

长春理工大学

硕士学位论文

图像配准与拼接技术研究

姓名：林璐

申请学位级别：硕士

专业：模式识别与智能系统

指导教师：段锦

20100301

摘 要

图像拼接技术就是将两张或者两张以上存在重叠部分的图像，经过匹配对准后，构成一幅大型的无缝高分辨率图像的技术。图像拼接的两个核心问题是图像配准和图像融合，而图像配准是图像融合的基础。图像配准依据的是图像之间的相似度。图像的拼接过程一般由预处理、图像配准、图像融合三个步骤组成。近年来，图像配准与拼接技术越来越受到人们的普遍关注，该技术被广泛地应用到众多领域中。

本文阐述了图像拼接技术的发展现状，应用领域以及拼接方法，对图像拼接的两个主要问题，图像配准和图像融合分别进行了详细介绍。本文还对现有的图像拼接流程和图像配准方法进行了总结。对目前常用的图像融合方法也进行了详细的介绍。采用简单的边缘信息检测法实现了基准特征块的提取，利用相关系数法作为配准准则，使用金字塔式分层搜索策略进行块搜索，完成图像配准。图像融合上选择了线性加权的融合方法，使相邻图像间实现了平滑过渡，消除了拼接痕迹。实验证明，该种拼接算法减少了计算量，提高了图像拼接的精度和图像拼接的速度，具有较好的性能。

关键词：图像配准；图像拼接；图像融合；特征匹配

ABSTRACT

Image stitching technique is to overlap two or more parts of two images, through the match, after alignment to form a seamless large-scale high-resolution image technology. Image stitching of the two core issues is the image registration and image fusion. Image registration is the basis of image fusion. Image registration is based on the similarity between the images. Image stitching process is generally from the pretreatment, image registration, image fusion composed of three steps. In recent years, image registration and stitching technology, more and more people's attention, the technology is widely applied to many areas.

The article describes the status of the image stitching technology development, application areas and splicing method, the image mosaic of the two main issues, image registration and image fusion were described in detail. describes the origin of the image-splicing techniques, status, application areas and splicing method, the image mosaic of the two main issues, image registration and image fusion were described in detail. This is also the existing image stitching process and image registration methods are summarized. The current commonly used image fusion methods have also been introduced in detail. Used the simple edge information examination law to realize the datum characteristic block extraction, Use of correlation coefficient as a registration criterion. And use the pyramid hierarchical search block search strategy to complete the image registration. Image composition on the choice of adopting linear weighted method, so that adjacent achieved a smooth transition between images, eliminating the splice marks, improved image montage accuracy. Experiments show that the kinds of splicing to reduce the computational algorithms to improve the accuracy of image stitching and image stitching speed, with good performance.

Keywords: Image registration, Image stitching, Image composition, Feature matching

长春理工大学硕士学位论文原创性声明

本人郑重声明：所呈交的硕士学位论文，《图像配准与拼接技术研究》是本人在指导教师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

作者签名： 林琳 2010 年 3 月 25 日

长春理工大学学位论文授权使用授权书

本学位论文作者及指导教师完全了解“长春理工大学硕士、博士学位论文版权使用规定”，同意长春理工大学保留并向中国科学信息研究所、中国优秀博硕士学位论文全文数据库和 CNKI 系列数据库及其它国家有关部门或机构送交学位论文的复印件和电子版，允许论文被查阅和借阅。本人授权长春理工大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，也可采用影印、缩印或扫描等复制手段保存和汇编学位论文。

作者签名： 林琳 2010 年 3 月 25 日

导师签名： 段锦 2010 年 3 月 25 日

第一章 绪 论

1.1 引言

目前图像拼接技术现已成为计算机视觉和计算机图形学的研究焦点。图像拼接配准技术现已广泛应用于电子制造业、三维虚拟场景的模拟、数字视频会议、医学图像处理、机器人视觉、遥感航拍图像处理、数字图片的压缩等领域^[1]。给定某个场景的两张或两张以上互相之间有重叠部分的小尺寸图像,将两张或者两张以上存在重叠部分的图像,经过匹配对准后,构成一幅大型的无缝高分辨率图像的技术称为图像的拼接^[2]。由不同条件下得到的同一个图像或者不同图像间的信息数据之间存在不同程度上的差异,如平移、比例缩放、旋转以及几何畸变等,因此,我们需要对图像进行配准后才能进行下一步的处理过程,因此图像配准问题是图像拼接技术的最主要问题,也是图像拼接的核心问题,为了使图像有效的融合并且拥有令人满意的拼接效果,就必须找到相邻图像间重叠区域的具体所在和范围。比较两幅图像重叠部分所对应的颜色系统中灰度信息的相似程度,可以寻找到最佳匹配位置^[3]。而图像融合技术是为了把不同图像的数据信息组合后形一幅新的图像,从而达到改善图像效果的目的。图像融合可分为三个层次的融合:像素级融合、特征级融合和决策级融合。其中像素级融合是高层次的图像融合,它可以直接处理传感器采集的数据,得到融合后的图像,它也是目前图像融合研究领域的重点研究对象之一。

1.2 图像拼接配准研究背景及意义

当我们需要一幅大视角的大幅图像时,普通相机常常满足不了这一要求时,我们可以借助图像拼接技术实现这一要求。在对图像拼接处理时,相邻的图像之间必须有一定的区域重合。图像拼接就是确定相邻图像间在长和宽上的重叠程度,将这些重叠去掉,进行平滑处理后得到全景图像。

目前现有的图像采集设备虽然越来越先进,有的甚至可以做到采集全景图像,但是很难达到良好的视觉效果,并且需要高昂的成本做代价,但在人们日常的生产生活中,经常需要使用高分辨率的大幅图像,而普通的摄像设备很难满足这一要求。在进行航空拍摄时,由于距离有限,很难将大面积范围的场景拍摄下来。在对视频图像的编辑时,单纯依靠普通摄像设备很难实现这一过程。而利用大多数图像拼接处理软件对图像进行处理时,拼接效果往往不够理想,拼接的图像过多时需要反复的运作,工作量很大,费时费力,不能满足快速自动准确的拼接要求。快速自动高精度的实现高分辨率的图像配准拼接,是目前数字图像处理领域的研究热点。图像配准与拼接技术已经在计算机视觉领域中发挥了重要作用。

图像拼接技术可以快速拼接大尺度和超宽角度的静态和动态图像,这项技术的出现使普通的家用照相机和摄像机同样能够做到获取大尺寸大角度的图像和视频。而近年来图像拼接的发展趋势是如何利用普通的低端计算机自动生成大幅无缝的高清晰图像。在虚拟现实领域中,越来越趋向于基于图像的交互式研究,实现减少数据的传输量。因此,利用计算机和普通图像采集设备,实现图像拼接技术的应用前景十分广阔。

迄今为止,已经出现了大量的国内外图像配准研究成果,出现了很多图像配准方法。目前图像配准技术可以分为手工配准和自动配准。目前手工配准的方法已经被普遍应用,但是由于手动配准需要对图像提取大量的特征点,所有需要进行大量的运算,配准的时间也比较长。目前常用的配准方法有:

(1) 比值匹配^[4]: 先找到一个模板作为匹配模板,这一过程就是在第一幅图像的重叠区域中相邻的列上提取部分像素;然后在另一幅图像中搜索最佳匹配。该算法保证不了精度且计算量较大。

(2) 基于块的匹配^[5]: 同样需要找到用做匹配的模板,这里的匹配模板就是取第一幅图像处于重叠的部分,然后需要确定重叠范围,即在第二幅图像中搜索与匹配模板具有相近值的块。这种算法计算量很大但精度相对较高。

(3) 基于网格的快速匹配^[6]: 在基于块匹配的基础上,首先对模板进行粗略匹配,在水平方向或垂直方向每移动一个步长,计算对应像素点值的差的平方和,记录最小值的网格位置。然后,以此网格位置为中心进行精确匹配。搜索当前的最小值,每次移动的步长减为原来的 1/2,循环此过程,直到步长减少为 0,最后可以确定出最佳匹配位置。该算法虽然跟比值匹配和块匹配方法在运算速度上有所提高,但是很难达到精确匹配,因此这种算法的精度比较低。

目前的研究成果中,没有一种配准方法能够真正意义上完全涵盖所有的领域,以上这些方法大多是针对一定范围的应用领域。当需要快速准确地、保证高精度的配准图像以及进行图像分析时,就需要运用节省时间的自动化配准技术来实现图像的配准。因此研究并提出一种快速高精度的图像配准拼接算法具有十分重要的意义。

1.3 本文主要内容

本文第一章主要概述了图像拼接与配准技术的基本概念,介绍了图像配准与拼接技术的发展状况和应用的领域。第二章主要介绍了图像拼接技术的发展状况和算法基本原理、基本流程并且介绍了本文的主要工作内容。第三章介绍了图像拼接的首要步骤图像预处理的基本概念,介绍了两种图像校正方法:几何校正和灰度校正,还介绍了图像边缘检测的方法和几种常用的边缘检测算子。第四章介绍了图像配准技术的基本概念和配准流程,阐述了两类图像配准的方法:基于区域的配准方法和基于特征的配准方法,并详细介绍了几种基于这两类方法的配准

算法。第五章介绍了图像融合的基本概念和原理，并详细介绍了目前常用的融合方法。第六章采用边缘特征法提取特征，结合相关系数配准准则和金字塔分层搜索策略实现图像的配准，最后采用线性加权法融合配准后的图像，实现图像的平滑过渡。

第二章 图像拼接技术概述

2.1 图像拼接技术的发展状况

图像拼接技术就是将两张或者两张以上具有相同重叠区域的图像,经过特征匹配等一系列处理构成一幅大型的无明显拼接痕迹的高分辨率图像的技术。图像拼接的两个核心问题是图像配准和图像融合。图像拼接技术可以解决由于相机等成像仪器的视角的限制,导致不能生成大幅图像的问题。

图像拼接技术的研究现已成为计算机视觉和数字图像处理的研究热点。图像配准拼接技术越来越受到人们的普遍关注,该技术被广泛地应用到军事、科研生产和生活领域中。近年来,图像处理技术在医学领域也得到了广泛的应用,医学专家们也逐渐开始关注医学图像配准融合技术^[7-12]。医学图像融合可以将若干幅不同成像模式的图像结合起来,可以同时表达人体的多种信息,在对肿瘤的治疗上可以更为科学准确的诊断,并且具有无创伤、安全方便等优点,帮助了患者远离痛苦,受到了众多患者和医生的欢迎。目前图像配准技术已经广泛应用于临床诊断和治疗的各种成像技术,如超声波诊断和染色体分类等。CT 检测仪是目前图像处理在医学上应用最成功的例子,1971 年,英国科学家汉斯菲尔德成功地设计出名为 X 线电子计算机体层摄影机用于临床诊断。目前 CT 检测仪已经发展到了第五代,是医疗诊断中必不可少的设备。

国外从 20 世纪 60 年代就开始对图像配准领域进行研究,但直到上世纪 80 年代学者们才逐渐开始关注它。上世纪 90 年代初图像配准领域的研究成果开始大量涌现。90 年代末,科研工作者们已经解决了单模图像的配准问题,但多模态图像的配准问题却没有得到很好的解决。近年来,出现了许多有效的图像拼接方法,但是仍存在各种问题,没有一种方法能够避免所有外部因素的影响。1996 年,美国微软研究院 Richard Szeliski 教授提出了一种基于运动的全景图像拼接模型,该算法得到了业内人士的普遍肯定,被人们称为“经典算法”。四年后,Shmule Peleg, Benny Rousso 等人提出了自适应的图像拼接模型,实现了图像的自动拼接。Dai 等人运用几何特征导出投影变换,使变换更加精确,但是需要大量的计算,在图像特征抽取的问题上也比较困难,因此没有得到普遍应用,随后 Szeliski 等人又提出了一种平滑缝隙的方法,即对最小化图像的投影^[13]。

利用图像拼接技术进行的遥感图像配准:地理导航、数据融合、新传感器校准、地球资源调查、飞行器或行星探测、变化检测、基于内容的搜索、目标搜索、地图更新等^[14]。

目前大多数对遥感图像处理过程中都采用交互式的点匹配的方法选择特征点,因此计算多项式时需要大量分布均匀的特征点,这样才能避免产生配准误

差。通常情况下系统不能很好的自动处理遥感图像中大量的数据信息,因此这种方法需要耗费大量的人工和时间,所以实现自动图像配准的难度比较高,并不适用实时处理。因此自动化图像配准技术的研究比较热门,而对于基于特征点的自动图像配准研究则相对较少。

2.1.1 图像拼接算法

图像拼接可以基于图像的特征或灰度值,一般可以将图像拼接算法分为以下两个类型:

(1) 基于特征相关的拼接算法

基于特征的拼接方法是通过图像中的像素找出图像的特征,把图像的特征作为判断依据,搜索匹配图像重叠区域所对应的特征区域。

基于特征相关方法需要对待配准图像进行特征抽取和特征配准。从待拼接图像中抽取灰度变化明显的特征,在这些特征中,利用特征匹配算法提取出存在对应关系的特征对。引入边缘检测算子,对图像进行边缘检测和特征抽取,可以根据待配准图像之间的相同特征的几何关系确定配准参数。

(2) 基于区域相关的拼接算法

基于区域的拼接方法是基于图像灰度值信息的拼接,在待配准图像中选取一个区域,计算出与参考图像中的相同尺寸的区域,比较得出的灰度值差异后得到待拼接图像重叠区域间的相似度,从而确定了待拼接图像重叠区域的具体位置和范围。还可以计算两块重合区域对应像素点灰度值的相关系数,这样得到的相关系数与两幅图像的匹配程度成正比。基于区域的拼接还可以直接把每一像素点的灰度差值进行加权,但是这种简单加权容易导致拼接失败。还可以用傅立叶变换或相关函数等方法计算匹配参数。模板匹配是常见的最简单的基于区域的拼接方法。

比较基于特征拼接和基于区域的拼接:特征拼接是基于区域的拼接,区域拼接是基于像素特征的拼接。基于特征的拼接与基于区域的拼接方法相比,基于特征的拼接方法的相似度测量比较复杂,经常需要通过多种方法的结合来确定度量方法。另一方面需要事先选取出阈值,特征比较不容易产生变化,但计算量比较大,因此不适用于图像实时处理时的应用。当对特征差异比较不明显的图像进行拼接时,特征抽取的难度会大幅增加。

对于不同性质的原始图像需要应用不同的拼接算法:

(1) 基于仿射变换的拼接

仿射变换^[15]:将经过变换后的第一幅图像上的直线映射到第二幅图像后仍为平衡的直线。仿射变换可以分解为平移变换和线性变换,可以是扭曲、旋转、尺度、伸缩等方面的几何变换。可以用矩阵 A 表示,

$$A = \begin{Bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{Bmatrix} \quad (2.1)$$

$$A \text{ 可以分解为: } A = R(a)R(-b)DR(b) \quad (2.2)$$

其中 $D = \begin{Bmatrix} c_1 & 0 \\ 0 & c_2 \end{Bmatrix}$ 左上角的矩阵 D 为旋转部分, t_x 和 t_y 为平移因子, 有四个旋转,

一个水平方向平移和一个垂直方向平移这六个自由度。

(2) 柱面 / 球面 / 立方体全景图拼接

柱面全景图具有容易获得并且计算简单的特点, 它对水平方向的视场没有任何限制, 但对垂直方向上的视场有限制, 因此无法处理视点正上方和正下方的图像; 球面全景图则对视场没有任何限制, 因此产生的全景图视觉效果最好, 但容易出现两端画面歪斜变形的情况; 立方体全景图通常是由若干幅 90 度视角的图像组成的大幅整体图像。柱面、球面和立方体模型是投影变换时常用的运动模型

(3) 基于图像检索的图像拼接

把源图像分成若干小块, 对每个小块分别进行检索, 将它们与从图像数据库中找到与它最相近的图像进行替换, 从而形成新的图像。使用基于图像检索方法对进行图像进行拼接时需要运用同时运用图像数据库、模式识别、图像检索等技术^[16]。

(4) 基于重投影变换的拼接

重投影变换是中心投影的射影变换, 也叫透视变换, 进行图像校正时经常会使用到重投影变换。重投影变换的拼接^[17]时为了防止出现视觉差, 拍摄的景物必须近似为平面。如果不知道相机的运动方式, 或者不知道图像的顺序时可采用该种模型。如果将摄取到的图像映射到平行于地面的平面上, 就相当于把摄影机角度改变为垂直于地面向下看, 我们可以在这一平面上得到物体的真实形状。

2.1.2 图像拼接流程

图像拼接的一般步骤是^[18]: ①借助图像采集设备获取一系列场景图像; ②将获得的图像数字化; ③对图像进行预处理, 主要是几何校正, 消除参考图像和待拼接图像中明显的几何畸变; ④确定配准算法, 使用选定的配准算法对图像进行配准定位; ⑤图像融合, 消除拼接的痕迹, 使图像视觉达到令人满意的效果。图像拼接的过程一般又可以主要分为图像配准和图像融合两个重要过程。对于几何畸变比较严重的图像, 在配准前需要进行图像几何校正, 才能取得良好的拼接效果。图像配准的优劣决定了图像最终拼接效果的好坏。

图像拼接的过程一般由图像预处理、图像配准、图像融合三个步骤组成。拼接流程图如图 2.1 所示:

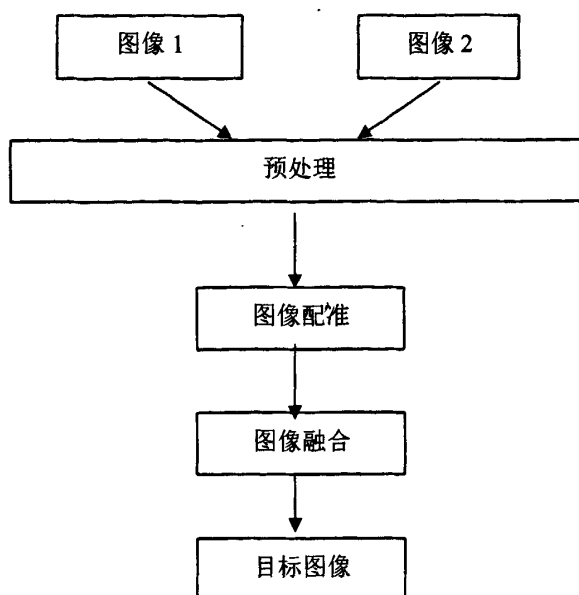


图 2.1 图像拼接流程图

首先获取待拼接的图像，可以使用普通图像采集设备，以三脚架固定设备底部，通过对观察点或焦距的变化拍摄获取图像；然后对图像进行预处理，让参考图像和待拼接图像不存在明显的几何畸变，从而去掉图像中不需要的信息，使图像质量能够满足图像配准的要求，主要是通过对图像进行几何畸变校正和噪声点抑制等处理实现的；接着进行图像配准，提取图像中的匹配信息寻找最佳匹配位置；最后进行图像融合，目的是为了消除拼缝实现整个拼接过程，主要是对边图像接缝处进行平滑处理。其中预处理不是必须的，可以根据待拼接图像的具体情况处理选择性的使用^[19]。

2.2 本文主要工作

(1)本文阐述了图像配准与拼接技术的研究现状和发展前景，总结了图像拼接配准概念、各类型算法以及原理，介绍了拼接和配准的流程，还介绍了图像融合技术的基本算法原理和基本流程。

(2)提出了利用边缘信息检测法进行基准特征块的选取，并引入阈值避免选到邻域孤立的特征点，从而提高图像拼接的精度。

(3)图像配准算法上采用改进了的相关系数法，通过对图像目标块的相关系数的计算，反映出当前像素块之间的相互关系。大大减少了计算量，节省了配准时间。

(4)在匹配搜索上，采用金字塔式分层搜索策略，由粗到细，由上到下逐步搜

索的方法。先从最低分辨率的分层开始,确定出粗匹配的位置,然后依次分层搜索下去,再进行精匹配,直到最大分辨率层,提高匹配速度。

(5)采用线性加权的图像融合方法,将原图像和目标图像的像素融合起来生成拼接平面上的像素,在两幅图的重叠区域按照一定权值将两幅图像的像素合成到新图中,实现了相邻图像间的平滑过渡,从而消除了拼接的痕迹。

(6)利用所选取的拼接方法,对特定图片进行拼接,检验其效果。

第三章 图像预处理

图像预处理是指对图像进行几何畸变校正, 边缘检测和噪声抑制等, 使参考图像和待拼接图像不存在明显的几何畸变, 是对待拼接图像进行特征抽取、图像分割和特征匹配等前所进行的处理。图像预处理的主要目的是消除图像中无关的信息, 恢复有用的真实信息, 简化数据并且增强信息的被检测性, 从而提高特征抽取、图像分割、匹配和识别的效果。

3.1 图像校正

图像校正是指对失真图像进行的复原性处理。引起图像失真的原因有: 几何畸变、带宽限制、成像系统的像素差等原因造成的图像失真; 由于图像采集设备拍摄角度引起的图像几何失真; 由于模糊运动、噪声引入、辐射失真等原因造成的图像失真。图像校正的基本思想是, 根据导致图像失真的不同原因, 建立相应的数学模型, 从待处理图像信号中提取所需要的信息, 进行导致图像失真的逆过程, 起到恢复图像的作用。实际上整个复原过程可以看作是设计一个滤波器, 使其能从失真图像中计算出真实图像, 根据选定的误差准则, 最大程度上去除杂质, 找到具有真实细节图像信息的图像。

图像校正主要分为两类: 几何校正和灰度校正。

3.1.1 几何校正

在不同条件下拍摄的图像, 常会发生几何失真, 出现歪斜变形的现象。

以一幅图像为基准, 去校正一幅以另一种方式摄入的图像, 以校正其几何失真和几何畸变就叫做图像的几何失真校正。

图像几何校正的思路: 是通过对无失真图像的某些像素点和畸变图像相应像素点的坐标进行比较, 找出他们的对应关系, 构造出多项式系数, 并把它作为恢复其它像素的基础。首先建立几何校正的数学模型; 然后计算出数学模型的各个参数值; 最后根据建立的数学模型对图像进行几何校正^[20]。最终目的是为了找到校正后图像中每个像素的亮度值。具体操作为:

①对图像像素进行坐标变换, 找到一种数学关系, 建立图像像素点坐标和参考图对应点坐标间的关系, 根据这一数学关系校正所选的图像各个控制点的坐标, 即在图像的行上逐点计算出校正后每个控制点所对应的原始图像的位置。

②计算灰度内插

用 $f(x, y)$ 表示几何基准图像的坐标系, 用 $g(x', y')$ 表示校正后的图像坐标系, 设两个图像坐标系之间关系如下:

$$x' = h_1(x, y) \quad y' = h_2(x, y) \quad (3.1)$$

$$\text{则有 } h_1(x, y) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} a_{ij} x^i y^j \quad (3.2)$$

$$h_2(x, y) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} b_{ij} x^i y^j \quad (3.3)$$

用二次型来近似，假设变换关系为：

$$\begin{aligned} x' &= a_0 + a_1x + a_2y + a_3x^2 + a_4xy + a_5y^2 \\ y' &= b_0 + b_1x + b_2y + b_3x^2 + b_4xy + b_5y^2 \end{aligned} \quad (3.4)$$

如果假设基准图像为 $f(x, y)$ ，失真图像为 $g(x', y')$ ，对于同一控制点它的灰度不变，则有如下关系 $f(x, y) = g(x', y')$ 。

3.1.2 灰度校正

灰度校正需要根据图像不同失真程度以及所选取的不同图像特征采用不同的修正方法。通常使用的方法主要有三种：

①灰度级校正：图像成像不均匀时使用这种方法，当图像成像不均匀时会使得图像明暗区分明显，校正后可以使图像灰度变的更加均匀。

②灰度变换：图像局部或全局曝光不充分的时使用该方法，变换后可以增强对比度。

③直方图修正：修正后可以使灰度分布满足后续处理要求，即使所需的图像特征更加明显。

(1) 灰度直方图

灰度直方图显示了不同灰度值的像素出现的次数，却不显示像素所在位置，表征了图像的一维信息；每幅图像只有一个与之对应的直方图，但不同图像却可能拥有同一直方图，即灰度直方图与图像之间的关系是多对一的映射关系；整幅图像的直方图是各子图的直方图之和。

如果图像的灰度级为 L ，则大小为 $m \times n$ 的灰度图像 $f(x, y)$ 的灰度直方图 $H[k]$ ， $k = 0 \dots L-1$ ，可以做如下计算获得：

$$(1) \text{直方图初始化: } \text{for}(k=0; k < L; k++) H[k] = 0; \quad (3.5)$$

(2) 统计：

$$\begin{aligned} &\text{for}(x=0; x < m; x++) \\ &\quad \text{for}(y=0; y < n; y++) H[f(x, y)]++; \end{aligned} \quad (3.6)$$

(3) 规格化：

$for(x=0; x < m; x++)$

$$for(y=0; y < n; y++) H[f(x, y)] / = float(m(n); \quad (3.7)$$

(2) 直方图均衡化

直方图可以逐点进行操作，从而使图像的灰度值发生改变，使每个灰度级的像素点数都相等，使直方图趋于平衡。直方图均衡化对于图像比较和图像分割起到十分重要的作用。

直方图均衡化变换：设灰度变换 $s = f(r)$ 为斜率有限的非减连续可微函数，输入图像的直方图为 $H_i(r)$ ，输出图像的直方图为 $H_o(s)$ ，均衡化变换可以将输入图像 $I_i(x, y)$ 转换成输出图像 $I_o(x, y)$ ，则由直方图的定义可知，图像经过灰度变换后输入图像 $I_i(x, y)$ 与输出图像 $I_o(x, y)$ 对应的像元相等：

$$H_o(s)ds = H_i(r)dr \quad (3.8)$$

变换函数 $f(r)$ 必须满足：

(1) $f(r) < 0 < r(1)$ 是单值函数、且单调增加；

(2) $0 < f(r) < 1$ ，且 $0 < r < 1$ 。

$r = f^{-1}(s)$ 为逆变换函数，同样需要满足这两个条件

它们保证了灰度级由暗到明的排列顺序，还将变换后像素的灰度级限制在了允许范围内。

3.2 边缘检测

边缘检测是图像处理和计算机视觉中的基本问题。图像的重要特征可以通过图像表面方向不连续，场景亮度变化，物质属性变化，深度上的不连续等变化反映出来。边缘检测的目的是为了找到并显示出数字图像中亮度比较变化明显的像素点。边缘检测是图像特征提取技术中的一个重要研究分支。

边缘是指图像周围像素屋顶状变化或像素灰度阶跃状变化的集合。边缘也可以被认为是图像中灰度发生明显变化区域的边界，边缘是图像中最基本的特征之一。边缘检测是通过边缘检测算子的计算找出物体的边缘，边缘检测算子是一组用于定位变化的图像强度函数，是种重要的局部图像处理方法。

边缘具有方向和幅度，水平的边缘像素值的变化比较平稳；垂直于边缘走向，像素值变化比较剧烈，可能呈现屋脊状：边缘大多在灰度值从增到减的变化转折

点上，它的二阶方向导数是在边缘处选取极大值或极小值；也可能呈现阶跃状：边缘的两边像素灰度值有明显差异，它的二阶方向导数在边缘处呈零交叉状。

经典的边缘检测方法是利用边缘检测算子进行图像边缘检测。常用的边缘检测方法有 Canny 算子，Sobel 算子，Laplacian 算子，Roberts 算子，Prewitt 算子等。

目前边缘检测分为三大类：

a. 基于一阶微分的边缘检测方法

对图像一阶导数进行检测确定图像边缘，使用的方法是将一阶微分算子同图像卷积相结合。其卷积过程为：

$$R = w_1 z_1 + w_2 z_2 + \dots + w_9 z_9 = \sum w_k z_k \quad (3.9)$$

其中 z_k 是与模板系数 w_k 相关联的像素灰度， R 为模板中心点的像素值。

基于一阶微分的边缘检测方法是通过比较典型模板的计算值，确定某个点是否在某方向的线上，以此达到检测目的。

基于一阶微分的边缘检测算子有：Sobel 算子，梯度算子，Roberts 算子，Prewitt 算子等。

b. 基于二阶微分的边缘检测方法

对图像二阶导数的零点进行检测，从而确定图像边缘，还可以通过二阶导数的正负来判断像素所在区域的明暗，使用二阶微分算子与图像卷积相结合进行检测。基于二阶微分的边缘检测方法，可以利用空间域的高通滤波器来检测孤立点，是针对点的检测方法。

基于二阶微分的边缘检测算子有：Laplacian 算子，LOG 算子等。

c. 基于曲面拟合的边缘检测方法

当遇到噪声比较严重的图像时，可用平面或高阶曲面来拟合图像中某一区域的灰度，求得外法线方向的一阶微分或二阶微分检测边缘，以此减少噪声对检测结果的影响，

用平面来拟合空间四邻的像素时，可将这一平面表示为：

$$p(x, y) = ax + by + c \quad (3.10)$$

其灰度值分别为： $f(x, y)$ 、 $f(x, y+1)$ 、 $f(x+1, y)$ 、 $f(x+1, y+1)$

$$\text{均方差为：} \varepsilon = \sum [p(x, y) - f(x, y)]^2 \quad (3.11)$$

利用式 (3.10) 和 (3.11) 可推导出：

$$a = \frac{1}{2} \{ [f(x+1, y) + f(x+1, y+1)] - [f(x, y) - f(x, y+1)] \}$$

$$b = \frac{1}{2} \{ [f(x, y+1) + f(x+1, y+1)] - [f(x, y) - f(x+1, y)] \}$$

$$c = \frac{1}{4} [3f(x, y) + f(x+1, y) + f(x, y+1)] \quad (3.12)$$

式中 a 为两行像素平均值的差分， b 为两列像素平均值的差分。

下面简单介绍几种边缘检测算子：

(1) 梯度算子

设函数 $f(x, y)$ 在 (x, y) 处的梯度为向量：

$$\nabla f(x, y) = \frac{\partial f}{\partial x} i + \frac{\partial f}{\partial y} j \quad (3.13)$$

表示为图像的梯度 $\nabla f(x, y)$ 所包含的灰度变化信息。记为：

$$e(x, y) = \sqrt{f_x^2 + f_y^2} \quad (3.14)$$

表示为 $\nabla f(x, y)$ 的梯度，其中 $e(x, y)$ 为边缘检测算。为了简化计算，可将 $e(x, y)$ 定义为偏导数 f_x 与 f_y 的绝对值之和：

$$e(x, y) = |f_x(x, y)| + |f_y(x, y)| \quad (3.15)$$

$$\text{梯度的方向角为：} \varphi(x, y) = \text{ctg} \left(\frac{\partial f_y}{\partial f_x} \right) \quad (3.16)$$

为了便于检测边缘点，需要选取适当的阈值 T ，对梯度图像进行二值化处理，则有：

$$g(x, y) = \begin{cases} 1 & \text{Grad}(x, y) \geq T \\ 0 & \text{其它} \end{cases} \quad (3.17)$$

(2) Laplacian 算子

Laplacian 算子为二阶导数算子，用于计算二维函数，并且不受方向的限制，所以计算量较小。Laplacian 算子可以借助模板计算数字图像的值，是图像进行微分计算和滤波处理的过程，它相当于一个高通滤波器，但是它的缺点是对噪声非常敏感，因此我们可以利用 Gauss 函数对图像进行平滑处理，然后用 Laplacian 算子求平滑后图像的二阶导数，以得到的零交叉点作为候选边缘，这就形成了 LOG 算子。LOG 算子利用了旋转对称的 LOG 模板与图像进行卷积，这样就确定了滤波器输出的零交叉点处。

二维函数 $f(x, y)$ 的拉普拉斯变换是一个二阶微分过程，可以定义为：

$$\nabla^2 f = \left[\frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial y^2} \right] \quad (3.18)$$

对其进行离散化处理可得到：

$$\nabla^2 f(x,y) = f(x+1,y) + f(x-1,y) + f(x,y+1) + f(x,y-1) - 4f(x,y) \quad (3.19)$$

将模板表示为数字形式；可以通过多种方法实现，中间像素的系数为负数，并且其周围像素的系数为正整数，因此系数之和必须等于 0。

(3) Canny 算子

Canny 算子^[21]是 Canny 于 1986 年提出的边缘检测算子，使用该算子检测出的边缘点必须尽可能位于实际边缘的中心，还要最大程度的抑制虚假的边缘，也就是必须精确的确定边缘的位置，同时还要有效的抑制噪声，使输出的信噪比为最大。Canny 边缘检测算法首先要用高斯滤波器对图像进行平滑处理，然后用一阶偏导的有限差分计算出梯度的幅值和方向，并且要非极大值的抑制梯度幅值，最后采用双阈值算法来检测和连接边缘。

(4) Sobel 算子

Sobel 算子是先做加权平均，然后再微分，即：

$$\begin{aligned} \Delta_x f(x,y) &= [f(x-1,y+1) + 2f(x,y+1) + f(x+1,y+1)] - \\ &\quad [f(x-1,y-1) + 2f(x,y-1) + f(x+1,y-1)] \\ \Delta_y f(x,y) &= [f(x-1,y-1) + 2f(x-1,y) + f(x-1,y+1)] - \\ &\quad [f(x+1,y-1) + 2f(x+1,y) + f(x+1,y+1)] \end{aligned} \quad (3.20)$$

Sobel 算子梯度有一定的抑制噪声的能力，在检测阶跃的边缘时，检测后得到的边缘宽度至少为两个像素。

(5) Roberts 算子

Roberts 算子是对角线两个像素间做差运算：

$$\begin{aligned} \Delta_x f(x,y) &= f(x,y) - f(x-1,y-1) \\ \Delta_y f(x,y) &= f(x-1,y) - f(x,y-1) \end{aligned} \quad (3.21)$$

Roberts 算子梯度是在以 $\left(x-\frac{1}{2}, y-\frac{1}{2}\right)$ 为中心上的连续梯度的近似。

(6) 阈值确定

计算出导数之后，接下来就要给出一个阈值来确定边缘的所在位置。为了减少原有噪声对图像的影响，可以提高阈值，阈值提高了能够检测出的边线就减少了，很容易从图像中挑出不相关的特性。但是，阈值过高将会遗失细小的线段。阈值过低，检测出的边线过多，同样增加了噪声对图像的影响。解决这一问题的一个常用的方法是带有滞后作用的阈值选择，该方法使用了不同的阈值去寻找边缘。首先使用一个阈值上限去寻找边线的起始点，找到起始点后在图像上逐点搜索边缘，当数值高于阈值下限时持续记录边缘位置，一旦数值低于阈值下限时停止。简单介绍几种阈值确定的方法：

①迭代法选取阈值

迭代法是建立在逼近的思想的基础上的，其步骤如下：

a. 分别求出图像的最大灰度值和最小灰度值，分别记为 Z_{MAX} 和 Z_{MIN} ，令初始阈值

$$\text{为： } T_0 = (Z_{MAX} + Z_{MIN}) / 2 \quad (3.22)$$

b. 根据阈值 T_k 将图像分割成两部分，分别求出两者的平均灰度值 Z_0 和 Z_B ；

c. 求阈值 $T_{k+1} = (Z_0 + Z_B) / 2 \quad (3.23)$

d. 若 $T_k = T_{k+1}$ ，则所得即为最终阈值；否则继续重复求 Z_0 和 Z_B ，直到满足 $T_k = T_{k+1}$ 时停止运算。

②由直方图灰度分布选择阈值

利用图像灰度统计信息的方法显示灰度分布，通常利用一维直方图阈值化，就是从直方图中选取合适的阈值进行图像分割处理即可。由直方图灰度分布选择阈值的方法简单并且常用。

③双峰法选择阈值

双峰法的基本原理是把图像看成是由前景和背景组成，在灰度直方图上前景和背景的像素灰度分布形成了波峰状。图像分割的阈值就处在双峰之间的最低谷处，这样就可以简单的计算出阈值。但是这种方法的缺点是容易丢失一些图像细节信息。

除此之外还有动态阈值，大津法选择阈值，全局阈值等阈值选取的方法。

第四章 图像配准算法原理

4.1 图像配准基本概念

图像拼接的第一步就是使多幅图像在空间域中的几何位置能够完全对齐，称为配准。图像拼接的第二步就是将配准后图像信息进行整合显示，称为融合。

图像配准就是把同一场景在不同时间、不同观察点、不同拍摄条件下获取的两幅或多幅图像进行匹配和重叠的过程。它已经被广泛地应用于计算机模式识别、遥感图像处理等领域。

判断图像之间的相似度是图像配准的依据。图像的采集方式不同，可以使用不同的方法进行配准，主要有以下四类^[22]：（1）模板匹配：将场景变换到模型上的配准；

（2）时间序列配准：对在不同时间内采集的图像之间进行配准；（3）多模态配准：对通过多个传感器采集到的图像进行配准配准。（4）多观察点配准：对基于不同的观察点处采集到的图像之间的配准。

图像配准基本可以定义为^[23]：两幅待配准的图像 A 和 B，分别用 $I_1(x, y)$ 和 $I_2(x, y)$ 表示，设 A 为参考图像，则 B 为待配准图像，那么两图像间的坐标变换的位置可以表示为：

$$I_2(x, y) = g(I_1(f(x, y))) \quad (4.1)$$

其中 g 是一维的灰度变换， f 是二维的空间变换，该过程就是匹配过程。图像配准最主要的问题是寻找最佳的特征空间^[21]。在机器人视觉识别的过程中，常常需要把不同传感器或同一传感器在不同时间、不同成像条件下对同一场景获取的两幅或多幅图像在空间上对准，该过程即为匹配过程。在遥感图像处理中需要把不同的传感器对同一景物的多光谱图像按照像素点对应配准，然后根据性质对这些像素点进行分类。配准对同一地点却在不同时间段拍摄的两幅照片时，需要确定出两幅照片中哪些像素点的特征发生了明显变化，这样我们就可以找出图中发生变化了的区域。

配准前选取连续图像中的一幅图像作为配准的基准，该幅图像就是我们所说的参考图，再选取另一幅图像做为搜索图。在选取参考图上以某一像素点做为中心选取一个子图像块，该子图像块我们称之为图像的匹配模板，使选取的匹配模板在搜索图上按一定步长移动，每位移一次，比较模板与搜索图中对应部分的相关系数，直到找到最佳的匹配位置时停止移动，图像配准多应用于多幅图像的合成处理。图像配准方法基本由以下四个部分组合^[24]：

（1）特征提取：从图像中提取出用于匹配的有用的特征。

（2）配准准则：用来确定图像特征空间的变换集，可以找出最佳变换用于匹配两幅图

像。

(3)搜索策略：搜索最佳变换参数所采用的正确算法，用于寻找下一个变换,解决了测试并搜索出最佳变换方式的问题。

(4)比较相似性：它决定了配准中的相关特性，可以对两幅图像之间的相似性进行衡量。

选择不同的特征提取方法、配准准则和搜索策略时，图像的配准方法也有所不同，可分为 4 类^[23]：基于弹性模型的匹配法、频域傅里叶法、基于全局准则的配准方法以及基于特征的匹配法。

4.2 配准流程

通常意义上的图像配准可分为以下四个步骤：(1) 提取有用的特征 (2) 对特征进行配对 (3) 选取正确的特征变换参数 (4) 进行特征变换。

首先对两幅图像进行特征提取得到特征点；然后在对特征进行相似性度量后找到匹配的特征点对，通过找到的匹配特征点对得到图像的坐标变换参数：最后由这些坐标变换参数实现图像的配准。图像的配准精度过低会直接导致拼接图像的效果失真。而特征提取是图像配准中的关键问题，特征匹配成功与否主要取决于特征提取的准确度。因此，需要找到一个能够对齐待配准图像中相互重叠部分的变换，这样图像经过融合后可以形成一个新的图像。只有找到能够很好的描述两幅配准图像之间映射关系的模型才能做到对图像实现精确配准, 对应的匹配模型由以下八个参数决定如下式所示^[26]：

$$p(x, y) = \frac{A \begin{pmatrix} x \\ y \end{pmatrix} + B}{C^T \begin{pmatrix} x \\ y \end{pmatrix} + 1}, \left(A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, C = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \right) \quad (4.2)$$

式中 A 为缩放旋转, B 为平移, C^T 则表示投影。

4.3 图像配准方法

图像配准是图像拼接中的关键环节，同时图像配准方法的选择也决定了图像拼接的算法。目前对图像配准研究可以把图像配准方法分为两类：

(1) 基于区域的方法

利用图像本身具有的灰度信息来度量图像的相似程度。简单的说就是，将待配准图像中的一个区域与同样大小的参考图像在的区域灰度上进行比较，取标准化交叉相关系数的最大值。还可以通过变换将图像从时域变换到频域再进行配准。

提取出待配准图像重叠区域中的特征像素点，通过评价函数计算出参考图像与待配准图像之间的相似度。基准图像与待配准图像的互相关函数可表示为：

$$C(u,v) = \frac{\sum_x \sum_y T(x,y)I(x-u,y-v)}{\left[\sum_x \sum_y T^2(x,y)\right]^{1/2} \left[\sum_x \sum_y I^2(x-u,y-v)\right]^{1/2}} \quad (4.3)$$

互相关系数达到最大值时，则两幅图像完全匹配。

除此以外进行图像配准的评价函数还有两幅图像像素点绝对值和最小和两幅图像灰度差值平方和最小：

当利用两幅图像像素点绝对值和的最小的评价函数来进行图像配准时其评价函数表示为： $E = \sum |I(x,y) - I'(x',y')|$ (4.4)

当利用两幅图像灰度差值平方和最小的评价函数来进行图像配准时其评价函数表示为： $E = \sum [I(x,y) - I'(x',y')]^2$ (4.5)

大多数基于区域的配准方法采用互相关技术或傅立叶变换技术来实现^[27, 30]，多用于配准有大面积重叠区域的图像，但是这种方法却不适用于处理灰度特征差异较大的两幅图像。基于区域的匹配主要是模板匹配和基于相位相关的匹配方法。

(2) 基于特征的方法

基于特征的方法首先要对待配准图像进行图像的特征提取和图像分割，这里图像分割技术应使用在特征提取和边缘检测上；再利用提取到的特征对两幅图像进行特征匹配，最后通过特征的匹配关系能够建立图像的配准映射关系。基于特征的方法往往会将特征点、直线段、边缘、及轮廓等作为特征提取出来。基于特征的方法^[28-30]适用于来自不同传感器图像的配准，它的主要过程就是提取待配准图像的共同特征结构。基于特征的方法基本步骤：

(a)特征提取：从两幅图像中提取特征集，这里特征集主要由若干个灰度变化明显的点、线等特征组成。

(b)特征匹配：对特征集进行匹配。在两幅图像对应的特征集中利用特征匹配算法尽可能的将存在匹配关系的特征对选择出来。

(c)变换模型：对匹配的特征点集使用最小二乘法得到变换参数，找到特征的多项式变换关系，使得特征点能够替代重叠区域。

(d)图像插值重采样：利用上述多找到的项式变换关系对图像进行几何校正，从而实现两幅图像之间各个像素的配准运算。对待配准图像进行高精度重建，这里通常使用三次样条函数插值。当图像被放大时，像素点也随之相应地增加，这一放大的过程就是插值过程，程序会自动选取有利于配准的像素作为增加的像素，所以图像视觉上比较平滑。目前常用的插值方法有三种：①最近邻插值：图像出现了马赛克和锯齿等明显走样的原因。但是最近邻插值法的优点是速度快。②样条插值：利用周围一定范围内的若干像素点，试图让插值的曲线变得更为平滑③双线性插值：这种方法计算量较大但图像效果最好，不会令像素值不连续，但会使图像边缘产生模糊现象，线性插值速度也相对较慢。

目前基于特征的图像拼接方法最为常见。基于特征的配准方法定位准确,且要首先进行特征提取,但是这一方法的缺点是过分依赖于图像特征,需要进行手动提取特征点,所以计算量非常大,容易产生很大的匹配误差。基于特征的配准包括基于特征点集的匹配和基于图像边缘信息的匹配算法。

4.4 基于区域的配准

4.4.1 金字塔分解法

图像金字塔配准法主要是把原始图像分解成多个不同空间分辨率的图像块,将较小尺寸,较低分辨率的图像块放在上面,而将较大尺寸,较高分辨率的图像块放在下面,这样就构成了一个金字塔形状,人们形象的称他为金字塔分解法。

分析上层图像后得到的信息可以对下层图像的分析作为依据,这样可以很大程度上简化分析过程并降低算法的复杂度。在进行空间搜索时,首先进行粗匹配,每在水平方向上或垂直方向上移动一个步长后,计算出相对应像素点灰度差的平方和,记录平方和为最小值时对应的网格位置。然后,以此位置为中心进行精匹配,将每次移动的步长减为原来一半,继续搜索平方和最小值,循环这个过程,直到步长为零,最终可以确定出最佳匹配位置。简单的说就是先将图像分成若干个大小为 $m \times n$ 的图像块,把每个图像块看做是一个像素进行配准。然后,把 m 的大小减半,继续把图像分块;以此类推,直到图像块的大小为一个像素为止。

4.4.2 相位相关法

相位相关法是基于区域的配准常用算法。将图像利用傅立叶变换使空域变换到频域以后再进行配准,是对相邻两幅图像间的位移、旋转和比例的变换。

设两幅待配准图像为 f_1 和 f_2 ,它们的傅里叶变换分别为 $F_1(u,v)$ 和 $F_2(u,v)$,则它们之间有如下关系:

$$F_2(u,v) = F_1(u,v)e^{-j2\pi(ux_0+vy_0)} \quad (4.6)$$

设 $f_2(x,y)$ 为 $f_1(x,y)$ 在 x 和 y 方向的空域上平移相关,平移量分别为 x_0 和 y_0 :

$$f_2(x,y) = f_1(x-x_0, y-y_0) \quad (4.7)$$

则 $f_1(x,y)$ 和 $f_2(x,y)$ 的互功率谱为:

$$\frac{F_1(u,v)F_2^*(u,v)}{|F_1(u,v)F_2^*(u,v)|} = e^{j2\pi(ux_0+vy_0)} \quad (4.8)$$

其中 F_2^* 表示 F_2 的复共轭, $e^{j2\pi(ux_0+vy_0)}$ 的反傅里叶变换是一个冲击函数。进行傅立叶逆变换,在 (x,y) 空间的 (x_0,y_0) 处将会形成一个单位脉冲函数,位置即为两幅被配准

图像间的相对平移量 x_0 和 y_0 。找出单位脉冲函数为最大值时对应的位置，就可以得到两幅图像大致的相对位移。

用傅立叶变换处理具有平移和旋转的图像时一般需要先进行旋转，再平移。相位相关法的优点是精确而简单，缺点在于使用相位相关法配准时图像间的重叠率如果低于 50% 时则无法实现图像匹配。

4.4.3 模板匹配法

模板匹配^[31]是在计算机视觉中常见并且有效的定位识别技术，它可以比较待配准图像和模板间的相似度，更加直观的反映出图像之间的相似度，这种方法比较简单，但计算量较大，定位准确率也比较低。

模板匹配法是在一幅图像中选取一个目标窗口作为模板，模板的大小通常为 5×5 或 7×7 ，然后找到它在搜索图中坐标所对应的位置，通常是通过对相关函数的计算来确定位置。将模板 T 在被搜索图 S 上平移，我们把搜索图中被模板覆盖的那块区域叫做子图。 i 和 j 为子图在被搜索图 S 上的坐标。把子图的中心点在被搜索图 S 中的坐标 (i, j) 称为参考点。模板与被搜索图的关系如图 4.1 所示：

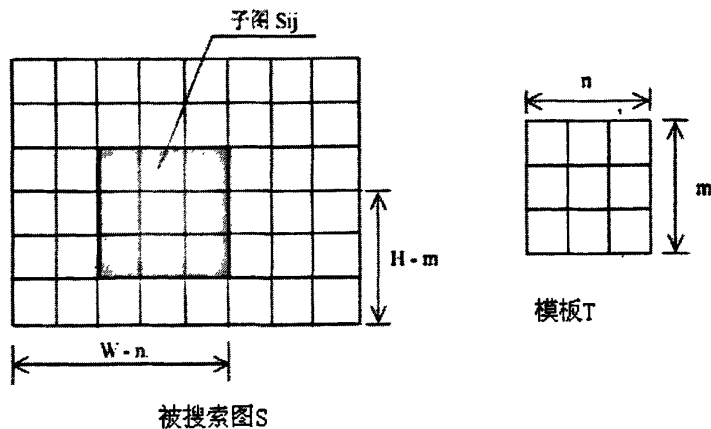


图 4.1 模板与被搜索图的关系

已知原始图像 $S(W, H)$ ，利用以下公式比较它们之间的相似度可以得到：

$$\begin{aligned}
 D(i, j) &= \sum_{m=1}^M \sum_{n=1}^N [S^{ij}(m, n) - T(m, n)]^2 \\
 &= \sum_{m=1}^M \sum_{n=1}^N [S^{ij}(m, n)]^2 - 2 \sum_{m=1}^M \sum_{n=1}^N S^{ij}(m, n) \times T(m, n) + \sum_{m=1}^M \sum_{n=1}^N [T(m, n)]^2
 \end{aligned} \tag{4.9}$$

其中 $\sum_{m=1}^M \sum_{n=1}^N [S^{ij}(m, n)]^2$ 为子图的能量， $\sum_{m=1}^M \sum_{n=1}^N [T(m, n)]^2$ 为模板的能量，子图的能

量和模板的能量都与模板匹配无关。 $2\sum_{m=1}^M\sum_{n=1}^NS^{ij}(m,n)\times T(m,n)$ 是模板和子图模板的互相

关值，它随 (i,j) 的变化而改变。可以通过比较 T 和 $S(i,j)$ 之间的相似性，完成模板匹配过程。模板 T 和子图 S 匹配时为最大值，我们可以将其归一化，能够得到模板匹配的相关系数，如下表示：

$$R(i,j) = \frac{\sum_{m=1}^M\sum_{n=1}^NS^{ij}(m,n)\times T(m,n)}{\sqrt{\sum_{m=1}^M\sum_{n=1}^NS^{ij}(m,n)\times T(m,n)}\times\sqrt{\sum_{m=1}^M\sum_{n=1}^N[T(m,n)]^2}} \quad (4.10)$$

当模板和子图完全匹配时，相关系数 $R(i,j)=1$ 在被搜索图 S 中完成全部搜索后，匹配目标块就是 R 的最大值所对应的子图。该种图像匹配法计算较量大，所以速度慢。为了简化运算使用另外一种算法来衡量 T 和 $S(i,j)$ 之间的误差，其公式为：

$$E(i,j) = \sum_{m=1}^M\sum_{n=1}^N|S^{ij}(m,n)-T(m,n)| \quad (4.11)$$

先计算出两个图像之间的向量误差，可以减少计算时间，然后根据匹配方向的不同，选取一个误差阈值 E_0 ，当 $E(i,j)<E_0$ 时，则匹配失败，当 $E(i,j)>E_0$ 时，则停止该点的计算，继续下一点的计算直到匹配成功。

改进的模板匹配算法是将模板匹配过程更改为进行两次匹配；第一次匹配为粗匹配；取 $1/4$ 的模板，在被搜索块上隔行隔列的进行匹配，也就是在原图 $1/4$ 的范围内进行匹配，这样计算的数据量被大幅减少，因此匹配速度也得到了大幅的提高。此外还需要提供适宜的误差阈值 E_0 可由以下公式得出：

$$E_0 = e_0 * (m+1)/2 * (n+1)/2 \quad (4.12)$$

其中 e_0 是所有像素点的最大平均误差， m 为模板的长度， n 为模板的宽度；第二次匹配为精匹配。在上次误差最小点 $(i_m i_n, j_m j_n)$ 的邻域内，进行搜索匹配得到最终匹配结果。

4.5 基于特征的配准

4.5.1 特征点匹配法

特征点表示了邻域局部特性的位置，标识了具有一定特征的局部区域位置。特征点可分为狭义特征点和广义特征点两种。狭义特征点是针对像素点本身定义的，如交

叉点、边缘点等。广义特征点则是基于区域定义的，可以是某一特征区域的某一相对位置或这一区域的中心，广义特征点可以认为是特征块。用于点特征提取的算子称为有利算子。基于特征点的匹配法可以降低因图像小角度旋转造成的误匹配。目前 RANSAC 算法和 Moravec 算法是特征点匹配法的两种常用算法。

(1) RANSAC 算法

RANSAC 算法是基于特征点匹配的典型算法。该算法非常适用于基于角点匹配的图像配准，不需要额外进行梯度运算。RANSAC 算法进行特征点匹配时首先需要随机抽取现在这幅图像与另一幅图像间的 8 组特征匹配点，使用 $p^F q = 0$ 计算图像的基础矩阵。使当现在这幅图像上的点 p 与另一幅图像上匹配点 q 满足外极线约束，计算 q 点到外极线 l_p 的距离 d ，再计算点 p 到外极线 l_p 的距离 a 。如果距离 d 与距离 a 的最大值小于规定的阈值，那么 p, q 点完成匹配，如果大于规定阈值则为误匹配，记下正确的匹配点数。接下来，使用所有正确匹配点重新构建基础矩阵 F 。最好在此基础上重新搜索正确的匹配点对，并重复使用正确匹配点重新构建基础矩阵 F 和重新搜索正确的匹配点对，直至匹配点对的数目达到理想状态。

(2) Moravec 算法

Moravec 算法是利用灰度方差提取特征点的算子。通过研究像素点水平，垂直和对角线方向上的邻近像素点的灰度变化情况，选择图像中灰度变化最为明显的点作为特征点，算法具体步骤如下：

①在参考图像 T 的重叠区域中选取 4 个区域，利用 Moravec 算子抽取出每个区域的特征点，计算出各像素的兴趣值，算法如下：

$$\begin{aligned} V_1 &= \sum_{i=-k}^{k-1} (g_{r+i,c} - g_{r+i+1,c})^2, & V_2 &= \sum_{i=-k}^{k-1} (g_{r+i,c+i} - g_{r+i+1,c+i+1})^2 \\ V_3 &= \sum_{i=-k}^{k-1} (g_{r,c+i} - g_{r,c+i+1})^2, & V_4 &= \sum_{i=-k}^{k-1} (g_{r+i,c-i} - g_{r+i+1,c-i-1})^2 \end{aligned} \quad (4.13)$$

找到其中的最小值作为该像素的兴趣值： $V = \min\{V_1, V_2, V_3, V_4\}$

②选取一个包含一定数量候选特征点的阈值，选择像素兴趣值大于所选取阈值的点作为候选特征点。

③选取局部为最大值的候选点作为有用的特征点。在窗口中留下找到的最大兴趣值的候选特征点，其余的候选特征点全部去除，则该像素点就可作为图像的一个特征点。

4.5.2 比值匹配法

比值匹配算法^[32]首先确定出模板，即在参考图像 T 的重叠区域的两列上分别提取出部分像素的比值，在搜索图 S 中进行搜索，然后找到最佳匹配位置，在搜索图 S 中，在相等间距的两列上从左向右依次取出部分像素计算其对应像素值的比值，将结果与模板逐次进行比较，最后得到的最小差值所对应的位置就是最佳匹配位置。比值匹配算法的具体步骤如下：

(1) 分别在参考图像 T 中的两列像素各取 m 个像素两列像素的间距为 L ，将这 m 个像素得到比值作为数组，由此形成了用于配准的比值模板。

(2) 从搜索图 S 中的两列像素各取出 $m+n$ 个像素，间距同样为 L ，再将 $m+n$ 个像素计算出的比值存入数组。

(3) 在搜索图 S 中用参考图像 T 中的比值模板寻找相应的匹配。首先比较垂直方向，记录下搜索图 S 中每个比值数组中的最佳匹配。接下来比较水平方向找到最小值即为全局最佳匹配。

4.5.3 特征块匹配法

图像配准有时也是运动估计问题。目前运动估计问题的配准普遍采用块匹配法，能够达到一定的精度，并能满足人们实时性的需要，我们可以通过硬件实现该方法。该方法是将序列图像中的每一帧分成 $m \times n$ 的非重叠子块，假定每个块具有相同的位移，根据一定的配准准则从相邻一帧图像中找出与当前这帧中最相似的块，也就是我们所谓的匹配块，而当前块与匹配块之间的相对位移则被称为运动矢量^[33]。

该匹配方法具体步骤是：把图像分为若干个非重叠的子块图像，设子图像是 $m \times n$ 的块，当两副待配准图像分别为 $f_1(x, y)$ 和 $f_2(x, y)$ 时，在 $f_1(x, y)$ 中大小为 $(M+2L, n+2L)$ 的搜索窗口中进行搜索得出它们的均方差和绝对差，计算公式如下表示：

$$\text{绝对差: } MAD(i, j) = \sum_{x=1}^M \sum_{y=1}^N |f_2(x, y) - f_1(x+i, y+j)| \quad (4.14)$$

$$\text{均方差: } MSD(i, j) = \sum_{x=1}^M \sum_{y=1}^N (f_2(x, y) - f_1(x+i, y+j))^2 \quad (4.15)$$

当被匹配的块与窗口内的块的绝对值或差平方为最小值时就完成了配准。子块图像的尺寸如果过大，匹配精度将会受到影响，反之，子块图像尺寸过小，搜索稳定的稳定性会受到影响，因此子块大小的选取至关重要。

第五章 图像融合技术

5.1 图像融合概念

图像融合是指来自同一不同传感器的两个或两个以上的图像或在同一时间的同一场景进行融合算法处理,使之形成一幅更精确更完整的新的合成图像的图像处理过程。简单的图像融合技术只能做到淡化拼接区域的像素值,无法做到无缝拼接,得到的图像往往视觉效果不令人满意。图像融合是对图像数据信息的融合。图像融合的主要目的是为了减少数据信息的不确定性,它通过处理多幅图像间冗余信息来提高图像的可靠性,还可以利用多幅图像间信息的互补提高图像的清晰度,此外图像融合可以起到图像增强、特征提取、去噪、目标识别和三维重构等作用。

图像融合分为像素级、特征级和决策级三个级别。它们的融合水平由低到高排列,一个多源信息图像的图像融合,可能会同时用到所有的融合级别。当外部条件要求不高时融合处理的对象可以是单一传感器在不同时间获取的图像。

(1) 像素级图像融合

像素级融合图像融合亦称为基于像素的图像融合,属于最低层次的融合,它是直接融合配准后的图像。

其优点是尽可能多的保持了图像数据。局限性在于:有大量的数据需要处理导致实时性差,并且有大量数据需要传输也导致了抗干扰能力差。

通过多幅图像进行像素级图像融合;可以增加图像中像素级的信息,它提供了其它特征级图像融合和决策级图像融合所不具备的细微信息。一般融合处理步骤分为:预处理、变换、综合和重构图像。像素级图像融合是现有的最重要,最根本的多传感器图像融合方法,该种融合获取的信息量最多,检测性能最好,但是难度却是最大的。

目前,像素级数据融合的主要方法回归模型法、HIS 变换法、代数法和小波变换法等。

(2) 特征级图像融合

特征级融合是一种中等水平的数据融合。特征融合首先特征提取来自传感器的图像,通常情况下,提取的特征信息应是像素信息的充分表示,这样会得到大量特征矢量,然后综合分析和处理这些特征矢量,从而实现图像的融合。通过特征级图像融合不仅可以增加了从图像中提取特征信息的可能性,还有可能获取到有利的复合特征。其优点是可以压缩大量的数据信息,有利于实时处理。缺点是精度不如像素级融合高。目前,特征级数据融合的主要方法有信息熵法、聚类分析法、表决法、加权平均法和贝叶斯估计法等。

(3) 决策级图像融合

决策级融合是一种高层次的融合。它是对每个图像数据进行属性说明后对其融合，它是三个级别的最终融合结果。决策级融合充分利用了来自各图像的出初步决策。它具有容错性很高，处理速度快，有很好的开放性，对图像数据的要求低等优点。缺点是决策级进行融合的代价较高，因为它对预处理及特征抽取的要求比较高。决策级融合是三级融合的最终结果，是直接针对具体决策目标的，它的融合结果直接影响到决策的水平。目前，常用的决策级融合的方法主要有神经网络法、逻辑模板法、专家系统及贝叶斯估计法等。

5.2 图像融合方法

图像融合的方法有很多，目前大多集中在像素级和特征级上进行的，像素级图像融合的方法分为两类：（1）空间域融合方法，它包括加权平均融合方法，像素灰度选大、选小融合方法，基于区域特征的融合方法等。（2）频率域融合方法，这种方法必须优先保持低空间分辨率图像的低频信息，并在此基础上尽可能多地提高融合结果的空间分辨率。频率域融合的方法主要有基于金字塔融合的方法，基于小波变换的方法等。图像融合一般情况下分为以下四个步骤：

①图像预处理：分别对从第一帧图像到最后一帧图像的原始图像进行滤波和配准处理，从而使被融合图像的每一个像素都得到对准。

②特征变换：找到每一幅图像的原始特征，也就是对所有帧经过预处理的图像进行对应的特征变换。目前普遍采用的特征变换方法主要是多分辨小波变换和金字塔算法。

③融合运算：对被融合图像的融合结果进行综合处理，获得最终的融合图像。

④逆变换：对前面计算所得的变换系数进行逆变换，得到最终的融合结果。

下面着重介绍几种常见的融合方法：

5.2.1 HIS 变换

HIS 分别代表图像的色调、亮度和饱和度。HIS 变换能够叠合来自不同空间分辨率的图像信息，多应用于遥感图像处理领域。它首先通过通过 HIS 变换将颜色系统的三个波段的低分辨率图像转化到 HIS 空间的光谱信息、亮度空间信息、饱和度这三个量上。然后，将一个波段的高分辨率图像做相应的对比度和拉伸处理，使之与 HIS 空间信息分量的均值和方差相一致，最后用拉伸后的高空间分辨率图像做为新的亮度分量代入 HIS 逆变换到原始空间中，把它同色调和饱和度进行 HIS 逆变换得到融合图像^[34]。这样获得的高分辨率图像会具有较高的空间分辨率，在色调和饱和度上也与源图像相同。

5.2.2 像素灰度选大融合法

像素灰度选大融合法只是简单的从待融合的图像中选择灰度值大的像素作为融合后的图像像素，因此适用范围比较小。

设待融合图像分别为 I_1 和 I_2 , 融合后图像为 I , 在融合时比较两幅图像中对应 (m, n) 处像素灰度值的大小, 以灰度值大的像素作为融合后图像 F 在 (m, n) 处的像素。可以表示为:

$$I(m, n) = \max \{I_1(m, n), I_2(m, n)\} \quad (5.1)$$

5.2.3 图像金字塔融合法

图像金字塔融合方法是一种多尺度多分辨率的融合方法, 通常情况下, 图像金字塔为图像序列, 它是在不同尺度、不同分辨率和不同分解层上分别进行的融合方式。在分解金字塔每一层的过程中, 都会产生一个多分辨率信号。因此差异金字塔可以被看作是输入图像的一个多分辨率的边缘表示^[34]。

图像金字塔融合法首先需要对每个源图像分别进行拉普拉斯金字塔形分解, 接下来对图像金字塔的各分解层进行有效的融合, 构建出新的拉普拉斯金字塔, 然后对它进行拉普拉斯逆变换使图像重构, 这样就完成了图像融合的全过程。

5.2.4 加权平均图像融合法

这是一种简单的图像融合方法, 广泛应用于图像融合领域, 设待融合图像分别为 I_1 和 I_2 , 融合后的图像为 I , 图像大小为 $M \times N$, 对两个待融合图像 I_1 和 I_2 的像素灰度值进行加权平均融合, 可以表示为:

$$I(m, n) = \omega_1 I_1(m, n) + \omega_2 I_2(m, n) \quad (5.2)$$

式中: m 为图像中像素的行 ($m=1, 2, \dots, M$), n 为图像中像素的列 ($n=1, 2, \dots, N$)

$$\text{这里的加权系数为: } \omega_1 = \frac{I_1(m, n)}{I_1(m, n) + I_2(m, n)} \quad (5.3)$$

$$\omega_2 = 1 - \omega_1 \quad (5.4)$$

如果满足 $\omega_1 = \omega_2 = 0.5$ 则图像 I_1 和 I_2 的融合为平均融合。

5.2.5 小波变换融合法

小波变换能够充分反映原始图像的局部变化特征。基于小波的融合是对每一个待融合图像进行小波变换处理, 由此建立融合图像的小波分解塔形, 对这一塔形的每一层逐层进行融合处理, 由于图像各分解层具有不同的空间分辨率, 因此我们需要选择不同的融合方法对图像进行融合, 计算后可以得到小波金字塔, 对小波金字塔小波逆变换处理, 可以重构图像, 这一重构图像就是图像最终的融合结果。它结合了低分辨率波段的和较高分辨率波段的分辨率。简而言之就是对每一幅待融合的图像分别进行小波变换, 将他们分别进行分解形成子图像, 逐

级融合待融合图像和分解出的子图像，得到所有的融合图像后对它们进行逆变换得到的图像即为最终的融合。

第六章 图像拼接算法实现

本文采用的图像拼接的算法基本流程如图 6.1 所示：

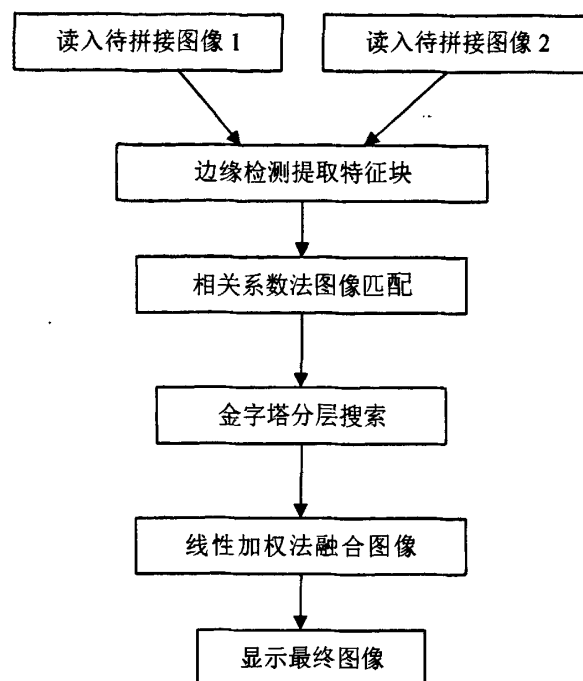


图 6.1 图像拼接处理基本流程

本文采用的图像匹配拼接算法的方法：首先读入两幅或两幅以上待拼接图 1 像，利用 Sobel 算子检测图像，以边缘为特征寻找基准特征块，进行特征匹配，找到重叠区域的匹配点，然后根据重叠区域匹配点的位置进行图像拼接，然后采用改进了的相关系数法作为配准准则，通过对图像目标块的相关系数的计算，找到最佳匹配位置，搜索策略上选择使用了金字塔式分层搜索策略，由粗到细，由上到下逐步搜索的，完成图像配准，减少了匹配点数。最后采用线性加权的图像融合法的融合方法，融合配准后的图像，实现相邻图像间自然平滑的过渡。

6.1 特征提取

图像配准的第一步就是要准确找到用于匹配的特征。本文以图像的边缘为特征寻找基准特征，采用边缘信息检测法对基准特征进行选取，同时为了避免出现多重匹配和选到无效的边缘信息，引入了一个特定阈值。图像的边缘信息是图像的主要特征之一，所选取的基准特征应该含有图像的主体信息。

首先对图像进行边缘提取，本文使用 Sobel 边缘检测算子，它能够计算图像亮度函数的梯度近似值。对图像使用 Sobel 边缘检测算子检测像素点可以计算出这点的梯度向量值。该算子由一个水平方向的 3×3 的矩阵，和一个竖直方向的 3×3 的

矩阵组成，可以通过卷积运算分别求出水平方向和竖直方向的亮度差分近似值。

以 W 代表原始图像， G_x 及 G_y 分别代表经横向及纵向边缘检测的图像，其公式如

$$\text{下: } G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} * W, G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * W \quad (6.1)$$

然后通过 $G = \sqrt{G_x^2 + G_y^2}$ 结合上式计算出梯度方向。

Sobel 算子是一种在边缘检测中常用的模板。Sobel 算子有两个，分别用于检测水平边缘和垂直边缘。由于 Sobel 算子加权了像素位置，因此边缘检测效果好于其他算子。Sobel 算子如图 5.2 所示：

Z_1	Z_2	Z_3
Z_4	Z_5	Z_6
Z_7	Z_8	Z_9

图 6.2 Sobel 算子

利用下面公式能够计算出梯度向量的方向角，可以构造出梯度图像，

$$G_x = (Z_7 + 2Z_8 + Z_9) - (Z_1 + 2Z_2 + Z_3) \quad (6.2)$$

$$G_y = (Z_3 + 2Z_6 + Z_9) - (Z_1 + 2Z_4 + Z_7) \quad (6.3)$$

$$\alpha(x, y) = \arctan(G_y / G_x) \quad (6.4)$$

其中 G_x 为 Sobel 算子的水平模板， G_y 为 Sobel 算子的竖直模板，即为上图标记为 Z_5 的水平方向梯度向量和垂直方向梯度向量， $\alpha(x, y)$ 为梯度向量的方向角。梯度的

计算需要将 G_x ， G_y 这两个分量引入公式 $G = \sqrt{G_x^2 + G_y^2}$ 后得到结果。

因为采用的 Sobel 算子为 3x3 模板，所以在图像边界处的像素仍然能够保持原来的灰度值。在图像的第 2 行 2 列到 n 行 n 列的范围内，用 Sobel 算子模板进行扫描计算，将当前像素和与之相邻的 8 个像素，分别与模板中位置相应的 9 个系数相乘，将这 9 个结果做累加计算，这样就得到了该方向的梯度。对两个不同方向的计算结果进

行比较，取其中最大的那个作为当前位置的灰度梯度。采用全局门限的方法，当某像素点的梯度值大于或等于设定的门限 T 时，规定该点的灰度值为 255，如果梯度值小于这一门限 T 时则灰度值为 0。图 6.4 为图 6.3 经过 Sobel 算子边缘提取后得到的图像：



图 6.3 原图像



图 6.4 边缘检测后效果图

经过边缘检测处理得到源图像的边缘信息图后，边缘信息的大小可以由图像中每个像素点的邻域边缘量来定义，设边缘信息图为 $E(x,y)$ ，邻域边缘量为 w 则有如下表示：

$$w(x,y) = \sum_{n=-L}^L \sum_{m=-L}^L E(x+m, y+n) \quad (6.5)$$

其中， $E(x,y)$ 为二值边缘的图像信息， L 为在图像中点 (x,y) 邻域边长的一半。利用下面公式 (6.6) 计算出 $D(x,y)$ 为最大值时所对应的点 (x,y) 作为基准点。为了避免后续的匹配过程中出现多重匹配，将以每一点的邻域边缘量与图像全局平均边缘量做减法，这样就可以找到该点所在区域的差异：

$$D(x,y) = |w(x,y) - h| \quad (6.6)$$

式中 h 为图像全局的平均边缘量。为了避免造成选到的点邻域边缘量过小，引入一个阈值，以邻域边缘量 w 的最大值的一半做为阈值，利用下列公式：

$$D(x,y) = x \times y \times |w(x,y) - h| \quad (6.7)$$

找出 $D(x,y)$ 值为最大值时所对应的点即为最终所求的基准点 (x,y) 。

6.2 图像配准

我们可以简单的将图像配准理解为，确定两幅待拼接图像的重叠区域，找出重叠部分，达到完成图像的拼接的目的。特征点能否精确匹配的关键在于所选用的匹配准则和搜索策略。

本文图像配准算法采用改进了的相关系数法，通过对图像目标块的相关系数的计算，反映出当前像素块之间的相互关系。

相关系数法的基本思想是，用 I_1 和 I_2 分别表示两幅待配准图像，模板大小为 $M \times N$ ，设 S 为搜索图像， T 为参考图像，则相关系数为 ρ 为：

$$\rho(u, v) = \frac{\sum_{x=1}^M \sum_{y=1}^N (T_{x+u, y+v} - \bar{T})(S_{x, y} - \bar{S})}{\sqrt{\sum_{x=1}^M \sum_{y=1}^N (T_{x+u, y+v} - \bar{T})^2} \sqrt{\sum_{x=1}^M \sum_{y=1}^N (S_{x, y} - \bar{S})^2}} \quad (6.8)$$

式中 $\bar{T} = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N T_{x, y}$ ， $\bar{S} = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N S_{x, y}$ ， ρ 满足 $|\rho| \leq 1$ ，

则相关系数 ρ 为最大值时所对应的位置即为图像最佳匹配位置。由于最佳匹配位置时需要对搜索次数和相关系数做乘积计算，为了减少计算量，可以简化公式 (6.9)，减少搜索次数和相关算法得到公式^[33]：

$$\rho(u, v) = \frac{\sum_{x=1}^M \sum_{y=1}^N T_{x+u, y+v} \times S_{x, y} - MNT \times \bar{S}}{\sqrt{\sum_{j=1}^M \sum_{k=1}^N T_{j+u, k+v}^2 - MNT^2} \sqrt{\sum_{j=1}^M \sum_{k=1}^N S_{j, k}^2 - MN\bar{S}^2}} \quad (6.9)$$

式中 $\sum_{x=1}^M \sum_{y=1}^N S_{x, y}^2 - MN\bar{S}^2$ 与均值 \bar{S} 为固定的常数，在第一次配准时可先计算出它

们的值，因此后面配准搜索时将计算出的数值直接带入公式 (6.10) 计算即可，这样大幅减少了计算量，又不会影响匹配的精度，节省了配准时间。

相关系数配准法中最大相关系数值所在位置即为最佳匹配位置，采用用一定的搜索策略，就可获得最佳匹配位置几何参数从而实现待拼接图像间的配准。

由于很多的图像特征匹配和灰度相似度测量方法都需要进行大量的运算，因此图像配准需要选择一种适合的搜索策略。搜索策略决定了空间中选择确定最理想的下一个变换。而在很多情况下，减少计算量是很重要的，误配准位置越多，那么计算量就越大。由于相关系数配准法的最佳匹配位置位于最大相关系数值所在位置，为了避免逐点计算，减少配准时间，本文采用由粗到细的金字塔式分层搜索策略，减少匹配点数。首先对图像进行分层处理，再进行由粗到细的匹配：

(1) 图像分层处理

分别对两幅待拼接图像进行粗提取，依次对每幅图像中 $n \times n$ 大小的块内像素灰度加权平均得到一个粗像素点从而构成第二层图像，使用同样的方法再构造出第三层图像，按此方法依次进行，直到只剩下一个像素点时停止提取，这样就可以构成以下序

列:

$$\begin{aligned} S_w &= \left(\frac{M}{n^w} \times \frac{N}{n^w} \right) \\ T_w &= \left(\frac{K}{n^w} \times \frac{L}{n^w} \right) \end{aligned} \quad (6.10)$$

式中: $W=1,2,\dots,I$, S_w 为 $M \times N$ 维的搜索图像序列, T_w 为 $K \times L$ 维的参考图像序列其中 I 表示分解出的实际层数, 具体的层数由参考图像和待配准图像的大小以及像素点数的情况决定。随着分层的不断增加, 所选择的粗匹配点也随之增加, 但图像的层数并不是越多越好, 如果分层太多, 反而会使计算量增加, 而且容易造成误匹配, 一般选 3~4 层效果较好。

(2) 由粗到细匹配

将这些序列图像叠加起来后看起来很像古埃及的金字塔形状, 因此人们称它为金字塔图像或分层结构图像。金字塔分层结构如图 6.5 所示:

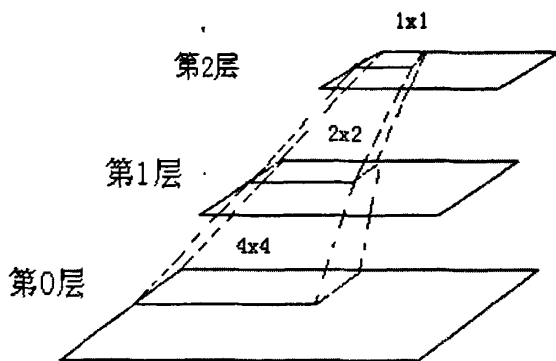


图 6.5 金字塔分层结构

把高分辨率, 大尺寸的子图像放在下层, 低分辨率, 小尺寸的子图像放在上层, 从而构成金字塔形。先对金字塔形自上而下进行粗匹配, 先从空间分辨率最低的那一层开始, 在每幅图像中取 2×2 大小的块, 依次求得每个块内像素灰度平均值, 这样就得到了一部分粗像素点, 由此能够确定出粗匹配的位置, 转到下一层时只在这些粗像素点中进行匹配, 每层图像得到多少个粗提取图像, 取决于粗提取的次数, 接下来向下逐层搜索下去, 直到搜索到最大分辨率层。最高层第 n 层上的图像大小为 $M/2^n \times M/2^n$, 搜索点数也为 $M/2^n \times M/2^n$ 个因此最底层第 0

层就是原始图像，在最低分辨率上，由于矩阵维数最小，像素点数最少，因此搜索速度较快。由于每一层搜索的范围越来越小，搜索点数变少了，大幅提高了图像配准的速度和精度。

6.3 图像融合

拼接几幅完成配准处理后的重叠图像时，如果只是简单地对几幅图像中重叠部分的图像数据进行融合，会使图像重叠区域出现明显的连接痕迹，在一定程度上造成图像边缘和轮廓的模糊。为了使拼接区域更加平滑，消除拼接缝，满足图像还原的真实性，需要对拼接区域进行必要的平滑过渡处理。在图像融合算法上选择了线性加权的融合方法，将源图像和目标图像的像素融合起来生成拼接平面上的像素。线性加权方法，即在两幅图的重叠区域按照一定权值将像素合成到新图中。

用一个窗口在图像重叠区域逐行移动，找到灰度值差异最小的像素点，把这个像素点作为图像拼接点，但是这样处理往往会由于邻近两行像素灰度值差异较大使图像出现拼接缝。为了限制邻近拼接点的灰度值差异和距离，引入一个阈值 T ，使最佳拼接点的范围保持在这个规定阈值内，阈值 T 的数值应选择 1-5 为宜。第一行拼接点用灰度值差异最小判断，与第一行选取出的拼接点同为一列点的左右距离均为 T 的区域，即为下一行拼接点的选取范围，从中选取一个最佳拼接点，以此类推，选取出每一行拼接点，这些拼接点就组成了拼接线，因为拼接点的灰度值差异均最小，由此明显的改善了接缝现象。得到了最佳拼接线由此我们可以设渐变因子为 λ ，且满足 $0 < \lambda < 1$ ，邻近的两幅图像重叠部分像素值分别为 $I_1(i, j)$ 和 $I_2(i, j)$ ，则重叠区域图像的像素值 $I(i, j)$ 可表示为：

$$I(i, j) = \lambda I_1(i, j) + (1 - \lambda) I_2(i, j) \quad (6.11)$$

式中， λ 为渐变系数，当 λ 的值由 1 渐渐变化到 0 时，左侧图像重叠区域由左向右慢慢向右侧图像重叠区域过渡，它与该点距重叠区域边界的距离有关。接下来设平滑后的像素值为 $I_3(i, j)$ ，引入公式 (6.11) 则 $I_3(i, j)$ 可用下式计算：

$$\begin{cases} I_1(i, j) = I_3(i, j) & (0 < l < L/3) \\ I_3(i, j) = \lambda I_1(i, j) + (1 - \lambda) I_2(i, j) & (L/3 < l < 2L/3) \\ I_2(i, j) = I_3(i, j) & (2L/3 < l < L) \end{cases} \quad (6.12)$$

$$\text{求得： } \lambda = \frac{3l - L}{L}$$

这里 l 为像素在重叠区域中水平方向上的位置， L 为水平方向上图像重叠的宽度。找出最佳拼接缝后，用公式 (6.12) 进行加权平滑处理，对重叠区域再进行过渡，实现相邻图像间自然平滑的过渡，这种方法相对于传统的方法，融合后的

视觉效果有了很大的提高。

6.4 试验结果

使用普通数码相机,在固定的地点进行拍摄,采集左右两端存在部分重叠的图像,相邻两幅图像间的重叠范围均在30%~50%之间。拍摄过程中镜头允许存在一定俯仰,相邻图像间允许存在一定范围内的亮度差。以图6.6-6.8为例对这三幅大小为640×480的图像使用本文算法进行拼接实验,图中红色圆圈表明了图像的特征点,用本算法匹配图像的时间约为8s,拼接时间为10s,融合后没有明显拼缝,由此可见本算法能够实现全景图的快速稳定生成。图6.9给出了拼接效果图,相邻图像间约存在约30%重叠区域。算法拼接时间与图像像素个数、事先选定的水平方向最大距离和垂直方向最大距离有关。从处理效果上看,当原始图像质量较好时,算法能基本做到准确无误地匹配拼接,并且边界平滑,无明显接缝痕迹,当图像获取时由于各种因素造成图像质量下降时,只要影响因素在一定范围之内,算法仍能进行准确地拼接。该算法具有一定的抗干扰性。



图 6.6 待拼接图像 1



图 6.7 待拼接图像 2



图 6.8 待拼接图像 3



图 6.9 拼接效果图

为了证明本实验算法的有效性，选取了 50 对不同大小的图像进行拼接实验，其中 43 对图像成功实现了拼接，7 对拼接失败，拼接时间均在 6-15S 之间，拼接成功率可达 86%。使用该算法时，如果图像重叠区域内容相差不大时，算法会受到干扰而导致拼接失败。

上述实验结果表明本文提出的算法可以实现能够实现多幅图像的无缝快速拼接，重影现象消除得比较干净，细节得到了较好的保留，拼接效果比较令人满意，具有良好的鲁棒性和较好的实用性。

结 论

图像拼接配准技术涉及到计算机视觉、图像处理以及一些数学工具等。现已广泛应用于电子制造业、三维虚拟场景的模拟、数字视频会议、医学图像处理、机器人视觉、遥感航拍图像处理、数字图片的压缩等领域

本文总结了图像拼接配准技术的研究现状和研究背景,对实现图像配准与拼接的具体方法进行了阐述,其基本流程为图像预处理、图像配准和图像融合三个步骤。本文采用简单的边缘信息检测法实现了基准特征块的提取,引入的阈值避免了选取到邻域孤立的特征点,从而提高图像拼接的精度,金字塔式分层搜索策略减少了搜索匹配点数,从而提高了运算速度,改进了的相关系数法大大减少了计算量,节省了配准时间,在图像融合上使用线性加权方法,将原图像和目标图像的像素融合起来生成拼接平面上的像素,实现了图像的平滑过渡;消除了拼接痕迹。实验证明,该种拼接算法大大减少了运算量,提高了图像拼接的精度和图像拼接的速度,具有较好的实用性。图像配准的难点在于:①多源信息图像配准精度;②更多的提高配准速度;③更多的提高大几何位移情况下的配准成功率;④怎样实现快速全自动配准。该方法的缺点在于过分依赖于图像特征,一旦特征提取和特征匹配时存在误差,那么得到的拼接图像会有很大的误差。

致 谢

首先，我要由衷的感谢在我撰写毕业论文的过程中，我的导师段锦老师所给予我的极大帮助，感谢他在百忙之中抽出时间给予我极大的支持，对我的实验工作以及论文撰写工作，段锦老师及其认真负责，对我在研究过程中遇到的每一个问题都能细致认真，不厌其烦的为我解答。他在理论与实践两方面的严格要求将使我受益匪浅。作为长辈，段锦老师的亲切随和赢得了我深深的尊敬。

另外，我还要感谢在研究生的学习过程中，各位任课老师所给予我的鼓励，感谢他们对我的教导。

同时，感谢被本文引用的文献的版权所有者，他们的研究成果为本文的研究提供了重要的基础和思路。

我还要感谢我的父母家人，是他们辛勤的为我提供生活保障，使我能够顺利地完成学业。...

最后，我要感谢所有在帮助过我的人。我所取得的成绩都与他们的帮助是分不开的。在我即将有幸的完成两年半的硕士学习生活之际，我要对所有帮助过我的人表示真挚的谢意

参考文献

- [1] Richard Szeliski Image Alignment and Stitching: A Tutorial Technical Report MSR_TR_2004-92
- [2] Heung Yeung Shum and Richard Szeliski, Panoramic Image Mosaics. Microsoft Adam Finkelstein and Marisa Range. Image Mosaics Technical Report:TR-574~98 Research, Technical Report ,MSR-TR_97~123
- [3] 侯舒维, 郭宝龙. 一种图像自动拼接的快速算法. 计算机工程. 2005. 08(3) .
- [4] 钟力, 胡晓峰. 重叠图像拼接算法. 中国图像图形学报. 1998, (5): 365~369
- [5] 汪成为. 灵境(虚拟现实)技术的理论、实现及应用. 南宁: 广西科学技术出版社, 1996
- [6] 李志刚. 边界重叠图像的一种快速拼接算法. 计算机工程. 2000, 26(5): 37~38
- [7] Hill DL, Hawkes DJ, Gleeson MH et al. Accurate frameless registration of MR and CT images of the head: applications in planning surgery and radiation therapy. Br J Radio, 1994, 191: 447
- [8] Zhang J, Levesque MF, Wilson CL et al. Multimodality imaging of brain structures for stereotactic surgery. Radiology, 1990, 175: 435
- [9] Parsai EI, Ayyangar KM, Dobelbrower RR et al. Clinical fusion of three-dimensional using bremsstrahlung SPECT and CT. J Nucl Med, 1997; 38(2): 319
- [10] Bettinardi V, Scardaoni R, Gilardi MC et al. Head holder for PET, CT, and MR studies. J Comput Assist Tomogr, 1991, 15: 886
- [11] Gilardi MC, Rizzo G, Savi A et al. Registration of multi-modal biomedical images of the heart. Q J Nucl Med, 1996, 40(1): 142
- [12] Pelizzari CA, Chen GT, Spelbring DR et al. Accurate three-dimensional registration of CT, PET, and/or MR images of the brain. J Comput Assist Tomogr, 1989, 13(1): 20
- [13] 王 晶, 游志胜, 钟 灵. 基于ROI块匹配的全景图像拼接鲁棒性方法. 计算机应用研究, 2007. 01(3)
- [14] 郑明玲, 刘衡竹. 遥感图像配准中特征点选择的高性能算法研究及其实现. 计算机学报 2004, 09(2)
- [15] 王 伟, 陆佩忠. 数字图像拼接技术. 小型微型计算机系统. 2006, 07(2)
- [16] Zhang Y, Nascimento MA, Zaiane, O R. Building image mosaics: an application of content based image retrieval. In: Proc. of the IEEE Intl. Confon Multimedia and Expo, ICM E. 2003.
- [17] Szeliski R, Kang S B. Direct method for visual scene reconstruction. In: IEEE Workshop on Representations of Visual Scenes, (Cambridge, MA) 1995, 26~33.
- [18] L-Brown, A Survey of Image Registration Techniques, ACM, Computing Survery, 1999, 24(4): 325~376
- [19] 刘楠. 数字图像的自动拼接与合成技术研究. 南京理工大学. 2004, 6(2)
- [20] 夏良正. 数字图像处理. 南京: 东南大学出版社, 2005: pp. 172~181
- [21] 王娜, 李霞. 一种新的改进Canny边缘检测算法. 深圳大学学报, 2005, 4(2), 149~152
- [22] 贺彬, 王国宇. 不同视点海底图像拼接算法. 微计算机信息. 2005, 12(2)
- [23] 徐正光, 田清, 张利欣. 图像拼接方法探讨. 微计算机信息. 2006, 22(2)
- [24] 曾文山, 李树山, 王江安. 基于仿射变换模型的图像配准中的平移、旋转和缩放. 红外与激光工程, 2001, 30(1): 17~20.
- [25] Mann S, Picard R W. Video or bits of the projective group: a simple approach to featureless estimation of parameters. IEEE Transaction on IP, 1997, 6(9): 1281~1295.
- [26] STONE H S, WOLPOV R. Blind cross-spectral image registration using prefiltering and Fourier-based translation detection. IEEE Trans, 2002, Geoscience and Remote Sensing, 2002, 40(3): 637~650.
- [27] LI H, MANJUNATH B S, MITRA S K. A contour-based approach to multisensor image registration. IEEE

- Trans ,1995 ,Image Processing ,1995 ,4 (3):320 ~ 333.
- [28] DAI X, KHORRAM S. A feature2based image registration algorithm using improved chaincode representation combined with invariantmoments. IEEE Trans ,1999 , Geoscience and Remote Sensing ,1999 ,37(5) :2351~2362.
 - [29] FLUSSER J, SUK T. A moment-based approach to registration of Images with affine geometric distortion. IEEE Trans. 1996,Geoscience and Remote Sensing ,1996 ,32(2) :382 - 387.
 - [30] Reddy B S, Charrer ji B N. An FFT based technique for translation, rotation, an dscale-invariant image registration. IEEE Transactions of Image Processing, 1996, 5(8): 1266~1271.
 - [31] Rovati F. Spatial-temporal Motion Estimation for Image Recon-struction and Mouse Functionality with Optical or Capacitive Sensors. IEEE Transactions on Consumer Electronics,2003, 49(3).
 - [32] 梁路宏, 艾海舟, 徐光佑等. 基于模板匹配与人工神经网络确认的人脸检测. 电子学报, 2001, 67(6): 744~747.
 - [33] 姜庆娟, 谭景信. 像素级图像融合方法与选择. 计算机工程与应用, 2003, 25(2): 116~120
 - [34] Richard Szeliski. Video mosaics for virtual environments. IEEE Computer Graphics and Applications, 1996, 16(2):22~30
 - [35] 朱永松, 国澄明. 基于相关系数的相关匹配算法的研究, 信号处理, 2003, 19 (6) : 531-534