中我们导出了摄像机的一般模型，如果我们考虑仅通过旋转关系生成的两个场景的特殊情况时，就会发现式中的第二个矩阵(外部参数矩阵)的第四列都为0，也就是不存在场景的位移描述元素。在这种情况下映射关系可以使用一个的矩阵*H*来描述，这个矩阵被称为单应矩阵(homography matrix)。这种情形下，同一个点在两个仅存在旋转关系的场景中的坐标对应关系可以表示为：



在齐次坐标中通过引入一个缩放因子*s*使这种关系得以保持。一旦矩阵*H*计算出来，一个场景中的所有点都可以通过式变换到另一个场景中，从而初步完成拼接工作。由于存在缩放因子*s*，所以可以将矩阵*H*的最后一个元素设定为1，这就将待定元素个数就从9个降低为8个，坐标和分别表示两个场景中的点坐标，此时的关系可以写为：



因为两个场景中的对应点也可以是已经筛选出的对应特征点，所以可以使用这些特征点来进行对矩阵*H*的估计。在估计的过程中可以使用RANSAC的策略，来提高对*H*矩阵估计的准确程度，因为经过两个步骤的特征点筛选并不能保证选出的特征完全正确匹配。RANSAC算法的特点是可以从一个包含异常值的数据集中找出满足一定可信度的子集用来计算模型。在这里模型描述为式，包含异常值的数据集就是经过筛选后的对应特征点集。要计算出*H*矩阵的8个元素至少需要将4个特征点对的坐标代入式获得9个不同方程式组成的方程组，因为是超定方程组，可以使用最小二乘法进行求解，从而估计出*H*矩阵。

应用RANSAC算法进行*H*矩阵估计的具体思路是：每次随机选择4个特征点对进行估计，把剩下的特征点对用来进行验证，验证关系正确的特征点对组成一个集合，称为支持集。支持集中的元素个数越多，也就说明所估计的*H*矩阵是正解的可能性越大。估计若干次后，选出支持集最大时所对应的*H*矩阵估计作为模型参数解。图像拼接的过程和结果如图5.3所示，其中图2经过矩阵*H*变换到了图1所在的平面（即第二行长图），再把图1叠加到其上就可以生成拼接图了，生成的拼接图中部明显存在拼接缝隙。

|  |  |
| --- | --- |
| 图1 | 图2 |
|  |  |
|  |  |
|  | |
|  | |
|  | |

图5.3 基于特征点的图像拼接

### 图像的融合

针对上节生成的拼接图存在明显缝隙的问题，在拼接时通过使用2.6.2.1节中描述的加权系数法中基于兴趣区域（ROI）的融合方法来消除。方法是分别找出图片重叠区域在图1和图2变换图中的位置（从最左边的特征点起到图1的右边界止），并用矩形框框出，在拼接这个区域时，随着重叠区域在水平方向上的变化，加权系数从1均匀的变化为0，在变化过程中执行融合算法：，。消除拼接缝隙后的效果如下图所示。



图5.4 基于兴趣区域的加权系数融合方法处理结果

### 生成全景图像

将上面的拼接融合方法应用到校正的鱼眼图像上，因为是多幅图像，每次处理两张，再将拼接好的图像与相邻的下一张图片进行拼接，直到所有图片都拼接完成一张图片，就可以得到一张全景图片。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 原始图片 |  |  |  |  |
| 校正图片 |  |  |  |  |
| 全景拼接 |  | | | |

图5.5 鱼眼图像的全景拼接过程

## 全景漫游的实现

### 全景图像的存储方式

拼接好的全景图一般按一定的格式进行保存，以便全景浏览器对其进行浏览。对于要实现的全景浏览来说，球面经纬度投影(equirectangular projection)是一种比较好的存放格式，因为它是全景球面的一种经纬度映射，这种格式比较标准，选它来作全景图的存储格式也是为了可以方便全景图与其它软件之间的数据交换。在给定球面模型的情况下，球面的经纬度和映射后的图像直角坐标之间的关系如下：



其中，是球面经度，是球面纬度，表示赤道南北的角度范围。是映射后图像的水平位置，是映射后图像的垂直位置。映射后的图像中心位置是原点。 当时，就变成经纬度直接投影了，本文中使用这种格式进行全景图的存储。由于,,所以全景图像的宽高比为，每一个像素点对应球面上一个经纬度的交叉点。因为全景图像是数字图像，只能记录离散像素值，所以对其浏览的清晰程度与全景图的像素分辨率有关。



图6.1 球面经纬度坐标与全景图像坐标之间的关系

上图是球面上一点的经纬度坐标与对应全景图上像素点图像坐标的对应关系，和分别表示全景图的宽和高（以像素个数计算）。通过分析我们可以得出以下的对应公式：



### 单位球面到视平面的重投影

要对全景图片进行漫游就需要一个浏览窗口来显示观察到的场景，这个窗口就是视平面，通过把单位球面可视部分映射到视平面上，就可以显示场景信息了。下面我们将建立单位球面参数坐标与视平面图像坐标的对应关系，因为球面的经纬度坐标与球面参数坐标本质上是一致的，所以可以通过简单的变换把其中一者转换成为另一者，有了球面参数坐标与视平面图像坐标的关系，相当于球面经纬度坐标与视平面图像坐标的关系也知道了。



图6.2 视平面与单位球面点间的映射关系

由图6.2所示，我们令视平面上（二维）的点*Q*的图像标坐标为,视平面的宽度为*w*，高度为*h*，图像的中心点坐标为，点*Q*与单位球的球心连线交球面于点*Qs*，考虑到单位球心*O*到视平面中心的距离为*d*，那么视平面上点*Q*的图像坐标转换成以单位球心为原点的三维直角坐标如下:



通过观察，我们发现点*Qs*与点*Q*的三维坐标间成一定比例，这可以通过三角形的相似性原理得出的，关系如下：



又由于，，所以点*Qs*的坐标可以计算得：



球面点坐标也可以用参数方程表示，所以存在下面的关系：



其中和已经于图6.2中标出。由于球面的经纬度坐标只是参数坐标的一种形式，所以我们可以直接把参数坐标看成是经纬度的表示，即，。联系公式就完成了视平面图像坐标与全景图像坐标的对应关系，整理如下：



其中，，，*w、h、d、W、H*可以根据需要设置，属于已知量。

若*w*、*h*的值一定，对于单位球心(视点)到视平面中心的距离*d*，改变它的值等价于改变视域的大小。令水平视域为，由图6.3可知，存在下面的关系：



水平视域一般控制在120º左右，以便与通常人眼的视角大小保持一致。



图6.3 水平视域与参数d之间的关系

### 视平面场景的运动变化调整

要漫游整个场景的不同部分，直观的想法是模拟摄像机在场景中的运动。摄像机在场景中的运动一般有三类：定点的摄像机绕坐标轴的旋转、摄像机镜头焦距的改变以及摄像机在三维空间中的平移。由于本文处理的全景图像是在固定视点下采集图像信息所形成的，没有包含场景的深度信息，因而不能实现对摄像机平移运动的模拟。摄像机镜头焦距的改变可以通过视平面视域的改变来模拟，也就是通过调节公式中的变量*d*来控制。至于摄像机的定点绕轴旋转，我们可以通过对单位球面的旋转来模拟。

在三维直角坐标系下，球面可以绕*X、Y、Z*轴三个方向进行旋转，因为我们通常把Z轴的指向定义为摄像机拍摄的朝向，通常情况下，这个方向的旋转不会发生，也不符合人们观察场景的习惯，所以系统中不考虑球面绕*Z*轴的旋转。以我们观察地球仪的实际经验来看，要浏览一个球上的任何位置，只需要两个正交轴向的绕轴旋转运动即可实现，因此，只要绕*X*轴*、Y*轴的旋转运动，通过这两种运动的配合，就可以浏览整个场景的信息了。

我们以图6.2的三维直角坐标系为参考坐标系，原点在球心位置处，根据空间几何的原理，绕*X*轴、Y轴的旋转运动可以分别用矩阵表示出来，基中AX、AY表示绕轴旋转的角度大小：





当浏览球面的某处时，绕两个轴向的旋转运动都会发生，这样的复合旋转运动可以表达为：。

考虑6.2节中所描述的映射关系，我们发现只要在映射过程中对球面上的点进行前述的复合旋转运动就可以等价的模拟球面的旋转运动，即在公式和之间作如下变换：



变换后的坐标代入公式，进行之后计算。与第三章中处理鱼眼图像的校正类似，视平面像素的生成需要进行插值计算，因为对应到全景图上的坐标不一定是整数的图像坐标，这里我们直接把浮点型的坐标值强制转换为整型的坐标值，简化了插值处理。

### 算法性能的讨论

直接从全景图重投影到视平面的绘制方式性能上有一定的损失，因为需要计算视平面上的每一个像素点的值，对每个点的计算都需要重新计算旋转矩阵的乘积，而且浏览的角度变化时，前面的所有计算过程需要重新进行一遍。这个过程中涉及到了很多的重复计算，是性能损失的主要原因，在浏览不同场景部分时，相邻的两个旋转场景之间有部分像素是重叠的，每一次显示其实可以不必全部重绘所有的像素，只把那些改变了值的像素替换成新计算出的值即可。为了进一步提高性能，可以将部分已经运算出的结果进行缓存，之后的计算如果能在缓存中找到就可以节省一部分计算量，从而提高性能。

### 全景漫游的实验结果

本文针对球形鱼眼全景图设计并实现一个全景浏览器，用于浏览鱼眼校正图拼接后的全景图像，下图为浏览窗口在全景图中漫游的过程，左边图中的白色部分是浏览窗口在全景图中对应的显示区域。

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

图6.4 全景浏览器在不全景图的不同区域进行浏览

## 总结与展望

### 总结

### 进一步的工作

参考文献

发表论文和参加科研情况说明

发表的论文:

张军, 王志舟, 杨正瓴. 单幅圆形鱼眼图像的校正[J]. 计算机应用, 35(5): 1444-1448.

参与的科研项目：

黄色网格线违章停车检测系统（天津市创新基金：13ZXCXGX40400）

车辆车型识别系统（天津市小巨人创新基金：2012-XJR21017）

致谢

感谢linuz，感谢MSVS,Microsoft Office, Viso感谢OpenCV，感谢GitHub