上海交通大學

SHANGHAI JIAO TONG UNIVERSITY

学士学位论文

BACHELOR'S THESIS



论文题目: ____多模态遥感影像自动匹配技术研究

| 学生姓名: | 黄翊航 |
|--------|--------------|
| 学生学号: | 517021910896 |
| 专业: | 测控技术与仪器 |
| 指导教师: | 毛义梅 |
| 学院(系): | 电子信息与电气工程学院 |

多模态遥感影像自动匹配技术研究

摘要

在遥感技术中,影像匹配是完成多源信息联合、实现影像优势互补的有效手段,在国防 军事安全、环境监测、图像处理等领域具有很高的应用价值。目前,随着影像数据的不断丰 富,仅依靠人工方法已无法满足大批量影像的匹配需求。而多模态影像间普遍存在大几何差 异和非线性辐射畸变,极大地制约了影像自动匹配质量的提升,研究难度较大。

本文概括性地梳理了多模态影像自动匹配技术的基本理论,分析了经典的匹配方法;针 对影像间的旋转差异,本文创造性地提出了一种基于双分支神经网络的多模态影像旋转校 正方法,通过构建旋转特征向量与神经网络模型实现影像旋转预测与校正;基于相位一致性 信息,本文还研究了影像最大矩图和最大值索引图的构建方法,分别应用于特征点提取与特 征描述过程;针对传统 Harris 算法在点分布与定位上的问题,本文提出了一套联合多尺度 空间、局域重心检测与局部最优提取方法的 Harris 特征点综合检测流程。

网络训练与测试结果证明,本文提出的影像旋转校正方法对多类型影像在多精度条件 下均达到高预测正确率,具有广泛有效性;对比试验与综合匹配性能测试证明,基于相位一 致性的特征匹配方法对非线性辐射畸变具有很强的抗性,尤其对于几何纹理清晰的影像可 以获得很高的匹配质量。

关键词: 多模态影像匹配, 旋转校正, 相位一致性, 双分支神经网络



RESEARCH ON AUTOMATIC REGISTRATION OF MULTIMODAL REMOTE SENSING IMAGES

ABSTRACT

In the deep research of remote sensing, image registration is an effective method to achieve the combination of multi-source information and realize the complementary advantages of images, which is greatly valuable in the fields of national defense and military security, environment monitoring, image processing and so on. At present, with the continuous enrichment of image data, only relying on manual methods can not meet the needs of a large number of registration tasks. Moreover, large geometric differences and nonlinear radiation distortion between multimodal images also greatly restrict the improvement of the quality of automatic image registration.

This paper summarizes the basic theory of automatic image registration, and analyzes several classical matching methods. Aiming at the rotation difference between images, this paper creatively proposes a multi-modal image rotation correction method based on a double-branch neural network, which realizes image rotation prediction and correction by constructing a feature vector of rotation and a useful neural network model. Based on the information of phase congruency, this paper also studies the construction methods of maximum moment map and maximum index map, which are applied to feature point extraction and feature description process respectively. Aiming at the problem of traditional Harris algorithm in point distribution and location, this paper proposes a set of Harris feature point detection process which combines multi-scale space, local gravity center detection and local optimal extraction.

The training and test results of the network show that the proposed image rotation correction method can achieve high prediction accuracy for multiple types of images under multi precision conditions, proved to have extensive effectiveness. The experiment of comparison and comprehensive matching performance both show that the feature matching method based on phase congruency has strong resistance to nonlinear radiation distortion, especially for the image with clear geometric texture.

Key words: multi-modal image registration, rotation correction, phase congruency, doublebranch neural network



目 录

| 第− | ・章 | 绪论 | - 1 |
|-----|-----------|-------------------------------|----------|
| | 1.1 | 课题背景及意义 | - 1 |
| | 1.2 | 国内外研究现状 | -2 |
| | 1.3 | 研究内容与章节安排 | - 3 |
| 第二 | 二章 | 多模态遥感影像匹配的基本理论 | 5 |
| | 2.1 | 引言 | 5 |
| | 2.2 | 影像匹配的定义与空间变换模型 | - 5 |
| | 2.3 | 影像匹配的基本模型与经典方法 | - 6 |
| | | 2.3.1 影像匹配的基本模型 | 6 |
| | | 2.3.2 基于局部信息的特征点提取方法 | 7 |
| | | 2.3.3 基于互信息的影像匹配方法 | 9 |
| | | 2.3.4 基于相位一致性的特征提取方法 | 10 |
| | 2.4 | 影像匹配的评价标准 | 11 |
| | 2.5 | 本章小结 | 12 |
| 第三 | 三章 | 基于双分支神经网络的影像旋转校正方法 | 13 |
| | 3.1 | 引言 | -13 |
| | 3.2 | 旋转特征同量构建 | 13 |
| | | 3.2.1 基于相位一致性的旋转向量构建 | 13 |
| | | 3.2.2 基于梯度特征的旋转回量构建 | 14 |
| | 3.3 | 双分文仲经网络楔型设计 | 14 |
| | | 3.3.1 伸经网络埋论基础 | 14 |
| | 2.4 | 3.3.2 双分文网络模型的构建 | 16 |
| | 3.4 | 模型训练与头验结果 | 17 |
| | | 3.4.1 | 1/ 10 |
| | | 3.4.2 模型顶侧头短 | 18 |
| | 25 | 5.4.5 具头影像测讽头拉 | 20 |
| ∽∏ | 3.3 □音 | 半早小纪 | 21 |
| ۶PE | 当早 / 1 | 坐了相位 我任的影像音伴苻征匹能力茲 | 22 |
| | 4 2 | 新音 | 22 |
| | 1.2 | 4 21 基于相位一致性较大化国的特征标准网 | .22 |
| | | 422 特征占提取策略 | .22 |
| | 4.3 | 基于最大值索引图的描述符构建 | 24 |
| | | 4.3.1 最大值索引图 | 24 |
| | | 4.3.2 特征描述符 | 25 |
| | 4.4 | 实验结果 | 25 |
| | | 4.4.1 特征匹配性能对比 | 25 |
| | | 4.4.2 特征匹配结果分析 | 26 |
| | 4.5 | 本章小结 | 28 |
| 第3 | 1章 | 总结与展望 | 29 |
| | 5.1 | 工作总结 | 29 |



.. .

多模态遥感影像自动匹配技术研究

| 5.2 工作展望 | 29 |
|------------------|----|
| 参考文献 | 30 |
| 插图索引 | 32 |
| 表格索引 | 33 |
| 谢辞 | 34 |
| 毕设研究期间已获得的科研成果目录 | 35 |



第一章 绪论

1.1 课题背景及意义

影像匹配是图像处理技术领域一个非常重要的基础研究,是实现图像信息融合的关键技术之一。影像间的特征匹配是指利用影像特征信息在两幅或多福影像之间集中检测可靠同名关系,实现影像空间位置对齐的过程^[1]。遥感影像则是一种由卫星所携带的传感设备按照某种特定机理成像并转换得到的数字影像。常见的遥感影像有可见光学影像、SAR(Synthetic Aperture Radar)影像、LiDAR(Light Detection and Ranging)影像、红外影像和栅格地图等多种类型,不同影像的来源与成像机理不同,导致影像间具有多光谱、多角度、多时相、多分辨率、大辐射差异等特点,统称为多模态遥感影像。



图 1-1 多模态遥感影像实例(可见光-左, SAR-右)

多模态遥感影像匹配目前仍然是一个很具挑战性的问题,最终的匹配质量受影像间差异 的影响很大,主要的差异类型包括旋转差异、尺度差异、平移变换和非线性辐射畸变。前三 类差异均为影像在几何层面的不同,统称为几何差异,主要由飞行器稳定性、飞行载体姿态 变化、地物运动等引起;辐射畸变指的是测量的地物光谱辐射率与真实光谱辐射率不一致的 现象,其成因主要为影像在成像过程中的传输干扰、噪声、传感器误差等。

影像的特征匹配除了存在许多关键问题亟待解决,具备重要研究价值外,其在国民生产 生活与经济建设的诸多领域也具有突出的应用价值,涵盖自然灾害与环境监测、多源联合定 位、城市规划与重大工程建设等^[2]。同时,作为图像处理的基础研究,影像匹配也被广泛应 用于目标识别跟踪、场景深度计算、三维重建、超分辨率重建、视觉导航、视觉定位、图像 拼接等众多图像处理技术中。

基于多模态遥感影像采集方式的特点,通常可将相关应用分为四大类别[3][4]:

(1)多角度影像匹配。表示同一目标的影像是从不同的卫星视点获取的。多幅不同视 角影像相互交叠,可以重建出被摄目标的实际二维或三维结构。匹配后的同一场景的不同视 角场下的多幅影像也可被应用于图像的拼接、三维结构重建、场景深度计算等领域。

(2)多时相影像匹配。该类影像涵盖了同一场景在连续或多个不同离散时间内的演变 情况,通过对不同时刻下场景影像的匹配,可以很好地实现场景监测与时变监测,如大范围 自然环境演变研究、城市目标规划等。另外,利用影像间的相关性与规律性,匹配影像也可 被应用于图像目标识别、运动跟踪、同时定位与建图(Simultaneous Location and Mapping, SLAM)等领域。

(3)多源影像匹配。不同遥感卫星载荷获得同一场景影像的成像机理不同,也带来了 针对同一场景的多种不同类型、不同描述角度的丰富数据信息。利用不同载荷获得的不同描述的多幅影像,通过匹配对齐可以实现影像信息融合,以此获得对目标更全面可靠的描述,

第1页共35页



从而实现影像信息利用最大化。相应配准影像可被具体应用于高空间与高光谱分辨率的影像 构建、特定任务的信息融合与信息互补等。该匹配影像还可应用于多源联合定位、超分辨率 重建等领域。针对某种单一模态的遥感影像数据而言,很难同时在图像分辨率、定位精度、 时效性、定位一致性等各个方面均有较优的表现,不同指标间的差异巨大,难以满足高精度 的精细化数据处理要求。而卫星不同载荷获得的多模态影像之间则具有优势互补的特点,如 光学影像分辨率高、可读性强、携带信息较多但定位精度较低,SAR影像定位精度较高而分 辨率较低,实现光学和SAR影像的精确匹配能够在很大程度上提升光学影像的定位精度,从 而获得高分辨率、高定位精度的双优影像数据,极大地提高遥感影像的利用价值而无需对卫 星成像设备提出更高的要求,具有良好的经济效益。

(4)模型匹配校正。对于计算机模拟场景研究,如模拟地图、数字高程模型(Digital Elevation Model, DEM)等,通过模拟场景影像与真实参考影像之间的匹配,可快速定位、精确识别、比较对应场景影像,为模型算法的优化与更精确表达提供便利。

综上,影像匹配技术无论在应用的广泛性还是算法的基础性上都具有强大的潜在利用价值。该项技术在过去半个世纪的快速发展也充分说明了相关领域的重要性。然而随着时代的 进步,数据信息日益丰富,很难有一种算法能很好地应用于各类影像匹配任务中,因此更需 要持续不断的研究来实现并拓宽算法的实际应用价值,这也是研究人员所热烈追求的目标。

1.2 国内外研究现状

国外研究学者最早在 1960 年代开始了影像匹配领域的研究,到 20 世纪末期,单一模态同源影像的匹配问题已经得到了基本解决。Zitova^[4]系统完整地总结了 2003 年以前的影像匹配经典方法,并对其进行了归类。

上世界 70 年代,初代研究学者首先对边缘特征在影像匹配中的应用进行了大量的研究。 许多研究将导数、二阶偏导数及其相关方法作为影像边缘提取的主要方法。Kitchen, Rosenfeld 等^[5]在 1982 年首先提出一种基于二阶偏导数的角点检测算法;在 1987 年,Förstner, Gulch 等^[6]重新采用影像一阶导数检测边缘特征,以增加算法对噪声的鲁棒性;在此基础上, Harris, Stephens^[7]在 1988 年提出了 Harris 算子;此后,Rosten 等^[8]从算法的计算速度出发, 提出了 FAST (Features from Accelerated Segment Test)算法,该方法在检测时效性上具有明显的优势。

除边缘检测外,部分学者也对相关方法及其变种的区域匹配方法进行了研究。其中相关 方法主要为互相关法(Cross Correlation, CC)。1993年,Hanaizumi, Fujimura等^[9]利用假定 相似性模型进行计算,提高了算法对旋转的鲁棒性;Lewis^[10]在1995年利用归一化互相关系 数(normalized cross correlation, NCC)进行匹配研究,用以减弱光照与灰度变化的影响。

与此同时,傅里叶变换与互信息法也被用于匹配研究。1986年,Bracewell^[11]提出相位 相关方法实现影像粗校准;Tzimiropoulos^[12]在2010年提出了一种基于FFT(Fast Fourier Transformation)的尺度不变配准方法;MI(Mutual Information,互信息)方法最初由Viola, Wells^[13]在1997年提出,该方法能够较好地解决影像间的非线性辐射差异。

上述早期的匹配方法在影像特征提取、减弱差异影响方面有了较好的成果,但是对多数 几何与辐射差异仍然非常敏感。到 2004 年,针对影像的几何差异,Lowe^[14]提出了经典的 SIFT (Scale Invariant Feature Transform)方法,在抵抗尺度与旋转差异、以及一定范围对比 度变化上取得了很大的突破,引起了行业的巨大反响,也是目前同源影像匹配任务中最常用 的特征点匹配方法之一。该方法首先构建了影像的高斯差分尺度空间,并在尺度空间中利用 邻域极值点方法提取具有尺度稳定性的控制点,并为控制点分配主方向,然后利用梯度直方 图构建影像的特征描述符,通过点对点的筛选完成特征点匹配。此后,不少研究学者对该方 法做出了改进,如 PCA-SIFT,加入主成分分析法进行降维加速处理、SURF (Speed-Up Robust Features,加速鲁棒特征,2008)^[15]实现快速鲁棒特征点检测、ASIFT (Affine-SIFT,2009)^[16] 等;针对 SAR 影像的具体特征,Dellinger^[17]在 2015 年提出 SAR-SIFT 算法,引入新的梯度 定义,提高对 SAR 影像乘性斑点噪声的鲁棒性。

国内关于多模态影像匹配的大部分研究工作是在 2000 年之后展开的,由于 SIFT 及其 众多改进算法在大多数同源影像中获得了较好的应用,因此国内的研究大多集中于多模态影

多模态遥感影像自动匹配技术研究



像的匹配算法中。多模态影像由于模式的复杂性,难度较大,同源影像的匹配方法无法有效地适用于多模态影像的配准,目前仍未得到很好的解决。

2001年,黄锡山,陈哲等人^[18]提出了模糊梯度方向的概念,将其应用于可见光-红外影像的匹配;陈煜等在2004年^{[19][20]}提出基于Sobel算子与模板匹配的快速匹配方法;2008年, 臧丽等人^[21]研究了小波变换在匹配中的应用,并结合NCC(Normalized Correlation Coefficient) 计算相似性度量;于盈等人^[22]在2011年提出了一种多模态影像多阶段配准算法,该算法可以抵抗一定的平移和旋转差异,并同时对可见光-红外影像对与可见光-SAR影像对进行了匹 配测试。

2016年,叶沅鑫等^[23]学者提出基于几何结构特征的 HOPC(Histogram of Orientated Phase Congruency)算法,给出适用于可见光与 SAR 影像的特征描述符并设计了快速模板匹配方法,在多模态影像的精匹配任务中达到较理想的效果;2019,南珂,叶沅鑫等^[24]学者研究了 深度卷积神经网络在匹配中的应用;2017年,李加元等^[25]提出 RIFT (Radiation Invariant Feature Transform)算法,充分利用相位一致性的优势特点,提高了算法的鲁棒性。

在影像匹配中,旋转差异是几何差异的一个主要部分。目前已经有部分关于旋转差异校 正方法的研究。David G. Lowe^[14]针对影像匹配的尺度不变性提出的 SIFT 算法及 H Bay 等学 者^[7]改进提出的 SURF 算法利用了赋予提取点方向特征的思想实现了算法的旋转不变性,但 由于相关方法对影像梯度的强依赖性,仅在特定的同源影像匹配任务中能够取得不错的表 现,应用于多模态影像匹配的效果不佳。周微硕等^[26]基于 SURF 算法中点方向特征的提取 生成图像的主特征方向,结合异常值剔除与双三次线性插值法进行异源影像的旋转校正,较 好地解决了可见光与红外遥感影像等的旋转变换问题,但该方法的基础仍然是图像的梯度特 征,难以完成具有大非线性辐射畸变的影像如 SAR 影像的高精度校正任务。随着地理空间 信息技术的进一步发展,叶沅鑫等学者在提出的基于结构特征的 HOPC (定向相位一致性) 匹配算法^[23]与 CFOG (方向梯度通道特征)算法^[27]中利用影像的辅助地理空间信息,如有利 多项式系数(RPC),基于有理函数模型与数字高程模型等对影像的几何差异进行直接校正,

该方法校正精度高,适应多类影像,但完全依赖于影像的 辅助信息,一旦影像像素点与真 实地理空间坐标点之间误差较高,相关方法将完全失去作用。Jiayuan Li 等^[25]提出的 RIFT (辐射不变特征变换)算法则通过构建最大值索引图(Maximum Index Map, MIM),根据 MIM 的旋转特性计算最小差异,对影像进行旋转校正,通过该方法的校正能够在一定程度 上减弱旋转差异的影响,提高影像的最终匹配质量,但该方法受到 MIM 计算过程中的滤波 器旋转角度的限制,实际校正精度不高,并未从根本上解决角度校正的问题。在图像处理的 其他相关领域,如图像识别、分类等,有研究学者将机器学习、熵理论等引入了图像的旋转 校正。CUI 等^[28]利用边界检测与熵最小化等方法寻找最佳旋转角。尹婉琳等^[29]则结合形状 特征与仿射变换进行人类指纹图像的旋转校正。Lee Donggu 等^[30]则提出了基于 CNN 的图像 旋转角校正算法,来提高图像识别率,但该方法的应用对象仍为同源影像。

1.3 研究内容与章节安排

在多模态影像中, SAR 影像具有高定位精度、具备全天候观测能力的突出优点, 但分辨率低于可见光影像; 可见光影像分辨率高且可读性强, 但定位精度低于 SAR 影像。通过可见光与 SAR 影像的配准融合生成联合影像,可以使可见光影像同时具有高分辨率与高定位精度,同时实现 SAR 影像的超分辨率重建。然而, SAR 影像主要通过接收斜角度方向的雷达回波来构建图像信息, 这就导致成像过程非常容易受到大气作用、光照条件的干扰而引起大辐射差异及斑点噪声。同时, 高楼、山地的背光部分由于没有直接的回波信息, 因此该部分图像利用插值算法生成, 也会造成很大的辐射畸变。

由于 SAR 影像与可见光影像之间的成像机理完全不同,两者之间差异大导致校正与匹配任务难度较大,适用于 SAR 与可见光影像的相关算法一般很容易推广至其他多模态影像的校正匹配任务中。因此,将具有较高的挑战性的光学、SAR 影像对作为本课题研究主要影像对,针对上述差异开展相关研究,并以最终的匹配精度与质量为目标优化算法匹配性能。



课题研究内容主要包括以下三个部分:

1. 研究光学、SAR 影像综合匹配算法。学习适用于多模态遥感影像匹配的可行性方案, 提出相应的匹配流程,构建一套完整的多模态影像自动综合匹配算法;

2. 研究几何畸变影像匹配问题。针对几何差异中具有关键影响的旋转差异,展开影像 间旋转差异校正方法的研究,以实现完整的旋转角预测及校正算法。研究有助于提取影像旋 转特征的梯度与相位一致性信息,以提高预测算法的针对性;研究预测影像间旋转差异角的 神经网络模型设计与优化,以实现预测角的高精度提取;

3. 研究非线性辐射畸变影像匹配问题。经典匹配算法与同源匹配算法对非线性辐射畸 变较为敏感,为解决该问题,本课题研究结合 log-Gabor 卷积的相位一致性特征匹配算法, 利用影像稳定结构信息精确提取特征点、构建鲁棒特征描述符实现光学、SAR 影像精配准。

本论文章节安排如下:

第一章 绪论,简要介绍多模态遥感影像匹配的研究背景;介绍目前匹配任务中的主要 难点;介绍影像匹配技术的国内外研究现状;

第二章 影像匹配的基本理论,解释影像匹配的定义并介绍常用空间变换模型;说明影 像匹配的常用方法与主要流程步骤;简要分析三种应用于影像匹配的经典方法;简要介绍影 像匹配的常用评价标准;

第三章 影像旋转校正方法,介绍旋转特征向量的构建方法;介绍神经理论相关基础与本文提出的神经网络 Rot-Net 的构建方法;展示实验测试结果与说明;

第四章 鲁棒特征匹配方法,简要介绍相位一致性最大矩图的构建原理;深入介绍改进的 Harris 检测方法在最大矩图中检测的计算流程;介绍最大值索引图的提出与特征描述符的构建方法;针对算法匹配正确率将本文方法与主要匹配方法进行比较,展示最终匹配结果;

第五章 总结与展望,总结本课题的主要研究内容,分析优势与不足,提出未来的研究 展望。



第二章 多模态遥感影像匹配的基本理论

2.1 引言

本章主要介绍影像特征匹配的基本概念和经典匹配方法。匹配的关键点在于寻找影像 间最优的空间变换函数,不同的空间变换模型有不同的特点。在选取合适的变换模型后,需 要通过以下步骤获得足量的精匹配点对:提取特征点、确定搜索区域、构建特征描述符和模 板窗口匹配,利用精匹配点对求解模型参数,完成影像映射。经典的影像匹配方法包括基于 影像特征的边缘检测方法,如 Harris 算法、FAST 算法等,以及基于互信息(MI)的模板匹 配方法和基于相位一致性的特征检测方法。均方根误差(RMSE)和匹配正确率(CMR)是 影像匹配的主要评价指标。

2.2 影像匹配的定义与空间变换模型

影像匹配是针对同一目标场景的两幅或多幅不同类型影像进行同名关系检索,实现影像空间 位置映射和过程。取两幅待匹配影像,将其中一幅设为参考影像,其灰度分布为*I*₁(*x*,*y*),另 一幅设为输入影像,灰度分布为*I*₂(*x*,*y*),其中(*x*,*y*)为影像任意像素点坐标。则存在如下式 (2-1)的映射关系:

$$I_1(x, y) = H(I_2(T(x, y)))$$
(2-1)

式中,H 为影像之间的灰度变换函数,T 为空间变换函数。通常灰度变换不是必须的,因此寻找最优的空间变换函数T 的过程就是影像匹配的过程。

常用影像空间变换模型主要包括刚性变换、仿射变换、投影变换与多项式变换,前三者 为线性变换,可以用矩阵式(2-2)来描述坐标(*x*,*y*)到(*x*',*y*')的变换关系:

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

其中, m₀、m₁、m₃、m₄为反映影像旋转与缩放变换参数; m₂、m₅表示影像的水平和 垂直位移; m₆、m₇表示影像的水平与竖直缩放。下面介绍几种常用的变换模型:

(1) 刚性变换

影像中任意两点的距离不随着变换发生改变,主要反映影像的平移、旋转、镜像关系, 在二维空间中的变换模型如下式(2-3)所示:

$$\begin{bmatrix} x'\\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta\\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x\\ y \end{bmatrix} + \begin{bmatrix} m_0\\ m_1 \end{bmatrix}$$
(2-3)

由式可知该模型有三个自由度,可以通过两组及以上匹配点对进行参数求解。

(2) 仿射变换

仿射变换可以保持影像的直线与平行关系,反映了平移、旋转、镜像、剪切与缩放变换, 其变换模型为式(2-4):

$$\begin{bmatrix} x'\\ y' \end{bmatrix} = \begin{bmatrix} m_0 & m_1\\ m_3 & m_4 \end{bmatrix} \begin{bmatrix} x\\ y \end{bmatrix} + \begin{bmatrix} m_2\\ m_5 \end{bmatrix}$$
(2-4)

该变换有六个自由度,求解模型要求三组及以上不在一条直线上的匹配点对。 (3)投影变换

第5页共35页



投影变换常用来表示视角的变化,变换前后仅保证线段不发生弯曲但会改变不同线段之间的平行关系,变换模型如下式(2-5)所示:

$$\begin{bmatrix} x'\\y' \end{bmatrix} = \begin{bmatrix} m_0 & m_1 & m_2\\m_4 & m_5 & m_6 \end{bmatrix} \begin{bmatrix} x\\y \end{bmatrix} + \begin{bmatrix} m_3\\m_7 \end{bmatrix}$$
(2-5)

投影变换共包含八个自由度,因此要求四组及以上不在一条直线上的匹配点对进行求解。 (4)局部线性变换

将影像的像素点两两相连可构成一张三角网格,局部线性变换对整幅影像而言是非线性的,但对网格中的单个三角形区域是线性仿射变换。

(5) 非线性变换

非线性变换模拟任意的函数形式对影像进行映射,可以用如下公式(2-6)表示:

$$(x', y') = F(x, y)$$
 (2-6)

非线性变换中多项式变换最为典型,表达式如下式(2-7)所示:

$$\begin{cases} x = a_{00} + a_{10}x + a_{01}y + a_{01}x^2 + a_{11}xy + a_{02}y^2 + \cdots \\ y = b_{00} + b_{10}x + b_{01}y + b_{20}x^2 + b_{11}xy + b_{02}y^2 + \cdots \end{cases}$$
(2-7)

对不同的影像匹配问题,应选择合适的变换模型,刚性与仿射变换常用于消除几何畸变 的粗匹配;投影变换适合多视角影像匹配与基于特征的影像精匹配;当发生局部变形时,则 应充分利用更多的匹配点进行局部线性变换校正。

2.3 影像匹配的基本模型与经典方法

本节依次介绍了影像匹配的基本模型和三种经典匹配方法:基于局部信息的特征点提取 方法、基于互信息的区域匹配方法和基于相位一致性的特征匹配方法。此三种方法是本课题 研究的基础,分别对三种方法的原理、步骤进行了分析。

2.3.1 影像匹配的基本模型



图 2-1 两种基本的匹配流程示意图

影像匹配的关键在于,选取了合适的空间变换模型后,确定最优的变换参数使得变换后 影像空间位置的匹配度最高。变换参数确定的基本方法包括参数推导法与参数寻优法。前者 先利用特定的相似性度量进行特征点匹配,再将匹配点对代入模型计算变换参数,多用于基 于特征的匹配算法中;后者则先初始化一组参数,通过不断优化产生新参数的方法提高匹配



相似性,直到满足相似性度量要求,常见于基于区域的匹配算法。两种方法的基本流程如上 图 2-1 所示:

根据参数推导法的基本概念,模型变换参数确定的前提是具有充分的精匹配特征点对, 特征点对的获取过程主要包含四个步骤:特征点提取、确定搜索区域、特征描述符构建和模 板窗口匹配。

(1)特征点提取。利用特征点检测算法获取参考影像中的稳定特征点,常选择角点与 边缘点作为影像特征点,这是因为这些点具有较为丰富的局部特征信息,有利于精确匹配。 算法对特征点在数量和质量上有较高的要求;

(2)确定搜索区域。获得参考影像特征点后,针对每个特征点,在输入影像相应像素点的邻域划定局部搜索区域。这一过程要求对待匹配影像进行粗配准,如无粗配准则需要利用辅助地理空间信息进行对应点的检索;

(3)特征描述符构建。针对每个特征点及其对应区域搜索点,根据设定模板窗口大小,利用局部特征信息构建特征描述符。特征描述符应具有局部显著性与稳健性,应能够较好地与其他搜索点的描述符相互区别,同时具有对多种差异尤其是非线性辐射差异的鲁棒性;

(4)模板窗口匹配。利用合适的相似性测度,如欧式距离、归一化互相关系数 (Normalized Correlation Coefficient, NCC)等,在搜索区域内滑动模板窗口进行逐像素匹配, 相似性最高的搜索点与原特征点组成匹配特征点对。然后利用 RANSAC (Random Sample Consensus)算法优化匹配结果,剔除误匹配点对,筛选出精匹配点对。



图 2-2 本课题影像匹配步骤流程图

2.3.2 基于局部信息的特征点提取方法

角点为影像的拐点,在现实中表现为房屋的拐角、十字路口等。角点通常通过局部窗口 在影像中滑动来检测,在角点附近的水平、垂直两个方向梯度均较大,与单方向大梯度的边 缘点相比,角点对旋转的稳定性更好,对光照、对比度的变化不敏感,常作为匹配的初始特 征点。1988年,在Förstner边缘检测算法的基础上,Harris,Stephens^[7]提出了Harris算子。 该方法首先通过对称自相关矩阵的特征值定义角点率,然后根据设定阈值筛选出显著特征角 点,该方法的普适性与稳定性较好,目前仍然得到广泛的应用。

Harris 算法的简要步骤介绍如下:

(1) 对影像*I*(*x*,*y*),在像素点(*x*,*y*)处引矩阵 M,其中 M 的特征值为影像自相关函数的一阶导数,表达式如下式 (2-8):

$$M = \begin{bmatrix} A & C \\ C & B \end{bmatrix}$$
(2-8)

其中,A、B、C的定义如下式(2-9)所示:



$$\begin{pmatrix} A = X^2 \otimes h(x, y) = I_x^2 \otimes h(x, y) \\ B = Y^2 \otimes h(x, y) = I_y^2 \otimes h(x, y) \\ C = XY \otimes h(x, y) = I_x I_y \otimes h(x, y) \end{cases}$$
(2-9)

式中, h(x,y)为自相关函数; X、Y近似表示原始图像灰度在 x 与 y 方向的梯度I_x、I_y; (2)设定一个小范围影像矩形窗口,将影像向某方向移动微小位移量(u,v),计算得对 应灰度改变量E(x,y)如下式(2-10)所示:

$$E(x, y) = \sum_{\substack{w_{x,y} \\ u,v}} w_{x,y} [I_{x+u,y+v} - I_{x,y}]^2$$

=
$$\sum_{\substack{x,y \\ u,v}} w_{x,y} [uX + vY + o(u,v)]^2$$

=
$$Au^2 + Bv^2 + 2C$$
 (2-10)

式中, w_{x,y}为高斯窗口在像素点(x,y)的权重系数;

(3) 将矩阵 M 代入上式, 可得:

$$E(x, y) = (u, v)M(x, y)^{T}$$
 (2-11)

该式得到的E(x,y)与局部自相关函数很接近;

(4)设矩阵 M 的特征值为λ₁和λ₂,他们分别与某像素点两个正交方向上的梯度正相关, 与影像灰度自相关函数的极值曲率成比例关系。E(x,y)是一个与λ₁和λ₂有关的二次多项式, 对应一个特征值为λ₁和λ₂的椭圆,该椭圆代表了邻域内所有点的梯度统计特征。



图 2-3 点特征与 Harris 检测子的关系

如图 2-3 所示,当某点各个方向梯度均较小时,说明两个特征值也较小,表现为半径较小的近似圆;当某点领域只有一个方向的梯度较大时,则有且仅有一个特征值较大,为扁椭圆;当某点各个方向梯度均较大时,则特征值也较大,近似为半径较大的圆。

在实际计算过程中,可以 Harris 角点率 R 代替特征值计算,当 R 大于设定阈值时,则 认为当前点为角点,表达式如下(2-12)所示:

$$R = detM - \alpha(traceM)^2$$
(2-12)

式中, $detM = \lambda_1 \lambda_2 和 traceM = \lambda_1 + \lambda_2 分别为矩阵的行列式和迹; \alpha为经验常数, 一般 取值在 0.04~0.06 之间。$

为更好地满足实时匹配要求,2006 年,E.Rosten 等^[8]从算法的计算速度出发,提出了 FAST 算法,该算法利用多次筛选领域像素特征来提高角点检测速度,在检测实时性上具有 明显的优势。

FAST 算法的具体过程为: 首先对影像的所有像素点,构建一个特征检测圆,其中圆心为该像素点,半径为3个像素,由此可以得到特征检测圆的圆周上共有16个,如下图2-4

第8页共35页



所示。给定阈值,依次比较上述 16 个点与圆心点之间的差异,只当有连续的 12 个(或 9 个) 像素点与圆心点的灰度差的绝对值大于阈值时,该点才可能为 FAST 特征点。



图 2-4 FAST 特征检测圆

由于考虑到在 16 个像素点中,如果存在 12 个连续大差异点,则其中必然包含了垂直与 水平方向圆周上的 4 个顶点中至少 3 个。因此为了进一步提高检测速度, FAST 提出了分割 测试的思想。首先考察上述特殊位置的 4 个点,如果有至少 2 个点不满足要求,则该点可直 接被判定为非 FAST 特征点。经过初步筛选后,对剩余的点再次实施完整 FAST 检测,最后 经非极大值抑制后得到最终特征点。

当然,经过 FAST 的筛选,除了符合要求的角点,孤立的噪声点同样会被检测出来,因此该方法时效性较佳,但稳健性一般。

2.3.3 基于互信息的影像匹配方法

互信息法最初由 Viola, Wells^[13]在 1997 年提出,该方法能够较好地解决影像间的非线性 辐射差异,在多模态影像配准中取得较好的效果,但计算量较大,对目标影像有一定的限制。

下面首先介绍熵的基本概念。在信息论中,熵是对信号中随即成分不确定性的衡量,而 数学意义上的 Shannon 熵则代表信号加权平均的概率分布,如下式(2-13)所示:

$$H(A) = -\sum_{a} p_A(a) \log p_A(a)$$
(2-13)

式中, $p_A(a)$ 为灰度影像A种灰度值a出现的概率; $-\log p_A(a)$ 表示某个信息发生的概率 越小,其贡献的价值越高。灰度分布越分散,则影像的熵越大。

当存在两个随机变量时,其信息的相关性能表示成联合熵,联合熵代表联合概率分布的 加权平均,其表达式如下(2-14)所示:

$$H(A,B) = -\sum p_{A,B}(a,b) \log p_{A,B}(a,b)$$
(2-14)

式中, p_{A,B}(a,b)为灰度对(a,b)出现在同一像素位置的概率。

互信息则由交叉熵为基础引申而来。在影像匹配中,互信息表示影像A、B相互重叠的 信息量大小,互信息值越大则两幅影像的相似性越大。互信息的表达式如下所示:

$$MI(A,B) = \sum_{a,b} p_{A,B}(a,b) \log \frac{p_{A,B}(a,b)}{p_{A}(a) \cdot p_{B}(b)}$$
(2-15)

互信息也可以根据影像的熵与联合熵来计算:

$$MI(A,B) = H(A) + H(B) - H(A,B)$$
(2-16)

由于互信息对影像间的相似度非常敏感,当相似区域的面积较小时,匹配结果很容易陷入局部极值而非全局最大值中,导致匹配的正确率下降。因此,Studholme 在 1999 年^[31]提出使用归一化互信息替代传统互信息的方法,以此减小区域大小的影响,表达式如下(2-17)所示:

第9页共35页



多模态遥感影像自动匹配技术研究

7)

$$NMI(A,B) = \frac{H(A,B)}{H(A) + H(B)}$$
(2-1)

2.3.4 基于相位一致性的特征提取方法

影像在经过二维频域变换后可以由不同幅值和相位的正弦谐波来表达。Oppenheim 和 Lim 通过实验证明了相位在人眼视觉对信息的获取中起到了至关重要的作用:

如下图所示,将两幅差异很大的影像在频域内的谐波幅值互换,组成两幅新的影像。从 结果可见,所观察到的新影像与提供相位信息的原图像十分相似,幅值信息在其中的贡献则 相对较小。这是由于,当影像各级谐波的相位分布较为均匀时,合成的谐波峰值被相互抵消; 而当相位分布趋于一致时,合成的谐波峰值相互叠加增强,使影像出现频率峰值,在空域中 表现为边缘、角点等特征。因此,影像的边缘、角点特征信息与频域相位的分布具有强关联 性,产生相位一致性(Phase Congruency, PC)的概念^[32]。



图 2-5 幅值交换实验示意图

对一维信号f(x)做傅里叶展开,得到式 (2-18):

$$f(x) = \sum_{n>0} A_n \cos(2\pi nx + \varphi_{n0}) = \sum_{n>0} A_n \cos(\varphi_n(x))$$
(2-18)

式中, A_n 为n阶谐波分量的幅值; φ_{n0} 为初相; $\varphi_n(x)$ 为信号点x处的相位。由该式得到一维信号相位一致性定义如下:

$$PC(x) = \frac{\sum_{n>0}^{\Delta A_n} \cos(\varphi_n(x) - \overline{\varphi}(x))}{\sum_{n>0}^{\Delta A_n}}, \varphi_n(x) \in [0, 2\pi)$$
(2-19)

式中, *φ*(*x*)为点 x 各阶谐波分量以幅值加权的相位平均值。各阶相位偏离平均值最小的 点即为相位一致性最大的点,由上式可得相位一致性的取值范围为(0,1]。由于式不便于计算, 也未考虑噪声、特殊值的影响,因此常利用其他方法对相位一致性进行近似。根据局部能量

第10页共35页



多模态遥感影像自动匹配技术研究

理论,局部能量峰值位置可以在一定条件下近似相位一致性的最大值位置。计算式如下:

$$PC(x) = \frac{E(x)}{\sum\limits_{n \ge 0} A_n}$$
(2-20)

其中, *E*(*x*)可由信号的一组正交变换计算得到, *F*(*x*)为信号的傅里叶变换, *H*(*x*)为希尔伯特变换,则*E*(*x*)可定义为式(2-21):

$$E(x) = \sqrt{F^2(x) + H^2(x)}$$
(2-21)

研究指出,小波函数是获取局部信息的有效工具,人眼的感光细胞可以理想化为一系列 Gabor 小波滤波器。针对二维影像,Kovesi 提出了利用 log-Gabor 小波计算相位一致性的方 法。二维 log-Gabor 滤波函数的形式如式(2-22)所示:

$$L(\rho,\theta,s,o) = \exp\left(\frac{-(\rho-\rho_s)^2}{2\sigma_\rho^2}\right) \exp\left(\frac{-(\theta-\theta_{so})^2}{2\sigma_\theta^2}\right)$$
(2-22)

式(8)中(ρ , θ)代表对数极坐标,s、o为滤波器的尺度与方向, ρ_s 、 θ_{so} 为中心频率对应参数, σ_{ρ} 、 σ_{θ} 为 ρ 、 θ 的带宽;

对影像每个像素点(*x*, *y*),其在 log-Gabor 滤波器的所有不同尺度 s、不同方向 o 下的相位一致性特征值*PC*(*x*, *y*)如下式(2-23)所示:

$$PC(x, y) = \frac{\sum_{s} \sum_{o} w_{o}(x, y) [A_{so}(x, y) \Delta \Phi_{so}(x, y) - T]}{\sum_{s} \sum_{o} A_{so}(x, y) + \xi}$$
(2-23)

式(2-23)中, $w_o(x,y)$ 为基于频响范围的权值函数,T为噪声阈值, ξ 为一个小量,运算符[·]表示当且仅当表达式为正时取本身,否则取 0,振幅分量 $A_{so}(x,y)$ 、相位偏差函数 $\Delta\Phi_{so}(x,y)$ 的定义分别如式(2-24)、(2-25)所示:

$$A_{so}(x,y) = \sqrt{E_{so}(x,y)^2 + O_{so}(x,y)^2}$$
(2-24)

$$A_{so}(x, y)\Delta\Phi_{so}(x, y) = (E_{so}(x, y)\bar{\phi}_{E}(x, y) + O_{so}(x, y)\bar{\phi}_{O}(x, y)) - |(E_{so}(x, y)\bar{\phi}_{O}(x, y) - O_{so}(x, y)\bar{\phi}_{E}(x, y))|$$

(2-25)

式 (2-25) 中
$$\bar{\phi}_E(x,y)$$
与 $\bar{\phi}_o(x,y)$ 分别由如下式 (2-26) ~ (2-28) 给出:

$$\phi_0(x,y) = \sum_s \sum_o O_{so}(x,y)/\mathcal{C}(x,y)$$
(2-26)

$$\overline{\phi}_E(x,y) = \sum_s \sum_o E_{so}(x,y) / C(x,y)$$
(2-27)

$$C(x,y) = \sqrt{(\sum_{s} \sum_{o} E_{so}(x,y))^{2} + (\sum_{s} \sum_{o} o_{so}(x,y))^{2}}$$
(2-28)

上述式(2-24)~(2-28)中*E*_{so}(*x*,*y*)、*O*_{so}(*x*,*y*)为滤波后的响应分量,定义如式(2-29) 所示:

$$\begin{bmatrix} E_{so}(x, y), O_{so}(x, y) \end{bmatrix} = [I(x, y) * L^{even}(x, y, s, o), I(x, y) * L^{odd}(x, y, s, o)]$$
(2-29)

式(2-29)中 $L^{even}(x, y, s, o)$ 与 $L^{odd}(x, y, s, o)$ 分别为二维 log-Gabor 滤波函数在空间域内的实部与虚部。设响应分量 $O_{so}(x, y)$ 在方向 θ 下的值为 $O_{so}(\theta)$,构建中间变量 a、b 如下所示:

$$a = \sum_{\theta} \left(O_{so}(\theta) \cos(\theta) \right) \tag{2-30}$$

$$b = \sum_{\theta} \left(O_{so}(\theta) \sin(\theta) \right)$$
(2-31)

则相位一致性方向角Ø(x,y)为

$$\phi(x, y) = \arctan(b, a) \tag{2-32}$$

与 Harris 等角点提取算子一样,相位一致性也可用于提取特征点,并且通过小波信号提 取的相位一致性特征点受噪声的干扰更小,对非线性辐射畸变有更强的鲁棒性。

第 11 页 共 35 页



2.4 影像匹配的评价标准

影像匹配常用评价指标有匹配精度、特征点匹配正确率、算法复杂度和算法鲁棒性。 匹配精度是最重要的指标,用于衡量变换后影像与参考影像的重合程度,包括像素点均 方根误差(RMSE)、影像相关系数等计算方法。其中,RMSE 通过计算两幅影像正确匹配点 对(或人工选取检验同名点对)之间的均方根误差得到,计算式如下(2-33)所示:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \left[(x_{2,i} - x_{1,i})^2 + (y_{2,i} - y_{1,i})^2 \right]}{N_{cho}}}$$
(2-33)

式中, N_{cho} 为计算匹配点对总数, $(x_{1,i}, y_{1,i})$ 和 $(x_{2,i}, y_{2,i})$ 分别为参考影像与输出影像的对应匹配点坐标。

匹配正确率(correct matching ratio, CMR)用来衡量特征描述符与相似性测度的可靠性, 计算式如下(2-34):

$$CMR = \frac{N_{cor}}{N}$$
(2-34)

式中, N_{cor}为剔除误匹配点后的正确匹配点总数; N 为匹配完成后产生的所有匹配点对 总数。

算法复杂度则通常以算法运行时间来衡量,匹配完成的时间越短,则算法性能越好。算法鲁棒性也可视为普适性,通过对多组不同类型的图像进行匹配,计算相应的正确率进行衡量。

2.5 本章小结

本章针对影像匹配的基本理论与经典方法做了简要介绍。首先解释了影像匹配的定义 并介绍了常用空间变换模型的表达式与特点;然后针对本课题所研究的匹配任务说明了影 像匹配的常用方法与主要流程步骤;接着简要分析了三种应用于影像匹配的经典方法: Harris 与 FAST 特征检测子、MI 匹配法、相位一致性信息,阐明了各自的原理与优缺点; 最后简要介绍了影像匹配的常用评价标准。



第三章 基于双分支神经网络的影像旋转校正方法

3.1 引言

在影像匹配的搜索区域检索步骤中,许多算法借助辅助地理空间信息来选取对应点。然 而,有些影像的地理空间信息并不可用或定位精度很差时,这些算法将在很大程度上失去原 有的作用。在几何差异中,由于卫星飞行姿态角的扰动与拍摄视角的变化,旋转差异的发生 频率较高。

本章针对多模态遥感影像间的大旋转差异,创造性地提出了基于双分支神经网络的遥感 影像旋转校正方法作为影像匹配的预处理过程,该方法能够在仅有简单影像数据而无需任何 辅助地理空间信息的条件下实现影像旋转差异粗匹配。

如下图 3-1 所示为算法整体流程示意图。本方法的主要步骤可简要概括为以下三步:首 先通过影像特征信息构建旋转特征向量;然后设计并训练一种用于预测多模态影像间旋转差 异角的双分支神经网络;最后将旋转特征向量输入网络,输出并微调预测旋转角,通过仿射 变换模型完成影像旋转校正。



图 3-1 校正算法整体流程图

3.2 旋转特征向量构建

在基于神经网络的影像旋转校正问题中,若直接使用原始影像或灰度影像作为网络输入 将在很大程度上增加网络训练的难度,尤其对于多模态影像的处理,要保证网络深度、参数 量达到一定的大小才能更好满足特征提取、整合的要求。例如使用 VGG、YOLO、ResNet 等 深度神经网络框架进行预测训练,而网络深度与复杂度的提高又会带来训练过程缓慢、难度 增大、容易过拟合等问题。因此,本方法首先提取影像的旋转特征向量作为网络的输入向量, 要求所提取向量能够充分保留影像的旋转特征信息,以此作为简化网络结构的基础。

3.2.1 基于相位一致性的旋转向量构建

相位一致性特征在本文的 2.3.4 节中做了详细的介绍,由式(2-23)与式(2-32)可分别 计算得影像每个像素点(*x*,*y*)的特征值*PC*(*x*,*y*)与方向角Ø(*x*,*y*)。

将基于相位一致性的旋转特征向量初始化为 $RVPC = (r_0, r_1, \dots, r_{359})$,计算过程如下:

- (1) 依次取影像每个像素点的PC(x, y)与Ø(x, y);
- (2) 将方向角向下取整的值设为 n;

(3) 更新向量
$$RVPC$$
中元素 r_n 、 r_{n+1} 的值如下式(3-1)。若 $n+1=360$,则 $r_{n+1}=r_0$;

$$\begin{cases} r_n = r_n + (n+1 - \phi(x, y)) \times PC(x, y) \\ r_{n+1} = r_{n+1} + (\phi(x, y) - n) \times PC(x, y) \end{cases}$$
(3-1)



(4) 完整遍历影像的所有像素点后,对向量*RVPC*做归一化处理,某元素r_j归一化计算 式为式 (3-2):

$$n_j = (r_j - r_m)/(r_M - r_m)$$
 (3-2)

式中,**r**_M与**r**_m分别为旋转特征向量*RVPC*中元素的最大值与最小值;**n**_j为归一化处理后的最终元素值。最终的旋转特征向量维数为1×1×360。

所构建影像的旋转特征向量将作为后续神经网络的输入。由于相位一致性特征对光照、 对比度的变化均不敏感,同时能很好地应对影像间的非线性辐射畸变,因此所构建的旋转特 征向量 RVPC(Rotation Vector of Phase Congruency)同样能够提高对辐射的鲁棒性,在最终 的校正精度上有较大的优势。

3.2.2 基于梯度特征的旋转向量构建

考虑到影像匹配的预处理过程的计算时间不应过大而对完整匹配时效性造成影响,虽然 相位一致性在对辐射的鲁棒性上有很好的表现,但其计算过程较复杂、时长较长,因此本节 继续针对影像基于梯度特征的旋转向量做进一步研究。

影像的梯度本质上是对二维离散函数的求导,对影像G(x,y),其梯度为一个向量如下式 (3-3) 所示:

$$\nabla G = \left[\frac{\partial G}{\partial x}, \frac{\partial G}{\partial y}\right]^{\tau}$$
(3-3)

计算该向量的幅值,得下式(3-4):

$$|\nabla G| = mag(\nabla G) = \left[\left(\frac{\partial G}{\partial x} \right)^2 + \left(\frac{\partial G}{\partial y} \right)^2 \right]^{\frac{1}{2}}$$
(3-4)

对该式做近似可得:

$$|\nabla G| \approx |G_x| + |G_y| \tag{3-5}$$

梯度的方向角计算式(3-6)为:

$$A(x,y) = \arctan\left(\frac{G_y}{G_x}\right) \tag{3-6}$$

基于梯度值 $|\nabla G|$ 与梯度方向角A(x, y),利用 3.2.1 节的步骤流程,同样可以计算得基于 梯度的影像旋转特征向量,记为 RVG (Rotation Vector of Gradient),维数为 1×1×360。

与 RVPC 相比, RVG 能显著提高计算速度,基本满足实时预测要求,因此本章 3.4 节 所给出的实验结果也以 RVG 为例。

3.3 双分支神经网络模型设计

3.3.1 神经网络理论基础

神经网络的基本结构如下图 3-2 所示。输入和输出层可以是一维或多维矩阵,全连接层 是隐藏层的主要结构之一,除此以外还有卷积层、池化层等基本神经元,根据实际应用的需 要可以选择不同的神经元类型。输入层不进行任何计算,承担传递信息的功能;隐藏层执行 计算,是连接输入与输出的桥梁;输出层经过计算后向外部传递信息。



第 14 页 共 35 页



在本章网络的设计中主要应用了全连接层。全连接层可以看作是向量或矩阵的特征空间 变换,它包含一个权重系数矩阵和一个偏置矩阵,运算过程如下图 3-3 所示。它可以通过改 变特征维数等整合输入模块的特征信息。加上激活函数后,全连接层理论上可以对任何非线 性变换进行模拟。



图 3-3 全连接层计算过程

激活函数是实现神经网络非线性过程的重要环节,必须为可微函数。常见的激活函数有 ReLU(线性整流)函数、Leakey ReLU(带泄露线性整流)函数、tanh函数和 sigmoid 函数 等。其中,ReLU能满足较好的性能,因此应用最为广泛,其函数表达式如下(3-7)所示, 本网络使用的激活函数也为 ReLU。

$$RELU(x) = \begin{cases} x, \ x > 0\\ 0, \ x \le 0 \end{cases}$$
(3-7)

ReLU 实际上是一个分段函数,可使正数响应值保持不变,使非正数响应值被抑制,全 部映射成零,因此也被称为单侧抑制。通过单侧抑制,使得卷积神经网络中的神经元具有了 稀疏激活性,稀疏后的模型能够更好地挖掘相关特征,拟合训练数据。

本章所构建的双分支神经网络的灵感来源于孪生神经网络(Siamese Network)和多层感知机(Multi-Layer Perceptron, MLP)。

MLP 又称为多层全连接神经网络,它的特点是隐藏层神经元全部由全连接层构成。MLP 的层与层之间每两个神经元均存在连接,前一层的输出为当前层的输入,在多层神经网络中,中间隐藏层可以认为是特征提取器。模型的表达能力在一定程度上随着网络层数的增加而提高,但相应的优化过程也会更加复杂和困难。

孪生神经网络^[33]又称为"连体"神经网络,其关键在于具有两个或多个输入分支,且分支 结构是权值共享的。在实现过程中,两个分支常为同一个网络结构。孪生神经网络最重要的 用途在于衡量多个输入之间的相似关系,因此它比较适用于处理两个具有"大相似、小差异" 的输入。网络的结构原理如下图 3-4 所示:



图 3-4 孪生神经网络逻辑结构

第 15 页 共 35 页



3.3.2 双分支网络模型的构建

基于 MLP 与孪生神经网络设计思想,本方法构建了一个全新的双分支神经网络 Rot-Net,用于预测多模态影像的旋转差异角。

Rot-Net 网络的基本构建方式如下表所示:

| 层名称 | 层位置 | 层性质 | 输出维度 | 激活函数 |
|-------------|-----|-----------|----------|------|
| Branch-FC0 | 分支 | 全连接层 | 1×1×1024 | ReLU |
| Branch-FC1 | 分支 | 全连接层 | 1×1×512 | ReLU |
| Branch-FC2 | 分支 | 全连接层 | 1×1×256 | ReLU |
| Branch-FC3 | 分支 | 全连接层 | 1×1×128 | ReLU |
| Branch-FC4 | 分支 | 全连接层 | 1×1×64 | ReLU |
| Dropout | 分支 | Dropout 层 | 1×1×64 | |
| FC5 | 主干 | 全连接层 | 1×1×96 | ReLU |
| FC6 | 主干 | 全连接层 | 1×1×180 | ReLU |
| Log-softmax | 主干 | 输出层 | 1×1×180 | |
| | | | | |

| 表3-1 | Rot-Net网 | 络基本 | 构建方式 |
|------|----------|-----|------|
|------|----------|-----|------|

根据表所示, Rot-Net 共包含两条分支与一个主干, 每条分支均由相同的结构层组成; 分支中共有五个全连接层依次连接, 每层的激活函数均为 ReLU 函数, 输出维度依次减小; 分支的最后一层加入 Dropout 函数防止过拟合并提高网络的泛化性能, 最后每条分支输出一 个 1×1×64 大小的向量编码。

两条分支的向量编码通过欧氏距离融合为一个 1×1×64 大小的主干向量,随后再接入两 个全连接层,激活函数同为 ReLU;最后的输出层为 Log-softmax 函数,该函数将向量元素 值转换为位置概率,输出向量维数为 1×1×180,对应 0~180°范围的整数角。为了更好地理解 本方法构建的网络模型,给出 Rot-Net 的基本结构图如下图 3-5 所示:



图 3-5 Rot-Net 基本结构模型

经 Rot-Net 网络预测,输出大小为1×1×180的向量,记作*Sout* = ($s_0, s_1, \ldots, s_{179}$),我 们规定元素 s_0 的前一位元素为 s_{179} 。

记向量*Sout*中元素最大值为 s_m ,位于向量中第m位,其中 $m \in [0,179], m \in Z$ 。计算 s_m 的前 89 位元素的和为 A_{m1} ,后 89 位元素的和为 A_{m2} ,舍去向前第 90 位元素(也即向后第 90 位元素)。设网络的训练精度为*acc*。为了更好地突出*Sout*最大元素的决定性作用,并结合其余元素对最终预测角进行微调,得出如下计算式:

$$Ang = m + \frac{(A_{m1} - A_{m2})acc}{2(A_{m1} + A_{m2})}$$
(3-8)

最终根据旋转预测角Ang对影像进行仿射校正,形成影像的粗配准。

第 16 页 共 35 页



3.4 模型训练与实验结果

实验平台采用 Windows10-64 位系统笔记本电脑,处理器为 Intel(R)Core(TM)i5-7300HQ CPU@2.50GHz 2.50GHz,主要算法使用 Python 语言编写。

3.4.1 数据集的构建

本章所给出实验结果均以基于梯度特征的旋转特征向量 RVG 为实例,下面将介绍网络数据集的构建过程。

Rot-Net 网络的训练与测试数据采用可见光与 SAR 配对影像公开数据集 SEN1-2^[34]。该数据集是基于欧洲航天局哨兵一号采集的 SAR 遥感影像与哨兵二号采集的光学遥感影像,经过预处理获得的,是第一个完整的多类型可见光与 SAR 配对影像集,原始影像像素大小为 256×256。

如下图 3-6 所示为 SEN1-2 数据集中某海岸线配对影像:



图 3-6 SEN1-2 数据集影像实例

网络的训练集、验证集的具体构建过程如下:

(1)首先筛选出 SEN1-2 数据集中完整包含各类场景的清晰可见光与 SAR 配对影像共 1000 张,确保影像的高质量与可靠性;

(2)设置5类精度标准:3°、5°、7°、10°、15°,3°精度表示旋转校正误差不超过±1.5°;

(3)对精度标准α,对每对可见光与 SAR 影像,将 SAR 影像在 180°范围内按α为单位 角旋转,最终获得[180°/α]张旋转 SAR 影像,其中符号[.]表示向下取整。可见光影像保持 不变;

(4) 原始可见光影像与每张 SAR 影像构成一个训练影像对,该影像对的 RVG 共同构成一条训练数据;

(5) 最终各精度训练数据如下表所示:

表3-2 各精度训练数据统计

| 精度 | 15° | 10° | 7° | 5° | 3° |
|-----|-------|-------|-------|------|----|
| 数据量 | 1.2 万 | 1.8 万 | 2.5 万 | 3.6万 | 6万 |

(6)获得训练数据后,按9:1的比例划分训练集和验证集。

为了验证网络的先进性与广泛有效性,还必须增加测试实验对网络进行检验。测试集的 构建方法如下:

随机选择 SEN1-2 数据集中的可见光与 SAR 影像。为保证影像涵盖多种类型,进行不 影响测试结果的广泛性筛选,即要求测试集包含有春、夏、秋、冬四个季节的世界各地城镇、 农田、山地和湖泊等多种类型影像。所选取影像构成 6 组测试组,每组包含 100 对影像。测 试集影像实例如下图 3-7 所示:



多模态遥感影像自动匹配技术研究



图 3-7 多类型测试集影像(可见光-上; SAR-下)

3.4.2 模型预测实验

本课题以 pytorch 为框架搭建网络,并借助 Google Colab 的云端服务器进行网络训练。 训练过程中,设置初始学习率*lr* = 0.001,每轮训练结束后学习率*lr*以 0.95 的下降率阶梯式 减小。优化器采用 pytorch 内嵌的 Adam 优化器,为防止过拟合,设置 weight decay(权值衰 减)为 0.95。训练损失函数为 pytorch 内嵌函数 NLLLoss,与网络输出层 Log-softmax 相匹配。

经过 50 轮训练后, Rot-Net 在各精度标准下的训练损失曲线、训练集与验证集的平均准确率(后文简称平均准确率)曲线分别如下图 3-8、图 3-9 所示:



图 3-9 各精度训练集与验证集平均准确率曲线



如图所示,针对各精度标准,Rot-Net均达到良好的训练效果。15°与10°训练集在20轮次训练后损失曲线趋向平缓,最终训练损失小于0.01,同时两者的最高训练准确率均达到99.5%,训练与验证的平均准确率同样超过99%;在7°训练集中,最高准确率达99.28%,平均准确率则为98.25%;而5°与3°训练集的平均准确率则分别为96.08%与92.27%。统计结果见表3-3:

| 精度 | 训练集 | 50 轮次后训练集 | 10 轮次 | 50 轮次 |
|-----|-------|-----------|--------|--------|
| | 最高准确率 | 与验证集平均准确率 | 训练损失 | 训练损失 |
| | (%) | (%) | | |
| 3° | 93.00 | 92.27 | 0.7353 | 0.2216 |
| 5° | 98.17 | 96.08 | 0.7075 | 0.0676 |
| 7° | 99.28 | 98.25 | 0.3294 | 0.0290 |
| 10° | 99.60 | 98.95 | 0.3268 | 0.0156 |
| 15° | 99.75 | 99.18 | 0.1049 | 0.0084 |

表3-3 各精度训练结果统计

训练结果显示,Rot-Net 在训练集与验证集上表现较好,本课题进一步利用 6 组每组包含 100 对影像的随机测试集进行了网络的广泛性测试实验。

实验过程如下:

取一组数据集,计算所有可见光影像的旋转特征向量;对所有 SAR 影像分别在 0~180° 范围内,以 3°、5°、7°、10°、15°为单位角做影像旋转,并计算旋转后的影像特征向量 RVG; 将可见光与旋转 SAR 影像的 RVG 输入 Rot-Net,使用与单位角相对应的网络权重进行角度 预测测试;重复 10 次预测过程,记录测试准确率并计算 10 次预测的平均准确率,其中预测 准确的判断标准为:对精度 α ,若影像真实旋转角为 β ,当预测旋转角 $\gamma \in (\beta - \alpha/2, \beta + \alpha/2)$ 时,认为预测正确,否则预测失败,例如,36°旋转,在 5°精度下的正确预测角范围为 (33.5°,38.5°)。

平均准确率如图 3-10 所示。可以看到,由于影像集本身存在一定的精度误差,3°精度测试的平均准确率在 75%到 85%之间;5°精度平均准确率已达 90%;对 7°、10°、15°精度测试实验,Rot-Net 在 6 组测试集上的平均预测准确率普遍高于 95%。



图 3-10 各精度 6 组测试集的平均准确率

第 19 页 共 35 页

3.4.3 真实影像测试实验

本实验包含四组真实影像,可见光影像来源于资源3号、Terra-SAR等; SAR影像来源 于高分3号、Google Earth等。在各训练精度条件下对影像进行6次随机旋转,检验Rot-Net 网络对影像差异角的预测结果。四组试验影像对如下图 3-11 所示:



(a) 试验一

(b) 试验二

(c) 试验三 图 3-11 四组真实试验影像(可见光-上; SAR-下)

(d) 试验四

测试结果如下表 3-4 所示:

| | 化3-7 兵大形队网队沿入 | | | | | | | |
|------------|---------------|-------|-------|------|-------|-------|--|--|
| 试验序号 | 旋转角度 | 15°精度 | 10°精度 | 7°精度 | 5°精度 | 3°精度 | | |
| | 27 度 | Ture | Ture | Ture | Ture | Ture | | |
| | 60度 | Ture | Ture | Ture | Ture | False | | |
| | 99度 | Ture | Ture | Ture | Ture | Ture | | |
| 2-17-10 | 117 度 | Ture | Ture | Ture | False | Ture | | |
| 风利亚 | 132 度 | Ture | Ture | Ture | Ture | Ture | | |
| | 149 度 | Ture | Ture | Ture | Ture | Ture | | |
| | | | | | | | | |
| | 9度 | Ture | Ture | Ture | False | False | | |
| 试验二 | 55 度 | Ture | Ture | Ture | Ture | False | | |
| | 78 度 | Ture | False | Ture | False | False | | |
| | 133 度 | Ture | Ture | Ture | False | Ture | | |
| | 149 度 | Ture | Ture | Ture | Ture | Ture | | |
| | 151 度 | Ture | Ture | Ture | Ture | Ture | | |
| | | | | | | | | |
| | 19 度 | False | Ture | Ture | Ture | False | | |
| | 35 度 | Ture | Ture | Ture | Ture | Ture | | |
| | 61 度 | Ture | Ture | Ture | Ture | False | | |
| 学硕二 | 102 度 | Ture | Ture | Ture | False | False | | |
| 瓜孙二 | 120度 | Ture | Ture | Ture | Ture | Ture | | |
| | 176 度 | Ture | Ture | Ture | Ture | Ture | | |

表3-4 真实影像测试结果



| | | | | | | 续表 3-4 |
|------|-------|-------|-------|------|-------|--------|
| 试验序号 | 旋转角度 | 15°精度 | 10°精度 | 7°精度 | 5°精度 | 3°精度 |
| | 2度 | Ture | Ture | Ture | False | False |
| | 19度 | Ture | Ture | Ture | Ture | Ture |
| | 22 度 | Ture | Ture | Ture | Ture | False |
| 试验四 | 42 度 | Ture | Ture | Ture | Ture | Ture |
| | 69度 | Ture | Ture | Ture | False | False |
| | 138 度 | Ture | Ture | Ture | Ture | False |

结果可以看出, Rot-Net 对非训练集样本依然具有较好的普适性, 四组试验影像预测正确率很高, 7°及以上精度共 72 次随机测试中, 预测正确 70 次。其中试验一影像具有全图丰富的几何纹理, 且噪声干扰较少, 所有精度测试效果较好; 试验二影像由于结构单一, 较难提取有效的特征信息, 5°及 3°精度测试结果不够理想; 试验三影像几何结构较为丰富, 5°及 以上精度结果较好; 试验四影像中心纹理丰富, 但边缘特征缺失, 且存在较大的斑点噪声, 5°及 3°精度测试结果同样还有一定的提升空间。

3.5 本章小结

本章的主要创新点在于:针对多模态影像中存在大旋转差异导致搜索区域检索困难、匹配算法失效的问题,创造性地提出了基于双分支神经网络的多模态影像旋转校正方法。本方法首先利用相位一致性信息、梯度特征等影像特征信息构建旋转特征向量 RVPC/RVG,然后设计了一种全新的双分支神经网络 Rot-Net 用于旋转角的预测,采用 SEN1-2 公开数据集与多卫星真实影像对网络进行训练和有效性测试。

实验结果充分表明,本章所给出的方法能够在多精度训练条件下,对多类型遥感影像及 非训练样本的的真实影像测试均能获得较优的预测正确率,该方法能够较好地保证后续高 精度匹配任务的顺利进行,具备一定的先进性与广泛有效性。



第四章 基于相位一致性的影像鲁棒特征匹配方法

4.1 引言

由于影像的相位一致性特征相比于其他特征信息在对抗辐射畸变上有突出的优势,因此 在多模态影像的匹配任务中获得研究者的青睐。本章研究了基于相位一致性的影像鲁棒特征 匹配方法,该方法首先利用了现有研究中提出的相位一致性最大矩图用于特征点的检测,随 后以最大值索引图为基础构建特征描述符,结合搜索区域检索与模板窗口匹配实现影像融 合。

针对当前研究中边缘检测算法可靠性不足、传统 Harris 算法点分布与定位问题,本课题 自主提出了基于多尺度局域角点重心与局部最优提取的 Harris 检测改进方法,并给出一套 完整的算法流程,以获得最终匹配正确率的大幅提升与对辐射畸变更强的抗性。

4.2 基于相位一致性最大矩图的特征点检测

4.2.1 基于相位一致性的最大矩图

由于 SAR 影像存在较大的 Speckle 噪声以及影像间的大非线性辐射畸变,传统特征点 检测算法很难在原始多模态影像中取得稳定的检测效果,如容易存在漏检现象或检测的特征 点局部信息不够丰富等。

通过公式(2-23)可以计算影像的相位一致性特征值*PC(x,y)*,获得较为精确的影像边缘信息层,但忽略了 log-Gabor 小波滤波器的方向变化对影像中每个像素点的相位一致性的影响。在 Kovesi 对相位一致性的分析中,可以利用一维 PC 公式计算得二维 log-Gabor 小波变换的每个方向对应的相位一致性特征,并计算其矩。通过矩分析算法可知,最大矩对应的轴垂直于主轴,最大矩的模的大小反映了特征的显著性。

相位一致性最大矩M的定义式如下(4-1)所示:

$$M = \frac{1}{2} \left(m_c + m_a + \sqrt{m_b^2 + (m_a - m_c)} \right)$$
(4-1)

式中, m_a、m_b、m_c的为定义的中间变量, 计算式如下(4-2)所示:

$$\begin{cases} m_a = \sum_o \left(PC(\theta_o) \cos(\theta_o) \right)^2 \\ m_b = 2 \sum_o \left(PC(\theta_o) \cos(\theta_o) \right) \left(PC(\theta_o) \sin(\theta_o) \right) \\ m_c = \sum_o \left(PC(\theta_o) \sin(\theta_o) \right)^2 \end{cases}$$
(4-2)

式中,o为滤波器的方向变量; *PC(θ_o)*为对应方向角的相位一致性特征。影像所有像素 点的最大矩M的集合构成了影像的最大矩图。最大矩图提取了影像的边缘信息,可用于检测 边缘特征。

4.2.2 特征点提取策略

本文 2.3.2 节详细介绍了 Harris 和 FAST 两种常用的特征点检测方法。FAST 检测子通 过在一定程度上损失精度的情况下赢得了较快的速度,但因为占影像匹配主要计算时间的是 模板匹配过程而非特征点检测,因此在影像模板匹配任务中并不适用;而 Harris 检测子难以 对抗影像光照与对比度的变化,且因对非线性辐射畸变具有敏感性,在多类型多模态影像匹 配中没有获得普遍的使用。

第 22 页 共 35 页



影像的最大矩图较好地消除了影像的非线性辐射畸变同时保留了边缘信息,因此在最大矩图中应用 Harris 算法进行特征点检测,能很好地实现"扬长避短"的目的。

在传统的 Harris 算法中,人工选取阈值是一个关键部分,Harris 值超过阈值的像素点将 被选取为特征点。因此,不恰当的阈值将很容易导致所检测特征点扎堆分布,而影像相对空 旷平坦的部分却缺少特征点或仅存在由噪声引起的孤立点,特征点扎堆分布则表达影像的特 征重复,不便与后续的匹配与观察,如图 4-1。





为解决该问题,同时提高特征点的检测精度,本课题针对影像匹配过程提出了改进提取 策略:基于多尺度局域角点重心与局部最优提取的 Harris 检测方法。

下面将介绍具体检测步骤:

(1) 首先获得影像的相位一致性最大矩图;

(2) 设基于像素点(x,y)的高斯函数G(x,y,σ_i)如下式(4-3) 所示:

$$G(x, y, \sigma_i) = \frac{1}{2\pi\sigma_i^2} e^{-(x^2 + y^2)/2\sigma_i^2}$$
(4-3)

分别取σ₁ < σ₂ < σ₃, 按下式 (4-4) 对 2.3.1 节中式 (2-9) 的系数A、B、C进行改进:

$$\begin{cases}
A(\sigma_i) = A \otimes G(x, y, \sigma_i) \\
B(\sigma_i) = B \otimes G(x, y, \sigma_i) \\
B(\sigma_i) = B \otimes G(x, y, \sigma_i)
\end{cases}$$
(4-4)

(3)分别由式计算得三个尺度下影像像素点的 Harris 值,并设定阈值筛选得三幅具有 Harris 检测点的影像;

(4) 对小尺度σ₁影像中的每个检测点,检查其在中尺度σ₂影像中对应像素点*n*×*n*邻域 内是否有检测点,若存在,则保留该小尺度影像检测点,否则删除该小尺度影像检测点;

(5)在第(4)步遍历完所有检测点后,对小尺度影像的每个剩余检测点,再次检查其 在大尺度σ₃影像中对应像素点*n*×*n*邻域内是否有检测点,若存在,则保留该小尺度影像检 测点,否则删除该小尺度影像检测点,所有剩余点为最终的多尺度影像 Harris 检测点;

(6)重新扫描第(5)步中的多尺度影像 Harris 检测点,若该点在自己的*m*×*m*邻域内 无其他检测点,则认为该点由噪声引起,予以去除;若存在且邻域内检测点数目大于等于 k, 则利用下式(4-5)计算重心坐标:

$$(x,y) = \left(\frac{\sum_{i=1}^{n} R_i x_i}{\sum_{i=1}^{n} R_i}, \frac{\sum_{i=1}^{n} R_i y_i}{\sum_{i=1}^{n} R_i}\right)$$
(4-5)

式中, (x, y)为该区域检测点重心坐标; (x_i, y_i)与R_i分别为该邻域内第 i 个检测点的坐标



与 Harris 值;

(7)最大矩图划分为*s*×*s*的影像小块,在每个小块中对检测点的 Harris 值进行排序, 取排名前 T 名的点作为该小块的最终检测点。最终获得影像的检测特征点共*s*×*s*×*T*个。

在上述步骤中,充分利用了小尺度影像检测点定位精度准而大尺度影像检测点可靠性高的特点,同时提高检测点的定位精度与可靠性;利用重心定位方法解决了点密集扎堆的问题; 局部分块阈值筛选的方法实现最终检测点的分布合理性。该提取策略获得的特征点稳定性 高、定位精度高、分布合理,能为影像最终匹配精度的提高带来较大的帮助。利用上述方法 获得的影像特征点在原始影像中的示意图如图 4-2 所示,为便于观察,设T = 5:



图 4-2 改进 Harris 算法检测点分布

4.3 基于最大值索引图的描述符构建

理论上,相位一致性对非线性辐射畸变有较好的抗性,利用相位一致性特征图层对影像 特征点进行描述将会达到更好的匹配效果。但研究发现,简单地利用相位一致性特征图层的 匹配结果很差,基本都是误匹配点。分析认为,相位一致性做了一次边缘信息筛选,保留的 信息量较少,并且其取值范围在0到1之间,大部分像素值接近于零,因此并不能用于鲁棒 的特征描述。本节研究了基于最大值索引图的特征描述方法,以实现影像鲁棒特征匹配。

4.3.1 最大值索引图

在本文 2.3.4 节中利用了 log-Gabor 卷积计算相位一致性信息,最大值索引图同样基于 该卷积序列来构建:

(1) 对影像*I*(*x*,*y*),利用式(2-29)卷积,可得结果*E*_{so}(*x*,*y*)与*O*_{so}(*x*,*y*);

(2)根据式(2-24)得到在滤波器s尺度、o方向下的幅值A_{so}(x,y);

(3)固定滤波器方向,计算像素点(x, y)在o方向下的综合卷积幅值 $A_o(x, y)$ 如式(4-6):

$$A_o(x, y) = \sum_{s=1}^{N_s} A_{so}(x, y)$$
(4-6)

式中, N_s为滤波器不同尺度总数;

(4)对像素点(x,y),将其在所有No个滤波方向下对应的综合卷积幅值编号为1~No;

(5)比较上述N_o个幅值,取最大幅值对应的编号作为该像素点的索引值,影像中所有像素点索引值的集合构成影像的最大值索引图。武汉地区 5m 分辨率影像及其最大值索引图如下图 4-3 所示:



多模态遥感影像自动匹配技术研究



图 4-3 武汉地区 5m 分辨率可见光影像及其最大值索引图

4.3.2 特征描述符

根据 SIFT 与 HOG^[35]的直方图统计思想, 在获得影像的最大值索引图后, 利用以下步骤 进行特征描述:

(1) 对参考影像的每一个特征点及输入影像搜索区域的每一个搜索点,选择一个以该 点为中心的 J×J 像素大小的局部影像块;

(2)使用标准差等于 J/2 的高斯函数为局部块中每个像素点分配权重;

(3) 再次划分局部影像块为 6×6 大小的子块;

(4) 取每个子块,构建一个 $1 \times N_o$ 维的向量,按如下评判标准更新向量的元素值:子块中像素点的索引值为 S_o , $1 \le S_o \le N_o$,则将向量 S_o 位置的元素值加一;

(5) 对子块向量归一化处理,局部影像块中所有子块向量连接为一个包含 6×6×N_o个元素的特征向量,该向量即为特征点的特征描述符。

4.4 实验结果

实验平台采用 Windows10-64 位系统笔记本电脑,处理器为 Intel(R)Core(TM)i5-7300HQ CPU@2.50GHz 2.50GHz,主要算法使用 MATLAB 编写。

4.4.1 特征匹配性能分析

为了验证基于相位一致性的可见光、SAR 特征匹配能力,本方案通过选择两组不同场 景的可见学、SAR 影像作为测试数据,在不同模板窗口大小下,以匹配正确率(正确匹配点 数量/总匹配点数量)为测度与 MI、HOG 以及 HOPC 等方法进行对比进行测试。其中第一 组影像为质量较好、畸变不严重的可见光和 SAR 影像。第二组中光学影像由于云雾干扰成 像模糊。





(a) 试验组一





(b)试验组二 图 4-4 试验影像对(可见光-左; SAR-右)





如图 4-5 可以看出,本文基于最大值索引图的特征匹配方法对于可见光、SAR 影像的非 线性辐射畸变有很强的抵抗能力。在试验一中因为可见光影像和 SAR 影像的成像畸变较少、 几何结构清晰,所以在模板尺寸不大的情况下也有很高的匹配正确率;MI 方法对于模板大 小非常敏感,小模板经常出现误匹配现象,虽然随着模板尺寸变大匹配成功率上升,但是其 匹配成功率还是远远低于本文算法;HOG 算法对模板尺寸以及影像质量要求较高,模板尺 寸越大、影像效果越清晰,性能越好。HOPC 算法的效果总体上略高于 HOG,但是对影像 清晰度也有较高要求。本文方法在试验一中对模板大小不敏感,在第二组试验中,由于可见 光影像比较模糊导致模板尺寸较小时无法正确匹配,但是随着模板尺寸提升本文算法效果迅 速提升,当模板尺寸达到 80×80 像素以后匹配成功率恢复到 90%。

4.4.2 影像匹配结果分析

为了验证基于相位一致性的多模态影像自动匹配算法的有效性,本文采用了 4 组不同 分辨率、不同场景的可见光、SAR、LiDAR 影像进行测试。第一组来自 TerraSAR 和 Google Earth 在城市地区的 3m 分辨率影像;第二组来自高分 3 号 SAR 波段数据和资源 3 号的 3m 分辨率光学数据;第三组来自航空光学影像 3m 分辨率和 LiDAR 的 2m 分辨率数据,第四 组来自 Google Earth 和 Google Map 的 2m 分辨率数据。需要说明的是本文所用传感器分辨 率为经过重新采样以后的分辨率并非传感器原始分辨率。为了对影像进行评定,本文用人工 确认的方式来验证算法匹配点的匹配精度,当匹配点误差小于 2 个像素便认为匹配成功。利 用匹配点计算的仿射变换的 RMSE 来表示匹配精度。



多模态遥感影像自动匹配技术研究

| | 表4-1 匹配试验影像数据 | | | | | | | |
|-----|---------------|------------|--------------|------------|--------------|--|--|--|
| 试验 | 参考影像 | 参考影像 | 输入影像 | 输入影像 | 影像特征 | | | |
| 编号 | 传感器 | 大小/分辨率 | 传感器 | 大小/分辨率 | | | | |
| 试验一 | TerrraSAR-x | 534×524/3m | Google Earth | 528×524/3m | 城市影像 结构丰富 | | | |
| 试验二 | 资源3号 | 562×535/3m | 高分3号 | 535×534/3m | 云雾下影 像模糊 | | | |
| 试验三 | 航空光学影像 | 550×550/3m | LiDAR 强度影像 | 550×550/2m | 影像有明 显噪声 | | | |



(a) 试验一



(b) 试验二



(c) 试验三 图 4-6 三组试验影像特征点匹配结果

第 27 页 共 35 页



| 多模态遥感影像自动匹配技术研究 |
|-----------------|
|-----------------|

| 表 4-2 试验匹配结果 | | | | | | | |
|--------------|-------|--------|--------|------|--|--|--|
| 试验编号 | 匹配点总数 | 正确匹配点数 | CMR | RMSE | | | |
| 试验一 | 219 | 210 | 0.9589 | 0.79 | | | |
| 试验二 | 133 | 121 | 0.9098 | 0.77 | | | |
| 试验三 | 77 | 68 | 0.8831 | 1.37 | | | |

如图 4-5 和表 4-2 所示,在第一组试验中,当影像的细节足够丰富,畸变足够小的时候 基于相位一致性的特征匹配方法达到了很好的匹配精度。可见本文设计的匹配算法能够很好 地抵抗影像的非线性辐射差异。在第二组试验中,由于云雾的干扰可见光影像边缘比较模糊, 其梯度分布和 SAR 影像有较大差异,匹配难度较大。本文算法在第二组试验中依旧可以成 功匹配,可以说明该算法对于梯度信息并不敏感,对影像的云雾干扰有一定的抗性。第三组 试验中,本文算法也可以更好抵抗 LiDAR 成像产生的噪声,达到很好的匹配效果,可以看 出本文方法除了对可见光、SAR 影像的非线性辐射畸变具有较强的抗性外,对其他多模态 影像畸变也具有一定的抗性。融合影像示意图如下图 4-6 所示



试验一

试验二

图 4-7 影像融合结果示意图

4.5 本章小结

本章在基于相位一致性信息的鲁棒特征匹配方法的基础上,针对 Harris 检测点分布不 均匀、定位精度及可靠性不高的问题进行了优化,首先利用最大矩图的多尺度空间对检测点 进行可靠性检查,筛选出一批具有较高可靠性的预选特征点;其次对检测点进行局域重心检 测,解决过多检测点扎堆分布与 SAR 斑点噪声引起孤立点的问题;再者利用局部最优提取 的方法,使检测点分布均匀,获得了较高精度的稳健可靠特征点。

在特征匹配阶段,基于相位一致性的最大矩图与最大值索引图的应用保证了较高的匹配 稳定性。将本章方法在匹配正确率上与 MI、HOG、HOPC 等算法进行对比,并进行三组匹 配测试实现,结果表明,本章方法在匹配精度与正确率上均有很大优势,尤其对于几何结构 明显的影像匹配质量更佳。



第五章 总结与展望

5.1 工作总结

完成多模态遥感影像的精确匹配,从而实现 SAR 与可见光影像的优势互补与信息联合 具有重要的研究意义,可以应用于国民生产生活、国防军事、图像研究等诸多领域。本文针 对多模态遥感影像的旋转差异与非线性辐射畸变两大难点问题,提出了一种基于双分支神 经网络 Rot-Net 的多模态影像旋转校正方法,以及基于相位一致性的影像鲁棒特征匹配方 法,实现了多模态影像在多精度下的旋转角校正,综合提升了相同条件下影像的匹配正确率 与匹配精度。

本文主要工作如下:

(1)介绍了影像匹配的原理、流程与经典方法,阐述了多模态影像匹配的研究背景、 难点问题与研究现状,构建了一套算法匹配流程。

(2)针对多模态影像的旋转差异,创造性地提出了一种基于双分支神经网络 Rot-Net 的多模态影像旋转校正方法,给出旋转特征向量的构建方法与 Rot-Net 的结构模型。通过模型训练过程与多组测试实验证明,该方法在 5°及以上的多精度条件下均能达到不错的预测 正确率,在无需任何辅助地理空间信息的情况下提升后续匹配算法的有效性。

(3)本文对于 RIFT 精匹配算法进行了优化,首先在特征匹配流程中加入区域检索与 模板匹配过程,保证了匹配点的定位精确性;然后针对传统 Harris 检测算法提取点分布不 均匀、密集扎堆分布和可靠性不高的问题,联合多尺度空间、局域重心检测和局部最优提取 的方法,构建了一套检测流程,提升了待匹配特征点的质量与均匀性。通过对比试验及多组 匹配试验证明,本文方法在匹配正确率上有很好的优势,对非线性辐射畸变有较好的抗性, 同时能抵抗多种影像噪声,尤其对几何结构明显的影像效果更为突出。

5.2 工作展望

本文的研究在以下方面仍有优化空间:

(1) 尽管算法取得很好的匹配正确率与精度,但因为相位一致性的计算过程较复杂、 模板匹配需逐像素计算等原因,算法计算时间较长,仍不能满足国防军事等领域实时匹配要 求。可以参考 HOPC 算法中的快速模板的思想对匹配过程进行加速;也可在从神经网络特 征匹配入手,利用网络可进行多线程计算的优势,提高算法速度。

(2)本文提出的旋转校正方法具备一定校正精度,但未能将大尺度差异纳入考虑中, 对该差异的校正作用不明显;该方法对影像的几何纹理信息较为敏感,当影像尺寸过小、几 何信息不足时,校正精度将下降。可以考虑对网络结构进行优化,加入多元神经元,以适应 更广泛的影像特征。

(3)本文匹配算法在特征描述过程中仅采用基于 log-Gabor 滤波器的响应信息构建最 大值索引图,所保留边缘信息不够多,不利于稳健局部特征描述符的构建。可结合相位一致 性特征值以及其他影像特征信息对最大值索引图进一步优化。

(4)针对滤波器参数、各类窗口、模板大小参数的研究不够完善,没有测试不同的计 算参数对匹配结果的影响。



参考文献

- [1] Szeliski R. Computer Vision: algorithms and applications[M]. Berlin: Spinger Science & Business Media,2010.
- [2] 李德仁.摄影测量与遥感的现状及发展趋势[J].武汉测绘科技大学学报,2000(01):1-6.
- [3] 李加元.鲁棒性遥感影像特征匹配关键问题研究[D].武汉:武汉大学摄影测量与遥感学 院,2018.
- [4] Barbara Zitová, Jan Flusser. Image registration methods: a survey[J]. Image and Vision Computing,20003,21(11).
- [5] Kitchen L, Rosenfeld A. Gray-level corner detection[J]. Pattern Recognition Letters, 1982, 1(2):95-102.
- [6] Förstner et al. A fast operator for detection and precise location of distinct points, corners and centers of circular features[C]. ISPRS, Intercom-mission Conference on Fast Processing of Photogrammetric Data,1987,281-305.
- [7] Harris C, Stephens M. A combined corner and edge detector[C]. Alvey vision conference, Manchester, UK,1988,15(50):147-151.
- [8] Rosten E, Drummond T. Machine learning for high-speed corner detection[C]. European Conference on Computer Vision(ECCV), Graz, Austria, 2006, 430-443.
- [9] Hanaizumi H, Fujimura S. An automated method for registration of satellite remote sensing images[C]. Proceedings of the International Geoscience and Remote Sensing Symposium(IGARSS), Tokyo, Japan, 1993, 1348-1350.
- [10] Lewis J P. Fast normalized cross-correlation[C]. Vision Interface, 1995, 10(1):120-123.
- [11] Bracewell R N. The Fourier transform and its applications[M]. New York: McGraw-Hill, 1996.
- [12] Tzimiropoulos G, Argyrou V, Zafeiriou S, et al. Robust FFT-based scale-invariant image registration with image gradients[J]. IEEE transactions on Pattern Analysis and Machine Intelligence,2010,32(10):1899-1906.
- [13] Viola P, Wells W M. Alignment by maximization of mutual information[J]. International Journal of Computer Vision, 1997, 24(2):137-154.
- [14] Lowe D G. Distinctive image feature from scale-invariant key-points[J]. International Journal of Computer Vision, 2004, 60(2):91-110.
- [15] Bay H, Ess A, Tuytelaars T, et al. Speeded-up robust features(SURF)[J]. Computer Vision and Image Understanding,2008,110(3):346-359.
- [16] Morel J M, Yu G. ASIFT: A new framework for fully affine invariant image comparison[J]. SLAM journal on Imaging Sciences,2009,2(2):438-469.
- [17] Dellinger F, Delon J, et al. SAR-SIFT: a SIFT-like algorithm for SAR image[J]. IEEE Transaction on Geoscience and Remote Sensing,2015,53(1):453-466.
- [18] 黄锡山,陈哲.景像匹配定位中的图像边缘检测算法研究[J].中国惯性技术学报,2001(01):25-31.
- [19] 陈煜,田裕鹏.图像融合中的图像配准方法研究[J].激光与红外,2004(01):34-36.
- [20] 陈煜.图像融合技术及其应用研究[D].南京航空航天大学,2004.
- [21] 臧丽,王敬东.基于互信息的红外与可见光图像快速配准[J].红外与激光工程



2008(01):164-168.

- [22] 于盈,程咏梅,等.一种异源图像多级配准算法[J].计算机仿真,2011,28(09):256-259+263.
- [23] Y. Ye, J. Shan, L. Bruzzone, et al. Robust registration of multimodal remote sensing images based on structural similarity[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, vol. 55, no. 5, pp. 2941-2958.
- [24] 南轲, 齐华, 叶沅鑫, 等. 深度卷积特征表达的多模态遥感影像模板匹配方法[J]. 测绘 学报,2019,48(6):727-736.
- [25] Jiayuan Li, Qingwu Hu, Mingyao Ai. RIFT: Multi-Modal Image Matching Based on Radiation-Variation Insensitive Feature Transform[J]. IEEE Transactions on Image Processing, 2019, vol. 29, pp. 3296-3310.
- [26] 周微硕, 安博文, 赵明, 等. 基于几何不变性和局部相似特征的异源遥感图像配准算法 [J]. 红外技术,2019,41(06):561-571.
- [27] 王蒙蒙, 叶沅鑫, 朱柏, 等. 基于空间几何约束和结构特征的光学影像与 SAR 影像配 准方法. 武汉大学学报(信息科学版). https://doi.org/10.13203/j.whugis20190354.
- [28] Cui Jiali, Cui Yanwei, Wang Yiding. Rotation Correction of DHV Images Using Entropy Minimization of Boundary Descriptor[C]. Proceedings of 2012 IEEE 11th International Conference on Signal Processing, IEEE Beijing Section, 2012:
- [29] Yin Wanlin, Ye Hua, Yang Yanlan. Rotation Correction and Classification of Fingerprint Image[J]. Computer Science and Application, 2014, 4(5).
- [30] Lee Donggu, Sun Young Ghyu, Kim Soo Hyun, et al. CNN-based Image Rotation Correction Algorithm to Improve Image Recognition Rate[J]. The Journal of The Institute of Internet, Broadcasting and Communication, 2020, 20(1).
- [31] Studhome C, Hill D L G, Hawkes D J. An overlap in variant entropy measure of 3D medical image alignment[J]. Pattern Recgnition,1999,32(1):71-86.
- [32] P. Kovesi. Phase congruency: A low-level image invariant[J]. Psychological research, vol. 64, no. 2, pp. 136-148, 2000.
- [33] S. Chopra, R. Hadsell, Y. LeCun. Learning a Similarity Metric Discriminatively, with Application to Face Verification, CVPR, IEEE, 2005.
- [34] M. Schmitt, L. H. Hughes, X. X. Zhu, "The Sen1-2 Dataset for Deep Learning in SAR-optical Data Fusion," ArXiv e-prints, Jul, 2018.
- [35] Dalal N, Triggs B. Histograms of oriented gradients for human detection, 2005, 886-893.



插图索引

| 图 | 1-1 | 多模态遥感影像实例 | 1 |
|---|------|--------------------------|-----|
| 图 | 2-1 | 两种基本的匹配流程示意图 | 6 |
| 图 | 2-2 | 本课题影像匹配步骤流程图 | 7 |
| 图 | 2-3 | 点特征与 Harris 检测子的关系 | 8 |
| 图 | 2-4 | FAST 特征检测圆 | 9 |
| 图 | 2-5 | 幅值交换实验示意图 | -10 |
| 图 | 3-1 | 校正算法整体流程图 | -13 |
| 图 | 3-2 | 神经网络基本结构 | -14 |
| 图 | 3-3 | 全连接层计算过程 | -15 |
| 图 | 3-3 | 孪生神经网络逻辑结构 | -15 |
| 图 | 3-5 | Rot-Net 基本结构模型 | -16 |
| 图 | 3-6 | 数据集影像实例 | -17 |
| 图 | 3-7 | 多类型测试集影像 | -18 |
| 图 | 3-8 | 各精度训练损失曲线 | -18 |
| 图 | 3-9 | 各精度训练集与验证集平均准确率曲线 | -18 |
| 图 | 3-10 | 各精度 6 组测试集的平均准确率 | 19 |
| 图 | 3-11 | 四组真实试验影像 | 20 |
| 图 | 4-1 | 传统 Harris 算法检测点分布 | -23 |
| 图 | 4-2 | 改进 Harris 算法检测点分布 | -24 |
| 图 | 4-3 | 武汉地区 5m 分辨率可见光影像及其最大值索引图 | -25 |
| 图 | 4-4 | 试验影像对 | -26 |
| 图 | 4-5 | 两组试验的匹配正确率与模板大小的关系曲线对比 | -26 |
| 图 | 4-6 | 三组试验影像特征点匹配结果 | -27 |
| 图 | 4-7 | 影像融合结果示意图 | -28 |



表格索引

| 表 3-1 | Rot-Net 网络基本构建方式 | 16 |
|-------|-------------------------|----|
| 表 3-2 | 各精度训练数据统计 | |
| 表 3-3 | 各精度训练结果统计 | 19 |
| 表 3-4 | 真实影像测试结果 | 20 |
| 表 4-1 | 匹配试验影像数据 | 27 |
| 表 4-2 | 试验匹配结果 | 28 |



谢辞

四年的认真努力在临近毕业的这一刻显得更有意义。有过初入交大的迷茫无知,也有过 求学探路的自信坚定,有爬上眉头就下不来的苦恼烦闷,也有藏入心底便抹不去的愉悦欣 喜……

在此,我想衷心感谢张海涛老师对我全方面的耐心教导与关心。张老师亦师亦友、和蔼 亲切、对待学术严谨细致,总能从细微之处发现问题的关键所在,在毕设阶段向我提供了很 多宝贵的建议,极大地鼓励了我在学术研究道路上的钻研热情。

我也非常感谢毛义梅老师的热心与支持。毛老师一直是所有同学们眼中的好老师,总为 我们着想、工作耐心细致,愿意在百忙之中热心地为我的毕设课题提供帮助。

我还要感谢李泽一学长对我的关心与在研究课题上的深入指导。泽一学长学术水平强 悍、知识储备丰富、在生活上也乐观积极,是我学术道路上的榜样。

感谢可爱喜人的小橙子,在我焦躁无力时总能与我一同冷静分析,简单生活因你而快乐,"TO BE PRICELESS"。

最后,感谢一直无条件支持我的爸妈,你们是我安心的源泉,也感谢身边陪伴的所有老师朋友,愿走上更高求学之路的自己能勇敢迎接挑战,继续前行!



毕设研究期间已获得的科研成果目录

一、学术论文

(会议论文)

Yi-Hang Huang, Ze-Yi Li, Haitao Zhang, et al. Rot-Net: A Rotation Angle Detector between Multimodal Images Using Siamese Network. OSA Imaging and Applied Optics Congress.19 July, 2021, Vancouver, Canada. (Accepted, EI 检索, 第一作者)

(期刊论文)

黄翊航,李泽一,张海涛,等.面向多模态遥感影像旋转差异的智能匹配方法[J].光学技术,2021.(已接收,第一作者)

Zeyi Li, Haitao Zhang, Yihang Huang. A Rotation-Invariant Optical and SAR Image Registration Algorithm based on Deep and Gaussian Features. Remote Sensing. (Accepted, SCI 检索)

二、发明专利

《面向非线性辐射畸变的稳健遥感影像特征点对提取方法》 (申请号: CN202011564394.5,公开号: CN112634335A)

《一种基于神经网络的多模态遥感影像旋转差异校正方法》 (申请号: CN202110530035.6,申请日: 2021.05.14)



RESEARCH ON AUTOMATIC REGISTRATION OF MULTIMODAL REMOTE SENSING IMAGES

Image registration is a very important and basic research in the field of image processing technology, and it is one of the key technologies to realize information fusion of different kinds of images. For definition, feature matching in the image registration task refers to the process of using feature information of images to detect reliable and homonymic relationship between two or more images and realize image spatial alignment. Remote sensing image is a kind of digital image which is formed and converted by the sensing equipment carried by satellite according to a certain mechanism. Common remote sensing images include optical image, SAR (synthetic aperture radar) image, LiDAR (light detection and ranging) image, infrared image, raster map, etc. Different sources of images and imaging mechanisms lead to multi-spectral, multi-angle, multi-temporal, multi-resolution and large radiation difference among images, which is called the multi-modal remote sensing image. Multi-modal image matching technology has high application value in national production and daily life, national defense and military, image processing and other fields. For example, multi-angle images can be used for 3D structure reconstruction, and multi-temporal images can be used for target recognition, SLAM and other fields, and registration between multisource images can realize information fusion, super-resolution reconstruction and simulation scene correction.

Among the multi-modal images, SAR image has the outstanding advantages of high positioning accuracy and all-weather observation ability, but its resolution is lower than that of optical image; Visible image has high resolution and readability, but its positioning accuracy is lower than that of SAR image. Through the registration and fusion of optical and SAR images to generate joint images, visible light images can have high resolution and high positioning accuracy at the same time, and the super-resolution reconstruction of SAR images can be realized. Because the imaging mechanism between SAR image and optical image is completely different, the large difference between them leads to the difficulty of correction and matching task. The correlation algorithm suitable for SAR and optical image is generally easy to be extended to other multi-modal image correction and matching task. Therefore, the optical and SAR image pairs of more challenging are selected as the main image pairs in this study.

Geometric difference and nonlinear radiation distortion are two main difficulties of multimodal image matching task, and the final matching quality is greatly influenced by the differences between images. Among them, geometric difference mainly includes scale difference, rotation difference and translation transformation, it is mainly caused by the stability of the aircraft, the change of the attitude of the flight carrier and the movement of the ground objects. Radiation distortion refers to the phenomenon that the measured spectral emissivity of ground objects is inconsistent with the real spectral emissivity. Its main causes are transmission interference, noise and sensor error in the imaging process. At present, there are many researches on image matching technology, and many matching methods have been proposed from different perspectives. For



example, Harris algorithm and fast algorithm based on edge detection; Cross-correlation method, Fourier transform method and mutual information method based on image similarity; SIFT algorithm, HOG algorithm, HOPC algorithm and RIFT algorithm based on depth feature, etc; At the same time, deep neural network is also used in high-precision image matching. Different methods have different suitable applications and some defects. In order to solve the differences mentioned above, this paper focuses on the correction method of rotation differences between images and the robust feature matching method to improve the resistance to nonlinear radiation distortion.

The main process of image matching method is as follows: Firstly, feature point detection. The stable feature points in the reference image are obtained by using the feature point detection algorithm, and the corner and edge points are often selected as the image feature points. Secondly, determine the search area. After obtaining the feature points of the reference image, rough registration is performed on the image to be matched, and then for each feature point, a local search area is delimited in the neighborhood of the corresponding pixel of the input image. Thirdly, feature descriptor construction. According to the size of the template window, the local feature information is used to construct the feature descriptor. Moreover, template window matching. Using appropriate similarity measures, such as Euclidean distance, normalized correlation coefficient (NCC), etc. the sliding template window in the search area is used for matching pixel by pixel, and the search points with the highest similarity together with the original feature points form as the matching feature points. Then, RANSAC (random sample consensus) algorithm is used to optimize the matching results, eliminate the false matching point pairs, and select the accurate matching point pairs. Furthermore, the appropriate transformation model is selected, and the model parameters are calculated by using those accurate matching point pairs. The common evaluation indexes of image matching include matching accuracy, feature point matching accuracy, complexity of algorithm and robustness of algorithm. Matching accuracy is the most important index to measure the coincidence degree between the transformed image and the reference image, including the root mean square error (RMSE) of pixels, image correlation coefficient and other calculation methods.

For more details, in order to solve the problems of the difficulty of searching region and the failure of matching algorithm the difficulty of searching region and the failure of matching algorithm caused by large rotation difference in multi-modal images, this paper creatively proposes a multi-modal image rotation correction method based on double-branch neural network. Firstly, the rotation feature vector, named as RVPC or RVG, is constructed by using unique image feature information such as the information of phase congruency and the feature of gradient. Then, a novel double-branch neural network, named as Rot-Net is designed to predict the rotation angle. A public image data set SEN1-2 is used to train and test the effectiveness of the network, and under the condition of 5-degree (\pm 2.5-degree) error, the highest training accuracy of precision is over 98%, and the average test accuracy reaches 90%. The experimental results fully show that the method presented in this study can get better prediction accuracy for multi-modal remote sensing image pairs under multi precision training conditions. This method can ensure the smooth follow-up of advanced high-precision matching tasks, and has certain advance and wide effectiveness.

According to the robust feature matching method based on the information of phase congruency, this study optimizes the problem of uneven distribution of Harris detection points, low positioning accuracy and poor reliability. Firstly, the multi-scale space of maximum moment graph is used to check the reliability of detection points, and a batch of pre-selected feature points with



high reliability are selected; Secondly, local barycenter detection is used to solve the problem of too many detection points narrow in a certain area and SAR speckle noise. Furthermore, by using the local optimal extraction method, the detection points are evenly distributed, and the robust and reliable feature points with high accuracy are obtained. In the stage of feature matching, the application of maximum moment graph and maximum index graph based on phase consistency ensures high matching stability. The method in this study is compared with MI, HOG and HOPC algorithm in matching accuracy, and three groups of matching experiments are carried out. For images with clear geometric texture, the matching RMSE is less than 0.8 and the accuracy is higher than 95%. The results show that the method in this chapter has great advantages in matching accuracy and the decreasing of matching error, especially for the image with obvious geometric structure.

To sum up, aiming at the difficult problems and the shortcomings of existing matching algorithms, this paper creatively proposes rotation correction method and robust matching method. Through many kinds of experiments, this method comprehensively improves the matching accuracy and decreases the matching error of images under the same conditions. However, there are still some problems in this research: for example, the calculation time of the algorithm is too long to meet the requirements of real-time matching in the field of national defense and military. Moreover, the scale difference has not been taken into account in this paper, and the rotation correction method is sensitive to image size. Furthermore, this paper does not test the influence of different filter parameters and template size parameters on the final matching results. Therefore, we need do further research on the shortcomings of the above algorithm in the future.