

电 子 科 技 大 学

UNIVERSITY OF ELECTRONIC SCIENCE AND TECHNOLOGY OF CHINA

专业学位硕士学位论文

MASTER THESIS FOR PROFESSIONAL DEGREE

论文题目 全景图像拼接关键技术研究

专业学位类别 工 程 硕 士

学 号 201022050306

作 者 姓 名 赵书睿

指 导 教 师 彭真明 教 授

分类号_____密级_____

UDC^{注1}_____

学 位 论 文

全景图像拼接关键技术研究

(题名和副题名)

赵书睿

(作者姓名)

指导教师_____彭真明_____教 授

电子科技大学_____成 都

(姓名、职称、单位名称)

申请学位级别_____硕士_____专业学位类别_____工 程 硕 士

工程领域名称_____电子与通信工程

提交论文日期_____2013.04_____论文答辩日期_____2013.05

学位授予单位和日期_____电子科技大学_____2013 年 06 月 30 日

答辩委员会主席_____

评阅人_____

注 1：注明《国际十进分类法 UDC》的类号。

THE KEY TECHNOLOGY OF PANORAMA MOSAIC

A Master Thesis Submitted to
University of Electronic Science and Technology of China

Major: Electronics and Communication Engineering

Author: Shurui Zhao

Advisor: Zhenming Peng

School : School of Optoelectronic Information

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及取得的研究成果。据我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得电子科技大学或其它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谢意。

作者签名：_____ 日期：_____ 年 _____ 月 _____ 日

论文使用授权

本学位论文作者完全了解电子科技大学有关保留、使用学位论文的规定，有权保留并向国家有关部门或机构送交论文的复印件和磁盘，允许论文被查阅和借阅。本人授权电子科技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，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

（保密的学位论文在解密后应遵守此规定）

作者签名：_____ 导师签名：_____

日期：_____ 年 _____ 月 _____ 日

摘 要

图像拼接技术是一种解决单幅相机视角不够问题的技术，它给人们的日常生活以及科学研究都带来了很大帮助。随着数字图像处理和微电子等技术的不断发展，全景图像已广泛应用于虚拟现实、计算机视觉、航空航天、军事应用、医学图像分析、遥感图像处理等多个领域。全景图像具有视角广阔，真实性高等优点，目前全景图像的拼接方法已经成为了虚拟现实和图像处理等领域研究的热点。

复杂场景的配准，图像畸变的校正，大曝光差异图像的拼接，动态场景图像的拼接都是图像拼接的难点，而且目前仍没有一套完善的针对大部分常用图像的拼接方法。本文主要针对全景图像拼接时畸变的校正、图像的配准、图像的融合、动态场景中融合鬼影的消除等关键技术进行研究，本文的具体研究内容及贡献如下：

(1) 研究了图像几何畸变的校正，包括常用的图像几何变换模型、全景图像投影模型和摄像头失真的校正。引入一种基于几何模型的图像失真校正方法，快速有效地实现了摄像头失真导致的图像几何畸变的校正。

(2) 分析了基于频域、基于灰度和基于特征的几类图像配准方法。比较了 Harris 角点，尺度不变特征变换和快速鲁棒性特征，确定了比较有效的特征提取方法。深入研究了特征点提纯和模型变换参数的估计，实现了图像的配准，对手持相机拍摄的图像的配准也能取得不错的效果。

(3) 针对动态场景中运动物体导致融合鬼影的问题展开了研究，总结现有几种最佳缝合线的搜索准则，针对它们存在的问题提出了一种改进的最佳缝合线。改进的最佳缝合线减小了曝光差异的影响并充分利用了相邻像素点间的相似性，同时提高了算法的速度。研究了多分辨率融合和泊松融合算法，利用傅里叶变换求解提高了泊松融合的速度。最后利用最佳缝合线与泊松融合实现了自然无缝的全景图像。

关键词： 畸变校正，图像配准，最佳缝合线，泊松融合，全景图像拼接

ABSTRACT

Image mosaic is a technology to solve the problem of the inadequacy of single camera's view perspective. It has brought a lot of help to people's daily lives and scientific research. With the development of digital image processing and microelectronics technologies, panoramic images have been widely used in virtual reality, computer vision, aerospace application, military application, medical image analysis, remote sensing image processing and other fields. The method of stitching panoramic image has become a hot research field of virtual reality and image processing.

The registration of complex scenes, image distortion correction, the mosaic of large exposure differences images and dynamic scene image mosaic are the difficulties of image mosaic. There is no holistic image mosaic method for the most commonly used images. This thesis focuses on distortion correction, image registration, image fusion and ghosting eliminating in dynamic scenes. The main research and achievements present as follows:

(1) This thesis studies image geometric distortion correction methods, including the common image geometric transformation model, panoramic image projection model and camera distortion correction. An image distortion correction method based on geometric model is used to achieve a quick and effective correction of camera geometric distortion.

(2) This thesis analysis the image registration method based on frequency domain, the image registration method based on gray scale and feature-based image registration method. To get an effective feature detection method the Harris corners, scale invariant feature transform and speeded up robust feature are compared. The feature points purification and estimation of model transformation parameters are depth researched. The registration of Hand-held camera image can get a good result by using the mehtod in this thesis.

(3) The method to solve the problem of fusion ghosting caused by moving objects in dynamic scene is studied. An improved optimal seam is proposed. The improved

ABSTRACT

optimal seam reduces the impact of exposure differences and takes full advantage of the similarity adjacent pixels. It also reduces the amount of computation. Multi-resolution fusion and Poisson fusion algorithm are deeply studied. The Poisson fusion solved by Fourier transform improves its speed. A seamless and natural panorama is achieved by using the image mosaic method in this thesis.

Keywords: distortion correction, image registration, optimal seam, Poisson fusion, panoramic image mosaic

目 录

第一章 绪 论.....	I
1.1 研究背景及意义.....	1
1.2 全景图像拼接研究现状及发展趋势.....	2
1.3 论文主要研究内容.....	5
1.4 论文结构安排.....	7
第二章 数字图像的获取和预处理.....	8
2.1 数字图像的获取.....	8
2.2 图像去噪.....	9
2.3 图像畸变校正.....	9
2.3.1 摄像头失真校正.....	9
2.3.2 常用几何变换模型.....	15
2.4 全景图像投影变换.....	17
2.4.1 立方体投影.....	17
2.4.2 球面投影.....	17
2.4.3 柱面投影.....	18
2.5 小结.....	20
第三章 基于特征的图像配准算法.....	21
3.1 概述.....	21
3.1.1 图像配准技术.....	21
3.1.2 基于频域的图像配准算法.....	22
3.1.3 基于灰度的图像配准算法.....	23
3.1.4 基于特征的图像配准算法.....	23
3.2 基于角点检测的图像配准.....	24
3.2.1 Moravec 角点检测算法.....	24
3.2.2 SUSAN 角点检测算法.....	24

3.2.3 Harris 角点检测算法	25
3.3 基于不变量的图像配准.....	26
3.3.1 SIFT 特征的提取	26
3.3.2 SURF 特征的提取	32
3.3.3 特征的匹配	35
3.4 特征提取算法性能分析.....	36
3.5 特征匹配方法.....	38
3.5.1 K-D 树特征点搜索方法	38
3.5.2 随机抽样一致算法原理	39
3.5.3 图像变换模型参数的估计	40
3.6 图像配准算法测试及分析.....	42
3.7 小结.....	44
第四章 全景图像拼接方法及系统实现.....	45
4.1 图像亮度调整.....	45
4.2 传统图像融合方法.....	46
4.2.1 直接平均法	46
4.2.2 加权平均法	46
4.2.3 距离权重法	46
4.2.4 对比度调制法	47
4.2.5 图像融合算法分析	47
4.3 融合鬼影的消除方法.....	48
4.3.1 最佳缝合线	48
4.3.2 多分辨率融合算法	59
4.3.3 泊松融合算法	61
4.3.4 实验及分析	66
4.4 全景图像拼接系统的设计与实现.....	67
4.4.1 全景图像拼接系统	67
4.4.2 基于图像的全景漫游	70
4.5 测试结果及分析.....	71

4.6 小结.....	74
第五章 总结与展望	75
5.1 全文工作总结.....	75
5.2 工作展望.....	76
致 谢	77
参考文献.....	78
攻读硕士学位期间取得的研究成果	83

第一章 绪 论

1.1 研究背景及意义

虚拟现实技术(Virtual Reality, VR)是一门新兴的综合信息技术,它综合了数字图像处理、多媒体技术、计算机图形学和传感技术等多个相关学科的内容^[1]。视觉是人们获取信息的主要来源,因此虚拟场景的绘制和漫游技术在虚拟现实系统中占据着极其重要的位置。基于图像的绘制(Image based Rendering, IBR)是目前比较常用的虚拟场景绘制方法。IBR 技术不同于计算机图形学中传统的三维几何模型绘制(Geometry based Rendering, GBR)重构场景时需要依赖于几何模型,而是利用一组图像序列拼接成为一幅全景图像,最终实现场景的漫游^[2]。图像本身的真实感被 IBR 技术充分利用,增加了虚拟现实的真实感,使得 IBR 技术有着非常光明的应用前景。全景图像拼接是 IBR 技术中最为关键的一部分,它直接影响着整个系统的性能。近年来全景图像的拼接已经成为了 IBR 研究领域研究的热点问题^[3-4]。

所谓全景就是指给人以三维立体感觉的实景全方位图像,此类图像一般具有三大特点:(1)能全面的显示较大范围内的信息;(2)全景图像往往是利用实际拍摄的场景图像拼接得到,因此可以非常好的保留原场景中的真实信息;(3)平面的全景图像却可以给人三维立体的感觉,让观察者可以感受到身历其境的感觉^[5]。所谓拼接就是指把多幅图像合成一幅无明显缝隙的大图像。在科研工作及日常生活中,超过人眼视角的高分辨率图像是经常会用到的,图像拼接的主要目的就是为了解决单幅图像视角不够的问题。随着图像拼接技术的不断发展利用普通的数码相机就可以得到视角足够大的图像,而不必使用价格昂贵又容易发生图像畸变的广角镜头。

目前全景图像拼接除了应用在 IBR 技术上外也已经广泛应用于航空航天^[6]、军事应用^[7]、医学图像分析^[8]、遥感图像处理^[9]等多个领域。航空航天领域是最初应用图像拼接技术的地方,它为人们获得更大场景的清晰照片提供了条件。在军事应用领域由于各种摄像设备的限制使得无法直接获得很大视角的图像。利用摄像设备拍摄多幅图像涵盖较大范围的场景,然后应用图像拼接技术就可以得到较大视角的图像从而获得更多的信息,这一点在提高军队战斗力方面起到了很大作用。在医学图像分析领域,医生无法通过一幅简单的显微镜或超声波得到的图像

进行诊断，全景图像的应用为这样的医学诊断及人体的结构和功能的研究提供了充分的信息。在遥感技术领域，人们通过图像拼接技术把拍摄得到的小视角遥感图像序列拼接成为一幅大视角图像，作为之后研究的基础。

从以上方面可以看出全景图拼接技术具有非常广泛的应用，已成为了虚拟现实、图像处理、计算机视觉等领域研究的热点，因此深入研究全景图像拼接技术具有重要意义。

1.2 全景图像拼接研究现状及发展趋势

早在计算机出现之前图像拼接技术就应用在实践中，人们把拍摄的图像拼接起来得到一幅视角更大的图像，这便是最初的手工图像拼接。随着电子信息技术的不断发展，现在所说的图像拼接往往指的是对数字图像的拼接。目前国内外有很多单位及个人致力于全景图像的生成和图像拼接算法的研究。如图 1-1 所示，图像获取、图像预处理、图像配准和融合这四步是图像拼接的整体流程，其中图像配准被视为图像拼接最为关键的一步。

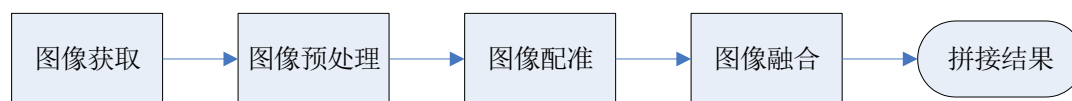


图 1-1 图像拼接基本流程图

根据变化域的不同图像配准算法可以分为两大类，一是基于频域的图像配准，另一种是基于空间域的图像配准。基于频域的图像配准方法主要利用傅里叶变换把图像先变换到频域再做处理。基于傅里叶梅林变换(Fourier-Mellin Transform, FMT)的图像配准算法^[10]就是一种比较典型的基于频域的图像配准算法，该算法的优点是速度比较快。Kuglin 和 Mines 提出的相位相关法^[11]以及 Castro 和 Morandi 提出的扩展相位相关法^[12]都是基于频域的图像配准算法。

基于灰度和基于特征的方法是目目前主要的两种基于空间域的图像配准方法。基于特征的图像配准算法是现在最为常用的图像配准算法。Harris 等人(1988)利用信号处理中的自相关函数提出了 Harris 角点检测算子，提取出对小角度旋转、平移保持不变的兴趣点^[13]。微软研究院的 Richard Szeliski(1996)利用 Levenberg-Marquardt 迭代算法在他提出的基于运动的全景图像拼接模型中计算相邻图像的几何关系，实现了图像的配准^[14]。而后他通过迭代求解摄像机的焦距进一步改进了他的算法，实现了发生几何变换的图像间的精确配准^[15]。该算法也成

为了图像拼接领域的经典算法,直到现在仍然有很多人研究 Richard Szelisk 的图像拼接理论。Kyung 等人(2000)针对在水平方向上旋转拍摄得到的图像,利用等距离匹配算法实现了一种基于水平移动模型的柱面全景图像拼接算法^[16]。该算法对图像的获取要求比较高,不适用于手持相机拍摄得到的图像的拼接。Brown 和 Lowe 等人(2003)提出了基于不变量技术^[17]的尺度不变特征变换(Scale Invariant Feature Transform, SIFT)算法^[18],并使用 SIFT 特征实现了图像拼接。该算法具有良好的鲁棒性,可以对旋转、缩放、仿射变换、噪音和光照变化保持不变性。SIFT 算法的提出又次掀起了图像拼接技术的研究热潮, SIFT 算法目前仍是图像拼接领域研究的热点,现在很多图像拼接应用的实现都是基于 SIFT 特征检测算的。Bay 等人(2006)提出一种利用积分图像及盒子滤波的基于不变量技术的快速鲁棒性特征(Speeded Up Robust Feature, SURF)检测算法^[19]。该算法在继承 SIFT 特征检测算法鲁棒性、抗干扰性的同时,提高了特征的提取速度。目前很多人致力于尺度不变特征的研究,同时也提出了很多改进的算法^[20-25],但是这些算法都是以牺牲 SIFT 或 SURF 某一特点来提高某点不足。Jungpil Shin(2010)提出了一种基于能量谱的鬼影消除算法^[26],该算法首先计算图像的能量谱以及灰度梯度来还原并放大图像拼接缝隙处的特征点,然后根据人眼更关注显著特征的特点来消除图像拼接后存在的鬼影现象。

国外的全景拼接技术无论是从理论研究方面还是产品设计都发展较快。现在国外研究全景图像拼接技术的科研单位主要有苏黎世理工大学,密歇根大学的虚拟现实实验室、华盛顿大学的人机交互实验室等,研究全景图像拼接的主要商业公司有 Google、Apple、Microsoft 等公司。目前国外已经出现了一些较为成熟的产品。加拿大的 Point Grey 公司设计了六摄像头全方位视频采集设备, Dodeca 公司设计了十二摄像头全方位视频采集设备。Apple 公司的 QuickTime VR 系统, Microsoft 公司的 Surrounding Video 是现在比较经典的全景图像拼接系统,但这些系统得到的全景图像仍存在画面感差、存在明显拼接痕迹等问题。Google 用全景摄像机采集城市的三维场景,这就是所谓的谷歌街景。它可以显示所选城市街道的 360 度全景,用户足不出户就可以浏览美国部分城市的全景照片。街景拍摄是由一个多镜头摄像机完成,它可以安装在汽车或者人力三轮车上。他的多镜头摄像机由 11 个摄像头构成,可以全方位拍摄各个方向的场景,再把多个摄像头拍摄的图像拼接成一幅全景图,图 1-2 为 Google 所用的摄像机。

国内对全景图像拼接技术的研究相对国外起步相对要晚一些,但也迅速发展了起来,很多科研机构 and 高校、企业都已经投入研究。国防科技大学、浙江大学、

北京航空航天大学和中科院软件所等是主要的几个国内研究全景图像拼接技术的科研单位。国内研究全景图像拼接的公司并不是特别多，只有少数几家公司在进行图像拼接的研究。目前国内大多数研究主要还是理论研究，只有一些还不够成熟的产品。国防科大开发了 HVS 系统^[27]，该系统可以将离散图像序列或是连续的视频，经过预处理、图像配准和图像融合等步骤后拼接成为一个 360 度的柱面全景图像，最后利用一个适当的虚拟空间模型把拼接得到的柱面图像绘制成一个虚拟的空间。用户可在在这个虚拟的空间中完成前进、后退、近看、远看、俯视、仰视、360 度环视等操作。浙江大学 CAD&CG 国家重点实验室提出了一种基于检测重叠区域及其边界的模板匹配的图像配准算法^[28]。该图像拼接算法原理简单易于实现，但是图像配准时计算量过大并且也不容易得到精确的图像配准结果。目前的这些全景图像拼接系统虽然取得了一些成果，但是还不够完善，仍然有待进一步的研究。



图 1-2 Google 多镜头摄像机

由于拍摄条件的不同导致采集到的图像具有不同特点，针对具有不同特点的图像需要利用某些特定的图像拼接算法来实现全景图像的拼接。立方体投影模型、球面模型、柱面模型是针对不同采集方式的三种主要的全景投影模型。唐琰等提出了一种全景图像绘制方法^[29]，该方法产生的就是一种基于立方体模型的全景图像。李学庆等设计的视景生成系统^[30]，华顺刚等提出的全景图像自动拼接算法^[31]，马向英等开发的虚拟漫游系统^[32]都是基于球面投影的全景图像。柱面全景图像的一个经典例子就是美国 Apple 公司的商业软件 QuickTime VR，该软件在模型中的关键位置点生成柱面全景图，最后在这些关键点之间跳转实现了基于柱面全景图像的场景漫游^[33]。

目前的图像拼接方法一般只能针对于某种与其对应采集方法获得的图像的拼接，还没有一种能够用于所有特点图像拼接中的图像拼接算法，一种可以适用于大部分常用图像的拼接算法是困扰着图像拼接领域的一大问题。正因为这一点现

在也不存在一种评价图像拼接的统一标准，目前图像拼接结果质量的评价一般仍是要依靠人眼的主观感知来评价。对于手持相机拍摄的图像往往因为抖动等原因使得图像拼接的难度大大增加，解决发生几何畸变及失真的图像的拼接问题也具有着重要的意义。大曝光差异图像的拼接往往会存在拼接痕迹，因此近年来图像拼接中的融合技术越来越受到人们的重视。如何消除图像拼接后的鬼影，尤其是当场景中存在运动物体时消除运动物体融合时产生的鬼影也是图像拼接领域的难点。全景图像往往需要大量的图像拼接而成，提高图像的拼接速度也是全景图像拼接的研究方向。处理的图像的质量直接影响着图像拼接的质量，高质量的待拼接图像自然相对容易实现图像的精确配准。研究如何获得高质量的待拼接图像也存在着一定的价值，也是未来图像拼接领域的一个研究方向。

1.3 论文主要研究内容

本文主要对全景图像的拼接中的几点关键技术进行研究，总结分析现有的图像拼接方法实现一种有效的全景图像拼接系统，本文的主要研究内容及贡献如下：

(1) 研究图像预处理，主要包括图像去噪、图像畸变校正、全景图像投影模型等。引入了一种基于几何模型的校正方法快速有效地还原了真实场景，降低了摄像头失真对全景图像的影响。

(2) 研究基于特征的图像配准方法，分析了 Harris 角点检测算法，SIFT 特征提取算法和 SURF 特征提取算法的优缺点，并比较了他们提取特征点的性能。深入研究特征点提纯和变换模型参数的估计，实现了多幅图像的全局配准，对手持相机拍摄获取的图像也能取得较好的配准效果。

(3) 研究几种传统融合方法，并针对动态场景融合鬼影的消除进行深入研究。总结分析现有的最佳缝合线搜索方法，提出一种改进的最佳缝合线准则，它减小了曝光差异的影响并充分利用了相邻像素点间的相似性，对特征点处的准则值加权减小误差点的干扰，同时该算法在速度上有所提高。研究多分辨率融合算法和泊松融合算法，利用傅里叶变换的方法求解泊松融合方程，提高了融合的效率。最后利用改进的最佳缝合线和泊松融合消除了动态场景中的融合鬼影，实现了图像的自然无缝拼接。

(4) 研究多幅图像的拼接方法及全景图像漫游方法，实现一种可以在水平方向旋转观察周围场景的全景浏览器。

图 1-3 为本论文的主要研究内容及简要技术路线图，它指明了本文的主要技术

路线。

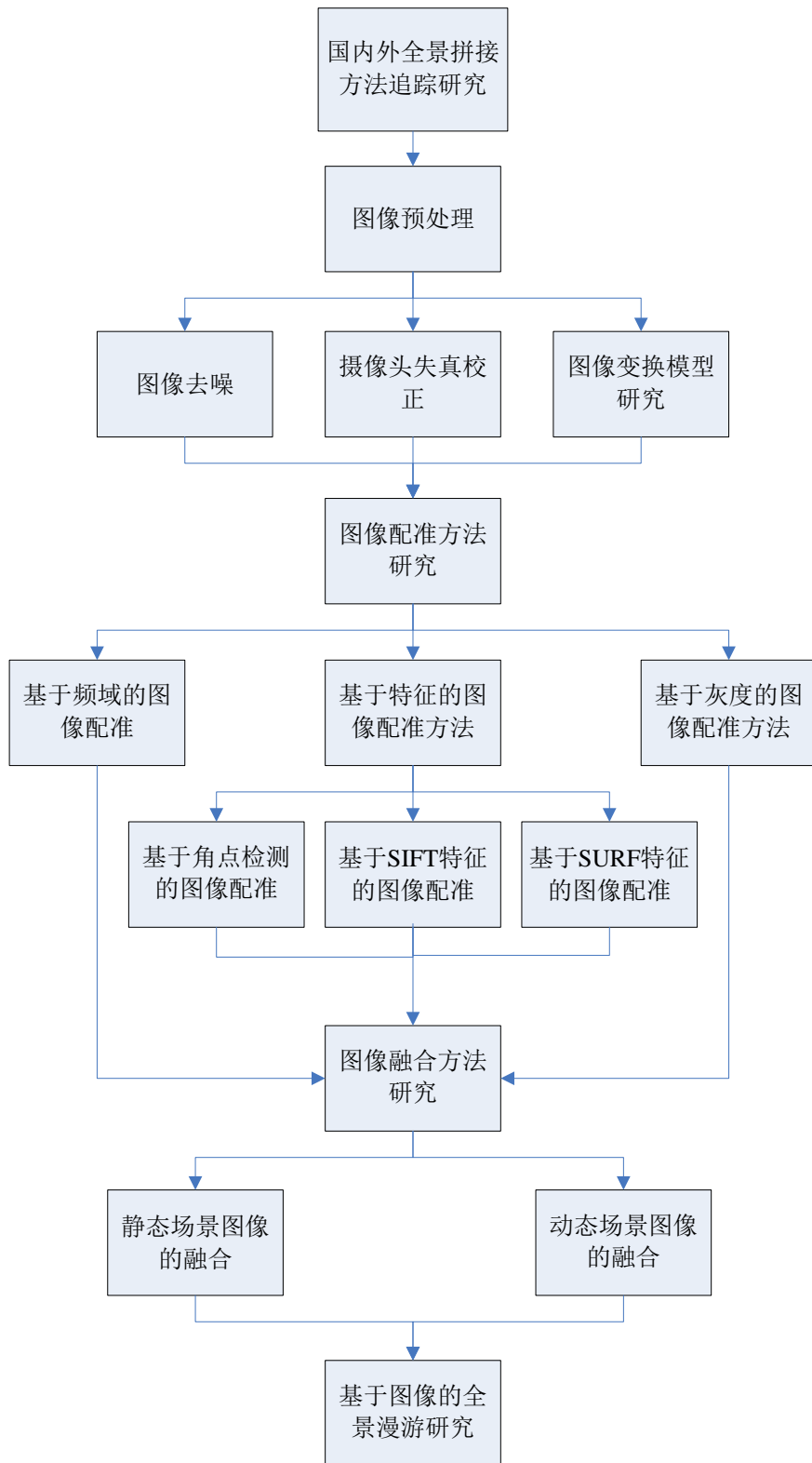


图 1-3 论文简要技术路线图

1.4 论文结构安排

本论文主要针对全景图像拼接中的图像校正、全景图的投影模型、图像的配准、图像融合以及融合鬼影的消除进行研究，论文的主要结构安排如下：

第一章，介绍了研究本课题的背景和意义，分析了图像拼接技术的研究现状及发展趋势，阐述了本文要解决的主要问题和取得的主要研究成果。

第二章，介绍了图像获取的几种基本方法和预处理的问题。图像的预处理中简要的介绍了图像的去噪方法，详细介绍了图像几何畸变的校正和针对不同采集方式采用的图像投影模型，并且引入了基于几何模型的图像失真校正方法。

第三章，介绍了研究基于频域、基于灰度和基于特征的图像配准方法。主要研究了基于特征的图像配准方法，详细介绍了基于角点检测的图像配准方法、基于 **SIFT** 特征的图像配准方法和基于 **SURF** 特征的图像配准方法。比较分析了 **Harris** 角点，**SIFT** 特征和 **SURF** 特征。并深入研究了特征点提纯和模型变换参数的估计。

第四章，主要讲述全景图像的拼接方法和系统的实现。该章分析了几种传统融合方法的优点与不足，深入研究了现有几种最佳缝合线的搜索准则，针对它们存在的问题提出了一种改进的最佳缝合线，介绍了多分辨率融合及泊松融合，最后阐述了全景图像拼接系统的设计方案并阐述了全景图像浏览器的设计方法。

第五章，总结了课题的主要工作，指出了取得的研究成果和不足，并对后续的研究工作提出了自己的观点和建议。

第二章 数字图像的获取和预处理

2.1 数字图像的获取

获取图像的方式不同必将导致输入图像的不同，从而需要根据不同的图像获取方式选择不同的图像配准方法。使用普通相机拍摄图像时，照相机的运动方式决定着图像的获取方式。使用普通相机拍摄时一般有旋转相机拍摄、平移相机拍摄和手持相机拍摄三种方式。

(1) 使用三脚架旋转相机进行拍摄：使用这种方法采集图像时三脚架放置在某一位置不动，照相机绕垂直轴旋转，每旋转过一定角度便拍摄一张图像，相邻的两种图像必须有部分重叠。因为相机固定不需要恢复太多参数，因此比较容易实现图像的拼接。但是由于这种旋转拍摄的图像序列不在同一平面上，若简单投影到同一平面上会导致图像发生扭曲等现象使图像质量下降影响整个图像拼接结果。旋转相机拍摄一般需要使用柱面投影或是球面投影来生成全景图像。这种拍摄方式也需要相邻的两幅图像有较大的重叠区域才能保证图像拼接的质量，一般选用不小于 50% 的重叠比例较为合适。而过大的重叠率会导致拼接全景图时需要更多的图像，增大需要处理的信息量。为了减小计算量采集图像时可以适当地使用广角镜头。

(2) 把相机固定在水平滑轨上，保持相机焦距不变，水平移动相机每隔一定距离拍摄一幅图像，最后得到有一定重合区域的图像序列。这种方法获取的图像都在一个平面上是比较容易实现图像的拼接。但是这种方法对拍摄条件的限制比较苛刻，物体和相机的距离也会影响图像的配准结果，而且这种拍摄方式视点放生改变不适宜用来生成 360 度全景。

(3) 手持相机拍摄，可以站在原地旋转相机进行拍摄，也可以水平移动进行拍摄，这种拍摄方式对拍摄器材的限制比较小只需要一台相机即可完成拍摄，是最易实现的拍摄方式。但是手持相机拍摄过程中相机的运动非常复杂，势必会发生难以预料的抖动，旋转角度或是平移距离都难以精确控制，给图像配准带来了很大的困难。但是研究这种相机运动方式复杂情况下拍摄的图像的拼接对图像拼接的广泛应用有着重要的意义。

这三种拍摄方式必须面对的同一个问题就是相邻图像间的光强变化，理想情况

下图像相同区域应该具有相同的光强，但是拍摄过程中光源的变化及相机与光源间角度的变化都是导致图像的相同区域之间具有不同光强的原因。使用多相机拍摄时不同摄像头的个体差别也会导致图像曝光的差异。使用合适的融合方法解决好图像获取中的曝光差异也是图像拼接需要研究的一个重要问题。

2.2 图像去噪

由于采集图像过程中环境的负责性和多样性，采集的图像中难免存在噪声点。在进行全景图像的拼接之前需要进行图像去噪的工作，假如不去除图像中的噪声点，在图像配准的过程中很可能会受到噪声点的影响，误认为噪声点是兴趣点导致效果不佳的图像拼接结果。所以说图像去噪处理是后面图像配准的基础，只有经过去噪处理后质量比较高的图像才能达到图像配准的要求。图像去噪的方法很多，也有很多学者一直致力于图像去噪算法的研究。中值滤波是基于排序统计理论的一种能有效抑制噪声的非线性信号处理技术。本文中只需要中值滤波去除比椒盐噪声就可以达到图像配准的要求，图像的中值滤波方法已经比较成熟，本文不再详述。

2.3 图像畸变校正

在图像采集的过程中相机的运动会导致图像发生几何形变，摄像头自身的问题也会使采集的图像无法还原真实世界中的场景，因此对图像的畸变进行校正是进行全景拼接前必不可少的一步。

2.3.1 摄像头失真校正

图像畸变校正就是对发生失真后的图像进行还原性处理。光学系统自身问题、成像器材的差异、运动模糊、引入噪声等都是导致图像发生畸变的原因。图像畸变校正就是要根据引起图像畸变的原因建立相应的数学模型还原失真的图像。

2.3.1.1 图像畸变分类

这里首先研究相机本身系统问题导致的失真。理想情况下的摄像头可以还原真实世界中的场景，但是由于拍摄方式的问题及摄像机本身光学和电子系统的误差会使获取的图像发生扭曲无法还原真世界中的场景，这种扭曲是不规则的，但

是这种扭曲通常情况下是以摄像头中心为对称轴成径向形态对称的，我们把这种畸变叫做径向畸变。径向畸变一般分为桶形失真和枕形失真两种，图 2-1 为径向畸变分类的示意图。

(1) 桶形失真：图像中心的放大倍数大于图像边缘的放大倍数，如图 2-1(b)所示，桶形失真是最为常见的失真方式。

(2) 枕形失真：图像中心的放大倍数小于图像边缘的放大倍数，如图 2-1(c)所示，枕形失真一般出现在望远镜中。

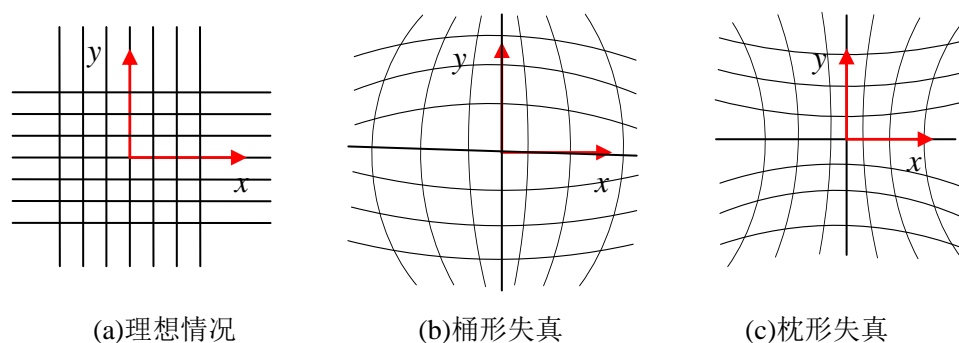


图 2-1 镜头失真分类

2.3.1.2 桶形失真校正

现在有很多的镜头失真校正方法，目前大部分的方法是根据实际的采集情况列出一系列的方程组，然后通过求解方程组得到图像的校正系数，最后实现图像的畸变校正^[34-36]。这种畸变校正方法的优点是可以得到精度较高的校正结果，但是该方法求解方程组的数学运算一般比较复杂，甚至某些参数无法精确求解只能近似估计。由于本文主要研究的是全景图像拼接，会有较多需要处理的图像序列，如果使用过于复杂的畸变校正方法会严重影响整个全景拼接系统的速度，所以本文使用一种基于摄像头光学系统等效球面模型^[37]的方法根据几何关系简单快速的校正失真。

对于一个球面光学系统，一般它产生的畸变具有以光轴为中心对称越靠近中心畸变越小的特点，因此可以用一个球面模型模拟图像畸变的产生，图像畸变的球形模型如图 2-2 所示。 P 是一个垂直于摄像头轴线的平面， P' 是它理想的像平面， S 是一个以焦距 f 为半径以镜头像方主点为圆心的球面，球面 S 与像平面 P' 相切于点 O ，点 O 为像平面 P' 的中心点。

设 P_1 是平面 P 上的一个点，如果不发生图像畸变 P_1 所成的像应该在像平面 P' 上的 P'_1 处。假设由于某些原因发生了桶形畸变，这时像平面 P' 不再是一个平面而

是一个曲面，为了方便实现图像畸变的校正把发生畸变后的像曲面近似为一个球面 S 。射向 P_1' 点处的光线与球面 S 相交于点 P_2'' ，图 2-2 中点 P_2' 为与点 P_2'' 在平面 P' 上对应的点，即 $OP_2' = OP_2''$ 。设 L_1 是点 P_1' 到 O 点的距离， L_2 是点 P_2'' 到点 O 的距离。根据图中的几何关系可以利用式(2-1)和式(2-2)计算出 L_2 和夹角 α 。

$$L_2 = 2f \sin(\alpha / 2) \quad (2-1)$$

$$\alpha = \arctan(L_1 / f) \quad (2-2)$$

根据式(2-1)和式(2-2)可得：

$$L_2 = 2f \sin\left[\frac{\arctan(L_1 / f)}{2}\right] \quad (2-3)$$

从式(2-3)可以看出被拍摄物体上的点距离其中心点越近 L_2 就越小，这符合桶形畸变的特点。

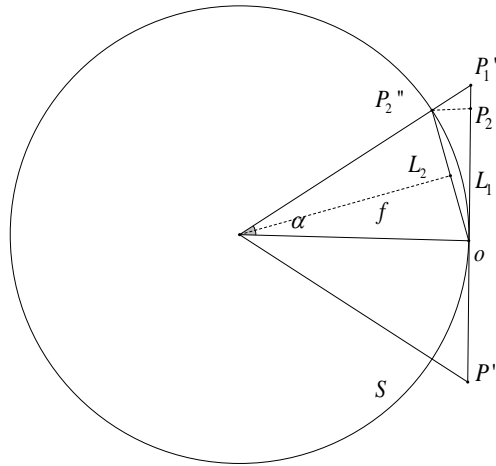


图 2-2 图像畸变球形模型

根据上述的畸变校正球形模型就可以得到畸变校正的公式。图 2-3 为图像畸变校正示意图， H 和 W 分别为图像的高和宽，点 O 为图像的中心点，其坐标为 $(W/2, H/2)$ ，点 P_2' 为物点 P 在发生畸变图像中点的位置，其坐标为 (x_2, y_2) ，点 P_1' 为物点 P 在校正后图像中的位置，设其坐标为 (x_1, y_1) 。解决这个畸变校正的问题就是要求出点 P_1' 的坐标。

点 P_2' 到图像中心点 O 的距离 L_2 为

$$L_2 = \sqrt{\left(x_2 - \frac{W}{2}\right)^2 + \left(y_2 - \frac{H}{2}\right)^2} \quad (2-4)$$

式中 β 为 P_2' 与图像的中心点 O 连线与虚线 l 的夹角

$$\beta = \arctan\left(\frac{y_2 - H/2}{x_2 - W/2}\right) \quad (2-5)$$

根据式(2-3)可以计算出校正后的距离 L_1

$$L_1 = f \tan\left(2 \arcsin \frac{L_2}{2f}\right) \quad (2-6)$$

然后就可以根据 L_1 计算出校正后的点 P_1' 的坐标

$$\begin{cases} x_1 = W/2 + L_1 \cos \beta \\ y_1 = H/2 + L_1 \sin \beta \end{cases} \quad (2-7)$$

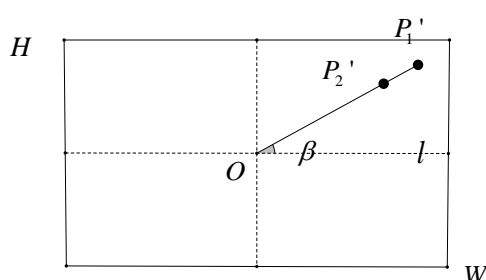


图 2-3 图像畸变校正示意图

从式(2-6)和(2-7)可以看出摄像头失真的校正过程只跟图像采集时的焦距有关，而且简单容易实现不需要复杂的迭代运算，可以大大降低畸变校正的运算量。对图像焦距的估计将在后面的 2.4 节中做详细的介绍。把图像上的每一个点根据上述的计算过程进行操作就可以实现桶形畸变的校正。对于枕形畸变分析方法跟桶形畸变分析方法类似，这里不再过多叙述。

2.3.1.3 图像插值方法

在图像校正过程中畸变图像需要映射到新的坐标上，新图像上的点有可能不存在与之对应的点，所以需要对其进行图像插值。最近邻插值法、双线性插值法和三次卷积插值法是最常用的三种插值方法。

(1) 最近邻插值法：又叫做零阶插值法，是一种最简单的插值方法。此方法取与待插值像素点距离最近的像素点的灰度值作为其灰度值。最近邻插值方法具有计算量小，简单容易实现等优点，但是它的精度较低，在图像边缘容易产生锯齿现象。

(2) 双线性插值：该方法又称为双线性内插或一阶插值，其核心思想就是在两个方向分别进行一次线性插值。设四个已知像素点 A、B、C、D 的坐标分别为 (x_1, y_1) 、 (x_1, y_2) 、 (x_2, y_2) 、 (x_2, y_1) ，他们的灰度值分别为 $f(A)$ 、 $f(B)$ 、 $f(C)$ 、

$f(D)$ ，未知点 M 的坐标为 (x, y) 。可以计算出点 E 和点 F 的灰度值 $f(E)$ 和 $f(F)$ 。

$$f(E) = \frac{x - x_1}{x_2 - x_1} (f(C) - f(B)) + f(B) \quad (2-8)$$

$$f(F) = \frac{x - x_1}{x_2 - x_1} (f(D) - f(A)) + f(A) \quad (2-9)$$

则点 M 的灰度值为：

$$f(M) = \frac{y - y_1}{y_2 - y_1} (f(E) - f(F)) + f(F) \quad (2-10)$$

线性插值插值的结果与插值顺序无关，但是双线性插值并不是线性插值，首先计算 x 方向再计算 y 的插值与先 y 在计算 x 方向的插值结果是不一样的。

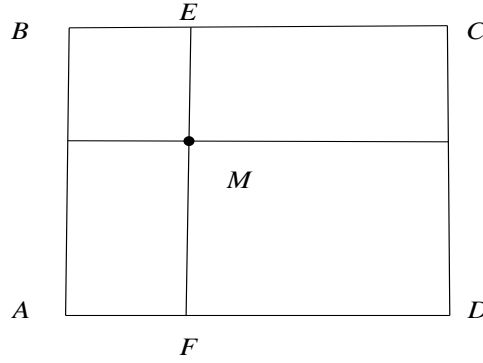


图 2-4 双线性插值

(3) 三次卷积插值法：想要得到更精确的插值结果就不能只考虑与待插值像素点直接近邻的四个像素点，而是要考虑与它相近邻的 16 个像素点对它的影响。如图 2-5 所示。

根据连续信号采样定理，如果对采样值使用式(2-11)的插值函数，就可以得到准确的函数恢复结果。

$$S(x) = \begin{cases} 1 - 2x^2 + |x|^3 & |x| < 1 \\ 4 - 8|x| + 5x^2 - |x|^3 & 1 \leq |x| \leq 2 \\ 0 & |x| \geq 2 \end{cases} \quad (2-11)$$

设待插值点 M 的坐标为 (x', y') ，其灰度值为 $f(x', y')$ ，设 $\alpha = x' - x$ ， $\beta = y' - y$ 。可以得到 $S(1+\alpha)$ 、 $S(\alpha)$ 、 $S(1-\alpha)$ 、 $S(2-\alpha)$ 、 $S(1+\beta)$ 、 $S(\beta)$ 、 $S(1-\beta)$ 、 $S(2-\beta)$ ，计算 $f(x', y)$

$$f(x', y) = S(1+\alpha)f(x-1, y) + S(\alpha)f(x, y) + S(1-\alpha)f(x+1, y) + S(2-\alpha)f(x+2, y) \quad (2-12)$$

同理可得 $f(x', y-1)$ 、 $f(x', y+1)$ 、 $f(x', y+2)$ ，最后可以计算出点 M 的灰度值 $f(x', y)$

$$f(x', y) = \frac{1}{9} [f(x', y-2) + f(x', y-1) + f(x', y) + f(x', y+1) + f(x', y+2) + f(x'+1, y-2) + f(x'+1, y-1) + f(x'+1, y) + f(x'+1, y+1) + f(x'+1, y+2)] \quad (2-13)$$

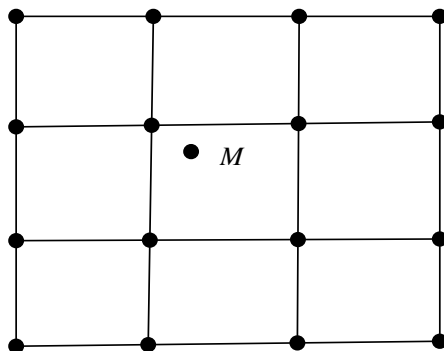


图 2-5 三次卷积插值

这三种插值方法各有优缺点，最近邻插值法最为简单，容易实现计算量最小，但是插值后得到的图像往往存在明显的不连续性。三次卷积插值法具有增强边缘的效果，但是计算量过大。双线性插值法克服了近邻插值法不连续的缺点，计算量较三次卷积插值法要小，但是图像边缘会受到平滑的作用导致图像模糊。在实际的应用中为了减小计算量又能获得较好的插值效果常常采用双线性插值法。

2.3.1.4 失真校正结果及分析

图 2-6(a)为一个大小为 756×420 发生了桶形畸变的图像，图 2-6(b)为利用基于几何模型的摄像头畸变校正方法实现校正的结果，在 CPU 主频为 2.4GHZ，内存为 2G 的硬件条件下，在 Windows 系统下以 VC6.0 作为编程平台进行测试。图 2-6(b)为图像畸变校正结果，畸变校正所用的时间为 0.12 秒。从图(b)中可以看出发生畸变后的曲线已经校正成直线，从直观的视觉出发此校正结果已经基本还原了真实场景。这种基于几何模型的校正方法不需要迭代，只需要对图中的每一个像素点进行相应的几何变换就可以实现畸变的校正。而且这种畸变校正方法只与焦距有关，无需估计过多参数。这种可以还原真实场景信息，又不会耗费过多时间的图像校正方法正是全景图像拼接的所需要的。



(a) 原图



(b) 校正结果

图 2-6 图像畸变校正

2.3.2 常用几何变换模型

一般情况下在拍摄场景时，相机并不总是安装在三脚架上的，可以手持相机拍摄，因此相机的运动不能通过平移运动、镜头缩放、水平扫动、垂直扫动和旋转运动中单一的一种方式描述相机运动的情况，而是需要使用上述几种运动方式的合成来描述相机的运动方式。当相机运动方式比较复杂时，相机拍摄的图像序列，尤其是图像序列间的重叠部分可能存在着较大的差异。例如在这些图像序列的重叠部分可能存在坐标的不统一和光强的变化等问题，这些问题都是由于相机的运动和光源的变化引起的。因此只要确定了图像序列间的空间变换关系才能完成这些具有重叠部分的图像序列的拼接。为了确定图像序列的空间变换关系，需要确定图像对应关系的模型。平移变换模型、刚性变换模型、仿射变换模型和投影变换模型是目前常用的几种图像变换模型，图 2-7 即为几种几何变换模型示意图。

利用齐次坐标可以把上述的几何关系用矩阵的形式描述出来：

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = M \times \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \\ m_6 & m_7 & 1 \end{bmatrix} \times \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2-14)$$

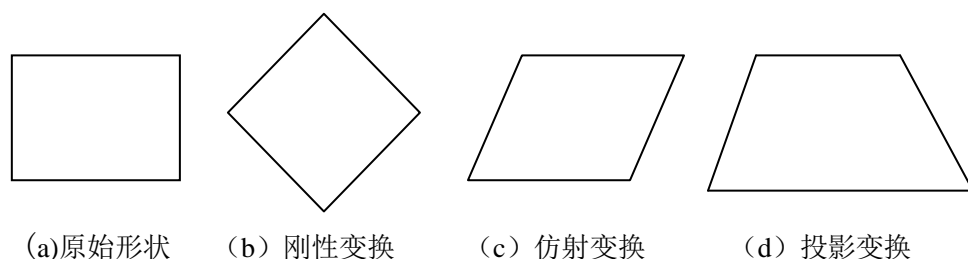


图 2-7 图像变换模型示意图

式(2-14)中 M 的各个参数的作用如表 2-1 所示

表2-1 变换矩阵 M 中参数作用

参数	参数作用
m_0, m_1, m_3, m_4	尺度和旋转量
m_2	水平方向位移
m_5	垂直方向位移
m_6, m_7	水平与垂直方向的形变量

对矩阵 M 进行简化就可以得到如下几种几何变换模型：

(1) 刚性变换模型：它用来描述相机绕光轴的旋转运动和平移运动。刚性变换不会使原来的图像发生扭曲。它的矩阵形式为：

$$M = \begin{bmatrix} \cos \theta & -\sin \theta & m_2 \\ \sin \theta & \cos \theta & m_5 \\ 0 & 0 & 1 \end{bmatrix} \quad (2-15)$$

式中 m_2 和 m_5 为平移量，角 θ 为图像旋转的角度。

(2) 仿射变换：它是用来描述相机的平移、旋转和缩放运动的。它的一个特点就是保持图像中的直线，即发生仿射变换的图像中直线仍为直线，互相平行的直线仍平行。它的矩阵形式为：

$$M = \begin{bmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \\ 0 & 0 & 1 \end{bmatrix} \quad (2-16)$$

(3) 投影变换模型：它可以描述相机的平移运动、镜头缩放、水平扫动、垂直扫动和旋转等运动，它是一种更为一般的形式，刚性变换和仿射变换是投影变换的一种特例。它具有 8 个自由度，矩阵形式如下：

$$M = \begin{bmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \\ m_6 & m_7 & 1 \end{bmatrix} \quad (2-17)$$

2.4 全景图像投影变换

全景图像时虽然是平面图像但可以给人一种立体感，所以我们需要一种投影模型来实现上述问题。常用的全景图像投影模型有立方体投影、球面投影和柱面投影三种^[38-40]。

2.4.1 立方体投影

立方体全景图是最为简单的多面体全景图，它以景物中心为固定视点观察整个场景，以视点为中心拍摄图像，并把空间中全部六个方位平面上的图像投影到立方体表面拼接成全景图像。这六个平面上图像的视角都是 90 度。由于立方体全景图的每个平面上的图像都是规则的正方体所以方便计算机存储和显示，而且可以实现水平和垂直方向全方位 360 度的全景图。基于立方体投影的全景相邻的平面是相互垂直的，所以在采集图像时拍摄相邻图像过程中相机的主光轴也要保持互相垂直。而且每一个平面上的图像要求视角是 90 度，所以这种立方体全景图对图像的获取和相机的准确定位要求都比较高，而且还要用到广角镜头，这就对图像采集带来了一定的难度。

2.4.2 球面投影

球面投影方法就是把拍摄的图像序列投影到一个球面上，通过计算图像映射到球面上的精度和纬度的参数实现全景图像^[41]，以球面的形式来存储全景图像。获取图像时分别以空间中水平和垂直两个方向为轴来旋转相机进行拍摄。

把所有的像素点投影到一个球体表面，球面投影的示意图如图 2-8 所示。设 f 为球体的半径，平面 S 为图像平面，点 P 和点 Q 在平面 S 上，点 M 和点 N 在球面上，并且点 P, Q 在球面上的投影点是点 M 和点 N ，点 C 为平面 S 的坐标原点，二维坐标系 (X_s, Y_s) 是以 O 为原点的球面坐标系。点 P 在球面坐标系下对应的点 M 的经纬坐标 X_s 和 Y_s 可以表示为：

$$X_s = f \arctan \frac{x}{f} \quad (2-18)$$

$$Y_s = f \arctan \frac{y}{\sqrt{x^2 + f^2}} \quad (2-19)$$

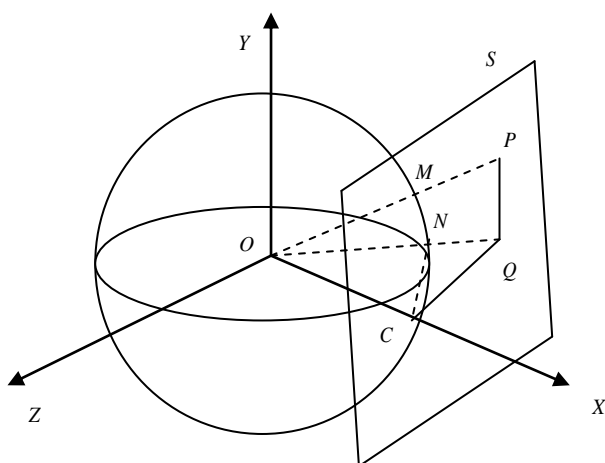


图 2-8 球面投影示意图

2.4.3 柱面投影

柱面全景是使视点固定，在水平方向选择拍摄得到一组相邻图像具有重合区域的图像序列，再对这组图像序列做柱面投影变换最后拼接成一幅柱面全景图像。

图 2-9 即为柱面投影的示意图。假设相机的所有运动都是在 $x-z$ 平面上发生的，把光轴与图像平面的交点作为图像的中心点。 I 为原始图像， W 、 H 分别为图像的宽与高，设 I' 为柱面图像， K 为投影坐标平面，柱面的中线点是投影坐标平面的原点，如图 2-9(a)所示。在图像 I 上任意取一点 P ，其坐标为 (x, y) ，柱面投影就是要计算出其在柱面图像上对应的点 P' 的坐标 (x', y') 。图 2-9(b)和图 2-9(c)是点 P 在 $x-z$ 平和 $y-z$ 平面上的截图。 r 为图像的焦距， θ 是每张图像所占的弧度角。我们可以得到柱面变换的公式

$$\begin{cases} x' = r \sin \frac{\theta}{2} + r \sin(\arctan(\frac{x - W/2}{r})) \\ y' = \frac{H}{2} + \frac{r(y - H/2)}{k} \end{cases} \quad (2-20)$$

$$k = \sqrt{r^2 + (W/2 - x)^2} \quad (2-21)$$

若一幅全景图像所占的角度是 360 度，相机拍摄时每次旋转的角度相同，那么每一幅图像对应的弧度可以近似表示为式(2-22)的形式。

$$\theta = \frac{2\pi}{n(1-\eta)} \quad (2-21)$$

式中 η 为相邻图像的重叠率， n 为拼接图像的张数，这样就可以近似的估计图像拍摄时的焦距：

$$r = \frac{W}{2 \tan(\theta/2)} \quad (2-22)$$

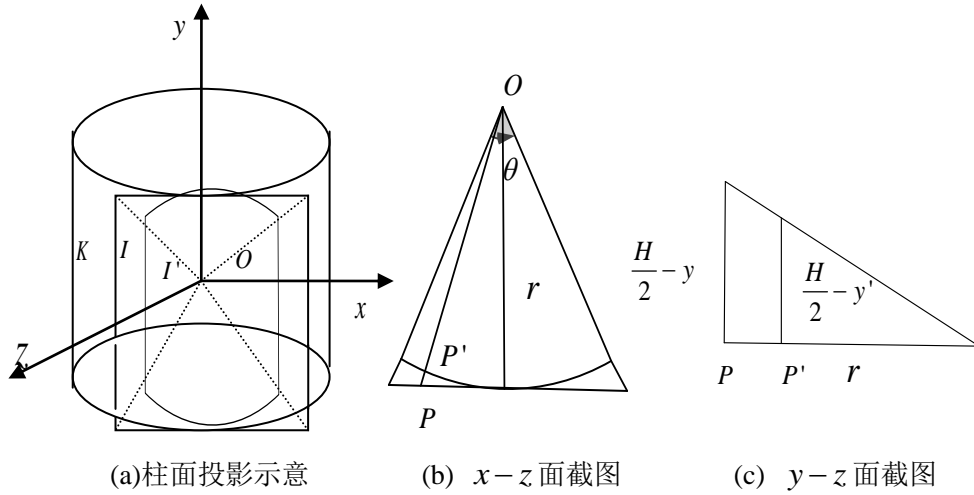


图 2-9 柱面投影变换示意图

对柱面投影公式变换可以得到柱面投影的反变换公式

$$\begin{cases} x = \frac{W}{2} + r \sin(\arctan(\frac{x' - r \sin \frac{\theta}{2}}{r})) \\ y = \frac{H}{2} + \frac{k(y' - \frac{H}{2})}{r} \end{cases} \quad (2-23)$$

通过柱面投影的反算法就可以把原图像投影到柱面图像上实现图像的拼接，避免了坐标不统一的问题，使拼接过程变得简单容易实现。

立方体投影、球面投影和柱面投影这三种投影模型针对不同的采集方式采集的图像来使用，这三种全景图的实现中各有利弊。立方体投影过于简单，实现的全景图像往往存在明显的边缘痕迹，使得全景图的真实性和完整性有所降低。球面投影方式扩大了全景图的视野，但是也增加了算法的复杂性不易实现。柱面投影相对简单，而且容易展开为一个矩形的平面图像，也不会降低全景图像的真实性和完整性。柱面全景图是目前比较常用的一种方式，本文中的全景图实现也是采用柱面全景的方

式。图 2-10 中(a)为一幅原始图像，图(b)就是一幅经过柱面变换后得到的图像，可以看出图像经过柱面变换后被投影到了另一个平面上，新平面为事先设定好的柱面。

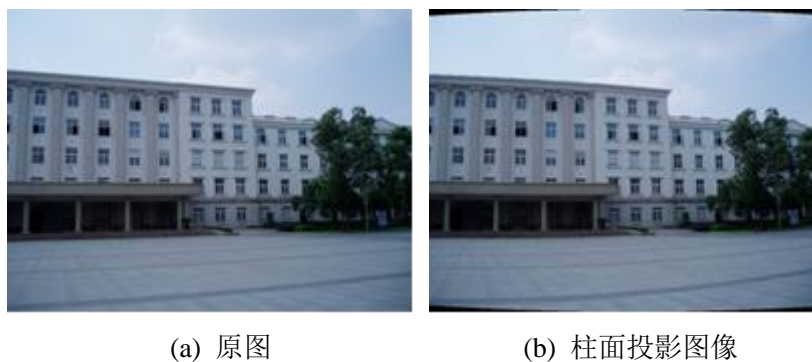


图 2-10 柱面校正实验结果

2.5 小结

本章首先详细介绍了旋转相机拍摄、平移相机拍摄和手持相机拍摄这三种图像采集方式。然后介绍了图像拼接的预处理过程，主要针对图像的去噪和图像畸变的校正，并深入研究了图像的畸变。引入了基于几何模型的摄像头失真校正方法，该方法可以简单快速的实现图像失真的校正。而后介绍了几种常用的图像几何变换模型。最后详细介绍了立方体投影、球面投影和柱面投影这三种全景图像的投影方式，为后面的全景拼接打下了基础。

第三章 基于特征的图像配准算法

3.1 概述

3.1.1 图像配准技术

图像配准可以认为是在不同时间或相同时间、从不同视角或相同视角对同一场景拍摄的两幅或者多幅图像进行空间域上的匹配过程^[42-43]。把两幅图像分别定义为已知的二维矩阵， $I_1(x, y)$ 和 $I_2(x, y)$ 分别表示两幅图像的灰度空间分布，则可以利用式(3-1)来表示两幅图像间的映射关系

$$I_2(x, y) = g(I_1(f(x, y))) \quad (3-1)$$

式中 f 为二维空间域的坐标变换， g 是一维图像灰度变换，所以图像配准的问题就是求解最优函数 f 和 g 的过程。

在理想情况下两幅图像的重叠部分完全相同，那么图像配准是一件非常容易的事情。然而在实际的图像配准过程中两幅图像往往是在不同条件下获取的，比如不同时间、不同的成像位置、不同的光源或是不同的传感器等等。这些条件使两幅图像的重叠区域不可能是完全相同的，所以说图像配准是一个相当复杂的过程。图像配准可以看为是以下几个方面的组合。

(1) 特征空间的选择：需要配准的图像存在着很多的可供选择的特征来实现图像的配准，如图像本身的灰度或亮度、图像的边缘、角点、直线、曲线、不变矩、重心等。

(2) 相似性度量的选择：它是图像配准过程中非常重要的一步，它直接决定着如何来确定图像的匹配位置。

(3) 搜索空间与搜索策略的选择：搜索空间就是指多个可能的配准位置的集合，在一般情况下减少测量数量和误匹配位置是非常重要的。通常情况下可以利用一些我们可以获得的信息来减少不可能的搜索子空间的数量。图像配准时选择合适的搜索策略也是对减小计算量非常有作用的。

要想改善图像配准算法的性能就需要从以上三个方面出发进行改进才能提高配准的精度提高配准的速度。

根据图像变换域的不同可以把图像配准算法分为基于频域的图像配准和基于

空间域的图像配准两大类方法。基于空间域的图像配准方法又可以分为基于灰度和基于特征的图像配准方法。

3.1.2 基于频域的图像配准算法

基于频域的图像配准算法就是首先将图像变换到频域再做相关的处理实现图像的配准。基于 FMT 的相位相关图像配准算法是一种典型的基于频域的图像配准算法，它是傅里叶相关算法的扩展。FMT 可以直接处理旋转和缩放的图像，它利用了图像的全部信息进行配准，克服了相位相关性噪声和频率噪声的影响，减小了几何畸变对图像匹配的影响，并且对于光照变换不敏感。

设 f_1 和 f_2 是具有平移和缩放关系的两幅待拼接图像，下面是基于 FMT 图像配准的步骤：

- ① 对图像 f_1 和图像 f_2 分别做傅里叶变换得到其频谱幅度 M_1 和 M_2 ；
- ② 对图像 f_1 和 f_2 的频谱幅度 M_1 和 M_2 做对数极坐标变换，利用相位相关法得到 f_1 和 f_2 的旋转角度 α 和缩放因子 σ ；
- ③ 根据②步中求出的转角度 α 和缩放因子 σ 对 f_2 进行傅里叶反变换，得到图像 f_3 ，再对 f_1 和 f_3 使用相位相关法求出平移量 (x_0, y_0) ；
- ④ 根据已经求出来的转角度 α 和缩放因子 σ 和平移量 (x_0, y_0) 构造图像 f_1 和图像 f_2 的坐标变换矩阵；
- ⑤ 选择图像 f_1 和图像 f_2 中的一幅作为参考图像，将另一幅图像根据坐标变换矩阵变换到参考图像的坐标系上，最后实现两幅图像的拼接。

在进行基于 FMT 的图像配准算法时多次利用了相位相关法来求解两幅图像的平移参数。当只存在平移的情况下相位相关函数会呈现单峰值性，但是在实际中，使用离散傅里叶变换(Discrete Fourier Transform, DFT)代替连续傅里叶变换，使得运动物体和噪声等因素对相位相关函数有着较大的影响，相位相关函数可能会在正确的峰值附近出现很多伪峰值^[44]，这样就会造成图像的变换矩阵估计不准确。由于实际应用时要使用 DFT，就需要对原始图像进行周期沿拓，若图像存在着较大的灰度差异频谱中必然会增加高频分量，多余的这些高频分量就会降低图像的配准的准确性。基于 FMT 的图像配准还用到了对数极坐标变换，这个变换过程往往需要引入插值，插值的准确性也会影响图像配准的准确性。

3.1.3 基于灰度的图像配准算法

基于灰度信息的图像配准算法一般通过图像上选定的区域来建立模版，并利用图像的统计信息作为匹配的相似性度量。基于灰度的图像配准算法主要有：块模版配准算法、四边模版匹配法和金字塔分层模版匹配等。互相关法和交互信息法是基于灰度信息的图像配准算法中主要的相似性度量算法。基于灰度信息的图像配准算法有着原理简单的优点，但是计算量大，而且对噪声光照都很敏感。

3.1.4 基于特征的图像配准算法

基于特征的图像配准算法通过提取和匹配图像的特征来实现图像的拼接，一般基于特征的图像配准算法需要分为下面三步：

- (1) 根据实际情况以及图像的特性来提取用于匹配的特征；
- (2) 对提取的特征点进行匹配；
- (3) 将匹配好的特征位置带入图像变换模型中，求解变换模型参数。

如何从变化情况复杂的图像中提取图像特征是基于特征的图像配准算法的关键。常用的图像配准特征有特征点、边缘、轮廓、直线段、闭合区域和统计特征等^[45]。

使用区域统计特征作为图像配准特征的前提是可以较好的分割出区域。它使用灰度值作为图像匹配的基元从而避免了提取特征的过程中导入的误差。它可以利用到图像中比较多的图像信息，因此复杂的图像可以使用区域统计特征来实现图像的高精度配准。但是它处理的信息量过大，计算复杂度高，而且受光照、灰度变化影响较大，对图像的尺度变化、旋转以及遮挡极为敏感。

边缘特征是一种常见的特征，它是一种图像的直观特征。以边缘作为匹配特征具有孤立边缘点位置偏差对边缘线段影响小的优点。边缘检测方法主要有：差分边缘检测方法、Roberts 边缘检测算子、Prewitt 边缘检测算子、Laplace 边缘检测算子、Canny 边缘检测算子、Sobel 边缘检测算子等。

基于特征点的图像配准算法可用于相机旋转角度和平移距离较大的情况，它是最为常用的特征提取方法。基于特征点的图像配准方法主要有以下几点优点：

- (1) 基于特征点的图像配准方法具有较强的抗噪声能力，并且它对图像的灰度变化、形变以及局部遮挡都有较强的适应能力，提高了图像配准的鲁棒性。
- (2) 图像中特征点的数目远小于图像像素点的个数，基于特征点的配准方法只需要处理特征点，减小了运算量提高了图像配准的速度。

(3) 特征点对位置变化比较敏感，具有较高的图像配准精度，适用于具有较复杂图像变换的模型。

3.2 基于角点检测的图像配准

基于角点检测的图像配准算法就是以提取的角点作为特征点，利用基于归一化互相关、基于像素差平方和等特征匹配方法进行匹配实现图像拼接。常用的角点有 Moravec 角点、SUSAN 角点和 Harris 角点^[46]。

3.2.1 Moravec 角点检测算法

Moravec 角点检测的基本思想是设计一个局部检测窗口，当窗口移动时计算窗口的平均能量变化超过阈值时窗口的中心点就提取为角点。

设 $I(x, y)$ 为图像，平移像 (u, v) 的灰度变换计算公式为：

$$E(u, v) = \sum_{x, y} W(x, y) |I(x+u, y+v) - I(x, y)|^2 \quad (3-2)$$

式中 $W(u, v)$ 是图像窗口，平移 (u, v) 表示了水平、垂直、对角线和反对角线四个方向。Moravec 角点检测算法简单速度快，但是它只计算四个方向上的角点响应，而且对强边界非常敏感。

3.2.2 SUSAN 角点检测算法

最小核同值区 (Smallest Univalve Segment Assimilating Nucleus, SUSAN) 角点利用一个基于像素临域的近似圆形模版，它对每个像素在模版邻域内计算角点响应函数，如果大于阈值且为局部极大值则认为此像素点为角点。以图像中非纹理区域上的任意一点为中心的模版存在一块与它亮度相同的区域，称为核同值区 (Univalve Segment Assimilating Nucleus, USAN) 区域。SUSAN 角点通过在不同位置时 USAN 区域的面积来判断当前点是区域内部的点、边缘点或是角点。通过将圆模版内的像素点与中心像素点比较得到的相似像素点的个数来确定 USAN 区域的面积，比较函数如下：

$$comp(x_0, y_0, x, y) = \begin{cases} 1 & |I(x, y) - I(x_0, y_0)| \leq t \\ 0 & |I(x, y) - I(x_0, y_0)| > t \end{cases} \quad (3-3)$$

式中 $I(x_0, y_0)$ 和 $I(x, y)$ 分别为模版中心像素点和待匹配像素点， t 为阈值，角点的

计算表达式为

$$R_N(x_0, y_0) = \begin{cases} j - n(x_0, y_0) & n(x_0, y_0) < j \\ 0 & n(x_0, y_0) \geq j \end{cases} \quad (3-4)$$

SUSAN 角点具有一定的抗噪能力，可以准确检测出角点，但是由于圆形模版的设计使得它在实现时非常困难。

3.2.3 Harris 角点检测算法

Harris 角点检测是一种基于灰度的角点检测方法，它通过计算像素点的曲率来实现角点检测。其基本思想与 Moravec 角点检测方法相似。设 $f(x, y)$ 是位于 (x, y) 处像素点的灰度值，该像素点移动 (u, v) 时期灰度变换为：

$$E_{u, v}(x, y) = \sum_{x, y} w_{u, v} |I(x+u, y+v) - I(x, y)|^2 \quad (3-5)$$

式中 (u, v) 的取值为：(1,0), (1,1), (0,1), (-1,1), (-1,0), (-1,-1), (0,-1), (1,-1)。Harris 角点利用自相关矩阵和微分算子来检测角点，微分算子反应了像素点在任意方向上的灰度变化。它的检测窗口为高斯函数，对图像平滑后再提取角点可以在一定程度上减小噪声的干扰。

$$E_{u, v}(x, y) = \sum_{u, v} w_{u, v} \left(u \frac{\partial I}{\partial x} + v \frac{\partial I}{\partial y} \right)^2 \quad (3-6)$$

式中 $w_{u, v}$ 是高斯窗位于 (u, v) 处的系数。

Harris 角点检测算法是最为常用的一种角点检测算法，它是角点检测中的经典。但是它只能在单一尺度下检测，角点检测的准确性也受阈值设置大小的影响。Harris 虽然利用了高斯函数，但是在实际的运用中高斯窗的大小不易确定，高斯窗选择过大会增加噪声的影响，若高斯窗旋转过大则会使角点的位置发生较大的偏移。

图 3-1 为 Harris 角点检测的结果，图 3-1(a)为对一个较简单棱角分明的图像做角点检测的结果，图 3-1(b)为对一幅实际拍摄获得的一幅较为复杂的图像做 Harris 角点检测的结果。可以看出 Harris 对于角点比较明显的图像设置正确的阈值可以成功的完成角点检测，对于较为复杂的图像阈值不易设置容易出现误检测点。但是在图像拼接时角点检测有着简单易实现的优点，在以建筑等为主要内容的简单棱角分明场景中较为适用，但不适合用于纹理复杂、形变较大、发生较大尺度变换的图像的拼接中。

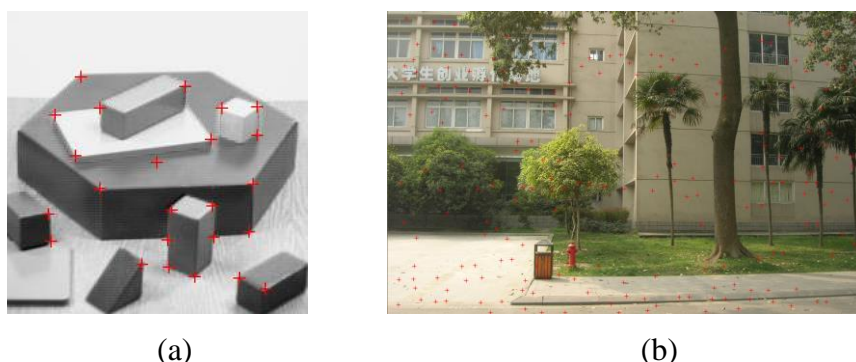


图 3-1 Harris 角点检测结果

3.3 基于不变量的图像配准

3.3.1 SIFT 特征的提取

尺度不变特征变换即 SIFT 对于图像的缩放、平移、旋转、仿射变换都具有不变性，而且对噪声和光照变化都具有良好的鲁棒性，所以基于 SIFT 特征的图像配准算法对于旋转和尺度变化较大的图像也具有较好的效果^[47]。SIFT 特征的提取可以分为检测尺度空间极值、确定特征点位置、确定特征点主方向和生成 SIFT 特征描述符四步。

3.3.1.1 检测尺度空间极值

高斯核是唯一的线性核，它是实现尺度变换的唯一变换核，因此尺度空间就是要利用不同尺度的高斯核对图像进行尺度变，然后在得到的不同尺度空间中提取特征。二维高斯核定义如下：

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (3-7)$$

式中 σ 表示高斯核的标准差，这里也称为尺度因子

设 $I(x, y)$ 是一个二维图像， $L(x, y, \sigma)$ 是图像 $I(x, y)$ 在不同尺度空间下的表示，它可以通过与高斯核的卷积得到：

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (3-8)$$

式中 (x, y) 表示图像 $I(x, y)$ 上的点， $*$ 表示在图像的 x 和 y 方向上的卷积。尺度因子 σ 的值越大图像就会变得越平滑，不同的尺度因子对应着图像不同的特征，因此说选择适当的尺度因子 σ 对建立合适的尺度空间有着重要的意义。

尺度空间的极值检测主要是构建高斯差值(Difference of Gaussian, DOG)和高斯差值金字塔,然后在 DOG 空间里检测极值来初步确定特征点的位置和尺度。DOG 空间的使用可以提高在尺度空间检测特征的效率。相邻尺度空间的差值即 DOG 用 $D(x, y, \sigma)$ 表示:

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (3-9)$$

式中 k 为常数,文中计算时取 $\sqrt{2}$ 。

实际提取尺度不变特征点时就要用到高斯差值金字塔,检测尺度空间极值的具体步骤如下:

(1) 建立高斯金字塔:将图像 $I(x, y)$ 与不同的尺度因子下的高斯核 $G(x, y, k\sigma)$ 做卷积,得到图像的不同尺度空间。获得的这一组图像作为图像金字塔的第一阶(Octave)。在同一阶中相邻两层尺度因子的比例为 k 。图像金字塔第二阶的第一幅图像根据金字塔第一阶的第一幅图像变换得到。首先求出第一阶第一幅图像的 2 倍尺度图像,然后对得到的新图像以 2 倍距离下采样,这样就得到了第二阶的第一幅图像。再利用不同尺度因子的高斯核做卷积得到第二阶的图像。以此类推就获得了高斯金字塔图像。高斯金字塔图像和 DOG 金字塔图像建立的过程如图 3-2。

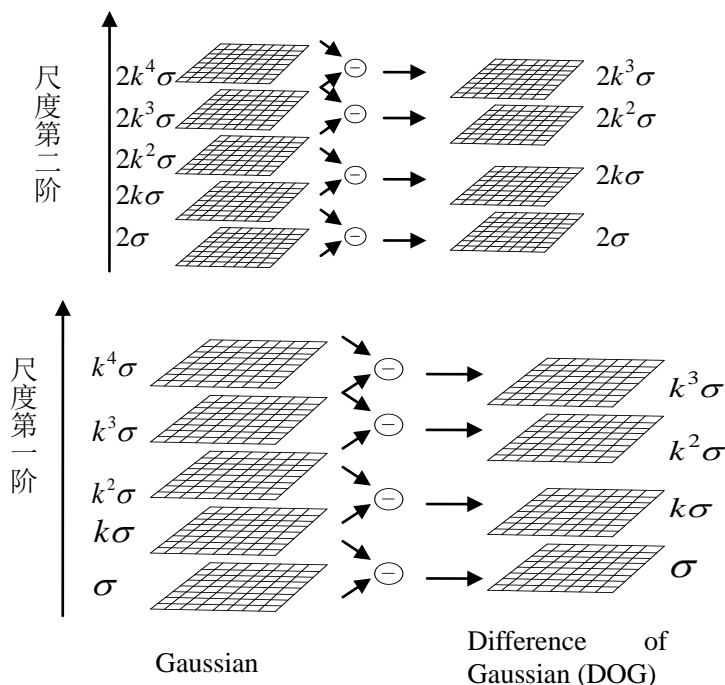


图 3-2 高斯图像与 DOG 图像的构造

(2) 建立 DOG 金字塔:对高斯金子塔中每一阶中相邻的高斯图像做差就得到

了 DOG 图像。对整个高斯金字塔进行如此的操作后得到的差值金字塔就是 DOG 金字塔。

(3) 把 DOG 尺度空间中中间层的每一个像素点与其在同一层相邻的 8 个像素点和其上一层相邻的 9 个像素点以及下一层相邻的 9 个像素点进行比较。如图 3-3，标记为叉号的像素点比其周围的 26 个像素点都小或者都大则记为一个局部极值点，并记录下它的位置和尺度。

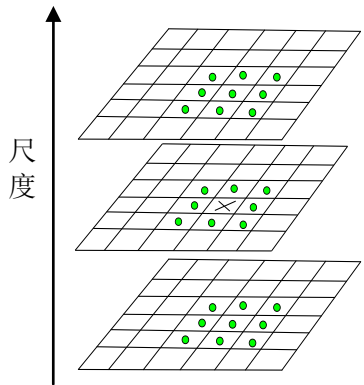
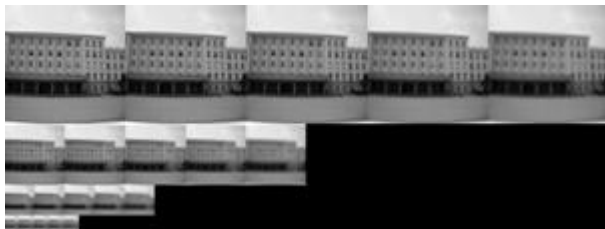


图 3-3 DOG 空间中的极值检测

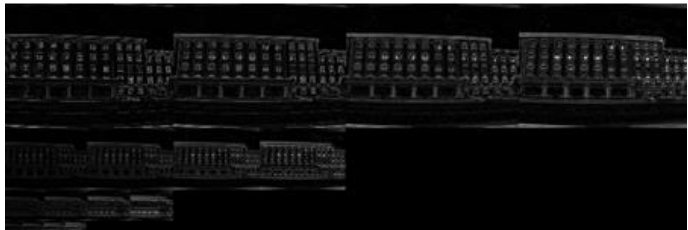
图 3-4(a)为用于检测 SIFT 特征点的原始图像，图(b)为根据(a)生成的 4 阶高斯金字塔图像，每一阶图像中包含 5 幅图像，图(c)为由图(b)为得到的 DOG 金字塔图像。



(a) 原图



(b) 高斯金字塔图像



(c) DOG 金字塔图像

图 3-4 SIFT 算法中生成的金字塔图像

3.3.1.2 确定特征点位置

检测到极值点后就要确定特征点的位置、尺度和曲率等信息。DOG 值对噪声和边缘比较敏感，需要将对对比度较低的和位于边缘的特征点过滤掉。首先利用泰勒级数把尺度空间方程 $D(x, y, \sigma)$ 展开：

$$D(X) = D + \frac{\partial D^T}{\partial X} + \frac{1}{2} X^T \frac{\partial^2 D}{\partial^2 X} X \quad (3-10)$$

式中 $X = (x, y, \sigma)^T$ 为样本点的偏移量。对式(3-10)求导，并令其等于零，可以计算出极值的精确位置 \hat{X}

$$\hat{X} = -\frac{\partial^2 D^{-1}}{\partial X} \frac{\partial D}{\partial X} \quad (3-11)$$

把式(3-10)带入式(3-11)可以得到：

$$D(\hat{X}) = D + \frac{1}{2} \frac{\partial D^T}{\partial X} \quad (3-12)$$

为了去除对比度较低的不稳定极值点，把 $|D(\hat{X})| < 0.03$ 的点视为低对比度点全部去掉。

一个 DOG 空间极值点的主曲率的大小跟其所在的位置有关，它在与边缘相互垂直处较小，而在边缘处较大，所以可以根据这一特点来去除位于边缘的极值点。设 H 是一个 2×2 的 Hessian 矩阵。

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (3-13)$$

H 的特征值与 D 的主曲率成正比，这里只需要计算出特征值之比就可以去除位于边缘的不稳定极值点。设 H 的最大和最小的特征值分别为 α 和 β ，他们的比值为 r ，则有。

$$Tr(H) = D_{xx} + D_{yy} = \alpha + \beta \quad (3-14)$$

$$Det(H) = D_{xx} D_{yy} = \alpha \beta \quad (3-15)$$

$$\frac{Tr(H)}{Det(H)} = \frac{(\alpha + \beta)^2}{\alpha \beta} = \frac{(r\beta + \beta)^2}{r\beta^2} = \frac{(r+1)^2}{r} \quad (3-16)$$

式中 r 为阈值，一般取 $r=10$ ，为了检测主曲率 D 是否小于 r ，只需要判断式(3-17)是否成立，若满足关系则保留特征点否则丢弃。这样就得到了 SIFT 特征点，图 3-5

就是一张经过 SIFT 特征检测得到的 SIFT 特征点。

$$\frac{Tr(H)}{Det(H)} < \frac{(r+1)^2}{r} \quad (3-17)$$



图 3-5 SIFT 特征点

3.3.1.3 确定特征点主方向

为了确保生成的 SIFT 特征描述符具有旋转不变性，需要利用特征点邻域像素的梯度特性为每个特征点指定一个主方向，在像素点 (x, y) 处梯度的模值和方向分别为：

$$m(x, y) = \sqrt{[L(x+1, y) - L(x-1, y)]^2 + [L(x, y+1) - L(x, y-1)]^2} \quad (3-18)$$

$$\theta(x, y) = \arctan((L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y))) \quad (3-19)$$

在以特征点为中心的邻域内采样，该特征点处的梯度方向利用梯度方向直方图来计算。直方图把 360 度的范围平均划分成 36 等份来统计，找到梯度方向直方图的峰值，这个峰值就是特征点的主方向。如果存在一个能量相当于主峰值 80% 的峰值，则将它作为特征点的辅方向。特征点可能不存在辅方向也可能存在一个或多个辅方向，这样也增强了算法的鲁棒性。

图 3-6 为生成的 SIFT 的方向，每个特征点的方向中较长的箭头表示特征点的主方向，较短的箭头表示特征点的辅方向，从图中可以确定每一个特征点必存在一个主方向，它可能有没有辅方向也可能有多个辅方向。



图 3-6 SIFT 特征方向

3.3.1.4 生成 SIFT 特征描述符

生成 SIFT 特征点的描述符时，首先要把坐标轴旋转与特征点的方向一致，以特征点为中心取一个 8×8 的邻域窗口。如图 3-7(a)所示，每一个小方格表示一个像素点，图中的圆圈是加权范围，箭头的方向是这个像素点的梯度方向，箭头的长度就是像素点梯度的模值。然后在每个 4×4 的方块上计算 8 个方向的梯度方向直方图，累加每个方向的值就生成了一个种子点，如图 3-7(b)。每一个特征点具有 4 个种子点，每个种子点有 8 个方向，这样就形成了一个 32 维的 SIFT 特征描述符。

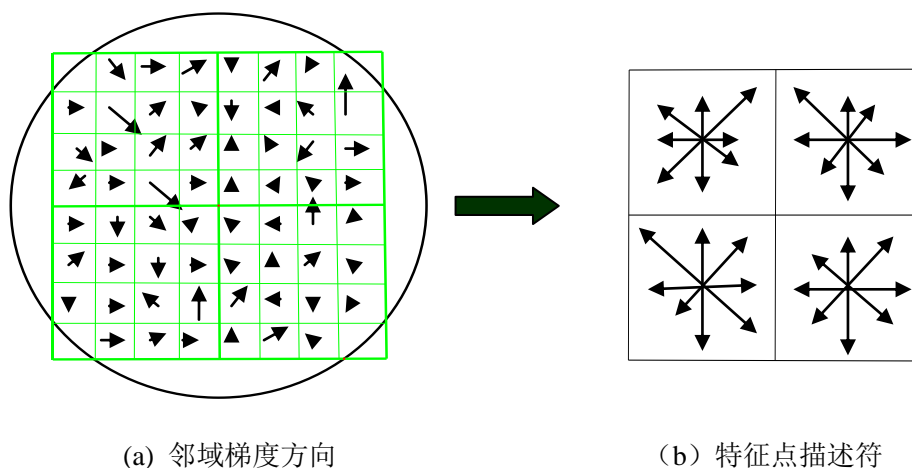


图 3-7 图像梯度及特征点描述符

在实际的应用过程中，为了增加匹配的稳定性可以以特征点为中心取一个 16×16 的邻域窗口，每个特征点使用 4×4 共 16 个种子点，这样每一个特征点生成了一个 $4 \times 4 \times 8$ 共 128 维的特征向量。应用时一般把 SIFT 特征向量归一化，这样可以减轻光照变化的影响。

3.3.2 SURF 特征的提取

SURF 算法就是一种加速的 SIFT 算法,它利用了盒子滤波、积分图像以及 Haar 小波实现了算法的加速

3.3.2.1 检测尺度空间极值

在 SURF 特征检测尺度空间极值时利用盒子滤波近似 SIFT 算法中的二阶高斯函数,并利用了积分图像进行卷积,这就是 SURF 算法和 SIFT 算法的本质区别^[48]。对于灰度积分图像来说在积分图像上任意一点 $X = (x, y)$ 的值为该点与图像原点形成的矩形区域中的所有图像的灰度值总和:

$$I_{\sum x} = \sum_{i=0}^x \sum_{j=0}^y I(i, j) \quad (3-20)$$

式中 $I(i, j)$ 表示位于原始灰度图像 (i, j) 处的灰度值

图 3-8 中积分图像中的矩形区域 ABCD 中的灰度值之和利用式(3-20)可以表示为式(3-21)的形式,区域内的灰度和的计算时间与矩形的面积无关,这种方法针对大尺度图像尤其有效。

$$\sum = A - B - C + D \quad (3-21)$$

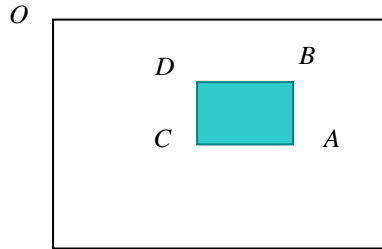


图 3-8 积分图像示意图

图 3-9 为盒子滤波的示意图,图(a)、(b)和(c)分别为 x 方向上的滤波、 y 方向上的滤波以及 xy 方向上的滤波。积分图像和盒子滤波的使用是 SURF 比 SIFT 算法速度快的主要原因。SIFT 算法构建尺度空间金子塔时,通过上一层的图像不断的下采样来调整下一层图像的大小。而 SURF 不需要对图像进行迭代下采样,它通过调整盒子滤波的大小,利用不同大小的盒子滤波与原始图像在不同方向上卷积来计算多尺度空间函数 D_{xx}, D_{yy}, D_{xy} 。

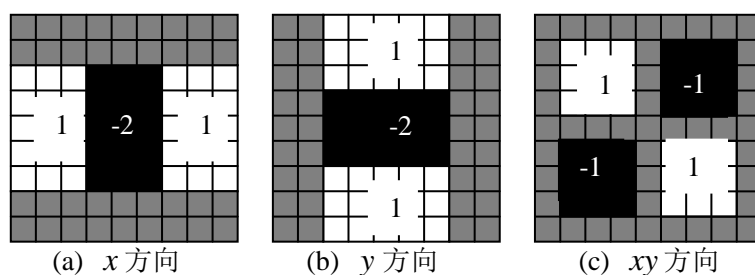


图 3-9 盒子滤波示意图

在 SURF 算法中尺度空间金字塔的最底层利用一个大小为 9×9 的盒子滤波得到，它对应二阶高斯滤波的 $\sigma = 1.2$ 。为了保证盒子滤波的结构不变，后面的盒子滤波大小至少要有 6 个像素步长的变换，即第一层为 9×9 第二层最小要为 15×15 大小。式(3-22)为当前大小的盒子滤波对应的高斯尺度值的计算公式

$$\sigma_{approx} = \frac{1.2}{9} N \quad (3-22)$$

式中 N 表示当前的盒子滤波尺寸。

假设要构建一个大小为 4 度 4 层的尺度空间，如图 3-10 所示，最低层的盒子滤波器的大小为 9×9 ，各层中盒子滤波器的大小变化分别为 6、12、24、48。

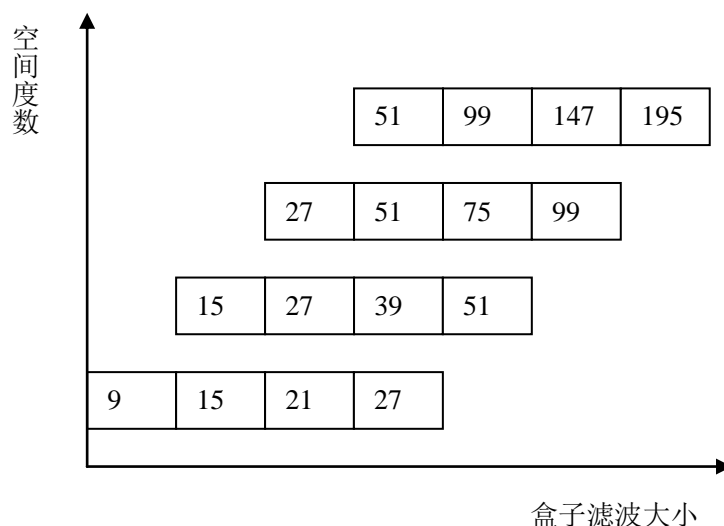


图 3-10 滤波器尺寸分布图

跟 SIFT 相同，在金字塔图像后把每个像素与其相邻的 26 个像素比较，比相邻的像素点都大或者都小的像素点才能成为候选极值点，并记录下它的位置和尺度。

3.3.2.2 确定特征点位置

得到候选极值点后需要确定特征点的位置和消除不稳定的极值点，在 SIFT 算法中利用 Hessian 矩阵的行列式 $\det(H)$ 来完成，而在 SURF 算法中利用一个近似 $\det(H)$ 的 $\Delta(H)$ 确定特征点的位置。

$$\Delta(H) = D_{xx}D_{yy} - (0.9D_{xy})^2 \quad (3-23)$$

如果 $\Delta(H)$ 为正就可以确定该点为一个极值点。去除不稳定极值点的方法与 3.3.1 小节中 SIFT 使用的方法相同，但是盒子滤波不存在边缘的影响，不需要去除边缘点。

3.3.2.3 确定特征点主方向

SURF 算法为了提高其鲁棒性利用了 Haar 小波，Haar 小波是一种简单的滤波器可以用来计算水平和垂直方向上的梯度。图 3-11 为 Haar 小波示意图，图(a)为 x 方向图(b)为 y 方向。

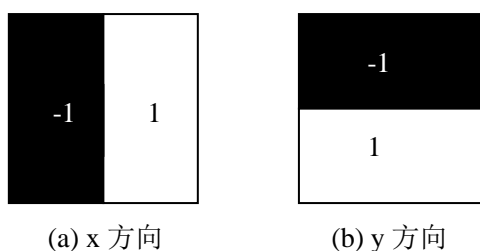


图 3-11 Haar 小波示意图

首先以特征点为中心，以 6σ 为半径做圆，在这个圆形区域内计算其中的点在 x 方向和 y 方向上大小为 4σ 的 Haar 小波响应，其中 σ 表示特征点所在的尺度空间的尺度值。然后对这些响应加权值，使距离特征点越近的响应具有越大的权重值，而距离特征点越远的响应具有较小的权重值。最后以 60 度范围作为一个区域，遍历整个圆形区域，可以得到 6 个扇形区域。把每个区域内的响应相加形成一个新的矢量，在 6 个新的矢量中找到模值最大的矢量，把其方向作为该特征点的主方向，如图 3-12 所示。这样就得到了 SURF 特征点的主方向，同时也保证 SURF 特征的旋转不变性。

3.3.2.4 生成 SURF 特征描述符

SURF 特征的描述符的构造与 SIFT 特征描述符的构造类似，首先要把坐标轴旋转到 SURF 特征的主方向，以特征点为中心构造一个 $20\sigma \times 20\sigma$ 的正方形窗口，

σ 为特征点所在尺度空间的尺度值。以 5×5 大小为一个小块，划分窗口，计算每个小块中采样点相对于主方向水平与垂直方向的 Haar 小波的响应，把每一个小块中的 25 个采样点的响应值以及他们的绝对值分别进行累加，所得结果记为 $\sum d_x$ ， $\sum d_y$ ， $\sum |d_x|$ ， $\sum |d_y|$ 。这样每一个子区域就形成一个 $4 \times 4 \times 4 = 64$ 维的向量，对其归一化除去亮度变化的影响就得到了 SURF 特征的描述符。对于大于 0 和小于 0 的 d_x 和 d_y 分别计算 $\sum d_x$ ， $\sum d_y$ ， $\sum |d_x|$ ， $\sum |d_y|$ 这样就可以把 SURF 特征描述符扩展为 128 维。

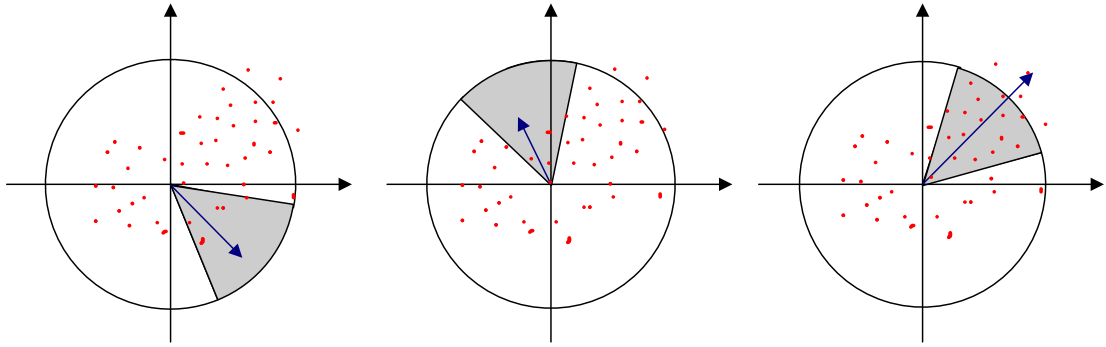


图 3-12 SURF 特征主方向的构造

3.3.3 特征的匹配

SIFT 特征和 SURF 特征的匹配方法基本相同，利用特征描述符之间的欧式距离作为特征点的匹配相似性度量准则是一种比较简单有效的特征匹配准则。设 p 和 q 是一对 128 维的特征点，他们的特征描述符为 Des_p 和 Des_q ，则这对特征点描述符之间的欧式距离为：

$$d = \sqrt{\sum_{i=0 \dots 127} (Des_p(i) - Des_q(i))^2} \quad (3-24)$$

搜索每一个特征点的欧氏距离最近邻特征点和次近邻特征点，设 p 为特征点，其欧氏距离最近特征点为 q ，次近特征点为 q' 。然后分别计算特征点 p 与特征点 q 的欧氏距离 d_1 以及特征点 p 与特征点 q' 的欧氏距离 d_2 。 r 为欧氏距离 d_1 与 d_2 的比值，若果 r 小于阈值 T ，则特征点对 (p, q) 是图像序列的一个正确匹配的特征点对。根据特征点的最近邻域与次近邻域距离的比值概率分布，比值 r 小于 0.75 时检测到正确匹配的特征点（内点）比检测到错误匹配特征点（外点）的概率大。经实验验证阈值取在 0.4 到 0.6 之间较好，为了保证能更多的去除外点，本文中阈值 T 取为 0.5。

3.4 特征提取算法性能分析

为了确定使用的特征提取算法，本文对 **SIFT** 特征、**SURF** 特征以及 **Harris** 角点进行比较。利用重复率可以来判别特征提取算法的优越性，所谓重复率就是指原始图像和变换图像重复的特征点数目与两幅图像中较小的特征点总数的比值。图 3-13 为比较特征提取算法的实验图像，图像实验图像来自于 Mikolajczyk^[49]。图中图(a)为发生光照变化的一组图像，图(b)为发生图像模糊的一组图像，图(c)为发生旋转和缩放变换的一组图像，图(d)为拍摄视点发生变化的一组图像。

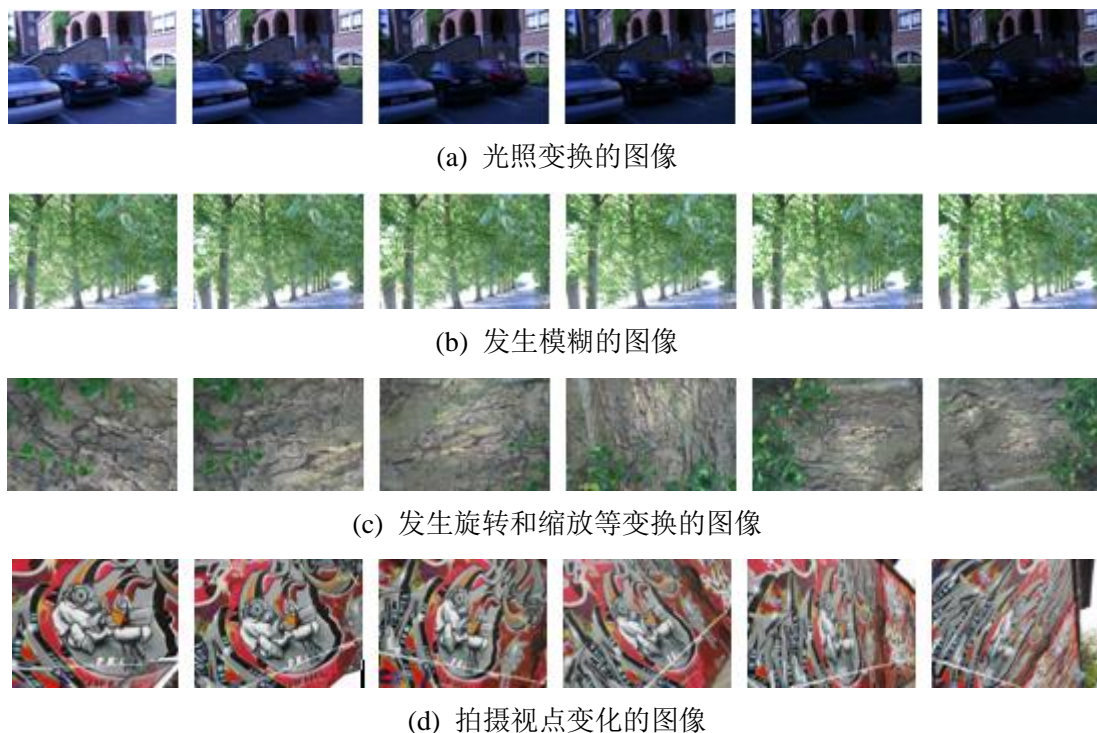


图 3-13 特征提取算法比较试验图像

图 3-14 至图 3-17 分别为图 3-13 中(a)到(d)第一幅图像与该组其余图像特征点重复率的测试结果。从图中可以看出 **SURF** 特征和 **SIFT** 特征性能差别不大，**SURF** 特征在视点变化、光照变化以及旋转缩放变换上略于 **SIFT**，而在模糊上略强于 **SIFT**。从图中还可以看出 **Harris** 角点对尺度变换和几何形变过于敏感。

表 3-1 为对图 3-13 中的(a)到(d)四组图像中的第一幅图像分别提取 128 维的 **SIFT** 特征点和 64 维的 **SURF** 特征点所用时间，实验在 CPU 主频为 2.4GHZ，内存为 2G 硬件条件下在 Window XP 操作系统中以 VC 作为编程平台的提取特征点。可以看出对同一幅图像提取特征，**SIFT** 算法所用的时间约为 **SURF** 算法的 3 倍。

在提取的性能方面 SURF 算法与 SIFT 算法相似，而速度明显要快。

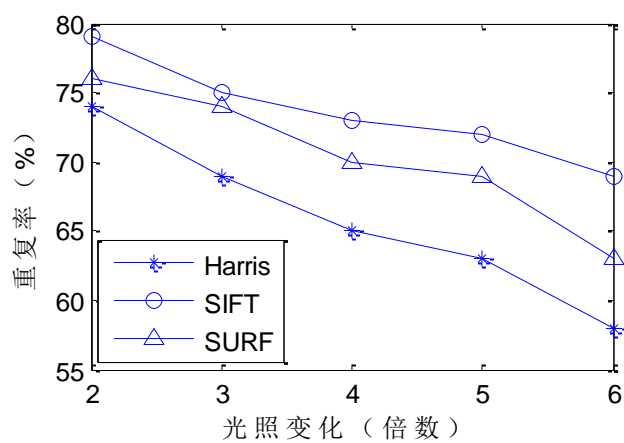


图 3-14 随光照变换的特征点重复率

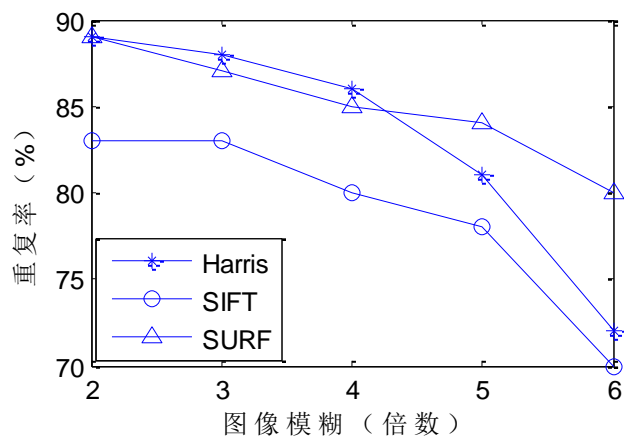


图 3-15 随图像模糊变换的特征点重复率

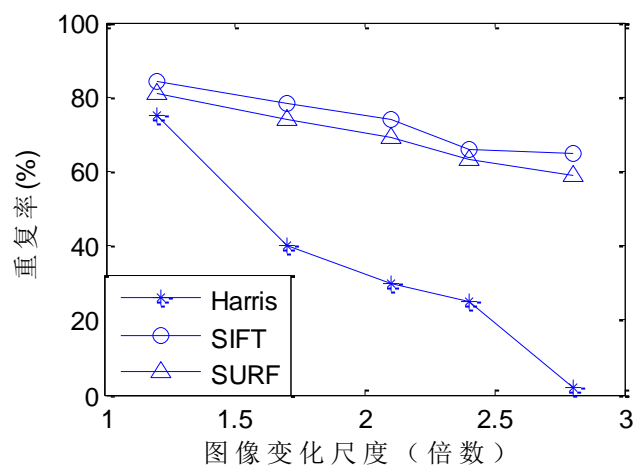


图 3-16 随尺度变换及旋转的特征点重复率

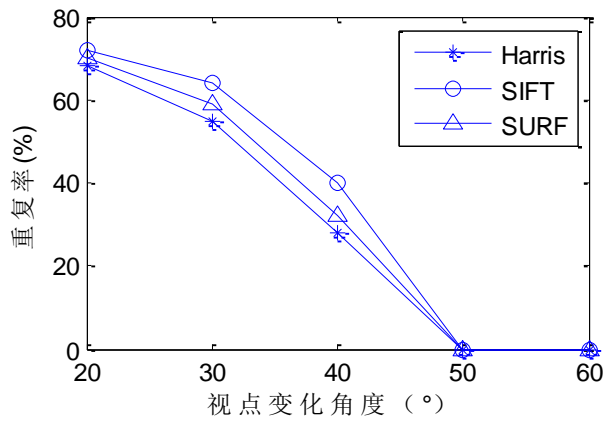


图 3-17 随视点变化的特征点重复率

表3-1 SIFT算法和SURF算法提取特征用时对比

图像	提取SIFT特征时间 (ms)	提取SURF特征时间 (ms)
图3-13(a)中第一幅图像	4344	1331
图3-13(b)中第一幅图像	10860	3543
图3-13(c)中第一幅图像	4438	1405
图3-13(d)中第一幅图像	4178	1278

3.5 特征匹配方法

在 3.4 节中提到可以利用特征点的欧氏距离最近点与次近点的欧氏距离比作为特征点匹配的方法，虽然该方法力求准确，但误差仍是不可避免的。为了降低误匹配特征点对的数量往往需要进行两次特征匹配。一般第一次匹配为初匹配初步确定特征点对，然后再进行一次精确匹配去除错误匹配的特征点对。

3.5.1 K-D 树特征点搜索方法

在进行基于特征点的最近邻与次近邻的欧氏距离比的特征点匹配时，首先需要搜索特征点的最近邻和次近邻特征点，搜索方法也是影响特征点匹配准确性的一个重要因素。基于优先队列的 K-D 树是多维二叉树的扩展，它可以快速有效的搜索到特征点的最近邻和次近邻特征点。

K-D 数的根节点下的子树分为左子树和右子树，其下一个子树是左子树还是右子树由识别器的值来决定。对于 SURF 特征而言它的识别器生成过程为：设特征点的个数为 n ，SURF 特征点描述符是一个 128 维的向量，首先要计算出 128 维

特征的方差，找出方差最大的那一维，记为 k 。然后找到 n 个特征点中第 k 维特征的中位数，记为 F_m ，它就是生成的识别器。

在划分左右节点时，首先把识别器对应的那一个特征点作为根节点，然后将剩余节点的第 k 维特征与 F_m 进行比较，若小于 F_m 则划分到左子树，否则就划分到右子树。再对所有特征点做上述操作，不断重复直到节点划分结束。由于识别器 F_m 是 n 个特征点中第 k 维特征的中位数所以左右子树的节点个数大体上是相等的。

在进行 K-D 树搜索时，不断的用识别器与特征点描述符相应维度的特征比较来缩小搜索范围。设 f_1 和 f_2 为两幅待匹配的图像，他们的特征点集合分别为 F_1 和 F_2 ，则 K-D 树的搜索过程如下：

① 根据 K-D 树节点的划分方法把集合 F_1 中的特征点写为 K-D 树的数据结构，记为 T_1 ，它的识别器为 F_m ，维度是 k 。

② 把 F_2 中的特征点的第 k 维特征与 T_1 中的识别器 F_m 比较。如果特征点的第 k 维特征大于 F_m 则继续与右子树进行比较，将左子树放入优先队列。如果特征点的第 k 维小于 F_m 则继续与左子树进行比较，右子树放入优先队列中。不断比较直到完成叶子节点的比较。

③ 再对优先队列中的节点进行比较，保留欧氏距离最小和第二小的节点，这两个节点就是要搜索的最近邻和次近邻特征点。

从进行 K-D 树的搜索过程可以看出 K-D 树的搜索策略避免了一定不必要的比较，从而快速有效的实现了最近邻和次近邻特征点的搜索。

3.5.2 随机抽样一致算法原理

使用任何一种特征匹配算法都难免存在误匹配的特征点对，消除匹配过程中的误匹配特征点的方法主要有霍夫聚类法、最小中位数法(Least Median of Square, LMedS)和随机抽样一致法(Random Sample Consensus, RANSAC)。RANSAC 算法是其中一种比较有效的方法，它是一种鲁棒性的数据拟合算法。它用某一种模式来估计数据，然后采用随机抽样的方法来剔除错误数据^[50]。假设集合 F 中存在 N 个数据，这 N 个数据由一个未知模型产生，RANSAC 要解决的问题就是求解这个未知模型。我们可以利用线性拟合来解释 RANSAC 算法的主要思想，如图 3-18 所示。首先设定一个判断集合中的点为某条直线内点的阈值，在图 3-18(a)的点中随机选择两个点构造成为一条直线，根据构造直线的内点集合计算出一条新的直

线。再根据阈值和新的直线求出新直线的内点集合，不断的进行这样的过程，直到内点的数量不再增加，这时获得的直线就是拟合出的最佳直线，如图(b)。

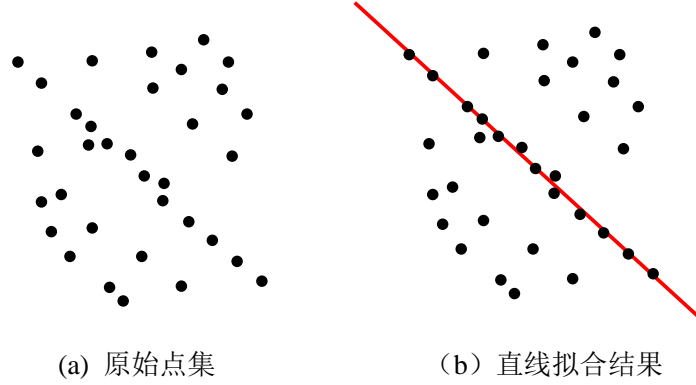


图 3-18 直线拟合示意图

在实际进行 RANSAC 特征点的提纯时，首先在集合 F 的全部 N 个数据中随机选择 m 个数据，然后根据之前选择的 m 个数据计算出模型 P ，最后计算模型 P 上小于阈值距离的点的个数 k 。不断重复上述步骤直到求出具有最大 k 值的模型。

3.5.3 图像变换模型参数的估计

对于全景图像的拼接来说需要使用全局配准方法来消除多幅图像拼接的累积误差，全局配准的方法主要有多帧调整对齐、线性方程求解、基于对应点的多帧调整和逐步扩大拼接法等。将图像投影到柱面图像坐标系再进行配准的方法就是一种逐步扩大的全局配准方法。在得到匹配的图像特征点对之后就需要计算出图像的变换模型参数来实现图像的配准。在前面的第二章中曾提到过图像间的变换模型大多数情况下可以利用投影模型来表示，八参数投影模型变换的齐次坐标可以利用式(2-14)表示。式(2-14)也可以展开，写成下面的形式：

$$\begin{cases} x' = \frac{m_0x + m_1y + m_2}{m_6x + m_7y + 1} \\ y' = \frac{m_3x + m_4y + m_5}{m_6x + m_7y + 1} \end{cases} \quad (3-25)$$

设 $m = [m_0, m_1, \dots, m_7]^T$ ，可以把式(3-25)写为以下形式：

$$\begin{cases} (x & y & 1 & 0 & 0 & 0 & -xx' & -yx')m = x' \\ (0 & 0 & 0 & x & y & 1 & -xy' & -yy')m = y' \end{cases} \quad (3-26)$$

模型变换参数矩阵 M 一共有 8 个参数，所以只要知道 4 对任意 3 点不共线的

特征点就可以求出模型变换参数矩阵 M ，关于 m 的线性方程组 可以写为：

$$Ax = b \quad (3-27)$$

$$A = \begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1x_1' & -y_1x_1' \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1y_1' & -y_1y_1' \\ & & & & & \vdots & & \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x_4x_4' & -y_4x_4' \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -x_4y_4' & -y_4y_4' \end{bmatrix}, \quad b = \begin{bmatrix} x_1' \\ y_1' \\ \vdots \\ x_4' \\ y_4' \end{bmatrix} \quad (3-28)$$

式(3-27)可以利用最小二乘法来求解，若果 $A^T A$ 可逆，则它的解为：

$$m = (A^T A)^{-1} A^T b \quad (3-29)$$

这里需要使用 RANSAC 算法来精确提纯特征点，特征点匹配的误差使用几何距离 d 表示，设 X_{1i} 和 X_{2i} 是一对匹配的特征点，特征点的匹配误差 d 可以表示为：

$$d = d(X_{1i}, X_{2i}) + d(X_{1i}', X_{2i}') = \|X_{1i} - MX_{2i}\| + \|X_{2i} - M^{-1}X_{1i}\| \quad (3-30)$$

当 d 大于阈值 th 时，这对特征点被认为是误匹配点，文中阈值 th 取为 2 个像素点。

理想情况下，RANSAC 算法要搜索出所有可能的组合情况，但是这样计算量过大。如果可以确定一个采样次数 N 保证计算变换模型参数矩阵时需要的 4 对匹配特征点均为内点的概率足够高就可以不必搜索出所有的可能组合。在实际情况中只要保证 4 对匹配特征点均为内点的概率 P 大于 95% 就可以了。采样次数 N 与 4 对特征点都为内点概率的关系为：

$$1 - P = (1 - \varepsilon^4)^N \quad (3-31)$$

式中 ε 为任意一对特征点为内点的概率，这里假设特征点为内点的概率与为外点的概率相等，即 0.5。这样就可以计算出随机采样次数了：

$$N = \frac{\log(1 - P)}{\log(1 - \varepsilon^4)} \quad (3-32)$$

利用 RANSAC 算法和八参数模型估计方法计算变换模型参数矩阵步骤如下：

- ① 在特征点对中随机抽取 4 对任意 3 点不共线的特征点对，如果存在 3 点共线的情况则重新选取，直到不存在 3 点共线的情况为止。
- ② 利用最小二乘法根据第①步选取的四对匹配特征点求出变换模型的参数矩阵。
- ③ 利用公式(3-30)计算出所有匹配点对的距离 d ，当 d 小于阈值 th 时，这对匹

配特征点为内点，并记录下内点的个数 n ，若 d 大于阈值 th 则此对匹配特征点为外点。

④ 重复①到③的步骤，直到内点的数目 n 足够大为止，把具有最大内点数目 n 的模型作为最后的计算结果，并记录下所有内点的位置。

图 3-19 为本文的中图像配准算法的整体流程图，首先利用 SURF 特征提取算提取特征，利用欧氏距离进行特征点的初匹配，然后根据 RANSAC 算法和 8 参数数模型估计计算出变换模型参数矩阵，最后对齐图像实现图像的配准。

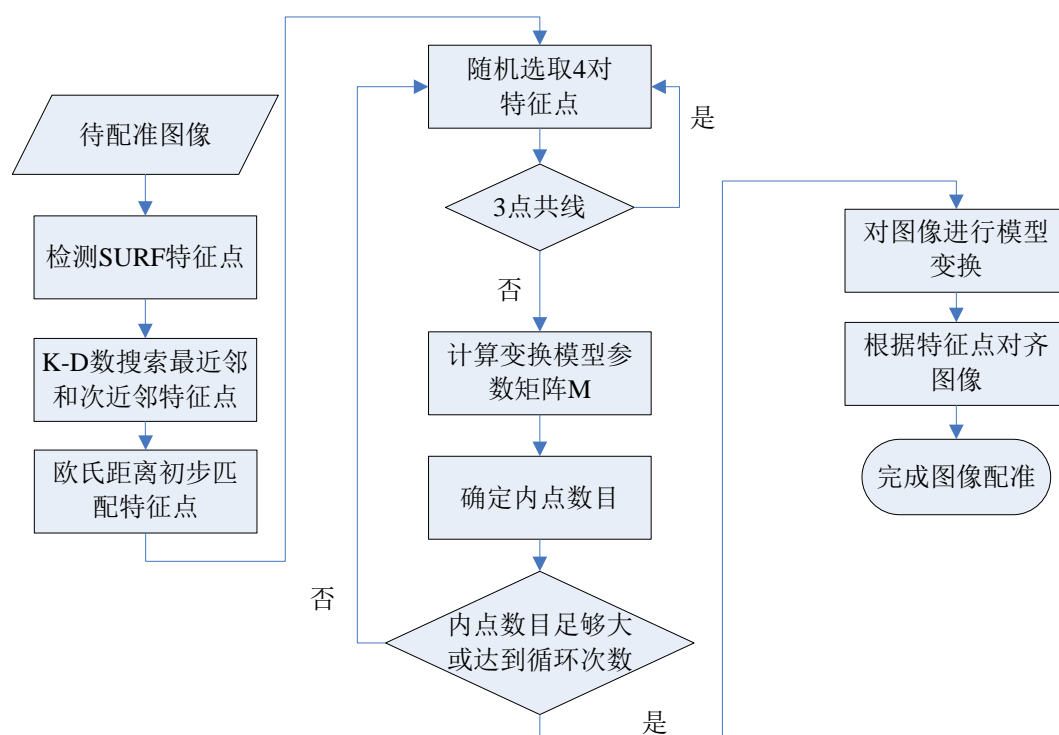


图 3-19 图像配准流程图

3.6 图像配准算法测试及分析

图 3-20(a)(b)为两幅通过手持普通数码相机拍摄得到的两幅图像，作为图像配准的原始图像，图(c)为利用欧氏距离匹配，再提纯特征点后的特征点匹配结果。

图 3-21 为利用文中的图像配准算法直接拼接的结果，从图中可以看到来自不同图像的两部分对齐比较准确可以连接成为一个物体，说明文中使用的图像配准方法有一定的可行性，对手持相机拍摄的图像也可以完成图像间的准确配准。

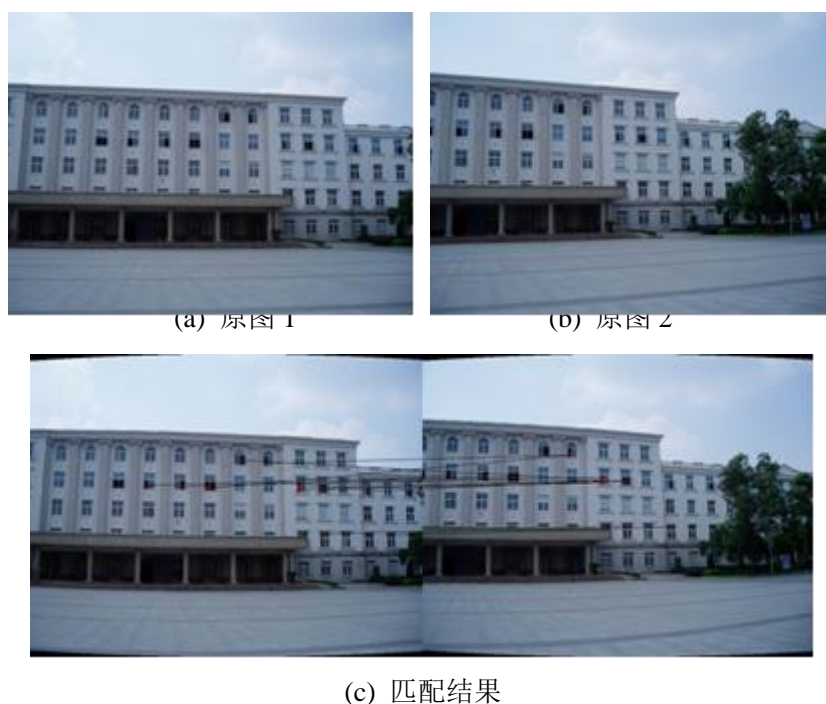


图 3-20 匹配的特征点



图 3-21 图像模型变换后拼接结果

差值图像可以用来评价两幅图像配准的准确程度，差值图像的计算公式为

$$diffI(i, j) = I_1(i, j) - I_2(i, j) \quad (3-33)$$

式中 $diffI(i, j)$, $I_1(i, j)$, $I_2(i, j)$ 分别为差值图像，原图 1 重叠区域和原图 2 重叠区域 (i, j) 处的像素点。图 3-21 为图 3-19(a)(b)配准后重叠区域的差值图像，差值图像的均值为 1.772，说明配准后两幅图像重叠区域差异不大，配准比较准确。

利用重叠区域的相关系数(Correlation Coefficient, CC)来定量的评价图像配准的准确性，相关系数的计算公式为

$$r = \frac{\sum_{i=1}^N (I_{1i} - \bar{I}_1)(I_{2i} - \bar{I}_2)}{\sqrt{\sum_{i=1}^N (I_{1i} - \bar{I}_1)^2} \sqrt{\sum_{i=1}^N (I_{2i} - \bar{I}_2)^2}} \quad (3-34)$$

式中 I_{1i} 和 I_{2i} 为重叠区域 1 和重叠区域 2 的第 i 个像素点， \bar{I}_1 和 \bar{I}_2 为重叠区域 1 的像素均值和重叠区域 2 的像素均值。



图 3-22 配准后重叠区域差值图像

配准后图像重叠区域的相关系数表征着他们的相关程度，对于两幅完全相同的图像来说他们的相关系数为 1，相关系数越接近于 1 说明他们越相似，通常大于 0.8 时认为两个变量有很强的线性相关性。配准后图像重叠区域的相关系数越接近于 1，说明配准的精度越高。

图 3-19(a)和(b)配准后重叠区域的相关系数为 0.972，结合重叠区域的差值图像可以看出本文的配准算法比较准确。

3.7 小结

本章主要研究了基于特征的图像配准方法，首先介绍了基于频域的、基于灰度的和基于特征的图像配准方法，分析了他们的优缺点。然后深入研究了基于角点检测和基于不变量技术的图像配准方法，并比较了 Harris 角点，SIFT 特征点以及 SURF 特征点的性能，证明了 SURF 特征具有较高的鲁棒性并具有较强的检测速度。最后研究了特征的匹配方法，利用 RANSAC 算法提纯特征点和八参数算法计算图像几何变换模型，实现了手持相机拍摄得到的图像的配准，通过实验证明本文的配准算法可以达到较高的精度。

第四章 全景图像拼接方法及系统实现

4.1 图像亮度调整

在实际拍摄过程中光照情况完全不变的场景是几乎不存在的，光源的变化，拍摄角度的变化，相机的自动对焦等因素都会导致拍摄的图像间存在曝光差异。当完成图像的坐标变换后，需要对图像进行一定的处理来减小曝光差异的影响。图像拼接中最常见的就是对图像进行融合，但是当图像中存在较大的曝光差异时直接融合的效果往往不佳。这就需要首先根据两幅图像的情况，进行亮度调整来减小图像间的曝光差异。

亮度调整最简单的方式就是根据两幅图像亮度值的比例模型进行调整， $P_1(x, y)$ 和 $P_2(x, y)$ 分别是两幅图像重叠区域中对应的像素点，对应像素点间只存在一个亮度差异的比例值 K ，则对应像素点间的关系式为：

$$P_2(x, y) = K \times P_1(x, y) \quad (4-1)$$

在 HSV 坐标下，分别统计图像 I 和图像 I' 在重叠区域内像素点亮度的累加值 V 和 V' ，根据两幅图像的光照强度累加得到一个光照强度的比例：

$$K = \frac{V}{V'} = \frac{\sum v(x, y)}{\sum v'(x, y)} \quad (4-2)$$

利用这个光照强度比例就可以对图像进行曝光差异校正。对于彩色图像来说需要把图像从 GRB 空间转换到 HSV 空间，只针对表示亮度的 V 分量进行计算，然后再转换回 GRB 空间。对于灰度图像，灰度值就相当于 HSV 空间中图像的亮度，所以只要把灰度当做亮度来处理即可。

亮度调整的另一种方法是对两幅图像中的像素点做线性估计：

$$P_2(x, y) = \alpha P_1(x, y) + \beta \quad (4-3)$$

式中 $P_1(x, y)$ 和 $P_2(x, y)$ 分别代表两幅图像重叠区域中对应的像素点， α 为图像间的增益， β 为图像间的偏执量。根据 RANSAC 算法思想可以求出 α 和 β ^[51]，就可以完成图像之间曝光差异的校正，这种线性亮度调整模型是比例模型的一般形式。

根据简单比例模型的亮度调整方法简单快速，根据线性估计模型利用

RANSAC 算法求解的亮度调整相对第一种方法精确，得到的亮度调整效果更佳但是由于 RANSAC 算法需要不断迭代计算速度较慢。这两种亮度调整方法都只是根据图像的整体亮度关系来求解，不可能精确到每个像素点亮度的准确调整。

4.2 传统图像融合方法

待拼接的图像序列之间往往存在着曝光差异等问题，虽然进行曝光差异校正后缩小了两幅图像的曝光差异，但直接对图像进行对齐仍会出现明显的拼接痕迹，降低全景图像的质量，一般要对对齐后的图像进行融合。图像融合又称图像平滑，目前在图像拼接中经常使用的几种比较简单直接的像素级图像融合算法有直接平均法，距离平均法，距离权重法，对比度调制法等。

4.2.1 直接平均法

直接平均融合方法将重叠区域的像素值直接取均值，如式(4-4)所示。

$$P(x, y) = \frac{P_1(x_1, y_1) + P_2(x_2, y_2)}{2} \quad (4-4)$$

式中 $P(x, y)$ 为拼接后图像在 (x, y) 处的像素值， $P_1(x_1, y_1)$ 为第一幅图像在 (x_1, y_1) 处的像素值， $P_2(x_2, y_2)$ 为第二幅图像在 (x_2, y_2) 处的像素值。

4.2.2 加权平均法

加权平均融合方法计算两幅图像重叠区域间对应的每个像素的加权平均值，如式(4-5)所示。

$$P(x, y) = \alpha P_1(x_1, y_1) + \beta P_2(x_2, y_2) \quad (4-5)$$

式中 $P(x, y)$ 为拼接后图像在 (x, y) 处的像素值， $P_1(x_1, y_1)$ 为第一幅图像在 (x_1, y_1) 处的像素值， $P_2(x_2, y_2)$ 为第二幅图像在 (x_2, y_2) 处的像素值。 α 和 β 为权重常数，并且满足 $\alpha + \beta = 1$ 。

4.2.3 距离权重法

距离权重融合方法是一种扩展的加权平均融合方法，该方法以一幅图像作为参考图像，距离的变化作为加权的权重值，如式(4-6)所示。

$$P(x, y) = \alpha P_1(x_1, y_1) + (1 - \alpha) P_2(x_2, y_2) \quad (4-6)$$

式中 α 为第一幅图像中的像素点 (x_1, y_1) 到边界像素的距离与到重叠区域几何中心的距离。

4.2.4 对比度调制法

对比度调制融合方算法通过图像间的对比度来实现图像的融合，图像的对比度利用一幅图像中的细节信息提取出来，然后利用对比度来调制另一幅图像的灰度值。它的实现步骤如下：

- ① 提取图像 $I_1(x, y)$ 的局部对比度，对比度提取公式如下：

$$C_L(x, y) = \frac{I_1(x, y) - \bar{I}_1(x, y)}{\bar{I}_1(x, y)} \quad (4-7)$$

式中 $C_L(x, y)$ 是图像 $I_1(x, y)$ 的局部对比度， $\bar{I}_1(x, y)$ 是图像 $I_1(x, y)$ 的低通滤波图像。

- ② 对局部对比度 $C_L(x, y)$ 进行归一化，其结果为 C_{LN}

$$C_{LN} = \frac{C_L(x, y) - \min[C_L(x, y)]}{\max[C_L(x, y)] - \min[C_L(x, y)]} \quad (4-8)$$

- ③ 对图像进行调制，把归一化的图像对比度 C_{LN} 与图像 $I_2(x, y)$ 相乘，得到融合后的图像。

$$I(x, y) = C_{LN}(x, y) I_2(x, y) \quad (4-9)$$

- ④ 对第③步得到的图像进行重新量化，使它与现实设备的动态范围(一般为 0~255)相匹配，这样就得到最后的图像融合结果。

4.2.5 图像融合算法分析

直接平均融合方法是最简单的融合方法，但是由于他过于简单，往往不具有良好的融合效果。对比度调制法可以针对灰度图像的融合可以实现两幅图像的无缝拼接，但是对于彩色图像效果不佳。加权平均法选择合适的权重时对两幅在重叠区域具有曝光差异的图像具有比较不错的效果，距离权重法针对图像拼接中的图像融合问题具有一定的效果，当只存在一般的曝光差异时它可以使拼接的图像显得自然、没有拼接痕迹，而且该方法简单快速容易实现，在一些图像拼接的工程应用上经常使用。图 4-1 是利用距离权重法进行的融合实验，图(a)为两幅图像配准的图像融合前的效果，可以看出存在着明显的拼接痕迹，图(b)是距离加权法

融合后的结果，看不出拼接痕迹实现了图像的无缝拼接。



图 4-1 距离权重法融合结果

但是当处理运动场景时，待拼接的图像中存在运动物体，直接利用简单的距离权重法融合，融合时运动物体有可能相互重叠在一起出现重影即鬼影现象，如图 4-2 所示。因此简单的距离权重法只能处理静态场景图像，对于动态场景需要对运动物体进行处理，来消除鬼影现象。



图 4-2 距离权重法直接融合带有运动物体图像

4.3 融合鬼影的消除方法

在图像拼接的结果中有时会存在着某个物体不完全重合的问题，这种图像拼接中的重影现象称为鬼影现象。根据鬼影现象的产生原因可以把鬼影分为两类，一种是由于图像配准不准确造成的配准鬼影，另一种是两幅图像融合时重叠区域的运动物体等导致的物体不能完全重合而出现的融合鬼影。

4.3.1 最佳缝合线

最佳缝合线是在运动场景中经常使用的分割运动物体的有效方法，最佳缝合线将两幅图像的重叠区域分成两部分，理论上的最佳缝合线应该位于两幅图像最为相似的地方，可以减少拼接后存在的锯齿现象，缝合线两边来自不同的图像，

可以在一定程度上避开场景中的运动物体，减少融合鬼影出现的概率。

4.3.1.1 最佳缝合线搜索准则

最佳缝合线是一种基于动态规划原理的最优化方法，理论上的最佳缝合线是两幅图像最相似的地方，所以最佳缝合线应该位于具有最小颜色差异的位置，并且缝合线上的像素点在几何结构上最相似。根据最佳缝合线的这两点要求，人们提出了最佳缝合线的搜索准则来求出最佳的分割线。

Davis 于 1998 年提出了一种基于颜色强度差异的缝合线搜索准则^[52]：

$$E(x, y) = E_{color}(x, y) = \frac{abs(I_1 - I_2)}{\max(I_1, I_2)} \quad (4-10)$$

式中 $E_{color}(x, y)$ 是两幅图像的重叠区域上的像素点的颜色差值。

这种搜索准则只包含了图像的颜色强度差异，当图像存在较大的曝光差异时，最佳缝合线的搜索可能不准确，缝合线两端无法对齐。

Duplaquet 于 1998 年提出了一种基于图像颜色差异和几何差异的缝合线的搜索准则^[53]：

$$E(x, y) = E_{dif}(x, y)^2 - \lambda E_{edge}(x, y) \quad (4-11)$$

$$E(x, y)_{dif} = \frac{1}{N_V} \sum_{i, j \in V} |I_1(x+i, y+i) - I_2(x+i, y+i)| \quad (4-12)$$

$$E_{edge}(x, y) = \min(g_1(x, y), g_2(x, y)) \quad (4-13)$$

式中 I_1 和 I_2 分别表示相邻的两幅图像， $E(x, y)_{dif}$ 表示两幅图像重叠区域的颜色差异， $E_{edge}(x, y)$ 表示两幅图像重叠区域的几何差异。图像的颜色差异 $E_{dif}(x, y)$ 用图像重叠区域内的像素点之差的均值来表示。 g_1 和 g_2 分别是两幅相邻图像的梯度，两幅图像的几何差异 $E_{edge}(x, y)$ 用两幅图像结构变化较小处的梯度来表示，式中 λ 为调节系数。

这种最佳缝合线的搜索准则考虑到了图像的颜色强度差异和几何差异，但是调节系数 λ 不易确定，而且还存在着其他的不确定因素使得最佳缝合线的求解比较困难。

Levin 于 2004 年提出了一种基于图像梯度差异的缝合线搜索准则^[54]

$$E(x, y) = E_{geometry}(x, y) \quad (4-14)$$

式中 $E_{geometry}(x, y)$ 是两幅图像重叠区域中对应像素点的梯度值之差。

这种最佳缝合线搜索准则只考虑了图像的梯度信息，容易出现无法避开运动物体的现象，而且也存在缝合线两端无法对其的问题。

方贤勇于 2005 年提出了一种改进的缝合线搜索准则，此准则结合了图像的颜色差异和几何差异^[55]：

$$E(x, y) = E_{dif}(x, y)^2 + E_{geometry}(x, y) \quad (4-15)$$

式中 $E_{dif}(x, y)$ 和 $E_{geometry}(x, y)$ 分别表示两幅图像重叠区域的颜色差异和几何差异。图像的几何差异 $E_{geometry}(x, y)$ 由两幅图像重叠区域的梯度之积来表示：

$$E_{geometry}(x, y) = Diff(I_1(x, y), I_2(x, y)) \quad (4-16)$$

式中 $I_1(x, y)$ 和 $I_2(x, y)$ 表示相邻的两幅图像， $Diff$ 表示他们在 x 方向和 y 方向的梯度值之差的积。

计算图像的梯度时通常采用的 Sobel 算子为：

$$S_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \quad S_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (4-17)$$

文献[55]通过实验证明了这种算子没有充分利用像素点周围的相似性，根据像素点中心对角线方向相邻像素点的相关性，提出一种新的计算梯度的算子：

$$S_x = \begin{bmatrix} -2 & 0 & 2 \\ -1 & 0 & 1 \\ -2 & 0 & 2 \end{bmatrix}, \quad S_y = \begin{bmatrix} -2 & -1 & -2 \\ 0 & 0 & 0 \\ 2 & 1 & 2 \end{bmatrix} \quad (4-18)$$

这种基于图颜色差异和几何差异的最佳缝合线搜索准则可以较好的避开运动物体，但是对于曝光差异较大的图像整个缝合线搜索的准确性不佳。

Mills 于 2009 年提出了一种含有曝光差异因子的最佳缝合线搜索准则^[56]

$$E(x, y) = E_{color}(x, y) + E_{geometry}(x, y) \quad (4-19)$$

$$E_{color}(x, y) = (1 - (|\ln \alpha| + |\beta|)^2) \frac{|I_1 - I_2|}{\max(I_1, I_2)} \quad (4-20)$$

$$E_{geometry}(x, y) = (|\ln \alpha| + |\beta|)^2 \frac{|g_1 - g_2|}{\max(g_1, g_2)} \quad (4-21)$$

式中 $E_{color}(x, y)$ 和 $E_{geometry}(x, y)$ 分别是图像的颜色差异和几何差异， I_1 和 I_2 分别是两幅图像的颜色强度， g_1 和 g_2 是两幅图像重合区域点 (x, y) 处的梯度值， α 和 β 是

曝光差异修正系数。

这种最佳缝合线搜索准则利用了曝光差异修正系数可以处理图像曝光差异较大的情况，但是它没有充分考虑到像素点周围的相似性， α 和 β 不易计算，同时也存在计算复杂度比较高的问题。

本文总结前人提出的几种最佳缝合线搜索准则存在的问题及优点，提出一种改进的最佳缝合线搜索准则：

$$E(x, y) = aE_{gray}(x, y)^2 + bE_{geometry}(x, y) \quad (4-22)$$

式中 $E_{gray}(x, y)$ 和 $E_{geometry}(x, y)$ 分别表示两幅图像重叠区域的灰度差异和几何差异， a, b 为灰度差异和几何差异的权重值。

为了尽量减小计算量用灰度信息来替代颜色信息，避免彩色图像需要在三通道上计算的问题，在一定程度上可以减小计算量。

在曝光差异较大的图像中，曝光差异校正的准确性比较低，所以图像的梯度差异可靠性更高，因此梯度差异的权重 b 被赋予更大的值。这里首先对个别曝光差异较大的图像进行一次亮度调整减小一定的曝光差异。考虑速度上的问题本文利用简单的比例模型进行亮度调整，而不用运算量较大的基于 RANSAC 思想的线性估计亮度调整方法。记录下亮度值累加的比例 K ，利用它来对灰度差异和几何差异加权。在不存在曝光差异时灰度差异和几何差异应该具有相同的权重，即亮度比例 $K=1$ 时，权重值 $a=b$ ，式(4-23)即为灰度差异和几何差异的权重值。

$$b = \left(\frac{1}{\sqrt{2}} + |\ln K| \right)^2, a = \begin{cases} 1-b, & b < 1 \\ 0, & b \geq 1 \end{cases} \quad (4-23)$$

式中 K 为两幅图像亮度调整的比例值，权重值 a, b 的趋势如图 4-3，亮度调整比例为 1 时 a, b 具有相同的值 0.5。

根据 Kwatra 提出的像素差异的方法^[57]计算 $E_{dif}(x, y)$ ：

$$E_{dif}(x, y) = \frac{|P_1(x, y) - P_2(x, y)|}{\max(P_1(x, y), P_2(x, y))} \quad (4-24)$$

式中 $P_1(x, y)$ 和 $P_2(x, y)$ 为两幅图像重叠区域对应的像素点。

图像的几何差异 $E_{geometry}(x, y)$ 为了能够充分利用像素点周围的相似性，利用式(4-18)的梯度算子计算图像的梯度差异。

$$E_{geometry}(x, y) = \frac{|g_{1x} - g_{2x}|}{\max(g_{1x}, g_{2x})} \frac{|g_{1y} - g_{2y}|}{\max(g_{1y}, g_{2y})} \quad (4-25)$$

式中 g_{1x} 和 g_{2x} 为两幅图像重叠区域对应像素点在 x 方向的梯度值, g_{1y} 和 g_{2y} 为两幅图像重叠区域对应像素点在 y 方向的梯度值。

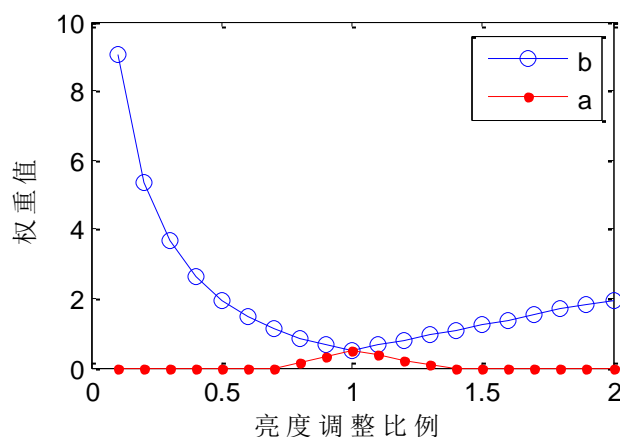


图 4-3 图像灰度差异与几何差异的权值

这种最佳缝合线搜索准则利用改进的算子计算几何差异使得最佳缝合线的搜索充分利用了像素点周围的相似性, 并且包含曝光差异因子根据曝光差异的大小来确定几何差异与颜色差异的权重, 增加了缝合线搜索的准确性。

4.3.1.2 最佳缝合线搜索方法

最佳缝合线根据动态规划原理的最优化搜索方法, 动态规划把多阶问题转化为一组一阶问题, 利用各个阶段的关系来逐个求解。在问题中可能存在多个解, 最佳缝合线就是要找出一条最优的缝合线。

搜索最佳缝合线时根据式(4-22)计算出图像重叠区域每一个像素点的准则值, 得到一幅差值图像。图 4-4 为搜索最佳缝合线的示意图。搜索缝合线时首先把差值图像第一行的每一个像素点作为一条缝合线的起始点, 把他的强度值作为缝合线的准则值, 选择下一行相邻的三个像素点中强度值最小的一个作为扩展方向, 不断的向下扩展, 最后选择准则值最小的一条缝合线作为最佳缝合线。

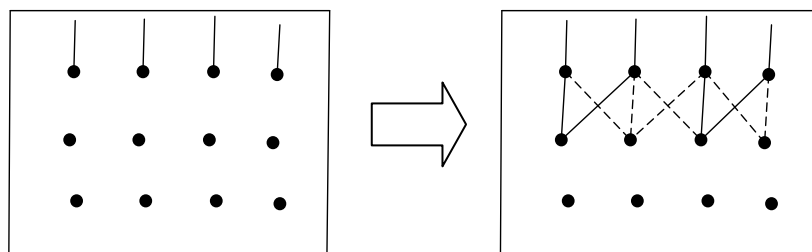


图 4-4 最佳缝合搜索示意图

两幅图像的重叠区域被最佳缝合线划分为两部分，每一部分的像素分别来自于不同的图像。因此，可以在一定程度上避开场景中的运动物体，减少鬼影出现的概率。但最佳缝合线也只是一个局部最优解，也要受到其他因素的干扰不能完全消除出现鬼影的可能性。在搜索缝合线时，其中的某一条缝合线可能存在一些较大的误差点，但是这条缝合线整体的准则值最小，最后被作为最佳缝合线^[58]。但是由于较大误差点的存在可能导致还会出现鬼影现象。当特征点匹配足够准确时，特征点对意味着分别在两幅图像中显示的同一个点，是两幅图像差异较小的点，适合作为缝合线的扩展方向^[59]。因此根据特征点的性质，可以尽量让特征点作为缝合线的扩展方向来减小个别误差点的影响。在前面已经利用了欧氏距离匹配和 RANSAC 特征点提纯的方法，已经能够比较准确的匹配特征点。所以给特征点处的缝合线搜索准则值设置一个大于 0 小于 1 的权值（文中取 0.5）来使缝合线中可以包含更多的特征点。在扩展缝合线时可以把原本的 3 个扩展方向增加至 5 个使缝合线中尽可能多的包含特征点。文中的最佳缝合线搜索步骤如下：

① 初始化：计算图像重叠区域的差值图像，把差值图像中的第一行的每一个点作为一条缝合线的起始点，根据式(4-22)计算它的强度值，并把它的强度值作为该条缝合线的准则值。

② 扩展：把每一条缝合线向下扩展，判断下一行相邻的 5 个像素点中是否存在特征点，如果存在特征点把其强度值加上 0.5 的权重，并计算其他点的强度值，找出其中强度值最小的一个点作为缝合线的扩展方向，并累加扩展点的强度值到该条缝合线的准则值。

③ 确定最佳缝合线：在生成的所有缝合线中找出具有最小准则值的一条作为最后的最佳缝合线。

4.3.1.3 实验及分析

如图 4-5 为对图像亮度的调整，其(a)和(b)为两幅待拼接的图像，(c)为以(a)为参考图像对(b)利用比例模型的亮度调整结果，可以看出图像(a)与(c)之间的曝光差异要小于图像(a)与(b)之间的曝光差异。

图 4-6 为利用不同方法对图像进行拼接的结果，图 4-6(a)为对图 4-5(a)和(b)进行直接拼接结果，图 4-6 (b)为对图 4-5(a)和(c)进行直接拼接的结果，可以发现进行亮度调整后拼接结果(b)中的拼接痕迹没有亮度调整前那么明显，可见图像曝光差异校正可以减小图像间的差异从而优化待拼接图像，降低了后面图像融合的难度。图 4-6(c)为利用文中的最佳缝合线搜索方法搜索最佳缝合线，在缝合线两边分别从

图 4-5(a)和(b)中取像素直接拼接的结果。图中的拼接痕迹更为不明显，只存在由于两幅图像曝光情况不完全一样产生的不是十分明显的痕迹。说明文中的最佳缝合线搜索方法可以找到两幅图像重叠区域的较为相似的地方。



(a) 原图 1



(b) 原图 2



(c) 亮度调整结果

图 4-5 曝光差异校正



(a) 两幅原图直接拼接结果



(b) 进行亮度调整后拼接结果



(c) 利用最佳缝合线拼接结果

图 4-6 图像拼接结果

利用最佳缝合线两侧相邻像素点的平均绝对误差(Mean Absolute Error, MAD)和均方根误差(Root Mean Square Error, RMSE)作为最佳缝合线精度的评定指标。最佳缝合线两侧像素点的 MAD 及 RMSE 越小说明缝合线两侧的像素点越相似, 拼接后的痕迹就越不明显, MAD 和 RMSE 的计算公式为

$$MAD = \frac{\sum_{i=1}^n \|I_{1i} - I_{2i}\|}{n} \quad (4-26)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n \|I_{1i} - I_{2i}\|^2}{n}} \quad (4-27)$$

式中 I_{1i} 和 I_{2i} 分别表示缝合线第 i 行左右两侧的像素值。

图 4-7 为利用不同方法搜索的最佳缝合线, 图(a)为利用文献[55]方法搜索结果, 图(b)为利用文献[56]方法的搜索结果, 图(c)为利用本文方法搜索最佳缝合线结果。



(a)



(b)



(c)

图 4-7 不同方法搜索的最佳缝合线

表 4-1 为图 4-7(a)到(c)最佳缝合线两侧像素点的 MAD、RMSE 及搜索缝合线所用的时间。可以看出本文最佳缝合线搜索时间最短，说明了本文的最佳缝合线搜索利用灰度值代替颜色值，利用简单的亮度校正方法计算曝光因子在一定程度上提高了最佳缝合线搜索的速度。表中的 MAD 和 MSE 说明了利用本文方法搜索的最佳缝合线位于两幅图像更为相似的地方，进而表明了本文提出的最佳缝合线搜索方法要优于文献[55]和[56]中的方法。

表4-1 最佳缝合线搜索方法对比

图像	MAD	RMSE	时间 (ms)
图4-7(a)	1.067	1.366	1132
图4-7(b)	0.982	0.991	2274
图4-7(c)	0.798	0.853	728

为了进一步证明本文中最佳缝合线搜索准则的可行性，对图 4-5(a)和(c)进行部分亮度调整并手动添加一个运动物体，使两幅图像的重叠区域中的一部分存在较大曝光差异，使曝光差异更容易影响到最佳缝合线的搜索，这样便于验证本文最佳缝合线搜索准则的优越性。合成的实验图像如图 4-8 所示。



图 4-8 最佳缝合线实验图像

图 4-9 为利用不同搜索准则对图 4-8 搜索最佳缝合线的结果，图(a)为利用文献[55]中的最佳缝合线搜索准则得到的缝合线，最佳缝合线两侧都存在运动物体，说明此缝合线并没有完分割出运动物体。图(b)为利用文中的最佳缝合线搜索准则得到的最佳缝合线，运动物体位于最佳缝合线一侧，这样融合图像中只会存在一个完整的运动物体避免了运动物体的影响。实验结果说明本文的最佳缝合线搜索方法可以减小图像曝光差异的影响，准确分割出运动物体，从而可以较好的消除运动物体导致的鬼影现象。



(a) 文献[55]方法搜索结果



(b) 本文方法搜索结果

图 4-9 最佳缝合线对比

图 4-10 (c)为利用文中的最佳缝合线搜索准则但不在特征点处加权对图 4-10(a)和(b)搜索最佳缝合线的结果，图中运动的人被分割在了缝合线的左侧。得到最佳缝合线后就可以知道缝合线在两幅原图中所处的位置，在图 4-10(a)的缝合线位置上任取一点 A，以 A 为对角线交点做一个矩形区域，对矩形区域中赋相同的像素值，得到的图像如图 4-11(a)所示。使矩形区域水平移动一定距离，并保持缝合线仍能穿过此区域，计算出它对应应在图 4-10(b)中的位置，把移动后的矩形区域置于图 4-10(b)中就得到了图 4-11(b)。由于原最佳缝合线通过矩形区域，而且区域中像素值相同，所以当搜索最佳缝合线时矩形区域重叠部分中的像素点在颜色上不存在差异，使该处的准则值变小，从而增加了最佳缝合线通过此处的概率。这样就快速的形成了一种误差点干扰缝合线搜索的特殊情况。



(a) 原图 1



(b) 原图 2

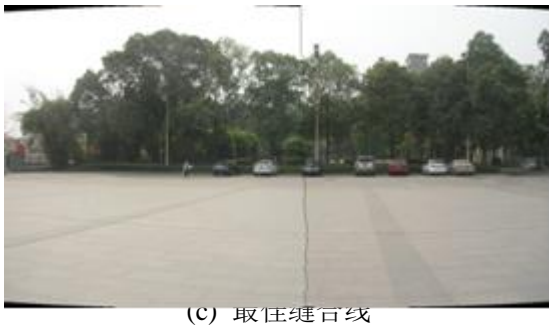


图 4-10 在特征点处不加权的最佳缝合线

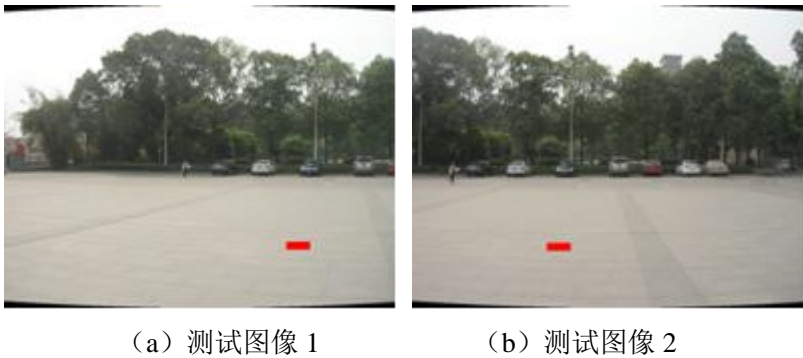


图 4-11 添加物体的测试图像

图 4-12 为利用文中的最佳缝合线搜索准则但不在特征点处加权对图 4-11(a)和(b)搜索最佳缝合线的结果，缝合线两侧均存矩形区域，说明缝合线的搜索受到了影响，无法实现融合鬼影的消除。图 4-13 为利用文中的最佳缝合线搜索方法得到的最佳缝合线，矩形区域只存在于缝合线的右侧，说明了本文对特征点处准则值加权的并增加扩展点数的最佳缝合线搜索方法可以在一定程度上减小误差点的干扰，提高了缝合线的鲁棒性。

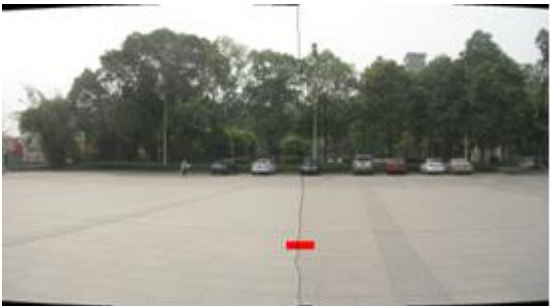
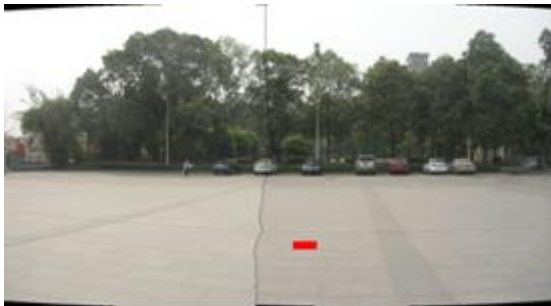


图 4-12 特征点处未加权搜索出的缝合线图



4-13 特征点处加权搜索出的缝合线

4.3.2 多分辨率融合算法

4.3.2.1 生成拉普拉斯金字塔

P. J. Burt 等人首先提出了多分辨率的思想，并应用到了图像拼接重叠区域的过度处理中^[59]。多分辨率融合方法把图像分解成一组带通图像，它在每一频带内利用不同的加权函数进行融合^[60-61]。高斯金字塔是一组低通滤波后的图像，设 I_k 和 I_{k+1} 是高斯金字塔中第 k 层和第 $k+1$ 层的图像。 I_k 的最高频率比 I_{k+1} 的最高频率高一倍，所以 I_k 和 I_{k+1} 的差值就是原始图像的一个带通图像。 I_k 的采样率比 I_{k+1} 高一倍，所以首先需要对 I_{k+1} 插值使其采样率提高一倍。

多分辨率融合要求图像重叠区域的长宽相等，但是通常情况下是不满足要求的，所以需要对图像进行外插。设 W 和 H 分别是图像的宽和高，则用来构建金字塔的图像满足式(4-28)的关系：

$$W = 2^N m + 1, H = 2^N n + 1 \quad (4-28)$$

式中 m 、 n 、 N 都是正整数，利用外插法扩展时需要先确定 m 、 n 、 N ，设

$$W' = 2^N m + 1, H' = 2^N n + 1 \quad (4-29)$$

求出满足 $W' \geq W, H' \geq H$ 的最小正整数 W' 和 H' ，然后设

$$a = (H' - H) / 2, b = (W' - W) / 2 \quad (4-30)$$

设外插后的图像为 I' ，就可以由式(4-31)求出插值后的图像，定义这种插值函数为 EXPAND 函数。

$$I'(i, j) = \begin{cases} I(i-a, j-b) & (a \leq i \leq a+H-1, b \leq j \leq b+W-1) \\ I(i-a, b-j) & (a \leq i \leq a+H-1, 0 \leq j \leq b) \\ I(i-a, 2W+b-j-2) & (a \leq i \leq a+H-1, j > b+W-1) \\ I'(2a-i, j) & (i < a) \\ I'(2H+2a-i-2, j) & (i > a+H-1) \end{cases} \quad (4-31)$$

设序列 $L_0, L_1 \cdots L_N$ 为拉普拉斯金字塔，则拉普拉斯金字塔可以表示为

$$L_k = \begin{cases} I_k - \text{EXPAND}(I_{k+1}) & k \neq N \\ I_N & k = N \end{cases} \quad (4-32)$$

然后可以得到一个由拉普拉斯金字塔图像构造源图像的递推公式：

$$\begin{aligned}
 I_N &= L_N \\
 I_{N-1} &= L_{N-1} + EXPAND(I_N) \\
 &\vdots \\
 I_k &= L_k + EXPAND(I_{k+1}) \\
 &\vdots \\
 I_0 &= L_0 + EXPAND(I_1)
 \end{aligned} \tag{4-33}$$

4.3.2.2 颜色融合算法

利用两幅图像的特征在不同的频带的差异求解插值区域大小和加权函数。设 I_{1k} 和 I_{2k} 是图像 I_1 和 I_2 的拉普拉斯金字塔第 k 层，它们的大小都为 (W_k, H_k) ，则有：

$$\begin{cases}
 dr = \sum_{i=0}^{H_k-1} \sum_{j=0}^{W_k-1} |I_{1k}(i, j, r) - I_{2k}(i, j, r)| \\
 dg = \sum_{i=0}^{H_k-1} \sum_{j=0}^{W_k-1} |I_{1k}(i, j, g) - I_{2k}(i, j, g)| \\
 db = \sum_{i=0}^{H_k-1} \sum_{j=0}^{W_k-1} |I_{1k}(i, j, b) - I_{2k}(i, j, b)|
 \end{cases} \tag{4-34}$$

$$\begin{cases}
 pr = dr / tr \\
 pg = dg / tg \\
 pb = db / tb
 \end{cases} \tag{4-35}$$

$$\begin{cases}
 tr = \sum_{i=0}^{H_k-1} \sum_{j=0}^{W_k-1} (|I_{1k}(i, j, r)| + |I_{2k}(i, j, r)|) \\
 tg = \sum_{i=0}^{H_k-1} \sum_{j=0}^{W_k-1} (|I_{1k}(i, j, g)| + |I_{2k}(i, j, g)|) \\
 tb = \sum_{i=0}^{H_k-1} \sum_{j=0}^{W_k-1} (|I_{1k}(i, j, b)| + |I_{2k}(i, j, b)|)
 \end{cases} \tag{4-36}$$

$$p = (pr + pg + pb) / 3 \tag{4-37}$$

式 (4-34) 中 dr, dg, db 分别代表图像 I_{1k} 和 I_{2k} 的 R、G、B 分量的差异。式(4-35)中 pr, pg, pb 分别代表图像 I_{1k} 和 I_{2k} 的 R、G、B 分量的相对差异。式(4-37)中 p 表示在第 k 个频带内图像 I_1 和 I_2 的特征差异，通过 p 可以得到式(4-38)确定颜色过渡区域的宽度 w

$$w = pW_k \tag{4-38}$$

融合时过渡区域宽度 w 要为奇数，并且图像 I_{1k} 和 I_{2k} 的颜色过渡区域要小于整幅图像的宽，就要求 $w \leq W_k - 2$ 。这样就可以得到第 k 个频带的融合公式：

$$S_k(i, j) = \begin{cases} I_{1k}(i, j) & 0 \leq j \leq W_k/2 - w/2 \\ coef_1 I_{1k}(i, j) + coef_2 I_{2k}(i, j) & W_k/2 - w/2 \leq j \leq W_k/2 + w/2 \\ I_{2k}(i, j) & W_k/2 + w/2 < j < W_k \end{cases} \quad (4-39)$$

4.3.3 泊松融合算法

4.3.3.1 泊松融合原理

图 4-14 中 S 表示一个图像定义域， Ω 是图像 S 上的一个封闭的子集， $\partial\Omega$ 是 Ω 的边界， f^* 是 S 上的已知标量函数， f 是 Ω 上的未知标量函数。 v 是一个指导场， g 是指导场的原函数。 f 可以通过模插值方法来求解。

$$\min_f \iint_{\Omega} |\nabla f|^2, \quad f|_{\partial\Omega} = f^*|_{\partial\Omega} \quad (4-40)$$

式中 $\nabla = \begin{bmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \end{bmatrix}$ 是梯度算子，式中的最小值需要满足欧拉—拉格朗日方程，则

它的解可以可以利利用带狄利克雷边界的拉普拉斯等式表示

$$\Delta f = 0, \quad (x, y) \in \Omega, \quad f|_{\partial\Omega} = f^*|_{\partial\Omega} \quad (4-41)$$

式中 $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ 是拉普拉斯算子。

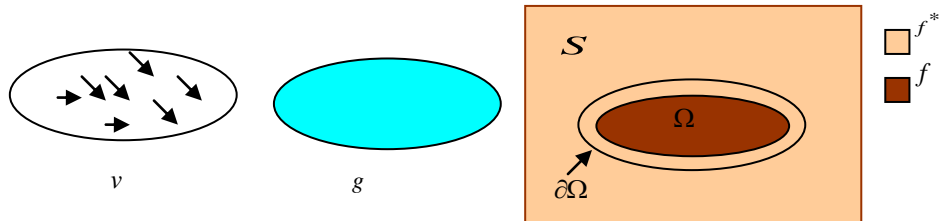


图 4-14 模插值和泊松融合

利用式(4-41)求解出的结果会变得模糊，泊松融合是一种利用指导场进行改进的求解方法，可以获得较好的效果，它利用指导场 v 使 f 和 v 具有相同的变化^[62]

$$\min_f \iint_{\Omega} |\nabla f - v|^2, \quad f|_{\partial\Omega} = f^*|_{\partial\Omega} \quad (4-42)$$

可以利用带狄利克雷边界的泊松方程来表示式(4-42)的解：

$$\Delta f = \text{div} v, \quad f|_{\partial\Omega} = f^*|_{\partial\Omega} \quad (4-43)$$

式中 $\text{div} v = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}$ 是 $v = (u, v)$ 的散度。

式(4-43)就是泊松融合的基本方程，当进行实际的图像融合时，把待合成图像目标区域的梯度场作为进行泊松融合的指导场，求解式(4-43)就能够把源图像融合到目标区域中，得到无缝自然的融合结果。

4.3.3.2 方程的求解

对彩色图像上的融合需要在 RGB 三通道上分别进行求解。对式(4-43)做有限差分离散化得：

$$\min_{f|_{\Omega}} \sum_{(p,q) \cap \Omega \neq \emptyset} (f_p - f_q - v_{pq})^2, \quad f_p = f_p^* \quad (4-44)$$

式中 v_{pq} 是 $v(\frac{p+q}{2})$ 在 $[p, q]$ 方向上的投影， $v_{pq} = v(\frac{p+q}{2}) \cdot \frac{\vec{pq}}{|\vec{pq}|}$ ，然后就可以得到相应的差分解

$$|N_p| f_p - \sum_{q \in N_p \cap \Omega} f_q = \sum_{q \in N_p \cap \partial \Omega} f_q^* - \sum_{q \in N_p} v_{pq} \quad (4-45)$$

根据式(4-41)把融合区域中的每一个像素写成式(4-43)的形式，并把它们组合到一起得到：

$$UX = V \quad (4-46)$$

式中系数矩阵 U 是对称的稀疏正定矩阵，这样就满足超松弛高斯赛德尔法(SOR)的求解条件，可以利用 SOR 法求解，式(4-46)的解可以表示为：

$$x_i^{k+1} = x_i^k + \frac{w}{u_{ii}} \left(v_i - \sum_{j=1}^{i-1} u_{ij} x_j^{k+1} - \sum_{j=i+1}^n u_{ij} x_j^k \right) \quad (4-47)$$

式中 w 为松弛因子。

由于图像矩阵往往很大，导致利用 SOR 方法直接求解的计算量过大，但根据泊松融合的原理每个像素点最多只有 4 个近邻点，所以在式(4-47)中只有 4 个未知数需要更新^[63]。设当对第 i 个像素点进行计算时它的下一行相邻的像素点编号为 s ，它的上一行相邻像素点的编号为 t ，根据式(4-45)可以将式(4-47)简化为

$$x_i^{k+1} = x_i^k + \frac{w}{u_i} (v_i + x_{i-1}^{k+1} + x_s^{k+1} + x_{i+1}^k + x_t^k) \quad (4-48)$$

式中 u_i 表示第 i 个像素相邻的像素点的个数。

简化后可以减小计算量提高融合速度，但是简化后仍需要迭代求解。Morel 等人指可以利用傅里叶变换的方法求解泊松融合^[64]。利用傅里叶变换求解时需要使边界与坐标轴重合，一般情况下是不满足条件的，把指导场扩展到整幅图像的梯度场可以解决这个问题。扩展后的指导场 V 可以表示为

$$V = \begin{cases} v & (x, y) \in \Omega \\ \nabla f * & \text{其他} \end{cases} \quad (4-49)$$

式中 (x, y) 表示图像上的像素点位置。设融合后的图像为 f' 则求解式(4-42)等价于求解式(4-50)

$$\min_f \iint_S |\nabla f - V|^2 \quad (4-50)$$

这样式(4-50)的解可以用带有齐次诺依曼边界的泊松方程表示

$$\Delta f' = \text{div} V, \quad \frac{\partial f'}{\partial n} \Big|_{\partial S} = 0 \quad (4-51)$$

式中 n 表示图像边界 ∂S 的法向。

式(4-51)具有常系数边界，只要扩展原始图像使其成为对称和周期的，就可以利用傅里叶变换来求解泊松融合方程。大小为 $J \times L$ 的区域中某点的 DFT 系数可以表示为：

$$\hat{f}'_{mn} = \sum_{j=0}^{J-1} \sum_{l=0}^{L-1} f'_{jl} e^{-i \frac{2\pi jm}{J}} e^{-i \frac{2\pi ln}{L}} \quad (4-52)$$

式中 $m=0, 1 \cdots J-1, n=0, 1 \cdots L-1$ ，则 f 的傅里叶级数公式可以由它的系数得到：

$$f'(x, y) = \frac{1}{JL} \sum_{m=0}^{J-1} \sum_{n=0}^{L-1} \hat{f}'_{mn} e^{i \frac{2\pi mx}{J}} e^{i \frac{2\pi ny}{L}} \quad (4-53)$$

式(4-51)经过傅里叶变换，他们的关系转换为：

$$\left(\left(\frac{2\pi m}{J} \right)^2 + \left(\frac{2\pi n}{L} \right)^2 \right) \hat{f}'_{mn} = \frac{2\pi i m}{J} \hat{u}_{mn} + \frac{2\pi i n}{L} \hat{v}_{mn} \quad (4-54)$$

式中 $V = (u, v) = \nabla F$ ， F 是定义域在 Ω 上的已知函数

$$F = \begin{cases} g & (x, y) \in \Omega \\ f * & \text{其他} \end{cases} \quad (4-55)$$

则有：

$$\hat{u}_{mn} = -\frac{2\pi im}{J} \hat{F}_{mn}, \quad \hat{v}_{mn} = -\frac{2\pi in}{L} \hat{F}_{mn} \quad (4-56)$$

这样就可以计算出 \hat{f}'_{mn} ，最后根据(4-48)就可以得到泊松融合方程的解。

4.3.3.3 图像融合实验及分析

图 4-15 为泊松融合及多分辨率的实验结果，在 CPU 主频为 2.4GHZ，内存为 2G 硬件条件下在 Window XP 操作系统中以 MATLAB 作为编程平台对融合算法进行测试。图 4-15 (a)和(b)分别为目标图像和源图，4-15 (c)为利用傅里叶变换求解的泊松融合将(b)中的月亮融合到(a)中的结果，从图中可以看出泊松融合可以将(b)中的指定区域无明显痕迹的融合到(a)中。图(d)为多分辨率融合的结果，可以看出泊松融合不存在多分辨率融合那样融合区域变暗的现象，效果要好于多分辨率融合。

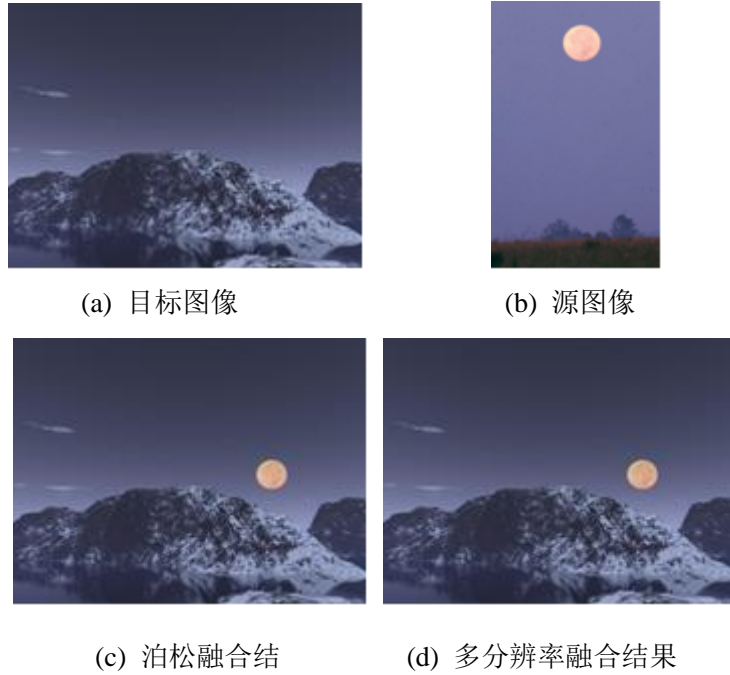


图 4-15 融合结果

均方误差(Mean Square Error, MSE)用来衡量融合图像与理想图像越接近程度，MSE 越小则图像处理的效果就越好，峰值信噪比(Peak Signal to Noise Ratio, PSNR)主要衡量融合图像的逼真程度，PSNR 越高则 RMSE 越小说明图像融合的效果和质量越好^[65]。PSNR 与 MSE 的计算公式为

$$MSE = \frac{\sum_{i=1}^M \sum_{j=1}^N (R(i, j) - F(i, j))^2}{MN} \quad (4-57)$$

$$PSNR = 10 \log \frac{MN(2^n - 1)^2}{\sum_{i=1}^M \sum_{j=1}^N (R(i, j) - F(i, j))^2} \quad (4-58)$$

式中 R 和 F 是大小为 $M \times N$ 的融合图像与理想图像, n 为一个像素点所占的二进制位数, n 通常取 8。

清晰度又叫做平均梯度, 它反映了融合图像中的微小细节反差与纹理变化特征, 同时也反映了融合图像的清晰度^[66]。平均梯度越大说明图像越清晰融合效果越好, 它的计算公式为

$$\Delta \bar{G} = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N \sqrt{\frac{\Delta F_x^2(m, n) + \Delta F_y^2(m, n)}{2}} \quad (4-59)$$

式中 F 是一个大小为 $M \times N$ 的图像, $\Delta F_x^2(m, n)$ 和 $\Delta F_y^2(m, n)$ 分别表示 F 在 x 和 y 方向上的差分。

表 4-2 为对图 4-15 中(a)(b)利用多分辨率融合、SOR 法求解的泊松融合以及傅里叶变换求解的泊松融合进行融合结果的 PSNR、MSE、清晰度 $\Delta \bar{G}$ 以及运行时间。可以看出泊松融合在逼真程度上不差于多分辨率融合, 在清晰度上泊松融合略好于多分辨率融合, 不会像多分辨率融合那样使融合区域变得平滑。傅里叶变换求解的泊松融合更具有速度上的优势, 所以本文选用傅里叶变换求解的泊松融合作为图像拼接中的融合方法。

表4-2 不同融合方法结果对比

方法	PSNR	MSE	清晰度	时间 (ms)
多分辨率融合	49.260	0.771	3.286	1138
SOR求解的泊松融合	49.705	0.696	4.452	552
傅里叶变换求解的泊松融合	49.594	0.714	4.452	343

4.3.3.4 图像拼接中的泊松融合

具体到图像拼接的融合时, 选择最佳缝合线一侧的图像进行融合, 例如位置关系为图 4-16(a)所示的两幅图像, 设左侧图像为图像 ABCD, 右侧图像为图像 EFGH。EICJ 为两幅图像的重叠区域, 曲线 MN 为对两幅图像搜索得到的最佳缝合线, 如图 4-16(b)所示。

融合时取图像 ABCD 中 EICJ 部分为目标图像, 取图像 EFGH 中 NMCJ 部分

为源图像，泊松融合的边界 IE、EJ 取图像 ABCD 中的像素，边界 JC、CI 取图像 EFGH 中的像素。融合时就是使融合区域具有与图像 EFGH 中的 NMCJ 部分具有相同的梯度场，最后根据泊松原来计算出泊松融合结果即可。当两幅图像存在较大曝光差异时，图像的右侧部分与重叠部分可能还会存在拼接痕迹，这时利用拼接结果与图像 EFGH 做一次距离权重法的图像融合，这样可以消除拼接痕迹，而不必以 ABCD 作为源图像，EFGH 作为目标图像再进行一次泊松融合，这样可以减少融合的计算量。

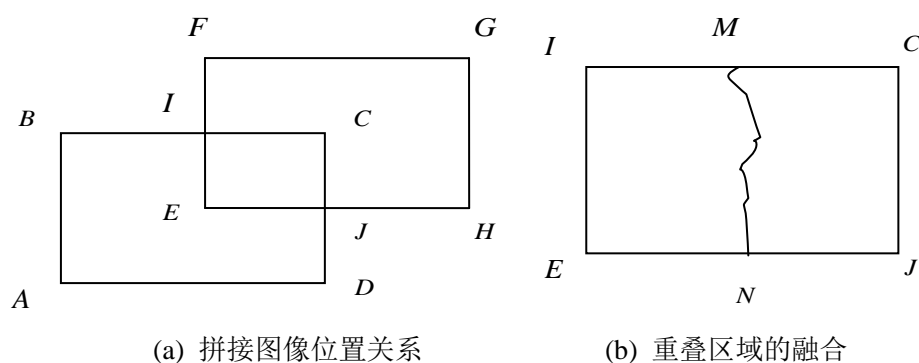


图 4-16 图像拼接中的泊松融合示意

4.3.4 实验及分析

图 4-17(a)和(b)为利用普通数码相机拍摄得到的两幅图像，将它们作为图像拼接的原始图像。图 4-17 (c)为利用文中方法搜索得到的最佳缝合线，它在缝合线的左侧从(a)中取像素，在缝合线右侧从图(b)中取像素，可以看出运动物体被分在了最佳缝合线的一侧。图 4-17 (d)为直接利用距离权重法融合的拼接结果，图中两幅原始图像中的运动物体重叠在了一起，存在着明显的鬼影现象。图 4-17 (e)为利用最佳缝合线和泊松融合得到的拼接结果，图中完全不存在融合鬼影现象，说明文中利用改进的最佳缝合线与泊松融合的方法可以有效的消除运动物体导致的融合鬼影现象，取得无缝自然的拼接图像。



(a) 原图 1



(b) 原图 2



(c) 取住缝合线



(d) 存在鬼影的拼接结果



(e) 消除鬼影的拼接结果

图 4-17 图像拼接结果

4.4 全景图像拼接系统的设计与实现

4.4.1 全景图像拼接系统

文中的全景拼接系统是在 Window 系统下以 VC6.0 作为编程平台实现的，图 4-18 为全景图像拼接整体流程图。首先读入图像，对图像进行中值滤波去噪后判断是否存在镜头失真，如果存在失真则利用基于几何模型的校正方法实现畸变校正，并做柱面变换投影到同一平面，利用 SURF 提取特征点，然后运用 RANSAC 算法和八参数算法得到图像的几何变换模型，这时就得到了图像的配准结果。当图像中存在运动物体时利用最佳缝合线和泊松融合相结合的方法实现融合鬼影的消除，不存在运动物体则直接利用距离加权法融合。

对于多幅图像的拼接，主要有帧到帧、帧到拼接图像、拼接图像到帧和拼接图像到拼接图像四种拼接方式，他们各有利弊需要根据实际情况选择正确的拼接方式。

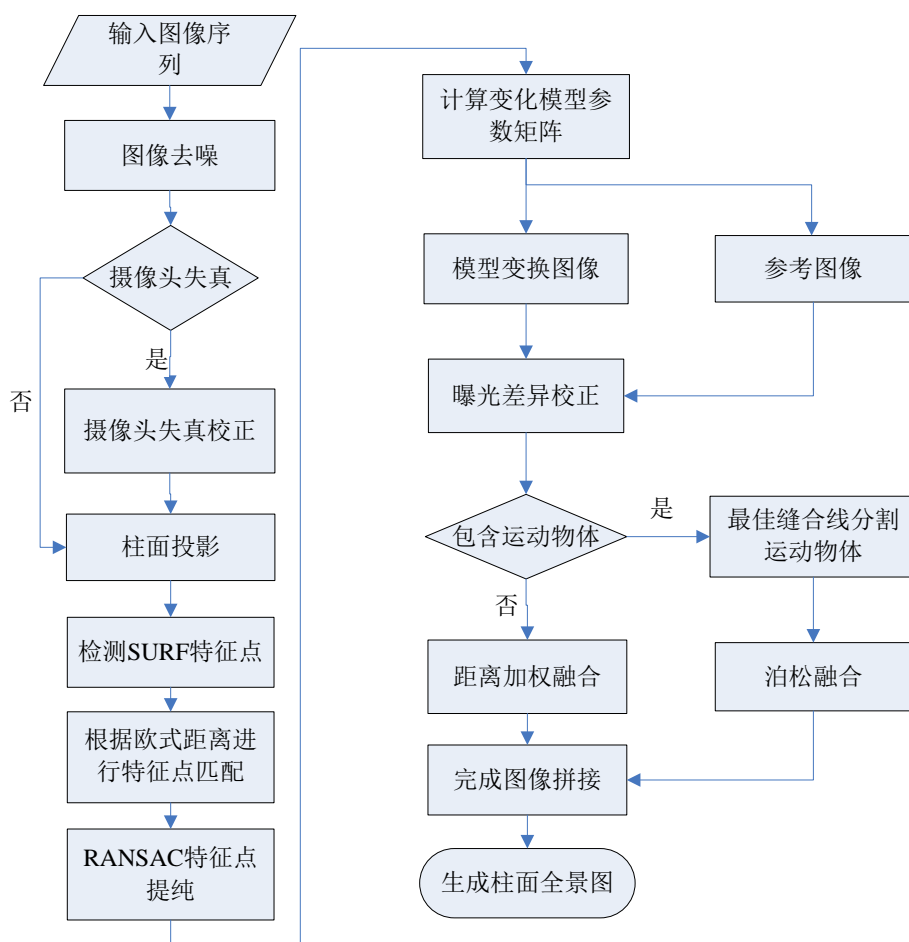


图 4-18 全景图像拼接系统流程

(1) 帧到帧的拼接方式：计算出相邻图像之间的变换模型参数，选择一幅图像作为参考图像，根据相邻图像之间的变换模型参数计算出其余图像与参考图像平面的变换参数，从而实现图像的拼接。这是一种静态的拼接方式，简单易实现速度快，但是因为变换参数及配准结果的累积误差不容易得到高质量的全景图像。图 4-19 为帧到帧的拼接方式示意图，图中 I_3 为参考平面，图像 I_1 到 I_3 的变换矩阵 M_{13} 由 M_{12} 和 M_{23} 计算得到，以此类推。

(2) 帧到拼接图像的拼接方式：将当前图像变换到已经拼接好的图像的参考系中进行拼接。这是一种动态的拼接方式，在精度方面有一定的提高但是，但是计算量也会随之提高，降低了图像拼接的效率。图 4-20 就是帧到拼接图像的拼接方式示意图。

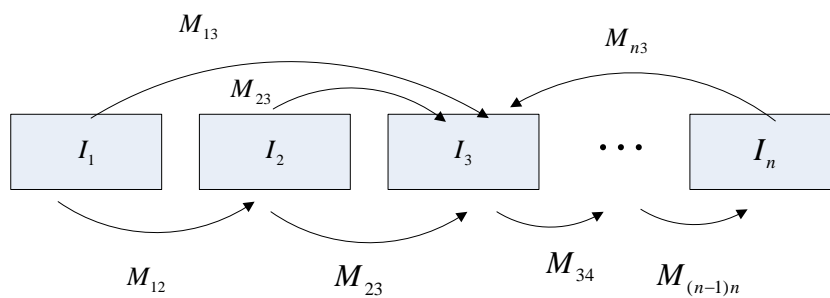


图 4-19 帧到帧拼接方式示意图

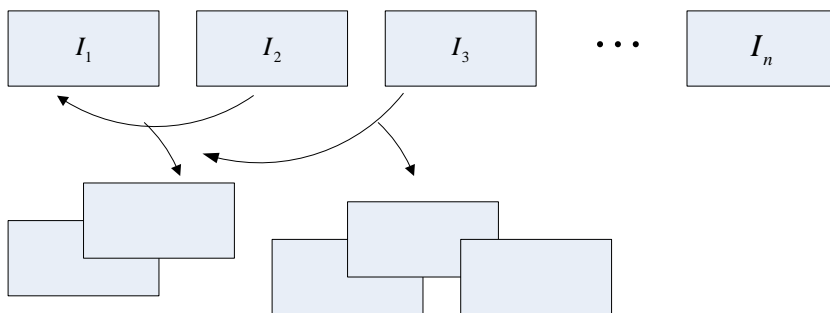


图 4-20 帧到拼接图像拼接方式示意图

(3) 拼接图像到帧的拼接方式：这种拼接方式与帧到图像的拼接方式类似，只是把当前图像作为参考图像。这种拼接方式不用每次重新计算变换参数，即能保证拼接效率又能在一定程度上提高拼接的准确程度。图 4-21 为拼接图像到帧的拼接方式示意图。

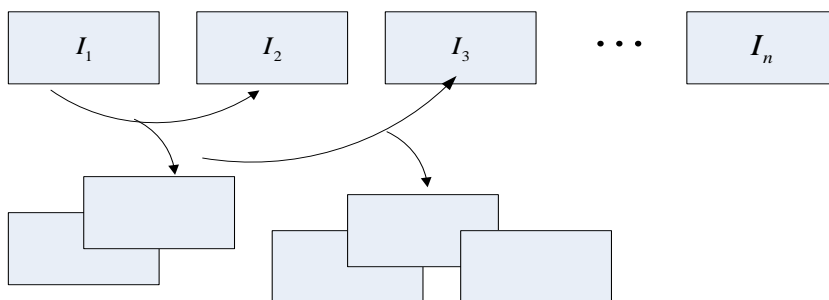


图 4-21 拼接图像到帧的拼接方式示意图

(4) 拼接图像到拼接图像的拼接方式：将图像序列进行分组，每组中按照一定的方式拼接，再将各组拼接得到的图像分组进行拼接，不断重复直到把所有图像拼接成为一幅图像为止。这是一种树形的拼接方式，这种方式相对前三种可以更大程度地降低累积误差。图 4-22 为拼接图像到拼接图像的拼接方式示意图。

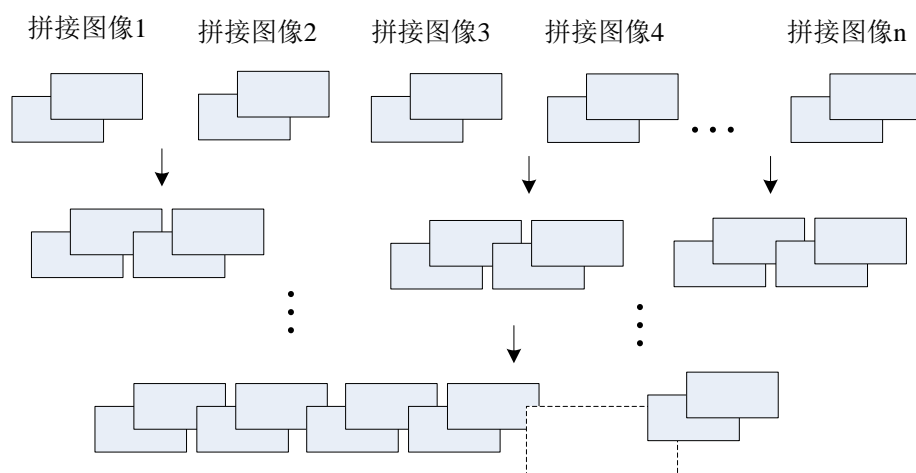


图 4-22 拼接图像到拼接图像的拼接方式示意图

为了实现手持相机拍摄图像的拼接本文选用精度较高的拼接图像到拼接图像的方式进行多幅图像的拼接。利用这种拼接方式首先对提取特征点后的图像序列进行分组，在每一组中计算变换模型参数，实现相邻图像间的配准。在每组中可以利用拼接图像到帧的方式进行拼接，这样可以适当减小计算量，为了提高拼接的效率也可以适当增加每组中的图像数量。

4.4.2 基于图像的全景漫游

全景图主要包括立方体全景、球面全景和柱面全景，2.4 节中已经对三种全景图像投影模型进行了分析，文中使用柱面模型投影生成柱面全景图。柱面全景是一种 IBR 技术即基于图像的绘制技术，柱面全景也被叫做柱面环境贴图，它可以在视点处在水平方向任意旋转观察场景，它获取图像简单，而且只在水平方向弯曲简化了图像变换的操作。全景浏览器就是要将拼接得到的图像映射到事先设定好的模型上，全景图是柱面的而用户观察到的场景是平面的，这个映射公式为：

$$\begin{cases} x_2 = f \arctan\left(\frac{x_1}{f}\right) \\ y_2 = \frac{y_1 f}{\sqrt{x_1^2 + f^2}} \end{cases} \quad (4-60)$$

式中 f 拍摄时相机的焦距， x_1, y_1 是在柱面全景浏览器中的坐标， x_2, y_2 是全景图展开后的坐标，即拼接得到的柱面图像。

这里实现的柱面全景浏览器是在 Window 系统下以 VC6.0 作为编程平台利用

OpenGL 实现的,它通过视角大小来估计拍摄图像时的焦距解决不同焦距图像显示的问题。该柱面全景图像浏览器允许用户点击鼠标旋转观察水平方向 360 度的场景。

4.5 测试结果及分析

图 4-23 为校园某处的图像拼接实验,图(a)为利用普通相机拍摄的一组待拼接的图像序列,拍摄时拍摄者手持相机,以其所在位置旋转拍摄。图(b)为使用文中的方法完成的柱面全景图像拼接的结果。从图(b)中看不出明显的拼接痕迹,实现了多幅图像的无缝拼接。



(a) 原图序列



(b) 拼接结果

图 4-23 全景拼接实验

图 4-24 是利用文中方法得到的三幅图像拼接结果, (a)、(b)、(c)三幅图像均 显得自然平滑, 配准准确不存在明显拼接痕迹。经过试验可以看出文中的算法可以

有效的实现多幅图像的无缝拼接。



(a)



(b)



(c)

图 4-24 全景拼接实验结果

全景图像配准的准确程度利用所有匹配特征点对的 MAD 和 $RMSE$ 以及所有重叠区域相关系数的均值来评价。由于全景图像拼接时存在着多处融合区域，我们无法利用单个区域的融合效果来评价整幅全景图的融合效果，把所有的融合区域看做是一个大区域来评价融合结果也是不太合适的，所以本文对每一个融合区域计算 $PSNR$ 和 MSE ，利用 $PSNR$ 的均值和 MSE 的均值来评价整幅全景图像的融合效果。

表 4-3 为图 4-23(b)和图 4-24(a)到(c)的评价指标。匹配特征点对的 MAD 和 $RMSE$ 说明图像配准的误差在可接受的范围之内，所有重叠区域相关系数接近于 1 说明本文的图像配准方法可以达到较高的精度。所有融合位置的 $PSNR$ 均值、 MSE 均值说明融合结果逼真程度较好没有较大的失真。表 4-3 内的测试数据说明了文中的全景图像拼接方法可以有效的实现全景图像的拼接。

表4-3 图像拼接结果评价

图像	MAD	RMSE	CC均值	PSNR均值	MSE均值
4-23(b))	3.604	4.045	0.953	46.652	1.432
4-24(a)	4.926	5.317	0.946	46.574	1.465
4-24(b)	5.452	6.136	0.951	46.137	1.568
4-24(c)	2.648	3.021	0.958	47.283	1.231

图 4-25 为实现的全景图像浏览器，它是一种基于柱面图像的全景漫游，图中显示的图像是利用本文的算法拼接得到的柱面全景图。在利用此全景浏览器观察场景时用户可以以拍摄点为中心水平旋转 360 度观察四周场景。图 4-25 (a)至(d)分别为观察者旋转约 90 度看到的场景。该全景显示系统使观察者仿佛真的身在场景中，给人一种身临其境的感觉。

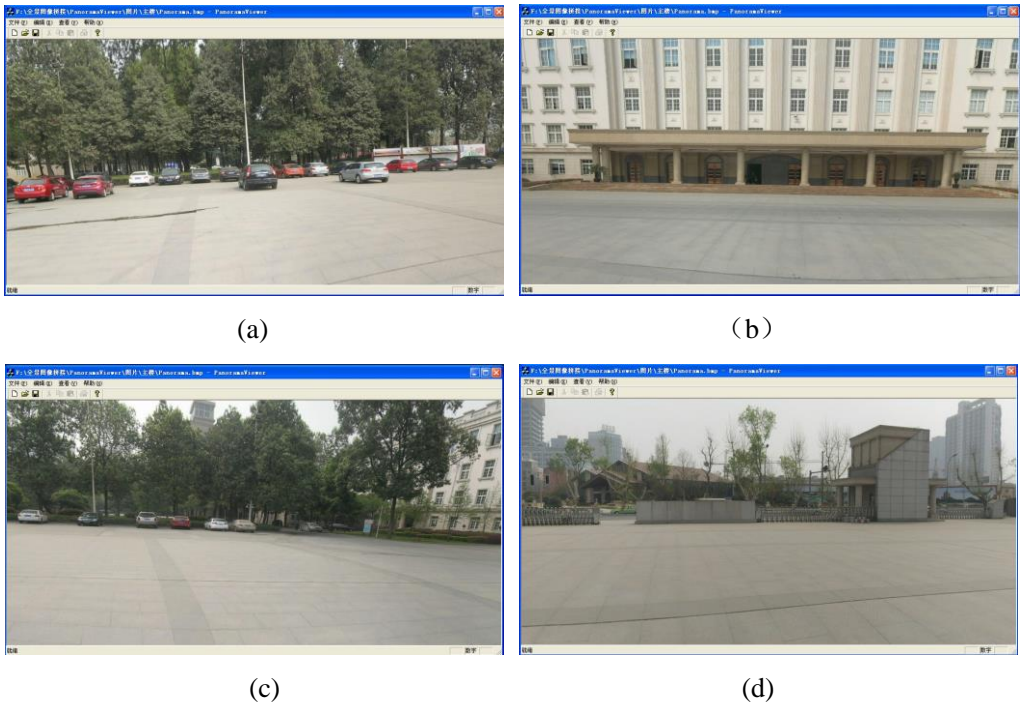


图 4-25 基于图像的全景漫游示意图

4.6 小结

本章研究了几种传统的图像融合方法并分析了他们的优点与不足，针对动态场景拼接使用最佳缝合线与泊松融合的方法实现了融合鬼影的消除。深入研究现有的几种最佳缝合线搜索准则，通过分析几种最佳缝合搜索准则提出一种改进的最佳缝合线搜索方法。它可以考虑到图像的曝光差异又可以充分利用图像像素点周围像素的相似性，并且对特征点处的缝合线准则值加权使缝合线中尽量包含多的特征点，减小了误差点的干扰。求解泊松融合方程时利用了傅里叶变换的方法提高了算法的速度。本章又介绍了全景图像拼接系统，实现了图像的全景漫游，通过大量实验证明了文中的全景图像拼接方法可以实现多幅图像自然无缝拼接。

第五章 总结与展望

5.1 全文工作总结

图像拼接技术就是为了解决了单幅图像视角不够的问题，全景技术目前已经广泛应用于虚拟现实、航空航天、军事应用、医学图像分析、遥感图像处理等多个领域。目前还没有一种拼接技术可以适用于大部分常用图像的拼接中。对于手持相机拍摄的图像往往因为抖动等原因使得图像拼接的难度大大增加，解决发生几何畸变及失真的图像的拼接具有着重要的意义。如何消除图像拼接后的鬼影，尤其是当场景中存在运动物体时消除运动物体融合时产生的鬼影也是图像拼接的一个难点。

本文针对全景图像拼接中存在的若干问题展开研究，通过对全文的回顾与总结，本文的主要工作及贡献有以下几点：

(1) 深入研究了图像的预处理，包括几种常用的几何变换模型、全景图像的投影模型和图像畸变校正。引入了一种基于几何模型的图像失真校正方法，可以快速有效的实现摄像头失真的校正，较为适用于待拼接图像序列的失真校正。

(2) 研究了基于角点检测的图像配准方法和基于不变量技术的图像配准方法。比较分析了 Harris 角点，SIFT 特征和 SURF 特征的性能，确定了比较有效的特征提取方法。同时深入研究了特征点提纯和变换模型参数的估计，实现了图像序列的全局配准，对手持相机拍摄图像的配准取得了不错的效果。

(3) 针对动态场景图像融合时运动物体导致的融合鬼影问题展开了研究，利用最佳缝合线与泊松融合消除了动态场景中的融合鬼影，实现了自然无缝的全景图像。研究了现有几种最佳缝合线的搜索准则，针对它们存在的问题提出了一种改进的最佳缝合线搜索准则，它减小了曝光差异的影响并充分利用了相邻像素点间的相似性，同时在一定程度上提高了算法的速度。改进的最佳缝合线搜索方法对特征点处的准则值加权减小了个别误差点的影响。求解泊松融合方程时利用了傅里叶变换的方法提高了泊松融合的速度。最后研究了全景图像浏览器的设计方法，实现了一种可以旋转观察周围场景的柱面全景浏览器。

5.2 工作展望

本文针对全景图像拼接中的畸变校正、图像配准以及图像融合方法等进行了研究，但因为时间原因和实验条件等限制，还有一些问题有需要进一步研究，以下几点是需要进一步研究的地方：

- (1) RANSAC 算法需要多次迭代运算速度慢，改进算法提高速度。
- (2) 简单的柱面全景浏览器只能视点固定的观察场景，可以研究一种变换视点的全景图浏览器，实现真正的全景漫游。
- (3) 研究一种多摄像头全景相机，并在硬件上实现全景图像拼接算法。
- (4) 拼接的实时性和复杂场景图像的配准仍需要进一步研究。
- (5) 研究一种适用于大部分常用图像的拼接方法。

致 谢

光阴似箭，岁月如梭，转眼三年的研究生生活就要过去，在此论文即将完成之际，我对这三年里给予我帮助与关怀的老师、同学、家人等表示由衷的感谢。

首先我要感谢我的导师彭真明教授，在他的悉心指导下本论文的研究工作才得以顺利展开。他渊博的知识和丰富的实践经验给予了我很大的帮助，他一丝不苟的治学态度、精益求精的科研精神以及无私忘我的工作作风深深地影响着我。彭老师不仅在科研方面给予我极大的帮助和指导，而且在生活中也给予我很大的关怀。他的谆谆教诲将是我一生的财富，在此，谨向我敬爱的彭老师表示最诚挚的敬意和由衷的感谢。

感谢陪伴我三年的同窗挚友：赖建宏、景亮、陈颖频、普婕、冯谦、袁园、谢春华、普艳香。是他们陪伴我度过了这三年的时光，在生活和研究工作中给予了我很大的关怀和帮助，使我受益匪浅。在这里我还要向“图形图像与信号处理应用研究室”的其他老师同学表示衷心的感谢。

我还要感谢我的父母，他们总是无微不至的关怀着我，无怨无悔的为我付出，辛辛苦苦的为我创造良好的生活与学习环境，他们的关怀与支持是我完成学业的动力。

最后要感谢所有帮助关心过我的人，同时也要向评审本论文的各位专家表示敬意与感谢！

参考文献

- [1] K. A. Yuksel, A. Yucebilgin, S. Balcisoy, et al. Real-time Feature-based Image Morphing for Memory-efficient Impostor Rendering and Animation on GPU[J]. Visual Computer, 2012, 29(2): 131-140
- [2] S. Wang, C. Lin, W. Tai. Compressing 3D Trees With Rendering Efficiency Based on Differential Data[J]. IEEE Transactions on Multimedia, 2012, 15(2): 304-315
- [3] R. W. Pazzi, A. Boukerche, J. Feng. A Novel Image Mosaicking Technique for Enlarging the Field of View of Images Transmitted over Wireless Image Sensor Networks[J]. Mobile Networks & Applications, 2010, 15(4): 589-606
- [4] 王海颖. 基于图像的虚拟漫游关键技术研究[D]. 北京: 清华大学, 2010,1-20.
- [5] 林婧. 基于序列图像的全景图像拼接计算研究[D]. 沈阳: 沈阳理工大学, 2010, 1-5.
- [6] 阮鹏. 多幅图像拼接算法研究[D]. 长沙: 中南大学, 2009, 1-8.
- [7] 钮永胜, 倪国强. 多传感器图像自动配准技术研究[J]. 光学技术, 1999, 1(1): 16-18
- [8] 蔡海洋, 付俊辉, 吕精巧. 医学图像三维拼接重建优化技术与仿真[J]. 计算机仿真, 2012, 29(7): 308-310
- [9] 梁栋, 颜普, 朱明, 等. 一种基于 NSCT 和 SIFT 的遥感图像配准算法[J]. 仪器仪表学报, 2011, 32(5): 1083-1088
- [10] Y. Keller, A. Averbuch, M. Israeli. Pseudopolar-based Estimation of Large Translations, Rotations and Scalings in Images[J]. IEEE Transactions on Images Processing, 2005, 14(1):12-22
- [11] C. D. Kuglin, D. C. Mines. The Phase Correlation Image Alignment Method[C]. Proceedings of IEEE International Conference on Cybernetics and Society, New York, 1975, 163-165
- [12] E. D. Castro, C. Morandi. Registration of translated and rotated images using finite Fourier transforms[J], IEEE Transactions on Pattern Analysis and Machine Intelligence, 1987, 9(5):700-703
- [13] C. Harris, M. Stephens. A Combined Corner and Edge Detector[C]. Proceedings of Alvey Vision Conference, Manchester, 1988, 147-151.
- [14] R. Szeliski. Video Mosaic for Virtual Environments[J]. Computer Graphics and Applications, 1996, 16(2): 22-30

- [15] R. Szeliski, H. Y. Shum. Creating Full View Panoramic Image Mosaics and Environment Maps[C]. Proc. of ACM SIGGRAPH, New York, 1997, 251-258
- [16] K. H. Jang, S. K. Jung, M. Lee. Constructing cylindrical panoramic image using equidistant matching[J]. Electronics letters, 1999, 35(20): 1715—1716.
- [17] M. Brown, D. Lowe. Invariant features from interest Point groups[C]. Proceedings of the 13th British Machine Vision Conference, Cardiff, 2002, 253-262.
- [18] M. Brown, D. Lowe. Recognizing Panoramas[C]. Proceedings of IEEE International Conference on Computer Vision, Washington, 2003, 1218-1225.
- [19] H. Bay, T. Tuytelaars, L. V. Gool. SURF: Speeded Up Robust Feature[C]. Computer Vision-ECCV 2006 Lecture Notes in Computer Science, Graz, 2006, 404-417.
- [20] O. S. Kwon, Y. H. Ha. Panoramic Video using Scale-Invariant Feature Transform with Embedded Color-Invariant Values[J]. IEEE Transactions on Consumer Electronics. 2010, 56(2): 792-798
- [21] S. G. Stanciu, R. Hristu, R. Boriga, et al. On the Suitability of SIFT Technique to Deal with Image Modifications Specific to Confocal Scanning Laser Microscopy[J]. Microscopy and Microanalysis, 2010, 16(5): 515-530
- [22] C. Valgren., A. J. Lilienthal. SIFT, SURF & seasons: Appearance-based long-term localization in outdoor environments[J]. Robotics and Autonomous Systems, 2010, 58(2): 149-156
- [23] B. Fang, C. Huo, C. Pan, et al. Registration of Optical and SAR Satellite Images by Exploring the Spatial Relationship of the Improved SIFT[J]. IEEE Geoscience and Remote Sensing Letters. 2013, 10(4): 657-661
- [24] H. Wang, Z. Peng, J. Liu, et al. Feature points detection and tracking based on SIFT combining with KLT method[C]. The 2009 International Conference on Optical Instrument and Technology, Proc. SPIE, Xi'an, 2009, 75062N
- [25] 王洪兵. 增强现实环境下基于视觉的高精度标跟踪技术研究[D]. 成都: 电子科技大学, 2010, 26-40
- [26] J. Shin, Y. Tang. The ghosting for Image Stitching with Automatic Content Awareness[J]. Pattern Recognition, 2010, 23(26):26-27
- [27] 张茂军, 钟力, 孙立峰等. HVS: 构造一个虚拟实景空间[J]. 自动化学报, 2000, 26(6): 736-740
- [28] 万华跟, 高曙明, 彭群生. VDVAS: 一个集成的虚拟设计与虚拟装配系统[J], 中国图像图形学报, 2002, 7(1): 27-35

- [29] 唐琰, 谷士文, 费耀平等. 全方位全景图像的一种映射方式[J]. 计算机工程, 2000, 26(8): 95-97
- [30] 李学庆, 孟祥旭, 杨承磊等. 一个基于球面映射的视景生成系统[J]. 系统仿真学报, 2001, 11(13): 129-133
- [31] S. Hua, Z. Ou, X. Wang. Constructing full view panoramic image based on spherical model[C]. Proceedings of SPIE, San Diego, 2003, 117-122
- [32] 马向英, 杜威, 袁晓君等. 基于图像的室内虚拟漫游系统[J]. 中国图像图形学报, 2001, 6(1): 86-90
- [33] S. E. Chen. Quick Time VR: an Image Based Approach to Virtual Environment Navigation[C]. Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques, New York, 1995, 29-38
- [34] 张全法. 平面物体在曲面状态下扫描仪图像的校正实验[J]. 光学技术, 2003, 29(4): 454-45
- [35] 宣海燕, 张峰, 申功勋. 大市场短焦距成像系统的快速校正[J]. 测控技术, 2007, 26(6): 28-30
- [36] 张淳, 秦世引, 赵丹培, 等. 基于星图匹配的小视场空间图像畸变校正[J]. 北京航空航天大学学报, 2008, 34(9): 1016-1019
- [37] 韩广良, 宋建忠. 一种基于畸变等效曲面的图像畸变校正[J]. 光学技术, 2005, 31(1): 122-124
- [38] Y. Kanazawa, K. Kanatani. Image mosaicing by stratified matching[J]. Image and Vision Computing, 2004, 22(2): 93-103
- [39] 漆驰, 刘强, 孙家广. 摄像机图像序列的全景图拼接[J]. 计算机辅助设计与图像学学报, 2001, 13(7): 605-609
- [40] 李萍, 周荣福, 刘苗. 柱面全景图单点投影及配准拼接算法[J]. 自动化技术与应用, 2011, 30(3): 36-39
- [41] 杨燕, 王雪瑞, 戴青, 等. 球面全景图像生成技术的研究[J]. 计算机应用与软件, 2007, 24(10): 164-165
- [42] 彭真明, 雍杨, 杨先明. 光电图像处理及应用[M], 成都: 电子科技大学出版社, 2008, 190-194
- [43] Z. Peng, H. Wang, G. Zhang, et al. Spotlight SAR images restoration based on tomography model[C]. The 2th Asia-Pacific Conference on Synthetic Aperture Radar, Proc. IEEE, Xi'an, 2009, 1060-1063

- [44] 冯宇平. 图像快速配准与自动拼接技术研究[D]. 长春: 长春光学精密机械与物理研究所, 2010:82-91
- [45] 彭真明, 张启衡, 魏宇星, 等. 基于多特征融合的图像匹配模式[J]. 强激光与离子束, 2004, 16(3): 281-285
- [46] 陈乐, 吕文阁, 丁少华. 角点检测技术研究进展[J]. 自动化技术与应用, 2005, 24(5): 1-5
- [47] 刘洁. 基于特征点的目标检测与跟踪方法研究[D], 成都: 电子科技大学, 2008, 27-39
- [48] 张瑞娟, 张建奇, 杨翠. 基于 SURF 的图像配准方法研究[J]. 红外与激光工程, 2009, 38(1): 160-165
- [49] K. Mikolajczyk. T. Tuytelaars, C. Schmid et al. A Comparison of Affine Region Detectors[J]. International Journal of Computer Vision, 2005, 65 (2): 43-72
- [51] A. F. Martin, C. B. Robert, Random sample consensus: a paradigm for model fitting with application to image analysis and automated cartography[J]. Communications of AMC, 1981, 24(6): 381-395.
- [51] 徐光著, 朱水莲, 封建军. 一种改进的动态场景拼接算法[J]. 电子科技, 2011, 24(7): 126-128
- [52] J. Davis, Mosaics of Scenes with Moving Objects[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Santa Barbara, 1998, 354-360
- [53] M. L. Duplaquet. Building Large Image Mosaics with Invisible Seam-Lines[C]. Proceedings of SPIE Aerosense, Orlando, 1998, 369-377.
- [54] A. Levin, A. Zomet. Seamless Image Stitching in the Gradient Domain[C]. 8th European Conference on Computer Vision, Prague, 2004, 377-389.
- [55] 方贤勇, 张明敏, 潘志庚等. 基于图切割的图像拼接技术研究[J]. 中国图像图形学报, 2007, 12(12): 2050-2056
- [56] A. Mills, G. Dudek. Image stitching with dynamic elements[J]. Image and Vision Computing , 2009, 27(10): 1593-1602.
- [57] V. Kwatra, A. Schodl. Graph cut Textures: Image and Video Synthesis Using Graph Cuts[J]. ACM Transactions on Graphics, 2003, 22(3): 277-286.
- [58] S. Zhao, Z. Peng, An effective method of eliminating ghosting in image mosaic[C]. Optoelectronic Imaging and Multimedia Technology II, Proc. SPIE, 8558, Beijing, 2012, 8558217.

- [59] 党建武, 宗岩, 王阳平. 基于 SIFT 特征的图像拼接优化算法研究[J]. 计算机应用研究, 2012,29(1): 329-332
- [59] P. J. Burt, E. H. Adelson. A Multiresolution Spline With Application to Image Mosaics[J]. ACM Transactions on Graphics, 1983, 2(4):217-236
- [60] 曾义, 彭真明. 一种基于视觉特征的多分辨率快速图像融合方法[J]. 成都信息工程学院学报, 2007, 24 (4): 510-512
- [61] 曾义. 基于小波变换与视觉特性的图像融合方法研究[D], 成都: 电子科技大学, 2007, 11-23
- [62] P. Perez, M. Gangnet, A. Blake. Poisson image editing[J]. ACM Transactions on Graphics, 2003, 22(3): 313-318.
- [63] 方贤勇. 图像拼接技术研究[D]. 杭州: 浙江大学, 2005, 49-52
- [64] J. M. Morel, A. B. Petro, C. Sbert. Fourier implementation of Poisson image editing. Pattern Recognition Letters[J]. 2012, 33(3): 342-348
- [65] 刘丽红. 多光谱图像融合及评价方法研究[D]. 成都: 电子科技大学, 2012, 44-52
- [66] 罗少鹏, 卢洵, 史建军. 一种基于清晰度指标的多聚焦图像融合评价方法[J]. 红外与激光工程, 2010,39 (5): 277-280

攻读硕士学位期间取得的研究成果

一、学术论文发表情况

- [1] **S. Zhao**, Z. Peng. An Effective Method of Eliminating Ghosting in Image Mosaic[C] Optoelectronic Imaging and Multimedia Technology II, Proc. SPIE, 8558, Beijing, 2012, 8558217 (EI: 20131516202219)

二、专利申请情况

- [1] 岳碧波, 彭真明, 袁园, **赵书睿**, 范弘毅, 谢春华, 赖建宏.基于角度梯度弹性阻抗的储层预测方法[P], 中国专利, 201110439293X, 2011.10

三、参与科研项目情况

- [1] ×××研究所合作项目, “基于图像特征的高精度目标跟踪方法研究及验证系统开发”, 研究年限: 2011.10-2013.10
- [2] 川庆钻探工程有限公司地球物理勘探公司合作项目, “叠前/叠后反演方法深化研究及模块集成”, 研究年限: 2011.7-2013.3