

多模态遥感图像匹配算法综述

续云龙^{1,2}

(1. 中国科学院空天信息创新研究院, 北京 100080; 2. 中国科学院大学 电子电气与通信工程学院, 北京 101400)

摘要: 多模态遥感图像匹配的目标为识别来自不同传感器或不同成像条件下图像间像素的对应关系。然而, 由于非线性辐射畸变、光照变化和噪声干扰等问题, 传统匹配算法往往在准确性和鲁棒性方面存在不足。近年来, 深度学习技术的发展为图像匹配提供了新方案, 基于深度神经网络的方法能够自动学习深层次特征, 减少对先验知识的依赖, 提高匹配精度, 尤其在处理多模态图像的非线性辐射差异时表现出明显优势。综述各种传统算法和深度学习算法, 定量评估各自的代表性算法。结果表明, 与传统算法相比, 深度学习算法在匹配精度、匹配点数量方面表现突出, 如LoFTR、RDFM和LightGlue等深度学习算法在匹配精度与运行时间之间取得了较好的平衡, 展现出较高的实用价值。最后展望深度学习技术在优化算法和提升计算能力方面的潜力, 预测其将在多模态遥感图像匹配中发挥越来越重要的作用, 为实际应用提供更高效、精准的技术支持。

关键词: 图像匹配; 图像特征; 多尺度; 多模态; 粗差去除

DOI: 10.11907/rjdk.251016

扫描二维码阅读全文:



中图分类号: TP391.4

文献标识码: A

文章编号: 1672-7800(2026)004-0199-14

Multi-modal Remote Sensing Image Matching Algorithms: A Survey

XU Yunlong^{1,2}

(1. Institute of Air and Space Information Innovation, Chinese Academy of Sciences, Beijing 100080, China;

2. School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 101400, China)

Abstract: The goal of multimodal remote sensing image matching is to identify the corresponding relationship between pixels from different sensors or images under different imaging conditions. However, due to issues such as nonlinear radiation distortion, lighting variations, and noise interference, traditional matching algorithms often have shortcomings in accuracy and robustness. In recent years, the development of deep learning technology has provided new solutions for image matching. Methods based on deep neural networks can automatically learn deep level features, reduce reliance on prior knowledge, and improve matching accuracy, especially when dealing with nonlinear radiometric differences in multimodal images, showing significant advantages. Summarize various traditional algorithms and deep learning algorithms, and quantitatively evaluate their representative algorithms. The results show that compared with traditional algorithms, deep learning algorithms perform outstandingly in terms of matching accuracy and the number of matching points. Deep learning algorithms such as LoFTR, RDFM, and LightGlue have achieved a good balance between matching accuracy and running time, demonstrating high practical value. Finally, looking ahead to the potential of deep learning technology in optimizing algorithms and enhancing computing power, it is predicted that it will play an increasingly important role in multimodal remote sensing image matching, providing more efficient and accurate technical support for practical applications.

Key Words: image matching; image features; multi-scale; multi-modal; outlier removal

0 引言

遥感图像匹配一直以来都是计算机视觉领域的一个重要研究课题, 旨在对两幅或多幅图像中表征相同地物的像素点进行映射, 进而实现不同图像中同一对象的像素级

匹配^[1-5]。图像匹配不仅是视觉处理中的基础任务, 而且是许多应用的关键前置步骤, 在灾害评估、变化检测、目标识别等领域具有不可忽视的重要作用^[6-8]。

多模态遥感图像匹配是指在进行图像匹配时, 所匹配的遥感图像对之间存在显著的差异, 具体包括成像机制差异、成像尺度差异、视角差异及几何畸变等^[9]。

收稿日期: 2025-03-04

作者简介: 续云龙(1998-), 男, 中国科学院空天信息创新研究院硕士研究生, 研究方向为多模态遥感图像匹配算法。

其中,成像机制差异通常是指源于不同传感器获取影像的差异,例如光学影像、红外影像和合成孔径雷达(Synthetic Aperture Radar, SAR)影像等。这些不同模态的图像在研究对象的物理特性表现上各有优势,合理利用这些影像能够更好地分析研究对象的特性。然而,不同成像机制下获取的图像,其像素灰度值的含义及所反映的物理信息存在差异,这不可避免地会引入图像间的非线性辐射畸变。成像尺度差异主要表现为不同分辨率下研究对象在影像中所占像素数量的变化。高分辨率影像能够捕捉到更多细节,研究对象占据的像素数更多;而低分辨率影像则可能丧失大量细节信息,研究对象所占据的像素数较少。视角差异通常是由于遥感卫星、无人机、车载平台等不同成像平台对同一地物的成像角度不同,导致影像之间的视角产生变化。几何畸变则是由于受成像平台的成像方式限制,在图像处理过程中,某些区域的影像可能无法准确反映真实地物信息。例如,摆扫式成像的传感器在影像边缘处由于俯视角度较大,容易出现几何畸变。

以上差异大大增加了多模态遥感图像匹配的难度,主要体现在以下几个方面:首先,遥感图像的成像机制众多,某些成像机制下获取的影像特征点稀缺,难以提取特征;其次,成像机制差异使得影像间存在显著的非线性辐射差异,如何准确描述特征并建立同名特征之间的对应关系成为一大挑战;最后,由于视角、尺度等因素的变化,同一地物在不同影像中的表现可能出现差异,这会进一步影响特征点的对应,进而影响匹配的效果与精度。

随着遥感技术的快速发展,遥感影像的数量与种类呈爆炸式增长,多模态数据综合应用越来越广泛,给不同模态影像间的匹配与对齐带来了前所未有的挑战。许多学者针对图像匹配中的特定场景和问题提出多种解决方案,从最初的手工特征描述方法到如今以数据驱动为核心的深度学习方法,图像匹配技术经历了数次重要变革,取得了显著进展,尤其是深度学习及大模型相关技术通过扩大预训练模型的参数规模和数据量显著提升了模型性能^[10-12]。本文系统梳理近年来多模态图像匹配领域的最新算法,对其进行分类,并对部分具有代表性的算法进行比较实验,深入分析其优缺点及适用场景,并基于此展望了该领域未来的发展趋势。

1 多模态遥感图像匹配方法

图像匹配研究最早可追溯到20世纪70年代的计算机视觉与图像处理领域,而遥感图像的处理则起步较晚。因此,许多遥感图像处理方法都源自计算机视觉领域,并在此基础上进行了迁移与发展。经过近50年的不断研究,当前主流遥感图像匹配算法可大致分为三大类^[13]:①基于区域的匹配;②基于特征的匹配;③基于数据驱动的匹配。具体如图1所示。

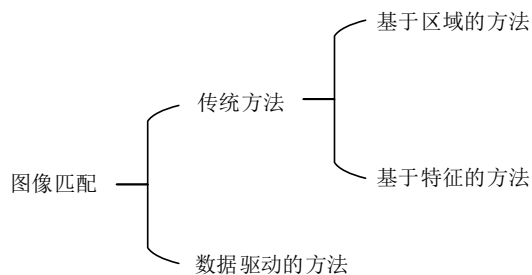


Fig. 1 The classification of image matching algorithms

图1 图像匹配算法分类

1.1 基于区域的匹配

在早期研究阶段,相关系数法作为一种经典且常用的技术手段,被广泛应用于实现图像间的匹配操作。这种方法通过构建相关性计算模型,评估不同小区域内灰度值的相似性,从而找到具有最高相似性的对应区域,实现图像匹配。通常情况下,这一过程需要将一幅图像划分为若干小区域作为模板区域,并在另一幅图像的搜索窗口内,利用同样大小的滑动窗口找到与模板区域相似性最大的部分。因此,这种基于空间域的匹配方法常被称为模板匹配。经典的相似性度量方法包括差平方和(Sum of Squared Differences, SSD)和归一化互相关(Normalized Cross-Correlation, NCC)^[14-15]。SSD算法原理简单,通过计算两个窗口内灰度值差异的平方和来衡量相似性,该方法在处理光照差异显著的图像时效果不佳;相比之下,NCC则通过计算两个窗口间像素灰度的互相关值来评估相似度,在一定程度上优化了光照差异的影响。例如,Emery等^[16]尝试使用互相关方法进行光学和被动微波图像的匹配,实现了自动化卫星姿态参数获取与无人工干预的图像导航。陆和平等^[17]从原理上分析了光学图像与SAR图像之间的差异,在光学图像中采用局部灰度方差,在SAR图像中采用灰度与局部方差的乘积作为基准,利用互相关算法实现了二者的匹配。尽管基于图像灰度或灰度变换的互相关方法具有步骤简单、易于实现的优点,但是这种方法完全依赖于图像的灰度值,且对灰度差异十分敏感,因此在处理非线性辐射畸变较大的多模态遥感图像时,往往难以获得良好的匹配效果。1997年,Maes等^[18]提出使用互信息作为相似性度量的多模态医学图像匹配方法。互信息主要用于衡量两幅图像区域之间的相互依赖关系,当图像几何位置对齐时,互信息的值最大。Chen等^[19]在小波变换后的遥感图像上进行互信息与互相关两种方法的比较,结果表明,互信息方法的匹配精度显著优于互相关方法,并且在应对噪声时具有更好的鲁棒性,在一定程度上能够更有效地抵抗图像间的非线性辐射差异。然而,归一化互信息方法在多模态遥感图像匹配过程中容易受到局部极值的影响,从而出现错误匹配现象。为此,Yu等^[20]提出一种基于灰度权重窗口的算法(Grayscale Weight with Window, GWW),通过从红外图像和光学图像中提取共同的强边缘特征来减少局部极值的影响,从而提升了归一化

互相关算法在红外与光学图像匹配中的效果。

基于区域的图像匹配方法通常不需要单独进行特征提取或特征描述,因为图像中的所有像素点均被视为潜在的图像特征点。其核心任务为构建合适的相似性度量方法,用于实现特征匹配。许多相似性度量算法受信号处理领域的启发,借助相关性计算来评估不同区域间的相似性^[21]。然而,区域匹配方法在实际应用中存在一些局限性:首先,由于成像机理的差异,多模态影像间的相似性较难准确度量,误匹配现象较多;其次,该类方法通常难以应对图像尺度变化和旋转变换。搜索窗口的定位通常依赖于两幅图像的初始像素位置,这一特性意味着匹配图像的分辨率和旋转差异不宜过大,在图像几何位置差异较大的情况下,往往需要增加搜索窗口尺寸,但这会显著增加计算复杂度。以上约束限制了区域匹配方法在处理尺度变化和分辨率差异较大的多模态图像时的灵活性和适应性。因此,尽管区域匹配方法相对简单,其较高的计算复杂度以及难以有效处理多模态图像间差异的缺陷,使其在某些实际应用中不如特征匹配方法灵活高效。表 1 为基于区域的匹配算法比较。

Table 1 Comparison of region-based matching algorithms
表 1 基于区域的匹配算法比较

方法	算法特点
SSD ^[14]	算法实现简单,难以抵抗图像灰度差异
NCC ^[15]	可以抵抗图像的灰度差异,计算量大,难以抵抗图像间非线性辐射畸变
MI ^[19]	一定程度上可以抵抗非线性辐射畸变,计算量大
GGW ^[20]	可以提取到光学和红外图像中的共同边缘

1.2 基于特征的匹配

在图像处理领域,提取图像中的显著特征是图像增强、分割等任务的关键步骤。许多学者受此启发,通过提取特征来实现图像匹配,从而有效降低算法的复杂度。随着相关理论的逐步发展与完善,基于特征的匹配方法逐渐成为图像匹配领域的核心体系之一,并得到广泛应用。

基于特征的匹配方法不再将图像视作单纯的像素集合,而是将其抽象为特征的集合,利用图像中具有显著性的元素来表征图像。这种方法通过从图像中提取关键的局部或全局特征,进而提高图像匹配的效率和准确性^[22]。与传统的基于区域的匹配方法不同,基于特征的匹配方式能够对图像中的结构信息进行有效提取与表示,进而实现更加灵活和高效的匹配过程。该类方法具有诸多优势:首先,其可以实现匹配过程的步骤分离,通过特征提取、特征描述、特征匹配等模块化操作,根据图像的具体情况进行自由组合,从而获得最佳匹配效果;其次,其能够有效减少计算量,特别是在大规模图像匹配任务中,使用局部特征能够显著提高处理效率;最后,其能够更加稳定地处理复杂的图像变换,如旋转、缩放、光照变化等。

1.2.1 特征提取

特征提取是指对图像进行预处理,提取其中具有显著

性的特征。在特征匹配方法中,点特征、线特征和面特征是 3 种常见的特征类型,它们各自有不同的特点及应用场景。点特征主要指图像中灰度变化剧烈或局部灰度极值的像素点,这些点通常具有较强的鲁棒性,适用于处理光照变化、噪声影响的图像。线特征着重提取图像中的边缘和纹理信息,通常用于表达图像的结构性特征。面特征则是指图像中灰度变化较小且具有一定局部连贯性的区域,通常通过区域分割方法得到,适用于尺度较大的匹配任务。目前,点特征匹配方法的研究最为广泛且应用最为成熟。表 2 为不同图像特征的比较。

Table 2 Comparison of different image features
表 2 不同图像特征比较

特征类别	特征特点
点特征	计算复杂度低,能够提取图像的细节或局部特征,低纹理时特征点提取难度大,对特征描述要求高
线特征	能够捕捉全局几何信息,相较于点特征稳定,计算复杂度高,对细节不敏感
面特征	最为稳定的图像特征,计算复杂度高,对细节不敏感

(1)点特征提取。图像特征提取与匹配技术经历了多年发展,涵盖了不同类型的特征点与算法,其中角点特征与斑点特征在图像匹配中尤为重要^[23]。角点特征是指图像中边缘相交的点,具有显著的灰度变化和较高的辨识度,常用于描述图像中的关键局部结构。斑点特征则是图像中的局部灰度极值点,能够在一定程度上抵抗灰度变化,适用于纹理较少的区域,弥补了角点特征的不足^[24]。Moravec^[25]提出一种角点检测方法,首先设计一个测度函数在图像邻域内移动,其次通过检测灰度差异的变化来寻找角点。随后,Harris 等^[26]提出的 Harris 算子通过分析滑动窗口内两个正交方向的灰度变化来检测角点,能够稳定地检测图像中边缘线相交处的特征点,然而,Harris 算子在角点定位精度上存在一定不足,尤其是容易受到噪声影响,导致误检。为了解决这些问题,张小洪等^[27]提出一种基于小波变换的多尺度 Harris 角点检测算法,有效解决了角点漏检、定位不准确和伪角点的问题。何海清等^[28]通过双重角点筛选进一步减少了伪角点的影响。Zheng 等^[29]提出的 Harris-affine 算法提升了仿射变形的鲁棒性。Ye 等^[30]构建了高斯差分空间(Difference of Gaussian, DoG)空间,在该空间内提取 Harris 角点,并结合 DoG 响应和非极大值抑制对特征点进行筛选,有效提升了后续匹配精度。Smith 等^[31]提出的 SUSAN (Smallest Univalued Segment Assimilating Nucleus)算法是常用的角点特征提取方法,其通过比较图像中某一点与其邻域点的灰度相似性来检测图像中的非平坦区域,进而识别角点。该方法简单易用,可同时检测边缘和角点,但其检测效果高度依赖于阈值的设定。为此,He 等^[32]在 SUSAN 算法的基础上引入梯度直方图技术,设计实现了自适应阈值 SUSAN 算法,显著提高了检测精度和计算速度。Rosten^[33]提出 FAST (Features from Accelerated Segment Test)算法,通过计算圆周像素的灰度

快速检测图像中的角点,具有明显的速度优势。Mair等^[34]提出的 AGAST(Adaptive and Generic Accelerated Segment Test)算法在 FAST 算法的基础上进一步提高了计算效率。尽管角点特征因其直观性而被广泛应用,但在某些复杂场景中,其匹配的稳健性往往不足。因此,许多研究者开始探索新的特征,其中斑点特征作为一种潜在的替代方案受到广泛关注。此外,David G. Lowe^[1]于 2004 年提出的尺度不变特征变换(Scale Invariant Feature Transform, SIFT)算法也成为经典的解决方案之一。SIFT 的核心思想为通过构建图像的 DoG 空间来检测其中灰度的局部极值点,并将这些点作为特征点。该算法虽然在图像特征提取方面取得了显著成果,但其计算量较大,运行速度较慢,限制了其实时应用。为解决这一问题,Bay 等^[35]提出 SURF (Speeded-Up Robust Features)算法,通过计算图像的 Hessian 矩阵行列式和积分图来检测极值点。该算法提高了特征提取速度,同时保持着较高的匹配精度。

此外,一些学者尝试通过对图像进行不同的预处理,或通过改变图像的特征空间,以提升特征提取算法的性能。例如,姚永祥等^[36]提出的顾及各向异性加权力矩与绝对相位方向(Histogram of Absolute Phase Consistency Gradients, HAPCG)算法通过构建图像的相位一致性空间,提升了 Harris 角点提取能力,使得特征点检索更加高效,能够提取到更多图像细节特征。Zhu 等^[37]指出相位一致性模型的最大矩和最小矩检测器在应用中可能导致多方向特征丢失,同时存在计算成本较高的问题,为解决这些问题,他们设计了一种结合多通道自相关策略与 log-Gabor 小波的特征点检测器(Rotation-invariant Maximum index map of the Log-Gabor, RMLG),该检测器能够有效提取图像的多方向和多尺度边缘结构信息。传统特征检测器通常只能检测角点特征或斑点特征中的一种。针对这一局限,MMPC-Lap (Multi-Modal Principal Curvature and Laplacian Detector)检测器实现了角点特征与斑点特征的同时检测,其不仅融合了两类特征的优点,而且显著增加所检测特征点的丰富性,为后续的特征匹配提供了更丰富的支持^[24]。

角点特征和斑点特征各有其优缺点。其中,角点特征具有较高的视觉辨识度,能够代表图像中的显著边缘或结构,但对光照、对比度变化敏感,并且在低纹理区域较难提取到足够的特征点。斑点特征能够在一定程度上抵抗灰度差异,适用于低纹理区域,且特征点的分布不再局限于边缘区域,补充了角点特征的不足。然而,斑点特征的计算量相对较大,且其视觉辨识度较低,通常难以提供与角点特征同等程度的显著性。表 3 为部分点特征提取算法的比较。

(2)线特征提取。线特征相较于点特征蕴含了更丰富的结构信息,通常通过提取图像的边缘或纹理特征来表达,具有较强的稳定性。即使在成像条件发生变化时,线特征仍能保持一定的稳健性。经典的图像边缘提取算子

Table 3 Comparison of point feature extraction algorithms

表 3 点特征提取算法比较

方法	算法特点
文献[25]方法	便于实现,无法反映多方向特征趋势,无法精确到亚像素
Harris算子 ^[26]	可精确到亚像素,对光照敏感,低纹理区域效果差,无尺度不变性
文献[27]方法	一定程度解决了Harris漏检,定位差,伪检测等问题
文献[28]方法	减少了Harris算法伪角点的情况
文献[29]方法	提升了Harris算法的仿射变换鲁棒性
文献[30]方法	提升了Harris算法的特征点精度
SUSAN ^[31]	无需计算图像梯度,低纹理区域效果差,依赖于人工阈值设置
文献[32]方法	改进SUSAN算法无需人工设置阈值
FAST ^[33]	计算速度快,光照鲁棒性差
AGAST ^[34]	改进FAST算法,使用自适应阈值,提升效率,具备了旋转不变性
SIFT ^[1]	稳定鲁棒性强,适用范围广,计算复杂度高,不适合实时应用
SURF ^[35]	计算速度快,鲁棒性不如SIFT
HAPCG ^[36]	增强了Harris提取特征点的数量和鲁棒性
RMLG ^[37]	优化了相位一致性空间角点提取的问题,可有效提取多尺度,多方向特征点
MMPC-Lap ^[24]	实现了角点特征与斑点特征的同时检测,增加了特征点的丰富性

包括 Roberts 算子、Prewitt 算子、Sobel 算子和 Canny 算子。其中,Roberts 算子通过使用 2×2 卷积核计算图像的一阶差分,对图像的对角线方向边缘特别敏感,但容易忽略水平和垂直方向的边缘^[38]。为了克服这一问题,Prewitt 算子被提出,其采用 3×3 卷积核,相比 Roberts 算子在噪声抑制方面表现更为出色^[39]。Sobel 算子在 Prewitt 算子的基础上对卷积核内的各个位置进行了加权处理,从而增强了边缘定位的精度^[40]。然而,以上方法在弱边缘的检测方面效果有限。Canny 算子在 Sobel 算子的基础上引入非极大值抑制技术和双阈值法,显著提高了边缘提取的精确度,并改善了边缘的连续性^[41]。

Etemadi^[42]从给定的边缘图生成像素链,并遍历这些链以提取线段和弧线,该方法无需进行参数调整,然而其在处理复杂结构的图像时易产生大量误检,从而影响提取结果的准确性与可靠性。针对上述问题,GIOI 等^[43]提出 LSD (Line Segment Detector)的算法,用于图像中的线特征提取,与传统方法依赖图像梯度值不同,LSD 算法以图像灰度梯度方向为基础,通过计算梯度方向来划分线的支撑区域,在一定容差范围内,该算法利用区域生长方法合并线支撑区域,从而获得线特征的分布情况。Akinlar 等^[44]提出的 EDLines 算法设计了一种高效的边缘检测方法 ED (Edge Drawing),通过生成稳定且连续的边缘像素链,并进一步应用最小二乘法进行直线拟合,成功避免了传统边缘检测中需要人工设置阈值的局限性,显著提高了自动化程度和适用性。尽管这些方法在消除加性噪声方面表现良好,尤其在光学图像中效果显著,但在 SAR 图像中效果较差,主要是由于算法难以有效抑制乘性噪声。为此,Guo 等^[45]提出 ROA 算子,旨在更有效地提取 SAR 图像中的边

缘信息。表 4 为部分线特征提取算法的比较。

Table 4 Comparison of line feature extraction algorithms

表 4 线特征提取算法比较	
方法	算法特点
Roberts ^[38]	计算简单,无法检测平滑边缘,不支持多方向检测,边缘粗糙
Prewitt ^[39]	平滑效果好于 Roberts,不支持多方向检测,边缘定位不精确
Sobel ^[40]	边缘定位准确性优于 Prewitt,不支持多方向检测
Canny ^[41]	边缘连续性好,能区分强弱边缘,支持多尺度扩展,对人工参数敏感
文献[42]方法	无需进行参数调整,复杂结构时会产生误检测
LSD ^[43]	实现了亚像素精度,能够区分线段的长度和方向,无法检测曲线特征,后期使用复杂,重复线段可能出现误检测
EDLines ^[44]	计算复杂度低于 LSD,边缘连续性好,依赖图像梯度信息,无法检测曲线特征,不适合大规模检测
ROA ^[45]	能够有效抑制乘性噪声,适合 SAR 图像线特征提取

(3)面特征提取。在图像匹配任务中,面特征通常是指从图像的局部区域提取的特征,这些区域一般是具有相似特征的封闭区域。相较于其他类型的特征,面特征所包含的信息更为丰富,不仅覆盖了区域内的纹理和点特征,而且包括区域的形状、颜色等属性。由于面特征通过整体区域的信息进行表达,在面对随机噪声时具有较强的抗干扰能力。面特征的提取通常依赖于图像分割算法,经典的分割算法包括阈值分割法、聚类法、边缘检测法和区域生长法^[46-49]。其中,阈值分割法、聚类法和区域生长法均基于像素灰度值的相似性来进行区域划分。然而,阈值分割法在精细图像分割任务中的应用受到限制,原因在于其不仅需要指定多个分割阈值,而且可能导致空间上不连续的区域被错误地划分为同一区域。聚类分析法在一定程度上可以避免这些问题,但当图像整体的像素灰度差异较小时分割效果较差。区域生长法则依赖于大量的初始“种子”点,且在图像整体灰度区分度较小时效果同样不佳。相对而言,边缘检测法通过提取图像中的边缘,并将边缘所围成的封闭区域视为面特征,表现出较好的分割效果。近年来,学者们提出许多新的方法来克服传统算法的不足,包括马尔可夫随机场、数学形态学等^[50-54]。例如,张雍吉等^[55]提出一种基于光学影像与 SAR 影像共享区域特征的配准方法,该方法首先提取影像的边缘特征,结合数学形态学得到封闭轮廓,从而形成闭合区域。随后区域重心被选为控制点,用于推导配准变换参数,实现光学影像与 SAR 影像的配准。李雨谦等^[56]则在水平集框架下,通过映射函数将 SAR 与光学影像的特征相结合,用于构建能量泛函模型,并采用水平集方法求解曲线演化方程,在实现影像分割的同时完成了光学影像与 SAR 影像的配准。表 5 为部分面特征提取算法的比较。

1.2.2 特征描述

在完成特征提取后需要对所提特征进行合理的数学表示,以便后续进行匹配。特征描述的目的为将提取到的特征转化为易于比较的形式,常用方法为通过向量、矩阵等方式对特征的局部信息进行编码。

Table 5 Comparison of surface feature extraction algorithms

表 5 面特征提取算法比较	
方法	算法特点
阈值分割法	简单易实现,对噪声敏感,缺乏自适应,不适合处理复杂图形
聚类法	可解释性强,对噪声敏感,需要人工指定簇数量,结果不确定
边缘检测法	效果稍好于灰度分类方法,边缘复杂时效果较差
区域生长法	简单易实现,对种子(初值)敏感,容易受噪声影响,边界可能不清晰
MRF	能有效考虑像素及邻域关系,分类效果较好,模型设计较为复杂,优化过程可能出现不收敛情况
数学形态学	便于与其他方法结合使用,可能导致画面细节丢失,处理复杂图像时效果有限
文献[55]方法	结合边缘检测和数学形态学获取封闭区域,再以封闭区域重心提取特征点
文献[56]方法	通过映射函数将 SAR 与光学影像的特征相结合,构建了能量泛函模型,并采用水平集方法求解曲线演化方程

(1)点特征描述。SIFT算法的提出是点特征描述发展历程中的一次重要突破。该算法通过特征点邻域的梯度分布构建描述符,从而对特征点进行表征。相比之下,以往算法主要侧重于区域匹配,更加关注区域灰度的统计相似性,较少专门进行特征描述。而 SIFT 描述符结合了 DoG 空间和特征点主方向,使其在旋转、尺度和光照变化下均具备较强的鲁棒性。

尽管 SIFT 算法在光学图像匹配中表现优异,但由于多模态遥感图像普遍存在非线性辐射差异,SIFT 描述符容易导致错误匹配。为此,一些学者提出基于 SIFT 的优化算法。例如, Ma 等^[57]提出的 PSO-SIFT 算法通过改进图像梯度的计算方法及特征点邻域的分布方式来增强描述符对非线性辐射差异的抗干扰能力。Xiang 等^[58]提出的 OS-SIFT 算法在光学和 SAR 图像上分别使用多尺度 Sobel 梯度、多尺度 Sobel 梯度和加权平均梯度比计算梯度,确保了光学与 SAR 图像梯度的一致性,实现了高分辨率光学与 SAR 图像的匹配。ORB (Oriented FAST and Rotated BRIEF) 算法在图像匹配任务中表现出较高的效率,一些研究针对 ORB 算法进行了改进,实现了同源遥感图像的匹配。例如,张云生等^[59]提出一种基于改进 ORB 算法的遥感图像自动配准方法,获得了与 SIFT 算法和 SURF 算法相当或者更高的配准精度,并在配准效率上有较大提高。相位携带了图像的主要几何结构和形态特征信息,因此频域处理方法逐渐被引入多模态图像匹配中。例如, Li 等^[60]提出的辐射变化不敏感特征变换 (Rotation-Invariant Feature Transform, RIFT) 描述符结合了相位一致性与最大索引图,在抵抗非线性辐射差异方面表现出明显优势。姚永祥等^[36]提出的 HAPCG 算法通过各向异性滤波和非线性扩散结合相位一致性模型构建了各向异性加权力矩图,该算法在抵抗非线性辐射差异方面表现优异,运算效率也较高。Zhu 等^[37]通过快速分配主方向并采用 DAISY 空间配置进行特征描述来生成 RMLG 描述符,解决了传统 Log-Gabor 卷积和直方图描述方法中存在的计算复杂度较高问题。Ma 等^[61]提出的归一化自相似区域描述符 (Modality Inde-

pendent Region Descriptor, MIRD)首先根据相邻图像块之间的欧几里得距离构建点式自相似区域描述符,以反映多模态图像的结构特性;其次采用线性归一化方法将其转化为模态独立区域描述符,从而有效区分多模态图像中的点、线、角和平面等结构特征。这一改进显著提升了图像匹配的精度与鲁棒性。Ye等^[62]通过扩展HOG(Histogram of Oriented Gradients)描述符和引入方向梯度通道特征,并利用这些特征的傅里叶变换进行相似度计算,从而实现多模态图像之间的稳健匹配。表6为不同点特征描述方法的比较。

Table 6 Comparison of point feature description methods

表6 点特征描述方法比较

方法	算法特点
SIFT ^[1]	具备尺度、旋转、光照不变性,特征点适应性强
PSO-SIFT ^[57]	比SIFT更具尺度旋转不变性,计算复杂度高
OS-SIFT ^[58]	增强了旋转不变性,计算复杂度较高
ORB ^[59]	计算效率高,对大角度旋转、尺度变化敏感,匹配能力不如SIFT
RIFT ^[60]	增强了光照不变性,匹配区分度更高,计算复杂度高,实时性差
HAPCG ^[36]	匹配区分度进一步提高,算法实现复杂
RMLG ^[37]	优化了直方图描述的复杂性问题
文献[24]方法	一定程度解决了多模态图像相位反转的问题
MIRD ^[61]	匹配区分度提升,更好地区分图像的几何特征
CFOG ^[62]	扩展HOG特征描述子,结合特征傅里叶变换进行相似度量,特征匹配阶段采用模板方式,具有较高的稳健性和效率

(2)线特征描述。线特征的描述方式主要分为几何描述、拓扑描述和像素描述3类。其中,几何描述通过线段的端点坐标、长度和方向等几何信息来表征线段。例如,Huffman和Shriwas等^[63-64]使用霍夫曼变换(Hough Transform)检测并提取图像中线特征的斜率和截距等信息,但这种方式一方面受制于人工阈值的设置,另一方面在高密度边缘区会产生大量误检测。拓扑描述则是基于线特征之间的拓扑关系表征线特征组,例如利用线段间的夹角和距离等关系进行描述。像素描述通过线段及其邻域的像素点信息构建描述符,这种方法将复杂的线特征描述问题转化为点特征描述问题,如通过对线段上的点应用SIFT或SURF特征点描述方法来表征线特征。在此基础上,Zhang等^[65]提出的LBD(Line Band Descriptor)算法在线段两端的线段带中统计梯度信息,将线段带划分为多个区域,并按照一定顺序将各区域的梯度信息汇总为描述向量,最终构建出线特征的描述符。该方法有效结合了线特征与点特征的优势,能够较好地应对图像间的尺度、旋转等变化。表7为部分线特征描述方式的比较。

(3)面特征描述。面特征的描述方式通常分为几何描述、纹理描述和像素描述3类。几何描述基于区域的几何信息表征面特征,较为经典的方法为Hu矩描述。1962年,Hu^[66]提出7个几何矩,并证明这些几何矩能够有效表征图像的形状特征,同时对旋转、平移和尺度变化具有较强的

Table 7 Comparison of line feature description methods

表7 线特征描述方法比较

方法	算法特点
Hough ^[63]	可以检测多种几何特征(直线、椭圆、圆等),需要人工指定阈值,可能出现误检测
拓扑描述	表达能力强,支持多尺度分析,算法实现复杂,计算复杂度高
LBD ^[65]	描述符信息丰富,鲁棒性强,计算复杂度高,不适合实时应用

鲁棒性。Hu矩的计算复杂度相对较低,因此特别适合处理大幅宽的遥感数据。然而,Hu矩也存在一定局限性,其对于形状复杂的区域描述能力不足,难以全面捕捉区域的细节特征,并且其计算依赖图像灰度信息,对噪声和多模态遥感图像的非线性辐射畸变较为敏感。纹理描述方法通过分析区域内的纹理分布情况来表征面特征,常用方法包括灰度共生矩阵(Gray Level Co-occurrence Matrix, GLCM)和局部二值模式(Local Binary Pattern, LBP)^[67-68]。GLCM通过指定偏移量和偏移角度生成不同的矩阵,并根据这些矩阵计算图像的全局纹理信息,常用纹理特征包括对比度、同质性等。GLCM提取的纹理特征较为直观,并能较好地保持在尺度、旋转等变换下的稳定性。然而,由于其需要构建和处理多个矩阵,计算复杂度相对较高,这在大规模图像分析中可能成为瓶颈。LBP的主要思想为对像素及其邻域像素的灰度关系进行二值化编码,生成LBP特征图,并进一步统计该特征图的直方图以表征面特征。与GLCM相比,LBP的计算过程较为简便,且能够较好地抵抗一定程度的光照变化。然而,LBP本身不具备旋转不变性,限制了其在复杂变换下的应用。为克服这一局限,后续学者在LBP的基础上提出旋转不变的LBP(Rotation-Invariant Local Binary Pattern, ROLBP),有效提升了其在旋转和尺度变换下的鲁棒性。像素描述则是通过对区域内的点特征进行描述,从而实现对面特征的描述。表8为不同面特征提取算法的比较。

Table 8 Comparison of surface feature description algorithms

表8 面特征描述算法比较

方法	算法特点
Hu矩 ^[66]	具有尺度、旋转、平移不变性,计算简单,对复杂几何特征区分度差
GLCM ^[67]	能够较为全面的获取纹理特征,鲁棒性强,算法中方向指定为固定角度,对特殊角度敏感性不够强
LBP ^[68]	计算效率高,不适合平滑区域,扩展性较差

1.2.3 特征匹配

图像匹配的关键任务之一为识别两幅图像中对应的同名特征点,构建有效的特征描述符是为了更精确地衡量特征点之间的相似性,而特征匹配即为检测特征描述之间的相似性。特征描述的主要思想为构建特征的描述向量,衡量两个向量相似程度的常用指标包括余弦相似度和欧氏距离等。这两种度量方式在评估描述符相似度时往往依赖于描述符的准确性,即当所构建的描述符不合适时,匹配过程容易产生错误。尤其是在多模态遥感图像中,由于存在显著的非线性辐射差异,描述符往往容易受到影

响,进而产生大量错误匹配点,影响匹配的整体效果。为解决这一问题,许多学者不仅致力于提出更为精确的特征描述方法,而且通过在匹配阶段减少误匹配点来提升匹配精度。例如,PSO-SIFT算法通过综合考虑特征点的尺度、匹配误差等因素,设计了PSOED度量方法来衡量相似性,从而有效减少了错误匹配点数量^[57]。李培等^[69]结合张量方向平行度与梯度互信息提出一种新的相似性度量准则(Tensor Orientation and Mutual Information, TOM),该方法在光学影像、LiDAR数据、SAR图像及地图数据的匹配中表现出色,具有较强的鲁棒性。

1.2.4 粗差去除

在图像匹配过程中,由于描述符自身的局限性以及多模态遥感图像的复杂性,错误匹配点的出现是不可避免的。粗差去除即是在初步匹配的基础上使用合适的方法筛选出正确的匹配点,去除错误的特征匹配关系,从而提高匹配精度。常用的误匹配点去除算法包括随机样本共识(Random Sample Consensus, RANSAC)和快速抽样一致性(Fast Sample Consensus, FSC)^[70-71]。这两种算法的目标均为从包含内点(正确匹配点对)和外点(错误匹配点对)的数据中筛选出内点,并剔除外点。其基本原理为通过随机采样最小点集来计算模型参数,利用该模型对所有点对进行验证,得到内点集合,并通过重复迭代该过程,最终得到最大一致性的内点集合,同时估计几何变换参数。然而,这两种算法在面对较复杂的匹配点对情况,例如内点比例较小、非刚性变换或匹配点对过多(遥感图像幅宽较大)时,存在效率和稳健性差的问题。为解决上述问题,Wang等^[72]提出一种基于局部离群因子(Guided Local Outlier Factor, GLOF)的方法。该方法首先通过多特征匹配构建初步的对应关系集;其次利用局部离群因子(Local Outlier Factor, LOF)来识别并剔除错误匹配。实验结果表明,GLOF在公开数据集上的表现优于对照方法,但其计算复杂度较高。Meng等^[73]通过构建有向最近邻图(K-Nearest Neighbor Graph, KNN)将特征点匹配问题转化为二分类问题,并利用KNN密度估计和SOC计算得到每个特征点的内点概率,从而去除异常值。虽然该方法在高异常值比例下有效提高了特征匹配的鲁棒性和准确性,但是其计算复杂度和效率仍需进一步改进。Ma等^[74]提出的渐进向量场共识(Progressive Vector Field Consensus, PVFC)方法通过插值构建合适的向量场来区分内点和外点,实验结果表明,该算法在大多数图像上的精度和召回率均超过90%。Feng等^[75]在改进RANSAC算法的基础上提出PNP-RANSAC算法,通过在单应矩阵估计中引入几何约束项,有效减少了由于图像间非刚性变换引起的匹配点误筛除问题。

为了实现更高精度的匹配,二重匹配方法被提出。该方法的核心思想为在初次匹配(粗匹配)后利用粗差去除算法获得图像的变换矩阵,然后基于该矩阵及其他信息进

行二次匹配(细匹配)。这种方式能有效提高匹配精度,去除更多误匹配点,但同时也会降低计算效率。例如,PSO-SIFT算法在进行初步匹配后,结合仿射变换模型、特征点所在尺度等因素构建了PSOED距离,从而衡量特征点间的相似性,在此基础上重新筛选内外点,最终提高了匹配精度。

总体而言,基于特征的方法通过对像素进行处理,提取图像的抽象表述,能有效应对图像中的复杂情况。该方法不仅提高了算法对特殊情况的适应能力,而且在面对图像中的部分遮挡或形变时展现出更强的稳定性。通过从局部或全局特征中提取信息,基于特征的方法能够在较为复杂的环境中展现较为鲁棒的性能,特别是在不完整或变形的图像中可保持较高的匹配准确度。

1.3 基于数据驱动的匹配

近年来,随着计算机硬件性能的显著提升,深度学习网络凭借其强大的数据学习能力,逐渐在各个领域展现出显著优势。与传统的特征匹配方法相比,深度学习方法在学习和表达图像的深层次特征方面具有极强优势,尤其是在多模态遥感图像匹配中,浅层特征往往难以有效应对非线性辐射畸变的影响。深度学习能够更好地捕捉图像间复杂的深层次关系,因此能够实现更高精度的匹配,尤其在面对图像变形、噪声及其他复杂因素时,展现出更强的鲁棒性和适应性。

根据网络结构的不同,深度学习图像匹配方法可以分为两类:一类是将深度学习网络与传统匹配方法相结合的单环节深度网络,这类方法通常只在图像匹配的某个或部分环节中应用深度学习,结构相对简单;另一类是完全依赖深度学习网络来实现匹配过程的端到端深度网络,这类方法的网络结构较为复杂,算法复杂度较高,适合解决系统性问题。

1.3.1 单环节深度网络

单环节深度网络通过替换图像匹配的一个或多个环节,从而更好地应对不同任务场景,满足特定任务需求。例如,Dusmanu等^[76]提出的D2-Net算法采用卷积神经网络(Convolutional Neural Network, CNN)和非极大值抑制(Non-Maximum Suppression, NMS)来检索特征点,并使用CNN特征图对特征点进行描述。该方法在处理日夜变化图像时表现出色,并在多个公开数据集上取得了优异成绩。Ma等^[77]提出一种结合CNN与传统局部特征描述的方法,首先通过CNN进行图像间的粗匹配,其次结合局部特征和空间关系对粗匹配结果进行精细调整,以实现更高精度的匹配。南轲等^[78]对Siamese网络进行了优化,通过去除池化层来提升特征信息的完整性和位置精度,并结合模板匹配方法实现了多模态遥感图像的高精度匹配。该网络采用一种创新的损失函数,通过最大化正样本与难负样本之间的特征距离来提高性能,但算法训练效率较低,并且对旋转、尺度变化明显的复杂数据效果较差。Cao等^[79]

基于 Siamese 网络提出一种融合相位一致性的孪生网络 PCNet (Phase Congruency Network), 用于挖掘多光谱与多模态图像间的相似结构, 从而提升匹配精度与稳健性。Yi 等^[80]将检测器、特征方向估计与特征描述统一到一个 Siamese 网络中, 实现了更好的整体性能。Zhang 等^[81]提出一种基于 Siamese 网络的全卷积网络 SFCNet (Separable Factorized Convolution Network), 使用最大化正样本与负样本之间特征距离的损失函数, 在光学、近红外、热红外、SAR 等图像匹配中表现优异。Li 等^[82]提出的 RotNet (Rotation Network) 神经网络首先分析和预测图像间的旋转关系, 其次利用 GPOG (Gradient-Phase Oriented Grid) 描述符完成图像匹配。实验结果表明, 该网络能有效抵抗 SAR 图像中斑点噪声、旋转、非线性辐射畸变等因素的影响。Sarlin 等^[83]提出的 SuperGlue 算法将特征匹配问题转化为可微分的优化问题, 通过构建图神经网络 (Graph Neural Networks, GNN) 来进一步提升匹配精度。Lindenberger 等^[84]通过优化 SuperGlue 算法中 Transformer 的性能, 提出更高效的 LightGlue 算法, 在重叠较多或变化较小的图像匹配任务中实现了更低的延迟和更优的计算效率。Zhang 等^[85]提出一种轻量级自监督匹配网络 LTFormer (Light-weight Transformer Network), 利用 Transformer 的子注意力机制生成特征描述符, 并引入三元损失函数 LTLoss (Light-weight Transformer Loss) 来优化特征匹配过程。该方法在光学与近红外图像匹配任务中表现出优异性能。

许多学者在特征描述阶段采用单一的深度描述符学习网络 (Deep Descriptor Learning, DDL), 但该网络难以同时提取对旋转和模态变化均具有不变性的局部特征。为此, Quan 等^[86]提出一种深度序数回归 (Deep Ordinal Regression, DOR) 方法, 通过对图像进行粗匹配来实现旋转校正。该方法在预处理阶段显著降低了多模态图像配准的复杂性, 从而有效提升了配准效果和精度。Liang 等^[87]借鉴 Transformer 中的自注意力与互注意力机制, 结合图像的局部特征与全局特征, 提出一种特征匹配网络 MatcherTF。该网络通过对检测器提取或随机选取的特征点进行特征增强, 以生成更具区分性的特征描述, 从而提升匹配算法的性能与稳健性。表 9 为部分单环节深度网络方法的比较。

此外, 一些学者提出通过深度学习实现多模态图像的风格迁移, 将不同模态的图像转换为同一模态, 以解决多模态图像匹配问题。例如, 余佩伦等^[88]提出的 SK-GAN 通过动态感受野获取多尺度信息的生成结构, 显著提高了生成图像的质量。尹梦晓等^[89]提出的 KCC-GAN 将 K-means 分割作为图像生成器的输入之一, 通过约束空间信息合成来提升生成图像质量, 在 SAR 与光学图像匹配中取得了良好效果。

1.3.2 端到端深度网络

端到端深度网络是指从输入到输出的整个过程均由

Table 9 Comparison of single-link deep network methods

表 9 单环节深度网络方法比较

方法	算法特点
D2-Net ^[76]	可以实现高精度匹配, 对图像内容提取自适应描述符, 训练时间相对较长
文献 ^[77] 方法	结合 CNN 与局部特征描述, 提升了匹配精度
文献 ^[78] 方法	去除 Siamese 网络中的池化层, 使用优化的损失函数, 增强了特征定位精度
PCNet ^[79]	结合相位一致性的孪生网络用于挖掘多光谱与多模态图像间的相似结构
文献 ^[80] 方法	将检测器、特征方向估计、特征描述都集成到一个 Siamese 网络, 提升了整体的性能
SFCNet ^[81]	基于 Siamese 网络提出 SfcNet, 在光学、近红外、热红外、SAR 等图像匹配中有不错的效果
RotNet ^[82]	能有效抵抗 SAR 图像中的斑点噪声及旋转、非线性辐射畸变等因素的影响
SuperGlue ^[83]	使用 GNN 实现相似性度量, 提升了匹配精度
LightGlue ^[84]	优化了 SuperGlue 中计算复杂度, 实现了更高效的匹配
LTFormer ^[84]	结合自注意力机制, 在光学与近红外图像的匹配任务中表现良好
DOR ^[86]	通过粗匹配降低多模态遥感图像匹配的复杂性, 提升了匹配的精度
MatcherTF ^[87]	结合了图像的全局特征和局部特征实现了更具区分度的特征描述

一个统一的模型或系统自动完成, 不需要人为的中间步骤或单独模块。例如, Ma 等^[90]提出由 3 个神经网络构成的多模态图像匹配框架 LMR (Learning for Mismatch Removal), 分别用于特征空间提取、特征匹配和误匹配点去除, 该框架在光学与 SAR 图像匹配中表现优异。Sun 等^[91]提出的 LoFTR (Detector-Free Local Feature Matching with Transformers) 算法首先在粗颗粒度下实现图像的检索与匹配, 其次在细颗粒度下进行更精细的匹配。为提高匹配精度, 该算法引入自注意力和互注意力机制, 以获得更细节的特征描述。此外, 该算法还设计了特定的搜索器, 用于在低纹理区域获取足够置信度的特征点。Zhang 等^[92]提出一种名为 RSOMNet (Robust SAR-Optical image Matching Network) 的多粒度特征对齐网络结构, 通过在多个粒度级别的特征图上实现特征对齐与匹配, 显著增强了对噪声和分辨率变化的鲁棒性。实验结果表明, RSOMNet 在光学图像与 SAR 图像之间的匹配任务中表现出良好的性能和泛化能力。Li 等^[93]提出一种针对多模态遥感图像的融合框架, 通过自注意力机制对参考图像和基准图像的特征表示进行融合, 并利用全连接层预测特征点的位移参数, 从而计算出图像间的透视变换矩阵, 实验验证了该框架在 RGB 图像、多光谱图像和 SAR 图像匹配中的有效性。Shen 等^[94]提出 RF-Net, 即在 LF-Net (Learning-based Feature Network) 的基础上构建感受野特征图, 并引入邻域掩码作为损失函数, 以优化特征提取与匹配过程。该网络在 HPatches 数据集上表现优异, 有效提升了匹配的准确性与稳健性。陈鹏等^[95]提出 DFL-Net (Distinguishable Feature Learning Network) 匹配网络, 通过 DFL-Det (Distinguishable Feature Learning Detector) 特征检索器与 DFL-Des (Distin-

guishable Feature Learning Descriptor)特征描述器实现了遥感图像中多尺度关键特征的精准检索与匹配。在 Potsdam、Toronto 和 Vaihingen 数据集上,DFL-Net 的匹配性能优于 SIFT 与 RF-Net 算法,展现出更强的稳健性与泛化能力。孙晓坤等^[96]针对高分辨率下 SAR 图像因视角偏移导致的匹配困难问题,提出一种多视角 SAR 图像匹配网络。该网络基于局部像素偏移模型提取图像特征,并通过自适应调整卷积核的采样位置来优化特征描述,显著降低了匹配过程中的平均误差,提升了匹配精度与鲁棒性。石添鑫等^[97]提出一种多模态匹配网络,通过相位一致性模型提取特征,从而减少模态差异的影响,并结合交叉注意力机制构建高效的特征描述符。该网络在 Landsat8 公共数据集上表现优异,显著提升了不同模态图像间的匹配精度与稳定性。李绩鹏等^[98]提出一种融合双注意力机制的深度网络,通过双注意力层增强了对特征点的检测与定位能力,同时结合双向相关运算改善了匹配精度。在此基础上,该网络还引入单应矩阵预测层,实现了图像间变换参数的直接预测,有效提升了匹配精度与效率。邹融平等^[99]提出一种伪孪生网络,分别通过不同的残差网络对光学图像与 SAR 图像进行处理,然后使用卷积互相关方法实现两者之间的特征匹配,有效提升了匹配精度。Cao 等^[100]使用预训练的 VGG (Visual Geometry Group) 网络提取图像的深度特征,并结合 4-D 密集特征提取方法实现了较为稳健的遥感图像匹配。表 10 为部分端到端深度网络方法的比较。

Table 10 Comparison of end-to-end deep learning network methods
表 10 端到端深度网络方法比较

方法	算法特点
LMR ^[90]	三个主体部分构成的深度网络框架,在光学与 SAR 图像匹配中表现优异
LoFTR ^[91]	提升了算法在低纹理区域的特征点数量,引入自注意力机制和二重匹配,能够实现更高精度的匹配
RSOMNet ^[92]	对噪声和分辨率变化有着不错的鲁棒性
文献[93]方法	在 RGB 图像、多光谱图像和 SAR 图像匹配中具有一定的效果
RF-Net ^[94]	使用邻域掩膜作为损失函数,在公开数据集中有不错的表现
DFL-Net ^[95]	通过 DFL-Det 和 DFL-Des 检索描述多尺度特征,在多个公开遥感数据集中表现优异
文献[96]方法	基于局部像素偏移模型提取图像特征,结合自适应卷积核完成特征描述
文献[97]方法	使用相位一致性提取多模态图像中的相似特征,并使用交叉注意力实现特征描述
文献[98]方法	通过双注意力层增强特征定位性能,双向相关运算改善匹配的精度问题,网络中单应矩阵预测层直接得到图像间变换参数
文献[99]方法	使用伪孪生网络分别对光学和 SAR 图像进行卷积处理,实现二者的相似特征表达,卷积互相关实现两者的特征匹配
文献[100]方法	使用预训练的 VGG 获取图像特征,使用 4-D 密集特征提取,实现了较为稳健的遥感图像匹配

与传统的区域匹配和特征匹配方法相比,深度学习方法不再过度依赖先验知识,而是通过大量数据的训练自动优化网络结构权重,从而使得匹配性能更加理想。同时,

深度学习方法在图像深层次特征的学习与表达方面具有显著优势,能够更好地理解多模态图像之间的特征对应关系,尤其在应对非线性辐射畸变和噪声等复杂因素的影响时展现出更强的鲁棒性和适应性。

2 算法比较实验

2.1 实验方法与评价指标

本文选取 12 种具有代表性的匹配算法,包含基于区域的匹配方法(MI、SSD),基于特征的匹配方法(CFOG、SIFT、RIFT、HAPCG)以及基于数据驱动的匹配方法(SuperPoint+D2_Net、SFCNet、SuperPoint+SupeGlue、SuperPoint+lightGlue、RDFM、LoFTR),在同一测试数据集上比较其匹配性能。实验根据原文中推荐的参数进行设置,深度学习相关方法均采用 CPU 运行。采用平均正确匹配数量(Mean Number of Correct Matching, MNCM)、平均匹配准确率(Mean Matching Accuracy, MMA)、平均算法运行时间(Mean Running Time, MRT)作为评价指标。计算公式为:

$$MNCM = \frac{1}{n} \times \sum_{i=1}^n CMP_i \quad (1)$$

$$MMA = \frac{1}{n} \times \sum_{i=1}^n \frac{CMP_i}{IMP_i} \quad (2)$$

$$MRT = \frac{1}{n} \times \sum_{i=1}^n RT_i \quad (3)$$

式中, n 表示图像对总数; CMP 表示正确匹配数量; IMP 表示算法初次匹配后获得的初始匹配数量; RT 为算法完成某一个图像对匹配任务的时间。统计正确匹配点的标准为单应矩阵变换后像素误差小于 3px。数据驱动算法运行时间不包含模型训练时间,只包含模型泛化时间。

2.2 实验数据

实验数据采用公开遥感图像匹配数据集的部分影像,涵盖光学—红外、光学—激光雷达图(LiDAR intensity:激光雷达强度图,LiDAR depth:激光雷达深度图)、光学—SAR、光学—地图 4 种模态,具体信息见表 11。数据集使用及下载可以访问 <https://github.com/yeyuanxin110/CFOG>。

2.3 实验结果分析

实验结果如表 12 所示。在基于区域的方法中,SSD 算法的表现较差,MI 算法效果较好,但在光学图像与 SAR 图像的匹配中,由于两者互信息存在一定的差异,相关性较弱,导致匹配效果不稳定。在基于特征的方法中,SIFT 算法虽然匹配精度不低,但正确匹配点数较少,主要是由于该算法在这几组模态之间不容易获取到足够数量的初始匹配点。CFOG 算法结合了一些模板匹配的优点,一定程度上保证了匹配点分布的均匀性,结合稳定性较强的特征描述,实验过程中有着不错的匹配精度和数量,但在光学图像与地图图像匹配时,CFOG 的表现略显不稳定,这可能是由于地图图像在灰度平滑区域缺乏足够的显著特征,导致基于模板的匹配方法容易引入较多错误匹配点。RIFT

Table 11 Experimental data information

表 11 实验数据信息

类别	组别	基准图像	待匹配图像
光学—红外	1	卫星: Sentinel-2 GSD*: 10 m	卫星: SDGSAT-1 热红外 GSD: 30 m
	2	卫星: Landsat 5 TM band 1 GSD: 30 m	卫星: Landsat 5 TM band 4 GSD: 30 m
	3	卫星: LiDAR intensity GSD: 2 m	卫星: WorldView 2 optical GSD: 2 m
光学—激光雷达	4	卫星: LiDAR intensity GSD: 2 m	卫星: LiDAR intensity GSD: 2 m
	5	卫星: LiDAR depth GSD: 2.5 m	卫星: LiDAR Airborne optical GSD: 2.5 m
	6	卫星: TM band 3 GSD: 30 m	卫星: TerraSAR-X GSD: 30 m
光学—SAR	7	卫星: Google Earth GSD: 3 m	卫星: TerraSAR-X GSD: 3 m
	8	卫星: Google Earth GSD: 3 m	卫星: TerraSAR-X GSD: 3 m
	9	卫星: Google Maps GSD: 0.5 m	卫星: Google Maps GSD: 0.5 m
光学—地图	10	卫星: Google Maps GSD: 1.5 m	卫星: Google Maps GSD: 1.5 m

注: GSD (Ground Sample Distance) 指地面采样距离。

算法和 HAPCG 算法均结合相位一致性模型实现特征提取, 更容易获取到多模态图像间的共同特征, 相比之下, HAPCG 算法在计算效率和匹配点数量上略胜一筹。在单环节深度网络中, D2_Net 算法效果不稳定, 多模态特征描述能力较弱, 原始训练参数在多模态图像匹配任务中难以提取有效特征, 出现较多错误匹配。SFCNet 算法整体表现不错, 在不同模态下均能保证较为稳定的匹配, 但在光学与地图实验中表现欠佳。SuperGlue 与 LightGlue 匹配效果差异不大, 后者效率相较前者有较为明显的提升。在端到端深度网络中, RDFM 算法表现出较高的匹配精度, 而 LoFTR 在匹配点数和效率上有一定优势。综合来看, 深度学习方法在匹配点数量、正确匹配点比例方面整体优于传统方法, 特别是 LoFTR、RDFM 和 LightGlue 等算法, 在匹配精度与运行时间之间取得较好的平衡, 展现出较高的实用价值。

2.4 小结

不同图像匹配方法有其优劣势, 可根据具体任务选择合适的方法。其中, 基于区域的方法可以确保匹配区域内包含足够数量的匹配点对。然而, 多模态图像间普遍存在非线性辐射差异, 使得区域间相似性度量的设计变得较为复杂。在大多数情况下, 相似性度量的构建必须结合具体任务需求与先验知识。此外, 该类方法通常需要针对图像的尺度、旋转等因素进行专门优化, 且需依赖遥感图像的地理信息进行初步对齐, 降低了其通用性。许多基于特征的匹配算法天然具备尺度不变性和旋转不变性, 在不同应用场景下展现出较强的适应性。该类方法灵活性较高, 可根据具体需求调整流程并加以扩展。然而, 特征匹配方法

Table 12 Comparison test result of image matching algorithms

表 12 图像匹配算法比较实验结果

分类	方法	MNCM	MMA/%	MRT/s	
基于区域的匹配方法	MI	47	44.63	47.8	
	SSD	23	37.62	39.6	
	SIFT	51	61.27	37.7	
基于特征的匹配方法	CFOG	153	71.23	41.7	
	RIFT	197	78.64	41.2	
	HAPCG	217	75.58	29.1	
	SuperPoint+D2_Net	91	51.23	42.4	
基于数据驱动的匹配方法	单环节深度网络	SFCNet	134	74.95	45.1
	SuperPoint+SuperGlue		157	71.94	48.6
		SuperPoint+light-Glue	162	72.16	37.5
		端到端深度网络	RDFM	193	81.37
		LoFTR	271	76.32	36.4

设计有两个关键点。首先, 特征提取的质量和数量对匹配算法的稳定性和最终效果具有重要影响, 若提取的特征过少, 可能导致匹配精度下降, 甚至匹配失败; 其次, 特征描述的设计至关重要, 这一过程直接影响误匹配的比例, 从而决定整体匹配性能。基于数据驱动的方法通过训练模型, 能够有效应对多模态遥感图像的复杂性, 还能适应多样化的任务需求, 从而探索出兼具高性能与强通用性的匹配策略。然而, 数据驱动方法的一部分核心在于高质量数据集的构建, 当前公开的多模态遥感图像数据集相对稀缺, 且大多集中在光学与 SAR 影像匹配任务上。这一现状极大限制了该类方法在多模态遥感图像匹配中的发展。

3 结语

多模态图像匹配在遥感领域中发挥着至关重要的作用。不同模态的图像可以提供不同的对地观测信息, 如 SAR 图像不受天气、云层、光照等因素的影响, 可以实现全天候的对地观测; 同时由于穿透性强, SAR 对应波段可以穿透植被、土壤表层, 获取到地表下的信息。红外图像能有效监测地物的热辐射信息, 更容易实现火灾监测等任务, 同时还能够实现植被指数的监测。

综合利用多种模态的遥感图像能更好地实现对地观测, 获取更多地物信息。例如, 土地利用检测目标为识别不同土地的使用性质, 如区分农田、森林、城市建筑等, 图像匹配精度对最终分类结果有很大影响。一般情况下, 土地利用检测对时效性要求不高, 可以选择 CFOG、RIFT、HAPCG 等特征匹配方法, 在数据充足时也可以考虑 LoFTR、RDFM 等深度学习方法。灾害评估需要光学、SAR、GIS 数据的配合, 如灾前灾后变化分析需要高精度匹配, 特别是在建筑损毁、道路淹没等评估任务中, 精细匹配直接影响灾害评估的准确性, 通常需要在短时间内处理大

量图像,时效性要求较高,在面对大规模结构性变化(如房屋倒塌、洪水)时,算法需要具备更高的精度和鲁棒性,因此时效性较强的HAPCG、LightGlue以及loFTR可以作为首选。变化检测对匹配精度要求较高,尤其在微小变化检测任务中,需要高精度的匹配算法来识别细微差异。该任务通常为长时间序列数据检测,对时效性的要求相对较低。此外,该任务需要较强的噪声容忍度。基于特征的匹配算法(如CFOG、RIFT等)可以很好地实现小范围变化检测,深度学习方法则更适合大范围变化检测。目标检测与识别需要高精度匹配,特别是在小目标检测以及复杂背景下,匹配的细致程度至关重要。特征点匹配在低光照和低分辨率条件下较为稳定,但实时处理能力有限,深度学习方法LOFTR、RDFM等可在复杂背景下实现高效匹配,兼顾时效性和匹配精度。

随着未来天地一体化观测技术的不断完善,发展高精度、高普适性的匹配算法仍然是遥感领域的核心研究方向。如今的一个研究重心即为提升匹配算法在多模态图像间的普适性。基于数据驱动的匹配方法可以通过学习数据来增强算法的适应性。近年来,大模型在自然语言处理以及图像识别方面取得突破性进展,在遥感图像匹配领域开展大模型相关研究前景广阔。

数据驱动相关方法的未来发展主要集中在数据集的构建与优化,以及超大规模神经网络模型的探索两个方面。首先,数据集的汇总与整理是关键环节。深度学习的训练依赖于高质量、代表性强的数据集,特别是大模型更需要追求数据的数量与质量,而当前遥感图像匹配相关数据资源相对有限,尤其是公开数据集主要集中在光学与SAR这两个模态之间,难以覆盖更广泛的多模态匹配需求,未来研究应更加重视公开数据集对算法发展的推动作用,积极构建并共享涵盖更多模态、复杂场景和多种地物类别的数据资源,以支撑更具通用性的匹配算法发展;其次,超大规模神经网络模型的发展是提升多模态匹配适用性和匹配精度的重要方向。通过构建更深层次、更具泛化能力的模型,可以更有效地学习不同模态之间的非线性映射关系,提升匹配的鲁棒性与准确率。

参考文献:

- [1] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. *International Journal of Computer*, 2004, 60: 91-110.
- [2] BROWN M, LOWE D. Invariant features from interest point groups[C]// *Cardiff: British Machine Vision Conference*, 2002.
- [3] MIKOLAJCZYK K, SCHMID C. A performance evaluation of local descriptors[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(10): 1615-1630.
- [4] JIA D, ZHU N D, YANG N H, et al. A review of image matching methods[J]. *Journal of Image and Graphics*, 2019, 24(5): 677-699.
贾迪, 朱宁丹, 杨宁华, 等. 图像匹配方法研究综述[J]. *中国图象图形学报*, 2019, 24(5): 677-699.
- [5] CHEN L, ROTTENSTEINER F, HEIPKE C. Feature detection and description for image matching: from hand-crafted design to deep learning[J]. *Geo-Spatial Information Science*, 2020, 24(1): 58-74.
- [6] ZHU B, YE Y X. Multimodal remote sensing image registration: a survey[J]. *Journal of Image and Graphics*, 2024, 29(8): 2137-2161.
朱柏, 叶沅鑫. 多模态遥感图像配准方法研究综述[J]. *中国图象图形学报*, 2024, 29(8): 2137-2161.
- [7] LIU H Q, LIU M, GONG Z C, et al. A review of image matching methods based on deep learning[J]. *Acta Aeronautica et Astronautica Sinica*, 2024, 45(3): 82-95.
刘海桥, 刘萌, 龚子超, 等. 基于深度学习的图像匹配方法综述[J]. *航空学报*, 2024, 45(3): 82-95.
- [8] DAWN S, SAXENA V, SHARMA B. Remote sensing image registration techniques: a survey[C]// *Berlin: International Conference on Image and Signal Processing*, 2010.
- [9] SUI H G, LIU C, GAN Z, et al. Overview of multi-modal remote sensing image matching methods[J]. *Journal of Geodesy and Geodynamics*, 2022, 51(9): 1848-1861.
眭海刚, 刘畅, 干哲, 等. 多模态遥感图像匹配方法综述[J]. *测绘学报*, 2022, 51(9): 1848-1861.
- [10] ZHAO W X, ZHOU K, LI J, et al. A survey of large language models[DB/OL]. <https://arxiv.org/abs/2303.18223>.
- [11] SHAO M, BASIT A, KARRI R, et al. Survey of different large language model architectures: trends, benchmarks, and challenges[J]. *IEEE Access*, 2024, 12: 188664-188706.
- [12] HE X Y, YU H, PENG S D, et al. MatchAnything: universal cross-modality image matching with large-scale pre-training[DB/OL]. <https://arxiv.org/abs/2501.07556>.
- [13] JIANG X Y, MA J Y, XIAO G B, et al. A review of multimodal image matching: methods and applications[J]. *Information Fusion*, 2021, 73: 22-71.
- [14] SHI J B, TOMASI C. Good features to track[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1994: 593-600.
- [15] ZITOVÁ B, FLUSSER J. Image registration methods: a survey[J]. *Image and Vision Computing*, 2003, 21(11): 977-1000.
- [16] EMERY W J, BALDWIN D, MATTHEWS D. Maximum cross correlation automatic satellite image navigation and attitude corrections for open-ocean image navigation[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2003, 41(1): 33-42.
- [17] LU H P, GAO L. Improved algorithm for multi-source remote sensing images based on cross correlation[J]. *Aerospace Control*, 2009, 27(2): 18-21, 25.
陆和平, 高磊. 基于互相关的多源遥感图像匹配的改进算法[J]. *航天控制*, 2009, 27(2): 18-21, 25.
- [18] MAES F, COLLIGNON A, VANDERMEULEN D, et al. Multimodality image registration by maximization of mutual information[J]. *IEEE Transactions on Medical Imaging*, 1997, 16(2): 187-198.
- [19] CHEN H M, ARORA M K, VARSHNEY P K. Mutual information-based image registration for remote sensing data[J]. *International Journal of Remote Sensing*, 2003, 24(18): 3701-3706.
- [20] YU K, MA J, HU F Y, et al. A grayscale weight with window algorithm for infrared and visible image registration[J]. *Infrared Physics & Technology*, 2019, 99: 178-186.
- [21] KONG Q Q, WU F C, FAN B. Deep learning-based image matching: methods, applications, and challenges[J]. *Journal of Computer Science*, 2024, 47(7): 1485-1520.

- 孔庆群, 吴福朝, 樊彬. 基于深度学习的图像匹配: 方法、应用与挑战[J]. 计算机学报, 2024, 47(7): 1485-1520.
- [22] YIN C Y, ZHI H H, LI H B. A survey of deep learning-based stereo matching methods[J]. Computer Engineering, 2022, 48(10): 1-12.
- 尹晨阳, 职恒辉, 李慧斌. 基于深度学习的双目立体匹配方法综述[J]. 计算机工程, 2022, 48(10): 1-12.
- [23] ZHANG J, JIANG W S. LiDAR point cloud and optical image registration: status and trends[J]. Journal of Geo-Information Science, 2017, 19(4): 528-539.
- 张靖, 江万寿. 激光点云与光学影像配准: 现状与趋势[J]. 地球信息科学学报, 2017, 19(4): 528-539.
- [24] ZHANG Y J, YAO Y X, WAN Y, et al. Histogram of the orientation of the weighted phase descriptor for multi-modal remote sensing image matching[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2023, 196: 1-15.
- [25] MORAVEC H P. Obstacle avoidance and navigation in the real world by a seeing robot rover[D]. Pittsburgh: Carnegie Mellon University, 1980.
- [26] HARRIS C G, STEPHENS M. A combined corner and edge detector[C]//Alvey Vision Conference, 1988: 147-152.
- [27] ZHANG X H, LI B, YANG D. A new Harris multi-scale corner detection method[J]. Journal of Electronics & Information Technology, 2007(7): 1735-1738.
- 张小洪, 李博, 杨丹. 一种新的Harris多尺度角点检测[J]. 电子与信息学报, 2007(7): 1735-1738.
- [28] HE H Q, HUANG S X. Improved fast Harris sub-pixel corner detection[J]. Journal of Image and Graphics, 2012, 17(7): 853-857.
- 何海清, 黄声享. 改进的Harris亚像素角点快速定位[J]. 中国图象图形学报, 2012, 17(7): 853-857.
- [29] ZHENG Y, ZHENG P. Image matching based on Harris-Affine detectors and translation parameter estimation by phase correlation[C]//2019 IEEE 4th International Conference on Signal and Image Processing (IC-SIP), 2019: 106-111.
- [30] YE Y X, WANG M M, HAO S Y, et al. A novel keypoint detector combining corners and blobs for remote sensing image registration[J]. IEEE Geoscience and Remote Sensing Letters, 2021, 18(3): 451-455.
- [31] SMITH S M, BRADY J M. SUSAN—a new approach to low level image processing[J]. International Journal of Computer Vision, 1997, 23: 45-78.
- [32] HE W, DENG X L. A modified SUSAN corner detection algorithm based on adaptive gradient threshold for remote sensing image[C]//2010 International Conference on Optoelectronics and Image Processing, 2010: 40-43.
- [33] ROSTEN E. Machine learning for high-speed corner detection[C]//European Conference on Computer Vision, 2006: 430-443.
- [34] MAIR E, HAGER G D, BURSCHKA D, et al. Adaptive and generic corner detection based on the accelerated segment test[C]//Heidelberg: European Conference on Computer Vision, 2010.
- [35] BAY H, ESS A, TUYTELAARS T, et al. Speeded-up robust features (SURF)[J]. Computer Vision and Image Understanding, 2008, 110(3): 346-359.
- [36] YAO Y X, ZHANG Y J, WAN Y, et al. Image matching of heterogeneous images considering anisotropic weighted moments and absolute phase direction[J]. Wuhan University Journal of Information Science, 2021, 46(11): 1727-1736.
- 姚永祥, 张永军, 万一, 等. 顾及各向异性加权力矩与绝对相位方向的异源影像匹配[J]. 武汉大学学报(信息科学版), 2021, 46(11): 1727-1736.
- [37] ZHU B, YANG C, DAI J K, et al. R₂FD₂: fast and robust matching of multimodal remote sensing Images via repeatable feature detector and rotation-invariant feature descriptor[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 1-15.
- [38] ROBERTS L. Machine perception of three-dimensional solids[D]. Cambridge: Massachusetts Institute of Technology, 1963.
- [39] PREWITT J M S. Picture processing and psychopictorics[M]. New York: Academic Press, 1970.
- [40] SOBEL I, FELDMAN G. An isotropic 3×3 image gradient operator[C]//California: Stanford Conference on Computer Graphics and Image Processing, 1990.
- [41] CANNY J. A computational approach to edge detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, PAMI-8(6): 679-698.
- [42] ETEMADI A. Robust segmentation of edge data[C]//International Conference on Image Processing and its Applications, 1992: 311-314.
- [43] GIOI R G V, JAKUBOWICZ J, MOREL J M, et al. LSD: a fast line segment detector with a false detection control[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(4): 722-732.
- [44] AKINLAR C, TOPAL C. EDLines: a real-time line segment detector with a false detection control[J]. Pattern Recognition Letters, 2011, 32(13): 1633-1642.
- [45] GUO Q, HE M, LI A. High-resolution remote-sensing image registration based on angle matching of edge point features[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11(8): 2881-2895.
- [46] PARK S H, PARK K K, JUNG J H, et al. ISAR imaging of multiple targets using edge detection and Hough transform[J]. Journal of Electromagnetic Waves and Applications, 2008, 22(2-3): 365-373.
- [47] WANG X H, WU L S, CHEN H W, et al. Point cloud feature line extraction based on regional clustering segmentation[J]. Acta Optica Sinica, 2018, 38(11): 66-75.
- 王晓辉, 吴禄慎, 陈华伟, 等. 基于区域聚类分割的点云特征线提取[J]. 光学学报, 2018, 38(11): 66-75.
- [48] XU Y Y, SHEN H B. Research status of biomedical image processing based on pattern recognition[J]. Journal of Electronics and Information Technology, 2020, 42(1): 201-213.
- 徐莹莹, 沈红斌. 基于模式识别的生物医学图像处理研究现状[J]. 电子与信息学报, 2020, 42(1): 201-213.
- [49] ZAITOUN N M, AQEL M J. Survey on image segmentation techniques[J]. Procedia Computer Science, 2015, 65: 797-806.
- [50] PREETHA M M S J, SURESH L P, BOSCO M J. Image segmentation using seeded region growing[C]//2012 International Conference on Computing, Electronics and Electrical Technologies (ICCEET), 2012: 576-583.
- [51] ZHAO J, AN D X. Improved Markov random field for building segmentation in SAR Images[J]. Remote Sensing Technology and Application, 2024, 39(2): 405-412.
- 赵佳, 安道祥. 改进马尔科夫随机场的SAR图像建筑物分割[J]. 遥感技术与应用, 2024, 39(2): 405-412.
- [52] LIU L, SHI Z G, SU H R, et al. Image segmentation based on higher-order Markov random fields[J]. Journal of Computer Research and Development, 2013, 50(9): 1933-1942.
- 刘磊, 石志国, 宿浩茹, 等. 基于高阶马尔科夫随机场的图像分割[J].

- 计算机研究与发展, 2013, 50(9): 1933-1942.
- [53] MINAE S, BOYKOV Y, PORIKLI F, et al. Image segmentation using deep learning: a survey[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(7): 3523-3542.
- [54] GAREA S A, DAS S. Image segmentation methods: overview, challenges, and future directions[C]//2024 Seventh International Women in Data Science Conference at Prince Sultan University (WiDS PSU), 2024: 56-61.
- [55] ZHANG Y J, FAN J X, DUAN L F. An image registration algorithm for optical and SAR images based on regional features[J]. *Journal of Hefei University (Natural Science)*, 2008(4): 37-40.
张雅吉, 范晋湘, 段连飞. 基于区域特征的光学图像与 SAR 图像配准算法[J]. *合肥学院学报(自然科学版)*, 2008(4): 37-40.
- [56] LI Y Q, PI Y M, WANG J F. Image registration of SAR and optical images based on level set[J]. *Acta Geodaetica et Cartographica Sinica*, 2010, 39(3): 276-282.
李雨谦, 皮亦鸣, 王金峰. 基于水平集的 SAR 图像与光学图像的配准[J]. *测绘学报*, 2010, 39(3): 276-282.
- [57] MA W, WEN Z L, WU Y, et al. Remote sensing image registration with modified SIFT and enhanced feature matching[J]. *IEEE Geoscience and Remote Sensing Letters*, 2017, 14(1): 3-7.
- [58] XIANG Y M, WANG F, YOU H J. OS-SIFT: a robust SIFT-like algorithm for high-resolution optical-to-SAR image registration in suburban areas[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(6): 3078-3090.
- [59] ZHANG Y S, ZOU Z R. Automatic registration method for remote sensing images based on improved ORB algorithm[J]. *Remote Sensing for Land & Resources*, 2013, 25(3): 20-24.
张云生, 邹峥嵘. 基于改进 ORB 算法的遥感图像自动配准方法[J]. *国土资源遥感*, 2013, 25(3): 20-24.
- [60] LI J Y, HU Q W, AI M Y. RIFT: Multi-modal image matching based on radiation-variation insensitive feature transform[J]. *IEEE Transactions on Image Processing*, 2020, 29: 3296-3310.
- [61] MA D, LAI H C. Remote sensing image matching based on improved ORB in NSCT domain[J]. *Journal of the Indian Society of Remote Sensing*, 2019, 47(4): 801-807.
- [62] YE Y X, BRUZZONE L, SHAN J, et al. Fast and robust matching for multimodal remote sensing image registration[J]. *IEEE Geoscience and Remote Sensing*, 2019, 57(11): 9059-9070.
- [63] HUFFMAN D A. A method for the construction of minimum-redundancy codes[J]. *Proceedings of the IRE*, 1952, 40(9): 1098-1101.
- [64] SHRIWAS R N, BODKHE Y, MANE A, et al. Overview of Canny edge detection and Hough transform for lane detection[C]//2024 OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 4.0, 2024: 1-5.
- [65] ZHANG L L, KOCH R. An efficient and robust line segment matching approach based on LBD descriptor and pairwise geometric consistency[J]. *Journal of Visual Communication and Image Representation*, 2013, 24(7): 794-805.
- [66] HU M K. Visual pattern recognition by moment invariants[J]. *IRE Transactions on Information Theory*, 1962, 8(2): 179-187.
- [67] WANG Q, HUANG W, ZHANG X T, et al. GLCM: global-local captioning model for remote sensing image captioning[J]. *IEEE Transactions on Cybernetics*, 2023, 53(11): 6910-6922.
- [68] HONG Y M, LENG C C, ZHANG X Y, et al. HOLBP: remote sensing image registration based on histogram of oriented local binary pattern descriptor[J]. *Remote Sensing*, 2021, 13: 23-28.
- [69] LI P, JIANG G, MA Q L, et al. Hybrid model-based multimodal image registration combining tensor and mutual information[J]. *Journal of Geodesy*, 2021, 50(7): 916-929.
李培, 姜刚, 马千里, 等. 结合张量与互信息的混合模型多模态图像配准方法[J]. *测绘学报*, 2021, 50(7): 916-929.
- [70] FISCHLER M A, BOLLES R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography[J]. *Communications of the ACM*, 1981, 24(6): 381-395.
- [71] WU Y, MA W P, GONG M G, et al. A novel point-matching algorithm based on fast sample consensus for image registration[J]. *IEEE Geoscience and Remote Sensing Letters*, 2015, 12(1): 43-47.
- [72] WANG G, CHEN Y F. Robust feature matching using guided local outlier factor[J]. *Pattern Recognition*, 2021, 117: 107986.
- [73] MENG F Y, LI X, PEI J H. A feature point matching based on spatial order constraints bilateral-neighbor vote[J]. *IEEE Transactions on Image Processing*, 2015, 24(11): 4160-4171.
- [74] MA J Y, MA Y, ZHAO J, et al. Image feature matching via progressive vector field consensus[J]. *IEEE Signal Processing Letters*, 2015, 22(6): 767-771.
- [75] FENG S D, GAO M X, JIN X W, et al. Fine-grained damage detection of cement concrete pavement based on UAV remote sensing image segmentation and stitching[J]. *Measurement*, 2024, 226: 113844.
- [76] DUSMANU M H, ROCCO I, PAJDLA T, et al. D2-Net: a trainable CNN for joint description and detection of local features[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 8084-8093.
- [77] MA W P, ZHANG J, WU Y, et al. A novel two-step registration method for remote sensing images based on deep and local features[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 57(7): 4834-4843.
- [78] NAN K, QI H, YE Y X. A multimodal remote sensing image template matching method based on deep convolution feature representation[J]. *Acta Geodaetica et Cartographica Sinica*, 2019, 48(6): 727-736.
南轲, 齐华, 叶沅鑫. 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.
- [79] CAO S Y, YU B, LUO L, et al. PCNet: a structure similarity enhancement method for multispectral and multimodal image registration[J]. *Information Fusion*, 2023, 94: 200-214.
- [80] YI K M, TRULLS E, LEPETIT V, et al. LIFT: learned invariant feature transform [C]//European Conference on Computer Vision, 2016: 467-483.
- [81] ZHANG H, NI W P, YAN W D, et al. Registration of multimodal remote sensing image based on deep fully convolutional neural network[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2019, 12(8): 3028-3042.
- [82] LI Z Y, ZHANG H T, HUANG Y H. A rotation-invariant optical and SAR image registration algorithm based on deep and gaussian features[J]. *Remote Sensing*, 2021, 13: 2628.
- [83] SARLIN P E, DETONE D, MALISIEWICZ T, et al. SuperGlue: learning feature matching with graph neural networks [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 4937-4946.
- [84] LINDENBERGER P, SARLIN P E, POLLEFEYS M. LightGlue: local

- feature matching at light speed[C]//2023 IEEE/CVF International Conference on Computer Vision (ICCV), 2023: 17581–17592.
- [85] ZHANG W, LI T T, ZHANG Y T, et al. LTFormer: a light-weight transformer-based self-supervised matching network for heterogeneous remote sensing images[J]. Information Fusion, 2024, 109: 102425.
- [86] QUAN D, WEI H Y, WANG S, et al. A novel coarse-to-fine deep learning registration framework for multimodal remote sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 1–16.
- [87] LIANG C B, DONG Y Y, ZHAO C J, et al. A coarse-to-fine feature match network using Transformers for remote sensing image registration[J]. Remote Sensing, 2023, 15(13): 3243.
- [88] YU P L, SHI Q, WANG H. Parallel generative network for infrared-to-visible image translation[J]. Journal of Image and Graphics, 2021, 26(10): 2346–2356.
余佩伦, 施伦, 王吟. 并行生成网络的红外—可见光图像转换[J]. 中国图象图形学报, 2021, 26(10): 2346–2356.
- [89] YIN M X, LIN Z F, YANG F. Image translation based on dynamic receptive field for adaptive multi-scale information fusion[J]. Journal of Electronics and Information Technology, 2021, 43(8): 2386–2394.
尹梦晓, 林振峰, 杨锋. 基于动态感受野的自适应多尺度信息融合的图像转换[J]. 电子与信息学报, 2021, 43(8): 2386–2394.
- [90] MA J Y, JIANG X Y, JIANG J J, et al. LMR: learning a two-class classifier for mismatch removal[J]. IEEE Transactions on Image Processing, 2019, 28(8): 4045–4059.
- [91] SUN J M, SHEN Z H, WANG Y, et al. LoFTR: detector-free local feature matching with transformers [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021: 8918–8927.
- [92] ZHANG H, YUE Y X, LI H J, et al. Shared contents alignment across multiple granularities for robust SAR-optical image matching[J]. Information Fusion, 2024, 106: 102298.
- [93] LI L Z, HAN L, DING M T, et al. Multimodal image fusion framework for end-to-end remote sensing image registration[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 5607214.
- [94] SHEN X L, WANG C, LI X, et al. RF-Net: an end-to-end image matching network based on receptive field [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 8124–8132.
- [95] CHEN P, BAO B Y, CHEN X. Remote sensing image-matching network based on multiscale feature fusion and importance ranking loss[J]. Laser & Optoelectronics Progress, 2024, 61(18): 276–284.
陈鹏, 包倍源, 陈旭. 基于多尺度特征融合与重要性排序损失的遥感图像匹配网络[J]. 激光与光电子学进展, 2024, 61(18): 276–284.
- [96] SUN X K, YUN Z K, HU C B, et al. End-to-end registration algorithm for high-resolution multi-view SAR images [J]. Journal of Radars, 2025, 14(2): 389–404.
孙晓坤, 贡泽楷, 胡槲彬, 等. 面向高分辨率多视角 SAR 图像的端到端配准算法[J]. 雷达学报, 2025, 14(2): 389–404.
- [97] SHI T X, CAO F Z, HAN K Y, et al. A multimodal remote sensing image matching network based on cross-attention mechanism [J]. Intelligent Computer and Applications, 2023, 13(8): 107–113.
石添鑫, 曹帆之, 韩开杨, 等. 一种基于交叉注意力机制多模态遥感图像匹配网络[J]. 智能计算机与应用, 2023, 13(8): 107–113.
- [98] LI J P, CHEN Y, WANG D Z. Simulation of remote sensing image registration for fusing attention and networks in network [J]. Computer Simulation, 2022, 39(7): 42–47.
李绩鹏, 陈颖, 王东振. 融合双注意力与深度神经网络的遥感图像配准[J]. 计算机仿真, 2022, 39(7): 42–47.
- [99] ZOU R P, ZHU B, WANG C Y, et al. Heterogeneous remote sensing image matching algorithm based on residual pseudo-Siamese convolution cross-correlation network [J]. Laser & Optoelectronics Progress, 2022, 59(12): 458–466.
邹融平, 朱斌, 王晨阳, 等. 基于残差伪孪生卷积互相关网络的异源遥感图像匹配算法[J]. 激光与光电子学进展, 2022, 59(12): 458–466.
- [100] CAO F, SHI T, HAN K, et al. RDFM: robust deep feature matching for multimodal remote-sensing images [J]. IEEE Geoscience and Remote Sensing Letters, 2023, 20: 1–5.

(责任编辑:尹晨茹)