

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2024)08-2137-25

论文引用格式: Zhu B and Ye Y X. 2024. Multimodal remote sensing image registration: a survey. Journal of Image and Graphics, 29(08):2137-2161

(朱柏, 叶沅鑫. 2024. 多模态遥感图像配准方法研究综述. 中国图象图形学报, 29(08):2137-2161)[DOI:10.11834/jig.230737]

多模态遥感图像配准方法研究综述

朱柏, 叶沅鑫*

1. 西南交通大学地球科学与工程学院, 成都 611756; 2. 高速铁路安全运营空间信息技术国家地方联合工程实验室, 成都 611756

摘要: 随着对地观测技术的不断发展, 从星载、机载和地面平台上众多的一体化立体观测设施被发射, 这些传感器设备可以动态提供不同空间、时间和光谱分辨率的多模态遥感图像, 只有充分利用各类多模态遥感图像才能有效地为自然资源管理、防灾减灾和环境监测等不同应用提供更可靠和全面的对地观测结果。但是由于不同传感器之间的成像机理不同, 多模态图像之间呈现显著的辐射差异、几何差异、时相差异和视角差异等, 给多模态遥感图像高精度的配准带来了巨大的挑战。为推进多模态遥感图像配准研究技术的发展, 本文对当前主流的多模态遥感图像配准方法系统性地进行了梳理、分析、介绍和总结。首先梳理了单模态到多模态遥感图像配准的研究发展演化过程; 然后分别分析了基于区域、基于特征和基于深度学习方法中代表性算法的核心思想, 并给出已开源代码的链接; 同时分享了现有公开的多模态遥感图像配准数据集, 介绍了数据集的详细内容 and 特点; 最后给出了现阶段多模态遥感图像高精度配准研究中所存在的一些问题和严峻挑战, 并对未来研究的发展趋势进行了前瞻性的展望, 旨在推动多模态遥感图像配准领域实现更加深入的突破和创新。

关键词: 遥感; 传感器; 多模态图像; 图像配准; 配准数据集

Multimodal remote sensing image registration: a survey

Zhu Bai, Ye Yuanxin*

1. Faculty of Geosciences and Engineering, Southwest Jiaotong University, Chengdu 611756, China;

2. State-Province Joint Engineering Laboratory of Spatial Information Technology for High-Speed Railway Safety, Chengdu 611756, China

Abstract: The advent of new infrastructure construction and the era of intelligent photogrammetry have facilitated the rapid development of global aerospace and aviation remote sensing technology. Numerous multi-sensors integrating stereoscopic observation facilities have been launched from spaceborne, airborne, and terrestrial platforms, and the types of sensors have also developed from traditional single-mode sensors (e. g., optical sensors) to a new generation of multimodal sensors (e. g., multispectral, hyperspectral, light detection and ranging (LiDAR), and synthetic aperture radar (SAR) sensors). These advanced sensor devices can dynamically provide multimodal remote sensing images with different spatial, temporal, and spectral resolutions. They can obtain more reliable, comprehensive, and accurate observation results than single-modal sensors through joint processing of spaceborne, airborne, and terrestrial multimodal data. Therefore, investigating multimodal remote sensing image registration has great scientific significance. Multi-level and multi-perspective Earth observation can be effectively achieved only by fully integrating and utilizing various multimodal remote sensing images. In order to promote the development of multimodal remote sensing image registration research technology, we systematically

收稿日期: 2023-10-17; 修回日期: 2024-03-19; 预印本日期: 2024-03-26

* 通信作者: 叶沅鑫 yeyuanxin@home.swjtu.edu.cn

基金项目: 国家自然科学基金项目 (42271446, 41971281)

Supported by: National Natural Science Foundation of China (42271446, 41971281)

sort out, analyze, introduce, and summarize the current mainstream registration methods for multimodal remote sensing images. We first sort out the research development and evolution process from single-modal to multimodal remote sensing image registration. We then analyze the core ideas of representative algorithms among area-based, feature-based, and deep-learning-based pipelines, while the contribution of the author team in the field of multimodal remote sensing image registration is introduced. Area-based registration (template matching) pipeline mainly includes two types: information theory-based and structural feature-based registration methods. The structural feature-based method consists of sparse structural features and dense structural features. From the perspective of the robustness and efficiency of comprehensive registration, dense-structure-feature-based methods have obvious effectiveness and advantages in handling significant non-linear radiation differences between multimodal remote sensing images and can meet many current application needs. By contrast, area-based registration pipeline generally relies on geo-referencing of remote sensing images to predict the rough range of template matching. Feature-based registration methods can be refined into three categories: feature registration based on gradient optimization, local self-similarity (LSS), and phase consistency. The feature registration of gradient optimization usually designs consistent gradients for specific multimodal images. The generalization of this type of method based on gradient optimization is generally poor, and it has difficulty maintaining the same performance on other types of multimodal images. The feature registration of LSS also has limitations, given that the relatively low discriminative power of LSS descriptors may result in the inability to maintain robust matching performance in the presence of complex nonlinear radiation differences. The feature registration of phase consistency has high computational complexity, and the registration process is generally time consuming. Feature-based registration pipeline utilizes the local spatial relationship between adjacent pixels to construct a high-dimensional information feature vector for each feature point. Compared with template matching methods, they usually face a heavy computational burden, and inevitable serious outliers are prone to occur in matching, especially in multimodal registration situations where scale, rotation, and radiation differences exist simultaneously. In general, the registration robustness of feature-based methods is not as stable as that of area-based methods. The deep-learning-based pipeline can be divided into modular and end-to-end registration methods. The most common strategy for modular registration methods is to embed deep networks into feature-based or region-based methods. This approach takes advantage of the complete data-driven and high-dimensional deep feature extraction capability of deep learning to generate more robust features or more effective descriptors or similarity measures, which improves the robustness of image registration. Modular registration methods can be subdivided into three categories: learning-based template matching, learning-based feature matching, and style transfer-based modal unification. Modular registration methods are easy to train and have strong flexibility, but it has difficulty avoiding the error accumulation problem that easily occurs in multi-stage tasks and may fall into local optimality. The end-to-end registration methods directly estimate the geometric transformation parameters or deformation field to achieve image registration by directly constructing an end-to-end neural network structure. The training objectives of the end-to-end network are consistent and can obtain the global optimal solution. However, some problems arise, such as high training difficulty and poor interpretability. Moreover, no complete and comprehensive database containing all types of multimodal remote sensing image pairs is available to date, and the lack of training and testing data greatly limits the development of deep learning-based registration methods. Furthermore, we share existing public registration datasets of multimodal remote sensing images, as well as supplement by a small number of registration datasets in the field of computer vision. Finally, the existing problems and challenges in the current research on high-precision registration of multimodal remote sensing images are analyzed. A forward-looking outlook on the development trend of future research is given, which aims to promote further breakthroughs and innovations in the field of multimodal remote sensing image registration.

Key words: remote sensing; sensors; multimodal images; image registration; registration datasets

0 引言

当前全球数字化发展日益加快,习近平总书记

多次就推动新型基础设施发展做出部署,强调要加强新型基础设施建设(李德仁等,2022)。随着众多遥感卫星和遥感设施的发射,天空地多传感器一体化立体观测技术得到了全面的发展,标志着新基建

和智能摄影测量时代的到来(邵振峰等,2023;孙伟伟等,2020;张永军等,2021),多模态遥感图像数据集的可用性出现了爆炸式增长(Li等,2022c)。这些多模态遥感图像通常是由不同传感器、不同成像机制在不同时期对同一地物观测获取的,如光学—红外图像(李云红等,2022)、光学—高光谱图像(苏红军,2022)、光学—SAR(synthetic aperture radar)图像(Kulkarni和Rege,2020)、光学—LiDAR(light detection and ranging)图像(Zhu等,2021)等,如图1所示。

通过多模态遥感数据联合处理可以得到比单模态传感器更可靠、更全面和精确的综合观测结果(陈锐志等,2019),因此开展多模态遥感图像配准研究具有重要的科学意义。同时多模态遥感图像配准是遥感领域大量协同处理和集成应用的前提和关键步骤,如图像融合(唐霖峰等,2023)、变化检测(林志康等,2023)和目标识别(姚红革等,2020)等,只有充分集成和利用各类多模态遥感图像才能有效地实现多层次和多方位的对地观测(Tang等,2023)。

由于不同传感器(红外、LiDAR和SAR传感器等)之间的空间几何构象和物理辐射机理不同(Ye等,2022a),以及不同传感器所搭载的成像平台(卫星、航空和地面平台)不同,多模态遥感图像之间往往存在显著“五大差异”(辐射差异、几何差异、尺度差异、视角差异和时相差异)中的一种或几种。

1)辐射差异。由于传感器间的空间几何构象和物理辐射机理不同,导致地物在不同成像机理和不同的波谱范围下的辐射特性有所不同,遥感图像上

会呈现出不同的灰度信息。

2)几何差异(几何畸变)。不同成像的方式会导致地物目标出现不同形式的几何变形,同时图像容易受传感器侧视角度、地表起伏等因素的影响,造成图像局部范围内存在不可预知的几何畸变。

3)尺度差异。由于传感器平台高度以及传感器拍摄的瞬时视场角的不同,所获取的图像往往具有不同的空间分辨率,反映出观测地物不同的细节信息。

4)视角差异。不同成像平台(卫星、航空和地面)的传感器由于拍摄角度、拍摄高度和视野范围的不同,导致同一场景的图像内容会因视角的不同而产生显著变化。

5)时相差异。在对地观测中,需要卫星传感器在不同的时间对地表的同一区域进行监测,而随着时间的改变,地表覆盖信息也会随之发生变化,导致多时相的图像间呈现出不一致的内容。

同时,不同成像平台传感器的图像质量容易受到天气、气候和环境的影响(冯蕊涛等,2021),噪声、云层以及模糊等干扰都是不可避免的问题,给多模态遥感图像高精度的匹配和配准带来了巨大的挑战。图2举例说明了上述多模态遥感图像配准所面临的挑战。

多模态遥感图像配准的核心是尽可能识别两幅或多幅遥感图像之间均匀分布的同名点对,目的是获得最优的几何变换模型(黄磊等,2023;Xiang等,2023)。假设参考图像和待配准图像(输入图像)分

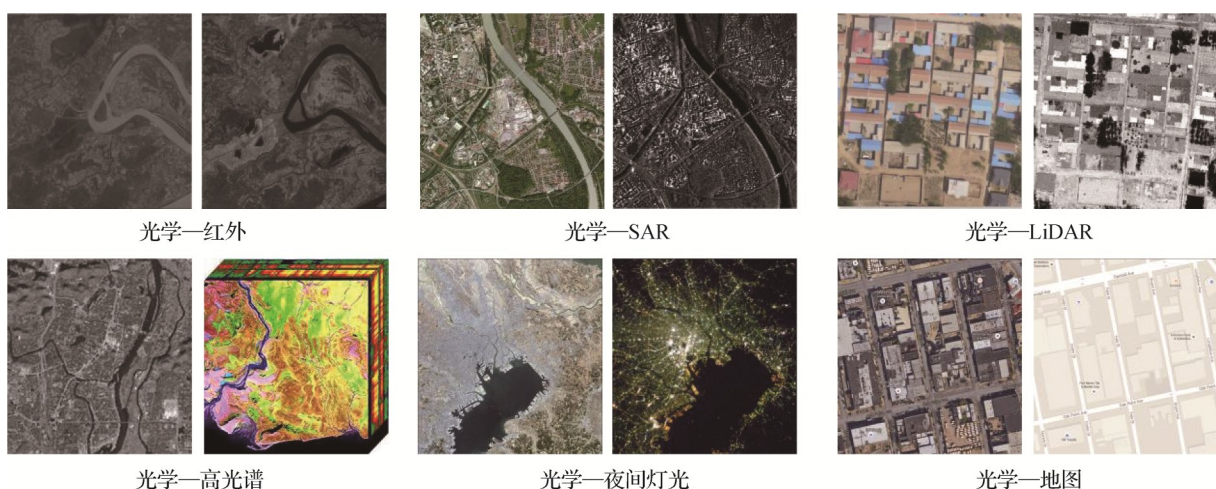


图1 多模态遥感图像

Fig. 1 Multimodal remote sensing images

别表示为 $I_1(x, y)$ 和 $I_2(x, y)$, 则两幅图像间对应像素点的几何映射关系可以定义为 $I_2(x, y) = g[I_1(f(x, y))]$, f 是二维空间几何变换函数, g 是一维灰度变换函数(Brown, 1992)。

传统的配准方法仅适用于具有线性辐射(强度)差异的单模态图像, 对于几何畸变和辐射差异显著的多模态遥感图像而言, 难以有效构建反映多模态图像间共有特征的特征检测和特征描述模型, 难以构建能有效克服非线性辐射差异的相似性度量。虽然国际上许多通用的主流遥感商业软件(如 ENVI、ERDAS 和 PCI)提供了基本的图像配准功能模块, 但对多模态图像的配准仍然采用传统的配准方法, 也

难以实现多模态遥感图像高精度的配准。近年来, 国内外众多研究机构(武汉大学、同济大学、西南交通大学、长安大学、西安电子科技大学、德国慕尼黑工业大学和旷视研究院等)的相关学者针对多模态遥感图像配准中的难点提出了许多高精度、高效率和高鲁棒性的多模态配准方法。

本文根据配准流程的不同, 全面系统地阐述当前主流多模态遥感图像配准方法的研究进展, 以方便研究人员能够更好地了解多模态遥感图像配准领域的研究现状。本文总体的结构框架如图 3 所示, 首先梳理不同类型配准方法(基于区域、基于特征和基于深度学习的方法)从单模态到多模态遥感图像

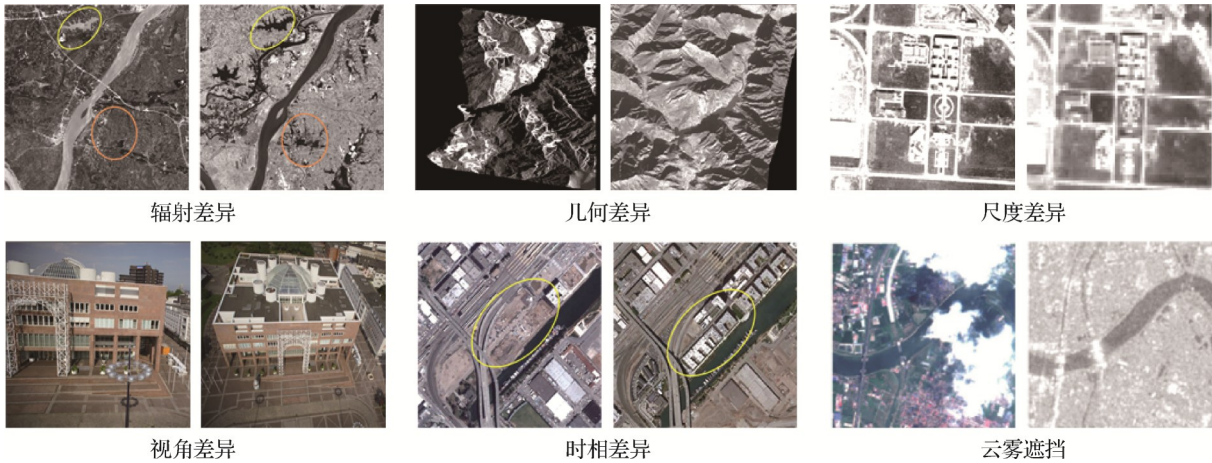


图2 多模态遥感图像配准面临的挑战

Fig. 2 The challenge of multimodal remote sensing images registration

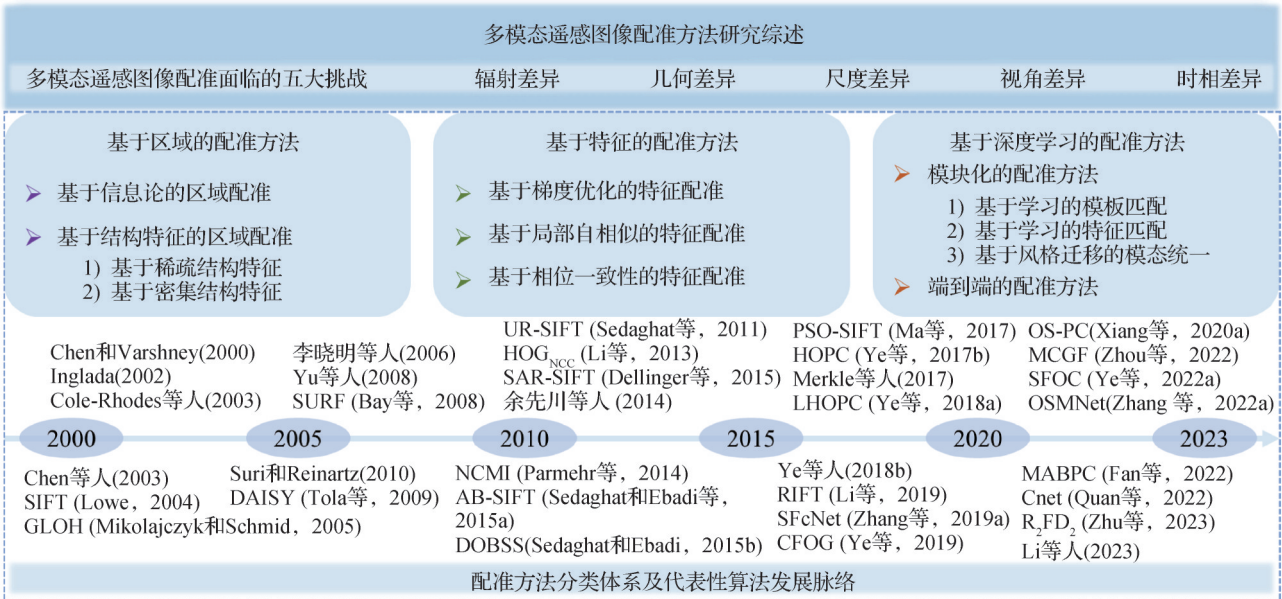


图3 本文结构框架

Fig. 3 Overall framework of this paper

配准的研究发展演化过程,并讨论上述3种不同范畴方法中代表性配准算法的特点和优势,同时给出其公开代码的获取渠道。然后简要介绍作者团队在多模态遥感图像配准领域的贡献,分享现有遥感领域公开的多模态图像配准数据集,并对数据集的特点进行分析。最后对全文进行总结,给出现阶段多模态遥感图像高精度配准研究中所存在的问题和挑战,并对未来研究的发展趋势进行了展望。

1 多模态遥感图像配准研究现状

根据遥感图像配准流程不同,传统的配准方法分为两大类,基于区域和基于特征的方法(叶沅鑫等,2014)。考虑到深度学习对深层特征有着强大的学习和表达能力,近年来发展了许多通过神经网络模型提取高维特征进行图像配准的方法(Jiang等,2021),基于深度学习的方法已经逐渐发展成多模态遥感图像配准领域不容忽视的方法。下面,本文将详细而全面地回顾上述3种不同类型的遥感图像配准方法。

1.1 基于区域的配准方法

基于区域配准方法的关键是选择合适的相似性度量,利用相似性度量和优化方法来准确估计几何变换参数,从而通过模板匹配的策略驱动整个配准过程的优化。因此,该种类型的配准方法也称为基于模板的配准方法。遥感图像与普通图像相比,最显著的特点是可以先通过先进的物理设备和导航设备获取遥感图像高精度的POS(position and orientation system)或RPC(rational polynomial coefficients)数据等先验信息(Shen等,2017;吴军等,2016)。通过利用这些先验信息得到的地理参考可以有效消除多模态图像间明显的尺度和旋转差异,为基于区域的方法预测模板匹配的粗略范围,因为遥感图像的地理参考信息在图像空间中的偏移量通常在几个像素到几十个像素之间(Jiang等,2015)。

基于区域的遥感配准方法根据图像表示域的不同分为基于空间域和频率域的方法。在空间域中,差平方和(sum of squared differences, SSD)、归一化互相关(normalized cross correlation, NCC)和互信息(mutual information, MI)是最具代表性的相似性度量(Chen等,2003;陆和平和高磊,2009)。在频域最广泛使用的相似性度量则是相位相关(Tong等,

2015),其可以利用傅里叶平移、旋转和缩放的性质在一定程度上解决图像之间的平移、缩放和旋转变化(Reddy和Chatterji,1996)。

但是上述传统的基于区域的配准方法通常通过对图像强度信息的相似度进行评估,寻找最优的几何变换参数,使所选的相似性度量达到最大或最小,从而驱动配准过程。因此,SSD、NCC和相位相关不能为多模态遥感图像提供鲁棒的配准结果,因为这些相似性度量对多模态遥感图像中普遍存在的显著非线性辐射差异的敏感性较强(樊建伟等,2023)。

在早期的基于区域的方法中,基于信息理论的相似性度量是多模态遥感图像配准首先广泛使用的方法,其中最具代表性的是互信息,它受到医学领域相关研究的启发(Wells III等,1996)。随后,相关学者陆续研究了一些新颖的解决方案,提升基于区域配准算法的准确性和效率,本文将其命名为基于结构特征的方法。表1对基于区域的多模态遥感图像配准代表性方法进行了综合分析,表中列出了每种方法的核心思想及已开源代码的链接。表中带*号的属于本文作者团队在多模态遥感图像配准领域的贡献。下面将对上述两类基于区域的配准方法进行详细回顾。

1.1.1 基于信息论的区域配准方法

基于信息论的方法因为其统计相关性在一定程度上可以抵抗非线性辐射差异而被应用于多模态遥感图像配准,展现了比其他传统指标(SSD、NCC和相位相关)更稳健的匹配性能。因此,在此基础上设计改进的信息论方法已被用于多种类型的多模态遥感图像配准,这些方法通常包括最常见的MI方法(Chen和Varshney,2000;Chen等,2003)及其改进算法(Cole-Rhodes等,2003;Suri和Reinartz,2010;Xu等,2016)。

Chen和Varshney(2000)提出了一种基于小波变换和MI结合的金字塔配准方法,他们的实验结果率先展示了MI在多模态遥感图像配准中的潜力。Inglada(2002)利用同一场景不同时期的两幅SPOT多光谱图像,进一步评估和比较了4种不同的相似性度量并得出结论,MI的综合配准性能是最好的,尽管图像间的辐射信息不相似。Chen等人(2003)首先研究了各种插值算法来估计MI的联合直方图,实验证明最近邻插值的性能是最有效的,并且他们提出了一种新的联合直方图估计算法,称为广义局

表 1 基于区域的多模态遥感图像配准算法归纳总结

Table 1 Overview and summary of region-based multimodal remote sensing image registration methods			
方法类型	代表算法	核心思想/结论/获取渠道	
基于信息论	Chen 和 Varshney(2000)	结合图像金字塔和 MI,对非线性辐射差异有一定的抵抗性	
	Inglada(2002)	评估和对比 MI 与其他相似性度量,证明了 MI 在传统相似性度量中综合配准性能是好的	
	Chen 等人(2003)	提出 GPVE 优化联合直方图计算,解决 MI 计算中某些条件下插值引起的伪像问题	
	Cole-Rhodes 等人(2003)	提出 SPSA 梯度策略优化 MI,加快 MI 迭代计算收敛的速度	
	Suri 和 Reinartz(2010)	基于 MI 的高分辨率图像配准的合适特征选择,提高算法的配准性能	
	NCMI (Parmehr 等,2014)	采用基于直方图的非参数方法进行联合概率密度函数估计,有效提高算法的配准性能	
	Xu 等人(2016)	使用 Kullback-Leiber 散度的对称形式作为相似性度量,提高在没有足够场景重叠区域情况下算法的性能	
基于结构特征	TOMI(李培 等,2021)	结合张量方向平行度与梯度互信息,对非线性辐射变化和噪声具有更强的稳健性	
	HOG _{NCC} (Li 等,2013)	结合 HOG 和 NCC,成功配准光学和 SAR 图像	
	稀疏结构特征	DLSS(Ye 等,2017a)*	提出基于形状属性的密集局部自相似描述; 公开代码链接: https://www.dropbox.com/s/o6f7j0igdiadlvs/DLSC%20publish%20code.rar?dl=0
		HOPC(Ye 等,2017b)*	提出相位一致性方向直方图 HOPC,已应用于全球范围的卫星图像处理; 公开代码链接: https://github.com/yeyuanxin110/HOPC
	HOMPC(Fu 等,2018)	利用 log-Gabor 滤波结合局部区域幅值和相位一致性信息,实现更鲁棒的配准	
	RLSS(Xiong 等,2019)	结合 Spearman 秩相关系数和 LSS,更好地描述图像的局部形状属性,实现更精确的配准	
	密集结构特征	CFOG(Ye 等,2019)*	提出逐像素结构特征匹配的框架(HOG, LSS, CFOG) ;公开代码链接: https://github.com/yeyuanxin110/CFOG
OS-PC(Xiang 等,2020a)		设计光学和 SAR 图像的不同多通道特征表示,实现快速鲁棒的配准	
AWOG(Fan 等,2021)		利用角度加权策略,将 CFOG 的梯度分布到两个最相关的方向上	
DCF-EPW(Xiang 等,2022)		提出基于扩张卷积特征和极向相位相关的新型模板匹配方法	
SFOC(Ye 等,2022a)*		结合多尺度的一阶和二阶梯度,构建更具判别性的图像结构特征,并开发由粗到精的多模态遥感图像配准系统;公开代码链接: https://github.com/yeyuanxin110/SFOC-Multimodal_Remote_Sensing_Image_Registration_System	
MAGD(Xiang 等,2023)		构建多方向各向异性高斯导数特征,同时贡献一个高精度的配准评估数据集; 公开代码链接: https://github.com/xym2009/OS-Eval	

注:“*”表示本文作者团队的工作。

部体积估计 (generalized partial volume estimation, GPVE),以解决 MI 计算中某些条件下插值引起的伪像问题。

Cole-Rhodes 等人(2003)研究了一种同时扰动随机逼近 (simultaneous perturbation stochastic approximation, SPSA)梯度策略以优化 MI,该策略已成功应用于多时相和多传感器遥感图像配准。Suri 和 Reinartz(2010)提出了一种基于 MI 的由粗到精自

动配准在密集城市地区获取的高分辨率 TerraSAR-X 和 Ikonos 图像。Parmehr 等人(2014)设计了一种新的相似度量,称为归一化组合互信息(normalized combined mutual information, NCMI),成功提高了光学图像与机载激光雷达数据的配准精度。Xu 等人(2016)提出了一种对称形式的 Kullback-Leibler 散度,即 Jeffrey 散度,并与其他基于改进互信息的相似性度量对比验证了其具有优异的配准性能。李

培等人(2021)利用张量方向平行度和梯度互信息构建了 TOMO (tensor orientation and mutual information) 相似度准则, 通过结合图像的梯度分布和结构方向对非线性辐射变化和噪声具有更强的稳健性。

总的来说, 这些基于信息理论的方法具有依赖于统计相似性计算联合直方图的特点, 在一定程度上适用于不同成像机制的多模态配准。但是基于信息论的方法最大的局限性在于所涉及的过程非常复杂和耗时, 这限制了其广泛的应用, 尤其是对于大尺寸多模态图像的配准。

1.1.2 基于结构特征的区域配准方法

根据是否为逐像素的结构特征, 本文将基于结构特征的方法分为基于稀疏结构特征的方法和基于密集结构特征的方法。这些基于结构特征的方法通过使用经典度量 (如 SSD 和 NCC 等) 来评估生成特征 (稀疏结构特征或密集结构特征) 的相似性, 以匹配出多模态遥感图像之间的同名点对。基于稀疏结构特征的方法通常是以相对稀疏的采样网格提取多模态遥感图像之间的几何结构特征 (Li 等, 2013; Ye 等, 2017a; Xiong 等, 2020)。

Li 等人 (2013) 引入梯度方向直方图 (histogram of oriented gradients, HOG) 来表达多模态图像的结构特征, 然后采用零均值 NCC 作为相似性度量来配准 SAR 和光学图像。Ye 等人 (2017a) 提出了基于形状属性的密集局部自相似 (dense local self-similarity, DLSS) 描述符来表示图像内的自相似性, 然后使用 NCC 评估 DLSS 描述符的相似性完成光学和 SAR 图像的高精度匹配。受 HOG 的启发, Ye 等人 (2017b) 通过扩展相位一致性模型来捕获图像的结构属性, 提出相位一致性方向直方图 (histogram of orientated phase congruency, HOPC), 并与 NCC 相结合在各种具有显著非线性辐射差异的光学、LiDAR、SAR 图像上成功验证。HOPC 已被应用于全球范围的卫星图像处理, 能够匹配出具有满意定位精度的同名点 (张永军 等, 2021)。基于上述研究, Fu 等人 (2018) 进一步提出了一种结合局部区域的幅值和相位一致性信息的描述子, 充分利用 log-Gabor 滤波器的多方向相位一致性和幅值信息实现了更鲁棒的配准性能。Xiong 等人 (2020) 利用统计学中的 Spearman 秩相关系数设计了 RLSS (rank-based local self-similarity), 其能够以一种可判别的方式描述图像的局部形状属性, 实现更加鲁棒的配准性能。

然而, 上述这些描述符 (HOG, HOPC 和 DLSS 等) 是由相对稀疏的采样网格生成的, 以捕获图像内部结构或形状特征。因此, 它们都是稀疏表示, 导致一些可用的和重要的结构信息被忽略, 进一步会降低这些描述符配准性能的准确性和鲁棒性。而基于密集结构特征的方法是通过使用逐像素描述符将图像的强度信息映射到高维空间像素级的特征表示来驱动配准过程 (Ye 等, 2019; Fan 等, 2021; Xiang 等, 2023)。

Ye 等人 (2019) 首先构建了逐像素的梯度方向通道特征 (channel features of orientated gradients, CFOG), 提出一种快速鲁棒的逐像素模板匹配框架, 然后利用快速傅里叶变换 (fast Fourier transform, FFT) 技术在频域加速 SSD 相似性度量的匹配。并且该匹配框架可以集成多模态图像其他像素级的特征表示。如 HOG, LSS (local self-similar) 和 SURF (speeded-up robust features)。Xiang 等人 (2020a) 针对相位相关无法有效应用于非线性辐射差异显著且几何差异较大的多模态遥感图像配准的问题, 分别利用 Sobel 算子和指数加权均值比 (ratio of exponential weighted average, ROEWA) 算子设计了光学和 SAR 图像的不同多通道特征表示, 进一步研究了两种鲁棒的解决方案, 提高三维相位相关在空间和频域的匹配性能。

Fan 等人 (2021) 利用角度加权的策略改进了 CFOG, 通过将梯度值分布到两个最相关的方向上, 以提升其匹配性能。Xiang 等人 (2022) 提出了一种基于扩张卷积特征和极向相位相关的新模板匹配方法 DCF-EPW (dilated convolutional feature-epipolar oriented matching window)。DCF 由具有多通道梯度的深度可分离扩张卷积构造, 即使地理参考信息有着较大的相对定位误差也能保持不变性, 因为扩张卷积的感受野很大。极向相位相关则提出沿极线延伸的矩形模板, 捕获比方形模板更多的重叠区域, 从而在光学和 SAR 图像间匹配出更多高精度的同名点。Ye 等人 (2022a) 注意到上述基于密集结构特征的方法忽略了图像的局部像素间关系, 缺乏多尺度信息并且没有利用图像的二阶梯度信息, 而二阶梯度在描述细节信息方面的性能优于一阶梯度。因此, 他们通过结合一阶和二阶梯度构建了 SFOC (steerable filters of first- and second- order channels) 描述符, 以提取图像间更具判别性的结构特征。并

利用FFT和积分图像对传统的NCC进行加速,在此基础上开发了一种由粗到精的多模态遥感图像配准系统,并提供了该配准系统的公开代码。Xiang等人(2023)构建了一种多方向各向异性高斯导数(multi-directional anisotropic Gaussian derivatives, MAGD)特征,并将其嵌入到具有几何不变掩膜的置信感知局部匹配中,能更有效地抵抗几何差异的影响;同时贡献了一个包含数百个多模态图像对的评估数据集。

对于基于结构特征的方法,基于稀疏结构特征的方法仅在设计的稀疏采样网格中进行计算,这意味着它们只能使用空间域的相似性度量来评估相似性识别同名点,而频率域的相似性度量则不适用,例如相位相关。另外基于密集结构特征的方法可以利用空间域和频域的相似性度量来进行匹配。综合考虑配准的鲁棒性和效率,基于密集结构特征的方法在处理多模态遥感图像间显著的非线性辐射差异时表现出明显的有效性和优势,能够满足目前许多应用的需求。

1.2 基于特征的配准方法

上文探讨的基于区域的方法只能应用于多模态图像之间只有几个像素或几十个像素偏移的情况,并且受旋转和尺度变化等的影响很大。虽然地理参考信息是遥感图像的固有特征,可以用来预测基于区域方法大致的匹配区域,但如果图像的地理参考信息缺失或定位误差过大(例如几百或几千像素),则基于区域方法的稳健性很难保证。

与基于区域的方法不同,基于特征的方法通常不依赖图像的地理参考信息作为初始约束条件。这些方法通常先检测图像之间显著的特征(如点特征、线特征和面特征),然后通过描述检测到的特征来识别同名点。因此,特征检测和特征描述是多模态特征匹配方法中最为关键的步骤,本文主要讨论回顾近些年遥感图像特征匹配发展中关于特征检测和特征描述部分的相关研究。

在遥感领域的特征匹配方法中,最常见的方法是提取具有高重复率的特征点,通过构建局部不变特征描述子来描述特征点的局部区域。在经典方法中,基于图像强度或梯度信息的特征点检测和描述是计算机视觉中最常用的方法。Moravec(Moravec, 1980)、Harris(Harris和Stephens, 1988)、DoG(Lowe, 2004)和FAST(features from accelerated segment test)(Rosten等, 2010)是最具代表性的特征点检测算子。

局部不变特征描述中最具代表性的是被广泛应用的SIFT(scale-invariant feature transform)(Lowe, 2004)、GLOH (gradient location and orientation histogram)(Mikolajczyk和Schmid, 2005)、SURF(Bay等, 2008)和DAISY(Tola等, 2010)。

遥感领域研究人员通常结合遥感图像的特征对上述经典的特征检测和描述算子进行改进。李晓明等人(2006)率先验证将SIFT用于遥感图像配准的可行性,Yu等人(2008)则利用Harris和SIFT在Quickbird、SPOT和TM遥感图像上结合小波金字塔的交叉匹配策略和不规则三角网进行匹配和纠正。Sedaghat等人(2011)基于稳定性和显著性约束提出UR(uniform robust)-SIFT,能够在尺度和图像空间均匀分布情况下提取更高质量的SIFT特征。余先川等人(2014)提出基于信号理论网格化的SG(signal gridding)-SIFT,能够获得比SIFT和UR-SIFT算法分布更加均匀的特征点和更高的配准精度。Sedaghat和Ebadi(2015a)进一步采用自适应直方图量化策略计算局部特征描述,构建了对局部几何畸变更鲁棒的局部特征描述子AB(adaptive binning)-SIFT。

由于上述特征检测和描述算子是利用图像的梯度信息或强度信息构建的,因此难以在同时存在明显几何畸变和非线性辐射差异的多模态遥感图像中获得高精度的匹配性能。近年来,相关学者在多模态遥感图像配准领域进行了大量的研究,并提出了一系列新的方法来克服多模态遥感图像之间的尺度、旋转、辐射变化和噪声干扰。

其中一类方法通过优化图像的梯度以提高描述子对非线性辐射差异的稳健性(Dellinger等, 2015; Xiang等, 2018; Yao等, 2022)。另外一类方法基于局部自相似利用对数极空间结构设计特征描述符,通过提取图像局部邻域内自相似的内部几何特征,以提升描述符对非线性辐射差异的抵抗性(Sedaghat和Ebadi, 2015b; Sedaghat和Mohammadi, 2019; Hong等, 2022)。还有一类方法利用相位一致性模型(phase consistency, PC)对光照和对比度变化不敏感的特性构建了一系列的基于PC的多模态特征配准算法(Ye等, 2018a; 姚永祥等, 2021; Zhu等, 2023)。

表2对基于特征的多模态遥感图像配准代表性方法进行了综合分析,表中列出了每种方法的核心思想及已开源代码的链接。下面将对上述3类基于特征的方法进行详细回顾。

表 2 基于特征的多模态遥感图像配准算法归纳总结

Table 2 Overview and summary of feature-based multimodal remote sensing image registration methods		
方法类型	代表算法	核心思想/结论/获取渠道
基于梯度优化	SAR-SIFT(Dellinger 等, 2015)	通过 ROEWA 算子计算比率梯度,提高对 SAR 中的散斑噪声的抵抗性; 公开代码链接: https://github.com/ZeLianWen/Image-Registration
	PSO-SIFT(Ma 等, 2017)	高斯尺度空间中的每个像素梯度的导数作为新的梯度,增强对强度变化的稳健性; 公开代码链接: https://github.com/ZeLianWen/Image-Registration
	何梦梦等人(2018)	Canny 和 Ratio of Averages 算子分别提取光学和 SAR 图像梯度,增强对非线性辐射差异的鲁棒性
	OS-SIFT(Xiang 等, 2018)	多尺度的 Sobel 和 ROEWA 算子为光学和 SAR 图像设计一致的梯度,融入多尺度信息增强配准的鲁棒性; 公开代码链接: https://github.com/xym2009/OS-SIFT
	CoFSM(Yao 等, 2022)	共现滤波构建尺度空间,并通过巴特沃斯滤波器优化图像梯度,增强多模态图像描述的鲁棒性; 公开代码链接: https://skyeearth.org/publication/project/CoFSM/
基于局部自相似	DOBSS(Sedaghat 和 Ebadi, 2015b)	相关值替代梯度来生成直方图构建自相似性描述符,捕获图像的内部几何布局,提升描述符鲁棒性
	HOSS(Sedaghat 和 Mohammadi, 2019)	结合自适应对数极空间结构和最大相关旋转索引图; 公开代码链接: https://www.researchgate.net/publication/328733655_HOSS_Descriptor
	OSS(Xiong 等, 2021)	基于自相似的对称性,利用偏移均值滤波快速计算自相似特征
	ASS(Xiong 等, 2022)	优化偏移均值滤波加速和细化特征计算过程,并采用局部统计加权差分运算来抑制相干散斑; 公开代码链接: https://gitee.com/xxin08/ASS
	MLSS(Hong 等, 2022)	结合最大值索引图和 LSS,将相同角度不同径向间隔的 LSS 向量相加
基于相位一致性	LHOPC(Ye 等, 2018a)*	构建 MMPC-Lap 检测器和 LHOPC 描述符,在复杂辐射差异的情况下保持稳定的配准性能
	RIFT(Li 等, 2020)	在最小和最大矩图上分别检测拐点和边缘点,构造最大索引图 MIM,提高特征检测的稳定性和特征描述的辐射不变性;公开代码链接: https://github.com/LJY-RS/RIFT-multimodal-image-matching
	HAPCG(姚永祥 等, 2021)	构建各向异性加权力矩图用于特征点检测,结合绝对相位一致性方向梯度和对数极坐标描述模板进行特征描述;公开代码链接: https://skyeearth.org/publication/project/HAPCG/
	MABPC(Fan 等, 2022)*	将非线性扩散率和空间一致性策略结合到哈里斯—拉普拉斯检测器中,采用自适应空间结构对多尺度相位一致性特征进行编码设计特征描述,提高了对几何和非线性辐射差异的鲁棒性
	R ₂ FD ₂ (Zhu 等, 2023)*	将多通道自相关策略与 log-Gabor 相结合,以检测具有高重复性和均匀分布的特征点,提出将具有旋转不变性的最大值索引图 RMIM 与 DAISY 的空间配置相结合,实现描述符的辐射和旋转不变性

注:“*”表示本文作者团队的工作。

1. 2. 1 基于梯度优化的特征配准方法

Dellinger 等人(2015)引入一种新的梯度定义,通过指数加权均值比 ROEWA 算子取代原始梯度计算的差分方法,获取对散斑噪声具有更为鲁棒的方向和幅度,并将 SIFT 算法的其余几个步骤应用于

SAR 图像配准(SAR-SIFT),实现了不同视角 SAR 图像的配准。Ma 等人(2017)重新为高斯尺度空间中的每个像素提出新的梯度定义(包括方向和幅度),引入结合关键点位置、尺度和方向的增强特征匹配方法 PSO(position scale orientation)-SIFT。何梦梦等

人(2018)根据多模态遥感图像来源的不同,设计不同的特征提取方法(Canny 边缘提取算子提取光学图像特征,Ratio of Averages 算子提取 SAR 图像特征),取得了比传统方法更高的配准精度。类似地,Xiang 等人(2018)提出的 OS-SIFT 根据光学和 SAR 图像的固有特性,分别使用多尺度的 Sobel 和 ROEWA 算子为光学和 SAR 图像设计一致的梯度,并构建 Harris 尺度空间来提取特征点和类似 GLOH 空间结构的描述符来识别同名点。Yao 等人(2022)提出的 CoFSM (co-occurrence filter space matching) 则利用共现滤波器设计尺度空间,同时通过巴特沃斯滤波器优化图像梯度以增强多模态图像描述的鲁棒性,减少非线性辐射差异的干扰。

1.2.2 基于局部自相似的特征配准方法

Sedaghat 和 Ebadi(2015b)使用相关值替代梯度生成直方图来计算主方向,构建的自相似描述符 DOBSS (distinctive order based self-similarity) 通过利用相关值获取图像的内部结构特征,成功配准多种传感器卫星图像对。Sedaghat 和 Mohammadi(2019)通过多个方向上计算的自相似值,设计了 HOSS (histogram of oriented self-similarity) 描述子,并且提出了一种结合自适应对数极空间结构的最大相关旋转索引 RIMC (rotation index of the maximal correlation) 图来实现旋转不变性,进一步提高了多模态配准的精度。Xiong 等人(2021)基于自相似的对称性,利用偏移均值滤波快速计算自相似特征,提出了一种面向自相似的特征描述符 OSS (oriented self-similarity); 他们进一步设计了优化的偏移均值滤波方法来加速和细化特征计算过程,并采用局部统计加权差分运算来抑制相干散斑,从而提取出 ASS (adjacent self-similarity) 特征(Xiong 等,2022),实验证明在配准鲁棒性、精度和效率方面均有很大提升。Hong 等人(2022)结合最大值索引图和 LSS,将相同角度不同径向间隔的 LSS 向量相加,构建 MLSS (max-index based local self-similarity) 描述符,实现了多模态遥感图像鲁棒的配准。

1.2.3 基于相位一致性的特征配准方法

Ye 等人(2018a)首先利用具有光照和对比度不变性的相位一致性模型进行特征检测和特征描述,特征检测阶段构建的 MMPC (minimum moment of phase congruency)-Lap 检测器结合相位一致性最小矩和自动尺度定位技术检测图像尺度空间中稳定的

特征点;特征描述符 LHOPC (local histogram of orientated phase congruency) 利用相位一致性的幅值和方向,并借鉴 DAISY 的空间配置计算局部不变性特征,在复杂辐射差异的情况下保持稳定的配准性能。Li 等人(2020)提出了辐射不变特征变换 RIFT (radiation-variation insensitive feature transform),首先在相位一致的最小矩图和最大矩图上分别检测拐点和边缘点;然后利用 log-Gabor 卷积序列构造最大索引图 MIM (maximum index map),提高了特征检测的稳定性和特征描述的辐射不变性。姚永祥等人(2021)提出结合各向异性加权力矩与绝对相位方向梯度特征直方图 (histogram of absolute phase consistency gradients, HAPCG) 的匹配方法,首先通过构建加权力矩方程获取各向异性加权力矩图用于特征点检测,然后结合绝对相位一致性方向梯度和对数极坐标空间结构进行特征描述,从而有效提升匹配的稳健性。Fan 等人(2022)首先将非线性扩散和空间一致性策略结合到哈里斯-拉普拉斯检测器中,以减弱模态变化对特征点检测的影响,然后采用自适应分箱空间结构对多尺度相位一致性特征进行编码,设计了 MABPC (multiscale adaptive binning phase consistency) 描述子,提高了对几何和非线性辐射差异的鲁棒性。考虑到基于相位一致性最小矩或最大矩的检测器可能会导致多方向特征的损失,因为其是相位一致性在不同方向上的加权响应,Zhu 等人(2023)提出的 R_2FD_2 将多通道自相关策略与 log-Gabor 小波相结合,构建了 MALG (multichannel autocorrelation of the log-Gabor) 检测器,以检测具有高重复性和均匀分布的特征点;进一步提出一种具有旋转不变性的最大值索引图 RMIM (rotation-invariant maximum index map),并将 RMIM 与 DAISY 的空间配置相结合,实现描述符的辐射和旋转不变性。

基于梯度优化的方法通常是特定的多模态图像设计一致梯度,该类型方法的泛化性普遍很差,难以对其他类型的多模态图像保持同样性能;基于局部自相似的方法也存在局限性,因为 LSS 描述符相对较低的判别能力可能导致在复杂非线性辐射差异情况下无法保持稳健的匹配性能;基于相位一致性的方法大多计算复杂度很高,配准过程总体比较耗时。

1.3 基于深度学习的配准方法

随着卷积神经网络 (convolutional neural net-

work, CNN)、全卷积神经网络(fully convolutional networks, FCN)、生成对抗网络(generative adversarial network, GAN)、孪生神经网络(siamese networks)和Transformer等深度学习技术的发展,这些深度学习技术对图像的深层特征有着优越的学习和表达能力,已经广泛应用于遥感摄影测量、分类和目标检测等多个领域(龚健雅和季顺平, 2018; 石争浩等, 2023; 叶沅鑫等, 2021)。

基于数据驱动的深度学习可以从大量训练数据中利用“学习—反馈—调整”的策略抽取出高层次的语义信息,因此深度学习在多模态遥感图像配准领域中引起了广泛的关注(眭海刚等, 2022)。许多研究学者已经尝试将其应用于多模态遥感图像配准,通过提取比传统手工特征(即浅层特征)更精细的共同特征以更好地适应多模态遥感图像间几何、辐射差异和噪声等因素带来的影响。

本文将基于数据驱动的深度学习配准方法分为两大类:一类称为模块化的配准方法,通过将深度学习技术融入传统配准流程中的某一阶段性任务,即利用神经网络(CNN、GAN和Siamese等)替代基于区域或基于特征方法中的某一模块形成新的配准流程;另一类称为端到端的配准方法,通过直接构建端到端的神经网络结构直接实现图像配准。

表3对基于深度学习的多模态遥感图像配准的代表性方法进行了综合分析,表中列出每种方法的核心思想及已开源代码的链接。下面将对上述两类基于深度学习的配准方法进行详细回顾。

1.3.1 模块化的配准方法

模块化的配准方法最常见的策略是将深度网络嵌入到基于特征或基于区域的方法中,利用深度学习完全数据驱动和高维深层特征抽取的优势生成更鲁棒的特征、更有效的特征描述符或相似性度量,从而提高图像配准的稳健性。

1)基于学习的模板匹配方法。大多数基于学习的模板匹配方法都是基于(伪)孪生神经网络架构(Merkle等, 2017; Zhang等, 2022c; Liu等, 2023)。这些方法的基本思想是通过训练一个(伪)Siamese网络来获得深度特征替代基于区域配准方法中的强度信息或浅层结构特征,然后评估输出特征映射之间的相似度来实现图像匹配,评估相似性可以通过神经网络、点积、卷积运算或传统的相似性度量来实现。Merkle等人(2017)和Zhang等人(2019a)都采用

Siamese全卷积网络提取深度共同特征,它们的主要区别在于前者使用扩展卷积来增加感受野大小,并通过点积计算产生相似图。而后者提出的SFcNet(siamese fully convolutional network)则使用卷积运算来连接输出特征对。Hughes等人(2018)设计了一个特殊的伪孪生网络架构,其中包含两个独立但相同的卷积流,用于并行处理SAR和光学图像,而不是权重共享的孪生网络,以适应不同图像的异构性质。Zhang等人(2022c)进一步提出了一种基于Siamese的深度密集特征网络架构DDFN(deep dense feature network)来提取深度特征,能够在高级信息抽象和低级信息保存之间取得更好的平衡。

考虑到池化操作可能造成的空间信息损失,南柯等人(2019)和Zhou等人(2022)采用全卷积架构生成与输入图像大小相同的特征图,从而使模型提取特征的空间信息更加完整,定位精度更高,并且后者构建基于伪孪生网络的多尺度卷积梯度特征MCGFs(multiscale convolutional gradient features)可以更好地适应非线性辐射差异。Zhang等人(2022a)提出的OSMNet(optical and SAR image matching network)在伪孪生网络架构下结合了多层特征融合和多频通道关注,并采用自适应加权损失来提高光学和SAR图像匹配的精度。Cao等人(2023)提出了基于孪生网络训练的相位一致性网络PCNet(phase congruency network)来增强多光谱或多模态图像的结构相似性,然后利用SSD作为相似性度量评估PCNet特征图的相似性实现配准。Liu等人(2023)使用log-Gabor滤波器提取图像的多方向相位结构信息,然后通过构建的多尺度融合MSF SiamUNet-7(multiscale fusion SiamUNet-7)网络提取相位结构卷积特征,实现大尺度局部纹理信息和小尺度全局结构信息的充分融合,实现了更高精度的光学和SAR图像匹配。

虽然这些基于学习的模板匹配方法通过深度学习强大的深层特征提取和高维特征表征能力提高了对多模态遥感图像间非线性辐射差异的抵抗性,但也无法适应多模态图像之间显著的尺度和旋转变化等。

2)基于学习的特征匹配方法。基于学习的特征匹配方法大多都使用基于图像特征的深度架构生成可训练的特征检测器或描述符以提高配准性能(Ye等, 2018b; 蓝朝桢等, 2021; Quan等, 2023)。

表3 基于深度学习的多模态遥感图像配准算法归纳总结

Table 3 Overview and summary of learning-based multimodal remote sensing image registration methods		
方法类型	代表算法	核心思想/结论/获取渠道
基于学习的模板匹配	Merkle 等人(2017)	Siamese 全卷积网络提取特征和点积相似度度量得光学和SAR 图像之间的偏移
	Hughes 等人(2018)	两个独立但相同的卷积流,而不是权重共享的孪生网络,以适应图像的异构性质
	南轲等人(2019)*	去除 Siamese 网络中的池化层和抽取特征来保持特征信息的完整性和位置精度
	SFCNet (Zhang 等, 2019a)	Siamese 全卷积网络的两个分支通过卷积运算连接,并构建新特征距离的损失函数
	Zhang 等人(2022c)	基于 Siamese 的深度密集特征网络架构 DDFN 提取深度特征
	MCGF (Zhou 等, 2022)*	构建基于伪孪生网络的多尺度卷积梯度特征 MCGFs,并使用 FFT 加速匹配
	OSMNet (Zhang 等, 2022a)	在伪孪生网络架构下结合多层特征融合和多频通道关注,并采用自适应加权损失;公开代码链接: https://github.com/zhanghan9718/OSMNet
	PCNet (Cao 等, 2023)	基于孪生网络训练的 PCNet 增强多模态图像的结构相似性,再利用 SSD 匹配
	MSF SiamUNet-7 (Liu 等, 2023)	构建多尺度融合 MSF SiamUNet-7 网络提取相位结构卷积特征,实现大尺度局部纹理信息和小尺度全局结构信息融合
	Ye 等人(2018b)	研究 CNN 中的全连接层特征和不同聚合大小的卷积层特征,并结合 SIFT 配准
模块化的配准方法	Dong 等人(2019)	用 DescNet 网络生成深层特征向量,为特征点描述生成更鲁棒的特征向量
	Ma 等人(2019)	结合深度特征和手工制作的局部特征进行精细特征匹配和变换估计以提高配准精度
	CMM-Net (蓝朝桢 等, 2021)	在 CNN 提取的多模态遥感图像高维特征图上进行特征点检测和描述; 公开代码链接: https://github.com/lan-cz/cnn-matching
	MAP-Net (Cui 等, 2022)	SPAP 模块集成全局和局部上下文信息,注意机制模块对网络生成的密集特征加权
	Li 等人(2022b)*	自监督检测网络检测高重复率特征点,交叉融合匹配网络进行跨特征描述和匹配
	Sim-CSPNet (项德良等, 2022)	构建 Sim-CSPNet 网络,快速提取包含深层和浅层特征的特征描述符
	M2DT-Net (Zhang 等, 2022b)	利用微调 CNN 模型学习不变性特征,结合距离约束与 DEGENSAC 进行误匹配剔除
	Cnet (Quan 等, 2022)	Cnet 利用注意力学习模块关注有意义的特征来增强特征表征
	Quan 等人(2023)	基于深度有序回归学习的旋转校正和基于深度描述符学习的精细图像配准
	MatcherTF (Liang 等, 2023)	利用具有自注意力和交叉注意力的 Transformer 构建局部特征匹配网络 MatcherTF
基于风格迁移的模态统一	Merkle 等人(2018)	利用 cGAN 生成伪 SAR 图像块,并改进基于浅层特征的 NCC 和 SIFT 等传统匹配
	Zhang 等人(2019b)	对图像视觉属性转移算法进行扩展,融合原始图像结构和纹理,消除模态差异
	宋智礼等人(2021)	基于 VGG 网络对图像的风格特征和内容特征进行抽取并重新组合以实现模态一致
	陈世伟等人(2022)	条件 GAN+小波多尺度特性增强边缘区域+互相关实现精确配准
	Zhao 等人(2022)	提出平滑循环一致生成对抗网络,并构造新的损失函数,提高迁移图像生成的质量
	陈彤等人(2022)	对原 CycleGAN 增加损失函数计算通路,引入本体映射损失实现对颜色信息的约束
	杨冰超等人(2022)	利用 GAN 将可见光转换为伪红外图像,再结合 SURF 和 PIIFD 算法实现特征描述
	Wang 等人(2018)	通过信息前馈和反馈的闭环直接学习待配准图像和参考图像与它们的匹配标签之间的映射,提出自学习方法解决图像训练数据量小和用迁移学习来减少训练计算量
	D-DenseNet (Li 等, 2022a)	设计权值共享的 D-DenseNet 网络提取多模态图像的特征,将图像配对后直接学习带配准图像块的4个角点相对于参考图像的位移参数,实现端到端可训练网络
	Hughes 等人(2020)	提出由特征提取、多尺度匹配和误差剔除网络3个顺序神经网络组成的端到端框架
端到端的配准方法	Li 等人(2021)	提出深度学习语义模板匹配框架,通过模板在参考图像上的空间位置概率映射获取变换参数;公开代码链接: https://github.com/liliangzhi110/semantictemplatematching
	MU-Net (Ye 等, 2022b)*	提出无监督网络 MU-Net,以多尺度方式直接学习变换参数,并基于结构相似性设计损失函数;公开代码链接: https://github.com/yeyuanxin110/MU-Net
	Li 等人(2023)	设计自注意力的融合框架获取图像的跨模态特征表示,将配准定义为用于回归规定的点位移函数以实现端到端匹配;公开代码链接: https://github.com/liliangzhi110/E2EIR

注:“*”表示本文作者团队的工作。

Ye等人(2018b)对CNN中的全连接层特征和不同聚合大小的卷积层特征进行了研究,然后将CNN特征与SIFT结合,并成功用于多传感器遥感图像配准。Dong等人(2019)利用DescNet深度卷积网络生成的深层特征向量代替传统手工设计的浅层特征,为特征点描述生成更具区分度的特征向量,显著增加了遥感图像配准任务中正确匹配点的数量。同样, Ma等人(2019)提出基于深度和局部特征的从粗到精两步配准方法,首先使用基于图像特征的深度架构(VGG-16)来近似空间关系。深度模型可以作为特征提取器,实现金字塔特征检测,在此基础上,结合深度特征和手工制作的局部特征用来进行精细特征匹配和变换估计以提高配准精度。

蓝朝桢等人(2021)借鉴了D2-Net特征提取思想,在CNN提取的多模态遥感图像高维特征图上进行特征点检测和描述,构建了一种稳健的多模态遥感图像匹配算法(cross modality matching net, CMM-Net)。Cui等人(2022)提出了一种基于空间金字塔聚集池化SPAP(spatial pyramid aggregated pooling)和注意机制的SAR与光学图像匹配方法MAP-Net, SPAP模块集成全局和局部上下文信息,而注意机制模块对网络生成的密集特征进行加权,提取出对不同模态图像间的几何和辐射变化具有更强适应性的高级语义特征,有利于跨模态图像的匹配。Li等人(2022b)提出“检测+匹配”框架,首先设计了一种自监督检测网络,在多模态图像之间生成相似的特征映射,用于检测高重复率的特征点,然后设计了一个交叉融合匹配网络,该网络通过全局优化和融合信息进行跨模态特征描述和匹配,实验证明其具有更鲁棒的特征检测和匹配性能。

项德良等人(2022)设计了一种Sim-CSPNet(siamese cross stage partial network)架构来快速提取包含深层和浅层特征的特征描述符,在不同传感器和不同分辨率的SAR图像上展示出优良的配准效果。Zhang等人(2022b)提出的M2DT-Net首先利用微调CNN模型学习具有高鲁棒性的不变特征,然后采用自适应特征描述子距离约束与DEGENSAC细化相结合的方法对得到的匹配进行过滤,实现了复杂场景下的各种遥感图像高精度匹配。Quan等人(2022)提出了一种用于多模态遥感图像配准的深度特征相关学习网络(deep feature correlation learning network, Cnet), Cnet利用注意力学习模块,在深度

卷积网络的基础上构建特征学习网络,通过关注有意义的特征来增强特征表征。Quan等人(2023)进一步提出了基于深度有序回归学习的旋转校正和基于深度描述符学习的精细图像配准来构建由粗到精的深度学习配准框架。在粗配准阶段,设计了一种有效的深度有序回归(deep ordinal regression, DOR)网络进行旋转校正, DOR网络将旋转校正任务转化为旋转有序回归问题,利用旋转序数之间的潜在关系,提高旋转估计的精度。在精细配准阶段,基于旋转校正后的图像,采用DDL(deep descriptor learning)网络处理图像模态变化。Liang等人(2023)利用具有自注意力和交叉注意力的Transformer构建了局部特征匹配网络MatcherTF,由于Transformer具有良好的全局关注特性,将基于深度学习的特征描述网络集成到特征匹配方法中,可以为传统检测器检测到的或随机选择的每个特征点生成更具判别性的描述符,从而提高特征匹配的性能。

基于学习的特征匹配方法通过深度学习技术强大的深层特征提取和高维特征表征能力可以提高特征检测器的重复率和特征描述子的鲁棒性。值得注意的是,目前多模态遥感图像配准领域发展的大多数方法都是基于深度学习的特征描述,对于特征检测的探索相对较少。并且一些基于学习的特征匹配方法没有针对旋转或尺度差异进行研究,因此不具备尺度和旋转不变性(蓝朝桢等, 2021; 项德良等, 2022)。

3) 基于风格迁移的模态统一方法。为了减弱多模态遥感图像之间显著非线性辐射差异的影响,生成对抗网络GAN等模型被引入,这类模型通过风格迁移实现图像的模态统一,再利用传统基于区域或者基于特征的配准方法实现模态统一图像的配准(Merkle等, 2018; 宋智礼等, 2021; 杨冰超等, 2022)。

这些方法首先通过训练进行图像到图像的转换,使待配准的多模态遥感图像具有相似的辐射信息和结构特征。Merkle等人(2018)利用条件生成对抗网络cGAN(conditional generative adversarial network)从光学图像中生成伪SAR图像块,然后基于浅层特征的NCC和SIFT得到了更高的匹配精度。Zhang等人(2019b)对图像视觉属性转移算法进行扩展,通过融合原始图像的结构和纹理,减弱多模态遥感图像之间的非线性辐射差异,有效提高了后续

采用传统特征方法进行匹配的精度。宋智礼等人(2021)基于VGG(Visual Geometry Group)网络架构对参考图像的风格特征以及待配准图像的内容特征进行抽取并重新组合,以实现图像的模态统一,然后通过图像分割的方法清除风格迁移图像中多余的纹理并使用SURF算法进行图像配准。

陈世伟等人(2022)提出一种基于风格迁移不变特征的光学与SAR图像配准算法,首先通过条件GAN训练SAR图像转换为光学图像的风格迁移网络并生成差异图,利用小波多尺度特性增强人工光学图像、光学基准图像和差异图的边缘区域,最后采用互相关作为相似性度量实现原始SAR图像与光学基准图像的精确配准。Zhao等人(2022)基于生成式对抗网络的图像风格转移思想提出了平滑循环一致生成对抗网络,通过构造新的生成式网络损失函数,提高迁移图像生成的质量,然后利用D2-Net和LoFTR实现风格迁移的航拍图像与原始航拍图像的匹配后将该变换关系映射到初始匹配的图像对中,有效提高了匹配算法的准确性和鲁棒性。陈彤等人(2022)依据生成对抗网络的风格转换思想,通过构建新的损失函数,设计了基于生成对抗模型的可见光和红外图像匹配方法,降低了多模态图像的匹配难度。杨冰超等人(2022)利用生成对抗网络将可见光图像转换为相应的伪红外图像,再结合SURF和PIIFD算法实现特征描述,提高了红外和可见光图像在尺度差异较大情况下的配准鲁棒性。

通过风格迁移实现图像模态统一的方法很大程度上依赖于训练的样本数量,训练数据集不足和图像覆盖的场景种类不全会影响图像模态转换的效果。因此,在对不同场景和不同类型的多模态图像进行模态转换时,效果都会存在差异,以至于影响后续配准结果的鲁棒性和精确性。

1.3.2 端到端的配准方法

越来越多的研究集中在直接估计几何变换参数或变形场,但主要应用于计算机视觉和医学图像配准领域(de Vos等,2019;Ma等,2021),一些研究将这一思想引入遥感图像配准(Wang等,2018;Hughes等,2020;Li等,2023)。

Wang等人(2018)构建的端到端架构将特征提取和特征匹配统一在一个网络中进行学习,通过信息前馈和反馈的闭环直接学习待配准图像和参考图

像与它们的匹配标签之间的映射,同时提出一种自学习方法解决遥感图像训练数据量小的问题,并采用迁移学习来减少训练阶段巨大的计算量。Li等人(2022a)开发具有权值共享的D-DenseNet网络以充分提取多模态图像的特征,将待配准图像与参考图像块进行配对后直接学习待配准图像块4个角点相对于参考图像的位移参数,实现端到端的可训练网络。Hughes等人(2020)提出一种SAR与光学图像稀疏匹配的端到端框架,该框架由3个按顺序使用的神经网络组成:第1个网络突出具有显著特征的区域;第2个是围绕特征空间相关函数构建的多尺度匹配网络;第3个是误差剔除网络直接估计匹配结果的质量,并剔除误差较大的匹配结果。

Li等人(2021)提出了一种用于遥感图像配准的深度学习语义模板匹配框架,该框架不通过特征提取和匹配优化,也不采用逐像素搜索来优化相似性度量,而是直接通过模板在参考图像上的空间位置概率映射来获得变换参数,能够在较短的时间内实现较高的匹配精度。Ye等人(2022b)提出了一种无监督网络MU-Net,以多尺度方式直接学习变换参数。MU-Net通过将多个深度神经网络(deep neural network, DNN)模型在多个尺度上叠加,形成一个由粗到精的配准框架,防止反向传播陷入局部极值,并基于结构相似性设计损失函数,使其适用于多模态图像间的几何和辐射变化。Li等人(2023)提出了一种多模态遥感图像匹配的融合网络,该网络首先设计了基于自注意力机制的融合框架用于获取输入图像和参考图像之间的跨模态特征表示,然后将配准定义为用于回归规定的点位移函数,并在网络内校正输入图像以实现端到端的匹配。

总的来说,模块化的配准方法虽然容易训练,灵活性较强,可根据不同的需求配合其他各具优势的模块构建更精确的匹配模型,但是这样会导致容易陷入局部最优的情况,因为该类型方法难以避免多阶段任务中易出现的误差累计问题。相较于模块化方法,端到端的配准方法通过构建端到端的神经网络结构替换图像匹配的整个过程,虽然通过训练时信息的不断反馈能够得到模型的全局最优解,避免陷入局部最优,但是存在整体训练资源消耗大、模型结构复杂和解释性差等问题。

2 多模态图像配准数据集

本节总结了遥感图像配准目前几个开源的数据集,表4介绍了这些公开数据集的详细信息,并给出了数据集的下载链接。这些数据集有的包括上万对的多模态图像,可用于深度学习相关研究(配准、融合、识别等)的模型训练;部分数据集仅有几十或几百对不同领域或不同覆盖场景下的多模态图像,可用于测试和评估算法。

1) SARptical 数据集(Wang 和 Zhu, 2018)是由 TerraSAR-X 和航空 UltraCAM 传感器获取的德国柏林 2009 年—2013 年间共 10 108 对 SAR 和光学图像组成,光学和 SAR 图像的空间分辨率分别为 0.2 m 和 1 m,像素均为 112 × 112 像素。图 4 展示了 SARptical 数据集中部分场景的图像。

2) SEN1-2(Schmitt 等, 2018)是由 Sentinel 系列卫星于 2017 年间拍摄覆盖全球五大洲四季不同地表区域的共 282 384 对 SAR 和光学图像组成,光学和

SAR 图像分辨率为 10 m,尺寸均为 256 × 256 像素。图 5 展示了 SEN1-2 数据集中部分场景的图像。

3) University-1652(Zheng 等, 2020)包含来自全球 72 所大学 1 652 座建筑的地面图像、Google Earth 模拟生成的无人机视角图像和卫星图像 3 种平台的数据。该数据集的每栋建筑平均有 1 幅卫星图像、1 幅无人机图像和 3.38 幅真实街景图像,数据源和成像视角丰富,并且具有明确的标注信息。图 6 展示了该数据集中部分场景的图像。

4) OS-dataset(Xiang 等, 2020b)是由 10 692 对 256 × 256 像素的高分辨率光学和 SAR 图像组成的匹配数据集。图像覆盖范围包括全球多个城市及其周边地区, SAR 图像来自于高分 3 号卫星,空间分辨率为 1 m,光学图像采集自 Google Earth 平台,并且重采样至 1 m 分辨率。图 7 展示了 OS-dataset 数据集中部分场景的图像。

5) QXS-SAROPT(Huang 等, 2021)数据集由 20 000 对 256 × 256 像素的高分辨率光学和 SAR 图像组成。QXS-SAROPT 数据集分别从高分 3 号 SAR 卫星和 Google Earth 中获得 SAR 图像和光学图像,空间分辨

表 4 图像配准公开数据集归纳
Table 4 Summary of image registration public datasets

数据集	图像类型	特点	下载链接
SARptical (Wang 和 Zhu, 2018)	光学— SAR	由 10 108 对 112 × 112 像素的 SAR 和光学图像组成,光学和 SAR 图像空间分辨率分别为 0.2 m 和 1 m	https://syncandshare.lrz.de/getlink/figixjRV9idETz-PgG689dGB/SARptical_data.zip
SEN1-2 (Schmitt 等, 2018)	光学— SAR	由 282 384 对 256 × 256 像素的 Sen-1 和 2 图像组成,光学和 SAR 图像空间分辨率为 10 m	https://mediatum.ub.tum.de/1436631
University-1652 (Zheng 等, 2020)	多种	包含来自全球 72 所大学 1 652 座建筑的地面影像、模拟生成的无人机视角影像和卫星影像 3 种平台的数据	https://github.com/layumi/University1652-Baseline
OS-dataset (Xiang 等, 2020b)	光学— SAR	由 10 692 对 256 × 256 像素的高分辨率光学和 SAR 图像组成,光学和 SAR 图像空间分辨率为 1 m	https://pan.baidu.com/share/init?surl=4bqaJhM-SZEy7EXcXVAc77w (code: vriw)
QXS-SAROPT (Huang 等, 2021)	光学— SAR	由 20 000 对 256 × 256 像素的高分辨率光学和 SAR 图像组成,光学和 SAR 图像空间分辨率为 1 m	https://github.com/yaoxu008/QXS-SAROPT
Multimodal_Image_Matching_Datasets (Jiang 等, 2021)	多种	涉及医学、遥感和计算机视觉领域 18 种不同类型共 164 对多模态图像,可用于测试和评估算法	https://github.com/StaRainJ/Multi-modality-image-matching-database-metrics-methods/blob/master/Multimodal_Image_Matching_Datasets/
SOPatch (Xu 等, 2023)	光学— SAR	基于 WHU-SEN-City、OSdataset 和 SEN1-2 共 3 个数据集,创建包含超过 65 万对的光学和 SAR 配准数据集	https://pan.baidu.com/share/init?surl=2yIhe-GOg6JTTsYAQfX7Pfg&pwd=1234
OSEval (Xiang 等, 2023)	光学— SAR	由数百对亚米级分辨率的光学和 SAR 图像组成,能够提供对配准算法公平且准确的比较	https://github.com/xym2009/OS-Eval

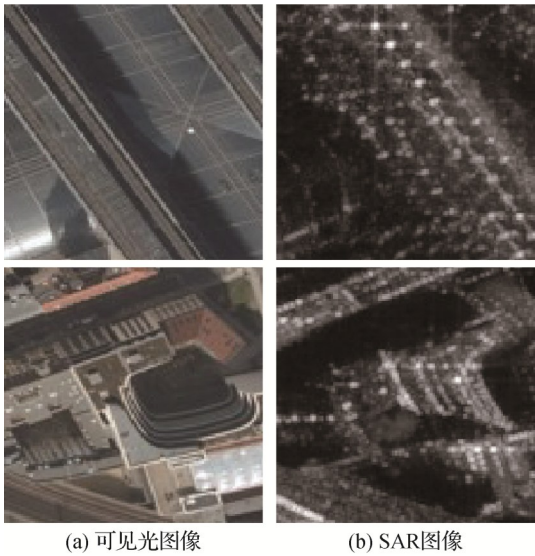


图4 SARptical数据集示例(Wang和Zhu,2018)
Fig. 4 Examples of SARptical dataset(Wang and Zhu,2018)
((a) visible images;(b) SAR images)

率都为1 m,这些图像覆盖了圣地亚哥、上海和青岛三个港口城市。图8展示了QXS-SAROPT数据集中部分场景的图像。

6) Multimodal_Image_Matching_Datasets (Jiang等,2021)数据集为从医学、遥感和计算机视觉领域相关数据集中选择部分原始图像对组成,涉及18种不同类型共164对多模态图像对。图9展示了该数据集18种不同类型的多模态图像对。

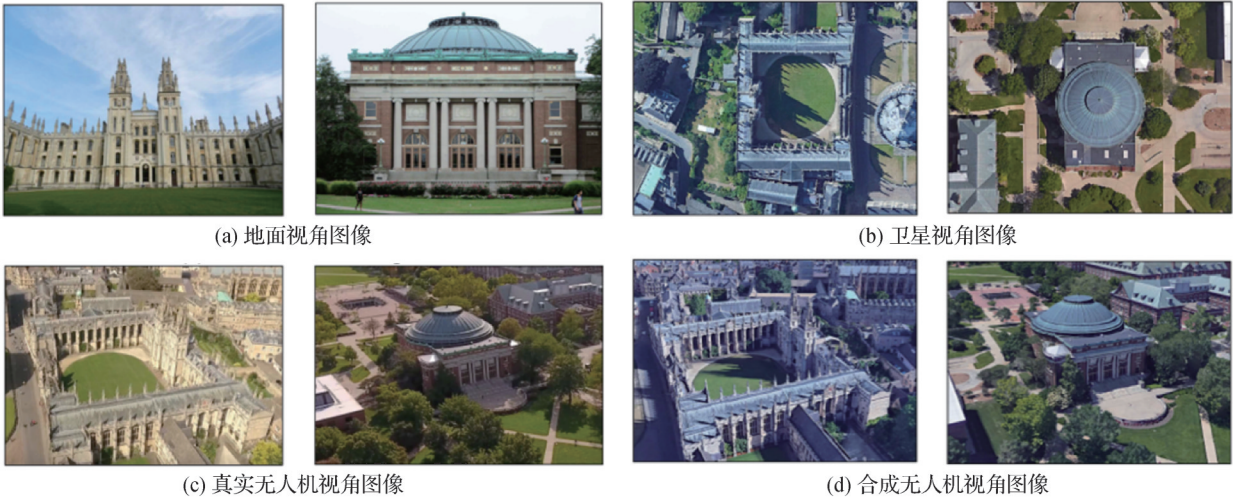


图6 University-1652数据集示例(Zheng等,2020)
Fig. 6 Examples of University-1652 dataset (Zheng et al. ,2020) ((a) ground-view images;
(b) satellite-view images;(c) real drone-view images;(d) synthetic drone-view images)

8)OSEval数据集(Xiang等,2023)包含由数百对亚米级分辨率的光学和SAR图像组成的评估

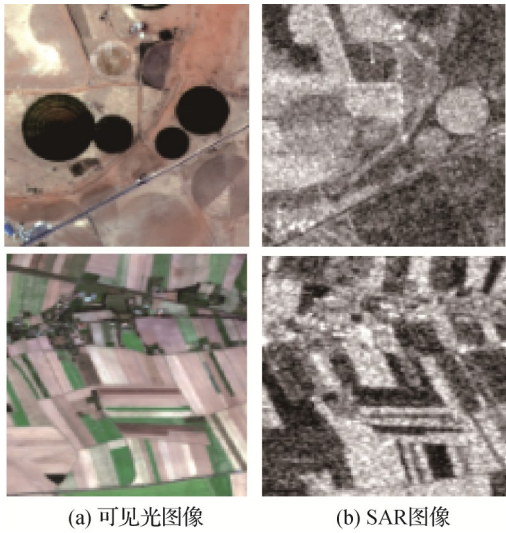


图5 SEN1-2数据集示例(Schmitt等,2018)
Fig. 5 Examples of SEN1-2 dataset(Schmitt et al. , 2018)
((a) visible images;(b) SAR images)

7)SOPatch(Xu等,2023)数据集是基于WHU-SEN-City(Wang等,2019)、OSdataset和SEN1-2这3个数据集的基础上进行进一步的精确配准,创建包含超过65万对的光学和SAR图像配准数据集。该数据集由多种卫星图像(Sentinel-1、Sentinel-2、高分3号等)组成,包含丰富的场景(山脉、湖泊、建筑物、农田和裸地等),能够训练出具有良好泛化性的模型。图10展示了该数据集中部分场景的图像。

数据集,具有各种视角并且涵盖丰富的场景,包括机场、住宅区和工业区,以及密集的城市地区。



图 7 OS-dataset数据集示例(Xiang等,2020b)

Fig. 7 Examples of OS-dataset dataset(Xiang et al. , 2020b) ((a) visible images; (b) SAR images)

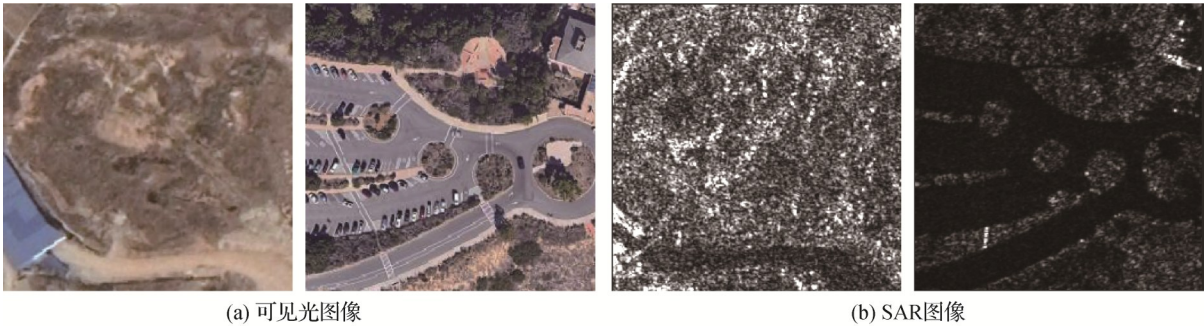


图 8 QXS-SAROPT数据集示例(Huang等,2021)

Fig. 8 Examples of QXS-SAROPT dataset(Huang et al. , 2021) ((a) visible images; (b) SAR images)

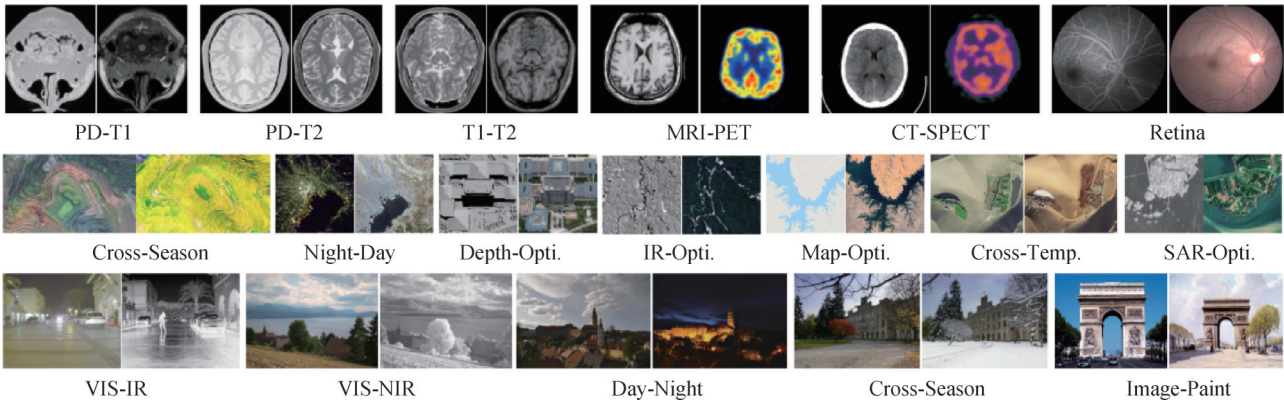


图 9 Multimodal_Image_Matching_Datasets数据集示例(Jiang等,2021)

Fig. 9 Examples of Multimodal_Image_Matching_Datasets dataset (Jiang et al. , 2021)



图 10 SOPatch数据集示例(Xu等,2023)

Fig. 10 Examples of SOPatch dataset(Xu et al. , 2023) ((a) visible images; (b) SAR images)

该评估数据集使用在光学和 SAR 图像中具有清晰结构的角点作为地面真值来定量评估光学和 SAR 配准算法,能够提供对传统算法公平且准确

的比较,以及对深度学习配准网络的泛化性评估。图 11 展示了 OSEval 数据集中部分场景的图像。

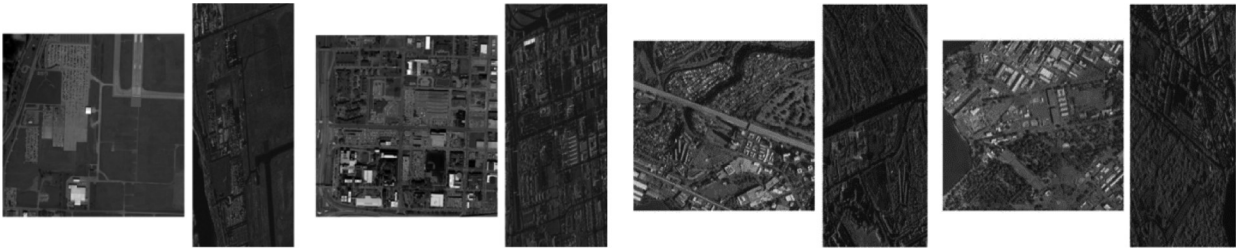


图 11 OSEval数据集示例(Xiang等,2023)
Fig. 11 Examples of OSEval dataset(Xiang et al. , 2023)

3 多模态遥感图像配准总结和展望

多模态图像匹配或配准在遥感领域的各种关键应用(包括图像融合、变化检测和图像拼接等)中起着至关重要的作用。在过去的几十年里,多模态遥感图像配准研究在很大程度上解决了因模态差异引起的非线性辐射差异和各种因素引起的几何畸变,配准的准确性和效率在不断提高。本文综合研究了现有的多模态图像配准方法,将其分为基于区域、基于特征和基于深度学习的配准方法3类。本文通过对3类不同类型配准方法的理论框架和代表性算法进行简要的介绍,对该领域的研究进行了全面总结,旨在为相关领域的研究人员提供更多关于多模态遥感图像配准的信息参考,为实现进一步的突破和创新奠定基础。

通过对前沿的多模态遥感图像配准方法的回顾,虽然不同类型的配准方法在性能上都取得了显著的进步,但不同的配准方法在实施上存在一定的局限性,并不适用所有复杂情况下的多模态遥感图像配准。因此,本文将针对不同类型的配准方法依次讨论这些问题。

基于区域配准方法的局限性主要表现在两个方面。首先,所有这些方法都依赖于一些先验条件预测近似的匹配范围,便于进一步实现后续的模板匹配过程。这些先验条件可以是多模态遥感图像的地理参考信息(RPC、POS数据等),也可以是一个粗略的映射模型(例如:仿射和投影变换模型),或是通过手动选择几个均匀分布的检查点获得。另一个重要的限制是基于区域的方法对输入图像和参考图像之

间的重叠区域有要求。如果输入图像与参考图像间的重叠区域过小,则配准结果会受到很大的影响,甚至导致失败。

基于特征的配准方法通过利用相邻像素之间的局部空间关系,为每个特征点构建高维信息特征向量,这些高维特征向量对所设计描述符的匹配性能至关重要。因此,特征匹配方法相比于模板匹配方法通常面临着沉重的计算负担,并且在假定的匹配中容易产生不可避免的严重异常值(误匹配),特别是在同时存在尺度、旋转和辐射差异的多模态配准情况下。一般来说,基于特征的方法配准鲁棒性不如基于区域的方法稳定。

基于深度学习的方法完全由训练数据驱动,利用其能够提取图像深层语义信息的能力可以有效地提高图像配准的性能。然而,为多模态遥感图像配准提出的基于学习的方法并不像计算机视觉和医学界那样丰富,大部分研究都是将该技术用于特征检测和特征描述。主要原因是多模态遥感图像由于场景多样、成像范围广以及混合噪声等因素,比自然和医学图像更为复杂,难以构建一个充分且具有代表性的数据集进行模型训练和测试。此外,由于较高的获取成本或不可避免的保密要求,只有少数遥感图像配准数据集是公开的。因此,遥感领域不同模态条件下的数据不足,目前还没有完整、全面的包含所有类型多模态遥感图像对的数据库,训练和测试数据的缺乏极大地限制了基于深度学习配准方法的发展。

在未来的研究中,随着天空地多传感器一体化立体观测技术的不断发展,发展智能化、精准化、实时化的多模态遥感图像配准方法是当前国际学术研

研究和工业应用的热点前沿。多模态遥感图像配准仍有许多未解决的问题,对于基于浅层特征的传统方法和基于深层特征的学习方法来说,研究智能鲁棒快速的多模态配准通用方法仍然是一个开放的问题,未来的发展将主要集中在以下几个方面。

基于区域的研究主要集中在解决多模态之间显著的辐射差异问题上,对于同时存在严重辐射畸变和复杂变形(如仿射变形)的多模态模板匹配研究相对较少。未来的研究可以重点探索变形相似性度量,将局部外观与结构特征结合起来,实现严重辐射变化和复杂变形下的多模态遥感图像可靠配准。另外,可以关注星载、机载和车载等在线实时处理,探索轻量化二进制结构特征表达模型并设计软硬一体的配准方法和装置,以满足战场态势动态监测等应急情况的实时性需求。

基于特征的研究更倾向于设计稳定的算子进行特征检测和特征描述,以在尺度、旋转、辐射差异下识别可靠的同名点。未来的研究可以从以下两个方面进行探索。一方面,研究人员可以专注于探索特征检测和描述的加速策略(如快速傅里叶变换技术),以提高特征匹配的计算效率;另一方面,未来的工作可以聚焦于探索异常值去除方法,从基于描述子相似性识别的假定匹配中更鲁棒地消除误匹配,以实现更精准化的多模态特征配准。

由于缺乏数据集,基于学习的研究难以获得具有良好泛化能力的网络模型。未来的工作应该更加注重提供更多的公开可用的数据集,可以通过基于开源卫星平台收集更多的多模态图像生成新的数据集来实现。此外,无监督学习是在有限数据集下获得更好配准结果的一种潜在方法,值得在多模态遥感图像配准领域进行更多的探索。同时,探索计算机视觉领域中更先进的概念与模型和不同模态数据的成像探测机理与遥感地学知识相结合,设计遥感图像专用的深层语义信息提取模型以实现更智能化的多模态遥感图像配准。

参考文献(References)

- Bay H, Ess A, Tuytelaars T and Van Gool L. 2008. Speeded-up robust features (SURF). *Computer Vision and Image Understanding*, 110(3): 346-359 [DOI: 10.1016/j.cviu.2007.09.014]
- Brown L G. 1992. A survey of image registration techniques. *ACM Computing Surveys*, 24(4): 325-376 [DOI: 10.1145/146370.146374]
- Cao S Y, Yu B N, Luo L, Zhang R M, Chen S J, Li C G and Shen H L. 2023. PCNet: a structure similarity enhancement method for multispectral and multimodal image registration. *Information Fusion*, 94: 200-214 [DOI: 10.1016/j.inffus.2023.02.004]
- Chen H M, Arora M K and Varshney P K. 2003. Mutual information-based image registration for remote sensing data. *International Journal of Remote Sensing*, 24(18): 3701-3706 [DOI: 10.1080/0143116031000117047]
- Chen H M and Varshney P K. 2000. A pyramid approach for multimodality image registration based on mutual information//*Proceedings of the 3rd International Conference on Information Fusion*. Paris, France: IEEE: MOD3/9-MOD315 [DOI: 10.1109/IFIC. 2000.862613]
- Chen R Z, Wang L, Li D R, Chen L and Fu W J. 2019. A survey on the fusion of the navigation and the remote sensing techniques. *Acta Geodaetica et Cartographica Sinica*, 48(12): 1507-1522 (陈锐志, 王磊, 李德仁, 陈亮, 付文举. 2019. 导航与遥感技术融合综述. *测绘学报*, 48(12): 1507-1522) [DOI: 10.11947/j.AGCS. 2019.20190446]
- Chen S W, Xia H, Yang X G and Li X F. 2022. SAR and optical image registration algorithm based on style transfer invariable features. *Systems Engineering and Electronics*, 44(5): 1536-1542 (陈世伟, 夏海, 杨小冈, 李小锋. 2022. 基于风格迁移不变特征的SAR与光学图像配准算法. *系统工程与电子技术*, 44(5): 1536-1542) [DOI: 10.12305/j.issn.1001-506X.2022.05.14]
- Chen T, Guo J F, Han X Z, Xie X L and Xi J X. 2022. Visible and infrared image matching method based on generative adversarial model. *Journal of Zhejiang University (Engineering Science)*, 56(1): 63-74 (陈彤, 郭剑锋, 韩心中, 谢学立, 席建祥. 2022. 基于生成对抗模型的可见光-红外图像匹配方法. *浙江大学学报(工学版)*, 56(1): 63-74) [DOI: 10.3785/j.issn.1008-973X. 2022.01.007]
- Cole-Rhodes A A, Johnson K L, LeMoigne J and Zavorin I. 2003. Multi-resolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient. *IEEE Transactions on Image Processing*, 12(12): 1495-1511 [DOI: 10.1109/TIP. 2003.819237]
- Cui S, Ma A L, Zhang L P, Xu M Z and Zhong Y F. 2022. MAP-Net: SAR and optical image matching via image-based convolutional network with attention mechanism and spatial pyramid aggregated pooling. *IEEE Transactions on Geoscience and Remote Sensing*, 60: #1000513 [DOI: 10.1109/TGRS.2021.3066432]
- de Vos B D, Berendsen F F, Viergever M A, Sokooti H, Staring M and Išgum I. 2019. A deep learning framework for unsupervised affine and deformable image registration. *Medical Image Analysis*, 52: 128-143 [DOI: 10.1016/j.media.2018.11.010]
- Dellinger F, Delon J, Gousseau Y, Michel J and Tupin F. 2015. SAR-

- SIFT: a SIFT-like algorithm for SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 53(1): 453-466 [DOI: 10.1109/TGRS.2014.2323552]
- Dong Y Y, Jiao W L, Long T F, Liu L F, He G J, Gong C J and Guo Y T. 2019. Local deep descriptor for remote sensing image feature matching. *Remote Sensing*, 11(4): #430 [DOI: 10.3390/rs11040430]
- Fan J W, Ye Y X, Li J, Liu G C and Li Y L. 2022. A novel multiscale adaptive binning phase congruency feature for SAR and optical image registration. *IEEE Transactions on Geoscience and Remote Sensing*, 60: #5235216 [DOI: 10.1109/TGRS.2022.3206804]
- Fan J W, Yuan X Y, Li B W, Qi C D, Li Y L and Lu T. 2023. Multi-modal image registration based on local self similarity structural features. *Journal of Xinyang Normal University (Natural Science Edition)*, 36(2): 295-299 (樊建伟, 袁喜悦, 李博文, 祁传达, 李艳灵, 路通. 2023. 基于局部自相似结构特征的多模态影像匹配. 信阳师范学院学报(自然科学版), 36(2): 295-299) [DOI: 10.3969/j.issn.1003-0972.2023.02.022]
- Fan Z L, Zhang L, Liu Y X, Wang Q D and Zlatanova S. 2021. Exploiting high geopositioning accuracy of SAR data to obtain accurate geometric orientation of optical satellite images. *Remote Sensing*, 13(17): #3535 [DOI: 10.3390/rs13173535]
- Feng R T, Du Q Y, Luo H, Shen H F, Li X H and Liu B. 2021. A registration algorithm based on optical flow modification for multi-temporal remote sensing images covering the complex-terrain region. *National Remote Sensing Bulletin*, 25(2): 630-640 (冯蕊涛, 杜清运, 罗恒, 沈焕锋, 李星华, 刘波. 2021. 基于光流校正的复杂地形区多时相遥感影像配准. 遥感学报, 25(2): 630-640) [DOI: 10.11834/jrs.20209280]
- Fu Z T, Qin Q Q, Luo B, Sun H and Wu C. 2018. HOMPC: a local feature descriptor based on the combination of magnitude and phase congruency information for multi-sensor remote sensing images. *Remote Sensing*, 10(8): #1234 [DOI: 10.3390/rs10081234]
- Gong J Y and Ji S P. 2018. Photogrammetry and deep learning. *Acta Geodaetica et Cartographica Sinica*, 47(6): 693-704 (龚健雅, 季顺平. 2018. 摄影测量与深度学习. 测绘学报, 47(6): 693-704) [DOI: 10.11947/AGCS.2018.20170640]
- Harris C and Stephens M. 1988. A combined corner and edge detector// *Proceedings of the 4th Alvey Vision Conference*. Manchester, UK: Alvey Vision Club
- He M M, Guo Q, Li A, Chen J, Chen B and Feng X X. 2018. Automatic fast feature-level image registration for high resolution remote sensing images. *Journal of Remote Sensing*, 22(2): 277-292 (何梦梦, 郭擎, 李安, 陈俊, 陈勃, 冯旭祥. 2018. 特征级高分辨率遥感图像快速自动配准. 遥感学报, 22(2): 277-292) [DOI: 10.11834/jrs.20186420]
- Hong Y M, Leng C C, Zhang X Y, Peng J Y, Jiao L C and Basu A. 2022. Max-index based local self-similarity descriptor for robust multi-modal image registration. *IEEE Geoscience and Remote Sensing Letters*, 19: #4505805 [DOI: 10.1109/LGRS.2022.3156622]
- Huang L, Zhu B, You J Q, Liao Y F and Ge X M. 2023. Registration of aerial images and LiDAR data via dense structural features. *Remote Sensing Information*, 38(1): 56-62 (黄磊, 朱柏, 游晋卿, 廖永福, 葛旭明. 2023. 利用稠密结构特征的航空影像与LiDAR数据配准. 遥感信息, 38(1): 56-62) [DOI: 10.20091/j.cnki.1000-3177.2023.01.008]
- Huang M Y, Xu Y, Qian L X, Shi W L, Zhang Y Q, Bao W, Wang N, Liu X J and Xiang X S. 2021. The QXS-SAROPT dataset for deep learning in SAR-optical data fusion [EB/OL]. [2023-10-17]. <https://arxiv.org/pdf/2103.08259.pdf>
- Hughes L H, Marcos D, Lobry S, Tuia D and Schmitt M. 2020. A deep learning framework for matching of SAR and optical imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 169: 166-179 [DOI: 10.1016/j.isprs.2020.09.012]
- Hughes L H, Schmitt M, Mou L C, Wang Y Y and Zhu X X. 2018. Identifying corresponding patches in SAR and optical images with a pseudo-Siamese CNN. *IEEE Geoscience and Remote Sensing Letters*, 15(5): 784-788 [DOI: 10.1109/LGRS.2018.2799232]
- Inglada J. 2002. Similarity measures for multisensor remote sensing images//*IEEE International Geoscience and Remote Sensing Symposium*. Toronto, Canada: IEEE: 104-106 [DOI: 10.1109/IGARSS.2002.1024955]
- Jiang X Y, Ma J Y, Xiao G B, Shao Z F and Guo X J. 2021. A review of multimodal image matching: methods and applications. *Information Fusion*, 73: 22-71 [DOI: 10.1016/j.inffus.2021.02.012]
- Jiang Y H, Zhang G, Chen P, Li D R, Tang X M and Huang W C. 2015. Systematic error compensation based on a rational function model for Ziyan1-02C. *IEEE Transactions on Geoscience and Remote Sensing*, 53(7): 3985-3995 [DOI: 10.1109/TGRS.2015.2388700]
- Kulkarni S C and Rege P P. 2020. Pixel level fusion techniques for SAR and optical images: a review. *Information Fusion*, 59: 13-29 [DOI: 10.1016/j.inffus.2020.01.003]
- Lan C Z, Lu W J, Yu J M and Xu Q. 2021. Deep learning algorithm for feature matching of cross modality remote sensing images. *Acta Geodaetica et Cartographica Sinica*, 50(2): 189-202 (蓝朝植, 卢万杰, 于君明, 徐青. 2021. 异源遥感影像特征匹配的深度学习算法. 测绘学报, 50(2): 189-202) [DOI: 10.11947/j.AGCS.2021.20200048]
- Li D R, Zhang G, Jiang Y H, Shen X and Liu W L. 2022. Opportunities and challenges of geo-spatial information science from the perspective of big data. *Big Data Research*, 8(2): 3-14 (李德仁, 张过, 蒋永华, 沈欣, 刘伟玲. 2022. 论大数据视角下的地球空间信息学的机遇与挑战. 大数据, 8(2): 3-14) [DOI: 10.11959/j.issn.2096-0271.2022012]
- Li J Y, Hu Q W and Ai M Y. 2020. RIFT: multi-modal image matching based on radiation-variation insensitive feature transform. *IEEE*

- Transactions on Image Processing, 29: 3296-3310 [DOI: 10.1109/TIP.2019.2959244]
- Li L Z, Han L, Ding M T and Cao H Y. 2023. Multimodal image fusion framework for end-to-end remote sensing image registration. IEEE Transactions on Geoscience and Remote Sensing, 61: #5607214 [DOI: 10.1109/TGRS.2023.3247642]
- Li L Z, Han L, Ding M T, Cao H Y and Hu H J. 2021. A deep learning semantic template matching framework for remote sensing image registration. ISPRS Journal of Photogrammetry and Remote Sensing, 181: 205-217 [DOI: 10.1016/j.isprsjprs.2021.09.012]
- Li L Z, Han L, Ding M T, Liu Z H and Cao H Y. 2022a. Remote sensing image registration based on deep learning regression model. IEEE Geoscience and Remote Sensing Letters, 19: #8002905 [DOI: 10.1109/LGRS.2020.3032439]
- Li L Z, Han L and Ye Y X. 2022b. Self-supervised keypoint detection and cross-fusion matching networks for multimodal remote sensing image registration. Remote Sensing, 14(15): #3599 [DOI: 10.3390/rs14153599]
- Li P, Jiang G, Ma Q L, Xue W F and Yang W H. 2021. A hybrid model combining tensor and mutual information for multi-modal image registration. Acta Geodaetica et Cartographica Sinica, 50(7): 916-929 (李培, 姜刚, 马千里, 薛万峰, 杨伟华. 2021. 结合张量与互信息的混合模型多模态图像配准方法. 测绘学报, 50(7): 916-929) [DOI: 10.11947/j.AGCS.2021.20200492]
- Li Q, Qu G Z and Li Z L. 2013. Matching between SAR images and optical images based on HOG descriptor//Proceedings of 2013 IET International Radar Conference. Xi'an, China: IET: 1-4 [DOI: 10.1049/cp.2013.0405]
- Li X M, Zheng L and Hu Z Y. 2006. SIFT based automatic registration of remotely-sensed imagery. Journal of Remote Sensing, 10(6): 885-892 (李晓明, 郑链, 胡占义. 2006. 基于SIFT特征的遥感影像自动配准. 遥感学报, 10(6): 885-892)
- Li Y, Wang L Z, Liu X L, Chu Q N and Yang X H. 2022c. A review of spatiotemporal super-resolution mapping for remote sensing data fusion. IEEE Journal on Miniaturization for Air and Space Systems, 3(1): 9-18 [DOI: 10.1109/JMASS.2021.3091837]
- Li Y H, Liu Y D, Su X P, Luo X M and Yao L. 2022. Review of infrared and visible image registration. Infrared Technology, 44(7): 641-651 (李云红, 刘宇栋, 苏雪平, 罗雪敏, 姚兰. 2022. 红外与可见光图像配准技术研究综述. 红外技术, 44(7): 641-651)
- Liang C B, Dong Y Y, Zhao C J and Sun Z G. 2023. A coarse-to-fine feature match network using Transformers for remote sensing image registration. Remote Sensing, 15(13): #3243 [DOI: 10.3390/rs15133243]
- Lin Z K, Liu W, Niu C Y, Gao G and Lu W J. 2023. Synthetic aperture radar image change detection based on difference image construction of log-hyperbolic cosine ratio and multi-region feature convolution extreme learning machine. Acta Optica Sinica, 43(12): #1228001 (林志康, 刘伟, 牛朝阳, 高贵, 卢万杰. 2023. 基于对数双曲余弦比差异图构造与多区域特征卷积极限学习机的合成孔径雷达图像变化检测. 光学学报, 43(12): #1228001) [DOI: 10.3788/AOS221491]
- Liu Y, Qi H and Peng S Y. 2023. Optical and SAR images matching based on phase structure convolutional features. IEEE Geoscience and Remote Sensing Letters, 20: #6007805 [DOI: 10.1109/LGRS.2023.3298687]
- Lowe D G. 2004. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2): 91-110 [DOI: 10.1023/B:VISI.0000029664.99615.94]
- Lu H P and Gao L. 2009. Improved algorithm for multi-source remote sensing images based on cross correlation. Aerospace Control, 27(2): 18-21, 25 (陆和平, 高磊. 2009. 基于互相关的多源遥感图像匹配的改进算法. 航天控制, 27(2): 18-21, 25) [DOI: 10.16804/j.cnki.issn1006-3242.2009.02.004]
- Ma J Y, Jiang X Y, Fan A X, Jiang J J and Yan J C. 2021. Image matching from handcrafted to deep features: a survey. International Journal of Computer Vision, 129(1): 23-79 [DOI: 10.1007/s11263-020-01359-2]
- Ma W P, Wen Z L, Wu Y, Jiao L C, Gong M G, Zheng Y F and Liu L. 2017. Remote sensing image registration with modified SIFT and enhanced feature matching. IEEE Geoscience and Remote Sensing Letters, 14(1): 3-7 [DOI: 10.1109/LGRS.2016.2600858]
- Ma W P, Zhang J, Wu Y, Jiao L C, Zhu H and Zhao W. 2019. A novel two-step registration method for remote sensing images based on deep and local features. IEEE Transactions on Geoscience and Remote Sensing, 57(7): 4834-4843 [DOI: 10.1109/TGRS.2019.2893310]
- Merkle N, Auer S, Müller R and Reinartz P. 2018. Exploring the potential of conditional adversarial networks for optical and SAR image matching. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11(6): 1811-1820 [DOI: 10.1109/JSTARS.2018.2803212]
- Merkle N, Luo W J, Auer S, Müller R and Urtasun R. 2017. Exploiting deep matching and SAR data for the geo-localization accuracy improvement of optical satellite images. Remote Sensing, 9(6): #586 [DOI: 10.3390/rs9060586]
- Mikolajczyk K and Schmid C. 2005. A performance evaluation of local descriptors. IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(10): 1615-1630 [DOI: 10.1109/TPAMI.2005.188]
- Moravec H P. 1980. Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover. Stanford, USA: Stanford University
- Nan K, Qi H and Ye Y X. 2019. A template matching method of multimodal remote sensing images based on deep convolutional feature representation. Acta Geodaetica et Cartographica Sinica, 48(6): 727-736 (南轲, 齐华, 叶沅鑫. 2019. 深度卷积特征表达的多模

- 态遥感影像模板匹配方法. 测绘学报, 48(6): 727-736 [DOI: 10.11947/j.AGCS.2019.20180432]
- Parmehr E G, Fraser C S, Zhang C S and Leach J. 2014. Automatic registration of optical imagery with 3D LiDAR data using statistical similarity. *ISPRS Journal of Photogrammetry and Remote Sensing*, 88: 28-40 [DOI: 10.1016/j.isprsjprs.2013.11.015]
- Quan D, Wang S, Gu Y, Lei R Q, Yang B W, Wei S W, Hou B and Jiao L C. 2022. Deep feature correlation learning for multi-modal remote sensing image registration. *IEEE Transactions on Geoscience and Remote Sensing*, 60: #4708216 [DOI: 10.1109/TGRS.2022.3187015]
- Quan D, Wei H Y, Wang S, Gu Y, Hou B and Jiao L C. 2023. A novel coarse-to-fine deep learning registration framework for multimodal remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 61: #5108316 [DOI: 10.1109/TGRS.2023.3306042]
- Reddy B S and Chatterji B N. 1996. An FFT-based technique for translation, rotation, and scale-invariant image registration. *IEEE Transactions on Image Processing*, 5(8): 1266-1271 [DOI: 10.1109/83.506761]
- Rosten E, Porter R and Drummond T. 2010. Faster and better: a machine learning approach to corner detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(1): 105-119 [DOI: 10.1109/TPAMI.2008.275]
- Schmitt M, Hughes L H and Zhu X X. 2018. The SEN1-2 dataset for deep learning in SAR-optical data fusion. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-1: 141-146 [DOI: 10.5194/isprs-annals-IV-1-141-2018]
- Sedaghat A and Ebadi H. 2015a. Remote sensing image matching based on adaptive binning SIFT descriptor. *IEEE Transactions on Geoscience and Remote Sensing*, 53(10): 5283-5293 [DOI: 10.1109/TGRS.2015.2420659]
- Sedaghat A and Ebadi H. 2015b. Distinctive order based self-similarity descriptor for multi-sensor remote sensing image matching. *ISPRS Journal of Photogrammetry and Remote Sensing*, 108: 62-71 [DOI: 10.1016/j.isprsjprs.2015.06.003]
- Sedaghat A and Mohammadi N. 2019. Illumination-robust remote sensing image matching based on oriented self-similarity. *ISPRS Journal of Photogrammetry and Remote Sensing*, 153: 21-35 [DOI: 10.1016/j.isprsjprs.2019.04.018]
- Sedaghat A, Mokhtarzade M and Ebadi H. 2011. Uniform robust scale-invariant feature matching for optical remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 49(11): 4516-4527 [DOI: 10.1109/TGRS.2011.2144607]
- Shao Z F, Cheng T, Li D R, Wu C Z and Guo S. 2023. Development and applications of aerial-ground collaborative mobile intelligent service platform. *Geomatics and Information Science of Wuhan University*, 48(5): 749-755 (邵振峰, 程涛, 李德仁, 吴长枝, 郭晟. 2023. 空地协同移动智能服务平台研发及典型应用. 武汉大学学报(信息科学版), 48(5): 749-755 [DOI: 10.13203/j.whugis.20220520])
- Shen X, Liu B and Li Q Q. 2017. Correcting bias in the rational polynomial coefficients of satellite imagery using thin-plate smoothing splines. *ISPRS Journal of Photogrammetry and Remote Sensing*, 125: 125-131 [DOI: 10.1016/j.isprsjprs.2017.01.007]
- Shi Z H, Wu C W, Li C J, You Z Z, Wang Q and Ma C C. 2023. Object detection techniques based on deep learning for aerial remote sensing images: a survey. *Journal of Image and Graphics*, 28(9): 2616-2643 (石争浩, 仵晨伟, 李成建, 尤珍臻, 王泉, 马城城. 2023. 航空遥感图像深度学习目标检测技术研究进展. 中国图象图形学报, 28(9): 2616-2643 [DOI: 10.11834/jig.221085])
- Song Z L, Zhang J Q, Xiong L and He H. 2021. Multimodal image registration algorithm using style transfer and feature points. *Remote Sensing Information*, 36(1): 1-6 (宋智礼, 张家齐, 熊亮, 何鸽. 2021. 利用风格迁移和特征点的多模态图像配准算法. 遥感信息, 36(1): 1-6 [DOI: 10.3969/j.issn.1000-3177.2021.01.001])
- Su H J. 2022. Dimensionality reduction for hyperspectral remote sensing: advances, challenges, and prospects. *National Remote Sensing Bulletin*, 26(8): 1504-1529 (苏红军. 2022. 高光谱遥感影像降维: 进展、挑战与展望. 遥感学报, 26(8): 1504-1529 [DOI: 10.11834/jrs.20210354])
- Sui H G, Liu C, Gan Z, Jiang Z J and Xu C. 2022. Overview of multimodal remote sensing image matching methods. *Acta Geodaetica et Cartographica Sinica*, 51(9): 1848-1861 (眭海刚, 刘畅, 干哲, 江政杰, 徐川. 2022. 多模态遥感图像匹配方法综述. 测绘学报, 51(9): 1848-1861 [DOI: 10.11947/j.AGCS.2022.20220126])
- Sun W W, Yang G, Chen C, Chang M H, Huang K, Meng X Z and Liu L Y. 2020. Development status and literature analysis of China's earth observation remote sensing satellites. *Journal of Remote Sensing*, 24(5): 479-510 (孙伟伟, 杨刚, 陈超, 常明会, 黄可, 孟祥珍, 刘良云. 2020. 中国地球观测遥感卫星发展现状及文献分析. 遥感学报, 24(5): 479-510 [DOI: 10.11834/jrs.20209464])
- Suri S and Reinartz P. 2010. Mutual-information-based registration of TerraSAR-X and Ikonos imagery in urban areas. *IEEE Transactions on Geoscience and Remote Sensing*, 48(2): 939-949 [DOI: 10.1109/TGRS.2009.2034842]
- Tang L F, Zhang H, Xu H and Ma J Y. 2023. Deep learning-based image fusion: a survey. *Journal of Image and Graphics*, 28(1): 3-36 (唐霖峰, 张浩, 徐涵, 马佳义. 2023. 基于深度学习的图像融合方法综述. 中国图象图形学报, 28(1): 3-36 [DOI: 10.11834/jig.220422])
- Tang L F, Xiang X Y, Zhang H, Gong M Q and Ma J Y. 2023. DIVFusion: darkness-free infrared and visible image fusion. *Information Fusion*, 91: 477-493 [DOI: 10.1016/j.inffus.2022.10.034]
- Tola E, Lepetit V and Fua P. 2010. Daisy: an efficient dense descriptor applied to wide-baseline stereo. *IEEE Transactions on Pattern*

- Analysis and Machine Intelligence, 32(5): 815-830 [DOI: 10.1109/TPAMI.2009.77]
- Tong X H, Ye Z, Xu Y S, Liu S J, Li L Y, Xie H and Li T P. 2015. A novel subpixel phase correlation method using singular value decomposition and unified random sample consensus. *IEEE Transactions on Geoscience and Remote Sensing*, 53(8): 4143-4156 [DOI: 10.1109/TGRS.2015.2391999]
- Wang L, Xu X, Yu Y, Yang R, Gui R, Xu Z Z and Pu F L. 2019. SAR-to-optical image translation using supervised cycle-consistent adversarial networks. *IEEE Access*, 7: 129136-129149 [DOI: 10.1109/ACCESS.2019.2939649]
- Wang S, Quan D, Liang X F, Ning M D, Guo Y H and Jiao L C. 2018. A deep learning framework for remote sensing image registration. *ISPRS Journal of Photogrammetry and Remote Sensing*, 145: 148-164 [DOI: 10.1016/j.isprsjprs.2017.12.012]
- Wang Y Y and Zhu X X. 2018. The sarptical dataset for joint analysis of SAR and optical image in dense urban area//IGARSS 2018 IEEE International Geoscience and Remote Sensing Symposium. Valencia, Spain: IEEE: 6840-6843 [DOI: 10.1109/IGARSS.2018.8518298]
- Wells III W M, Viola P, Atsumi H, Nakajima S and Kikinis R. 1996. Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis*, 1(1): 35-51 [DOI: 10.1016/S1361-8415(01)80004-9]
- Wu J, Rao Y, Hu Y J, Cheng M M and Peng Z Y. 2016. Automatic registration of single aerial image with LiDAR data based on "pin-hole" imaging simulation and iterative computation. *Journal of Remote Sensing*, 20(1): 80-93 (吴军, 饶云, 胡彦君, 程门门, 彭智勇. 2016. "针孔"模拟成像下的单航空影像与LiDAR点云配准. *遥感学报*, 20(1): 80-93) [DOI: 10.11834/jrs.20165070]
- Xiang D L, Xu Y H, Cheng J D, Hu C B and Sun X K. 2022. An algorithm based on a feature interaction-based keypoint detector and Sim-CSPNet for SAR image registration. *Journal of Radars*, 11(6): 1081-1097 (项德良, 徐益豪, 程建达, 胡黎彬, 孙晓坤. 2022. 一种基于特征交汇关键点检测和Sim-CSPNet的SAR图像配准算法. *雷达学报*, 11(6): 1081-1097) [DOI: 10.12000/JR22110]
- Xiang Y M, Jiao N G, Wang F and You H J. 2022. A robust two-stage registration algorithm for large optical and SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 60: #5218615 [DOI: 10.1109/TGRS.2021.3133863]
- Xiang Y M, Tao R S, Wan L, Wang F and You H J. 2020a. OS-PC: combining feature representation and 3-D phase correlation for sub-pixel optical and SAR image registration. *IEEE Transactions on Geoscience and Remote Sensing*, 58(9): 6451-6466 [DOI: 10.1109/TGRS.2020.2976865]
- Xiang Y M, Tao R S, Wang F, You H J and Han B. 2020b. Automatic registration of optical and SAR images via improved phase congruency model. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13: 5847-5861 [DOI: 10.1109/JSTARS.2020.3026162]
- Xiang Y M, Wang F and You H J. 2018. OS-SIFT: a robust SIFT-like algorithm for high-resolution optical-to-SAR image registration in suburban areas. *IEEE Transactions on Geoscience and Remote Sensing*, 56(6): 3078-3090 [DOI: 10.1109/TGRS.2018.2790483]
- Xiang Y M, Wang X Q, Wang F, You H J, Qiu X L and Fu K. 2023. A global-to-local algorithm for high-resolution optical and SAR image registration. *IEEE Transactions on Geoscience and Remote Sensing*, 61: #5215320 [DOI: 10.1109/TGRS.2023.3309855]
- Xiong X, Jin G W, Xu Q and Zhang H M. 2021. Self-similarity features for multimodal remote sensing image matching. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14: 12440-12454 [DOI: 10.1109/JSTARS.2021.3131489]
- Xiong X, Jin G W, Xu Q, Zhang H M, Wang L M and Wu K. 2022. Robust registration algorithm for optical and SAR images based on adjacent self-similarity feature. *IEEE Transactions on Geoscience and Remote Sensing*, 60: #3197357 [DOI: 10.1109/TGRS.2022.3197357]
- Xiong X, Xu Q, Jin G W, Zhang H M and Gao X. 2020. Rank-based local self-similarity descriptor for optical-to-SAR image matching. *IEEE Geoscience and Remote Sensing Letters*, 17(10): 1742-1746 [DOI: 10.1109/LGRS.2019.2955153]
- Xu W Y, Yuan X H, Hu Q W and Li J Y. 2023. SAR-optical feature matching: a large-scale patch dataset and a deep local descriptor. *International Journal of Applied Earth Observation and Geoinformation*, 122: #103433 [DOI: 10.1016/j.jag.2023.103433]
- Xu X C, Li X, Liu X P, Shen H F and Shi Q. 2016. Multimodal registration of remotely sensed images based on Jeffrey's divergence. *ISPRS Journal of Photogrammetry and Remote Sensing*, 122: 97-115 [DOI: 10.1016/j.isprsjprs.2016.10.005]
- Yang B C, Wang P, Li X Y, Li L L and Cao X F. 2022. Infrared and visible light image registration based on modal conversion and robust features. *Laser and Optoelectronics Progress*, 59(4): #0410013 (杨冰超, 王鹏, 李晓艳, 李亮亮, 曹小芳. 2022. 基于模态转换结合鲁棒特征的红外图像和可见光图像配准. *激光与光电子学进展*, 59(4): #0410013) [DOI: 10.3788/LOP202259.0410013]
- Yao H G, Wang C, Yu J, Bai X J and Li Y. 2020. Recognition of small-target ships in complex satellite images. *Journal of Remote Sensing*, 24(2): 116-125 (姚红革, 王诚, 喻钧, 白小军, 李蔚. 2020. 复杂卫星图像中的小目标船舶识别. *遥感学报*, 24(2): 116-125) [DOI: 10.11834/jrs.20208238]
- Yao Y X, Zhang Y J, Wan Y, Liu X Y and Guo H Y. 2021. Heterologous images matching considering anisotropic weighted moment and absolute phase orientation. *Geomatics and Information Science of Wuhan University*, 46(11): 1727-1736 (姚永祥, 张永军, 万一, 刘欣怡, 郭浩宇. 2021. 顾及各向异性加权力矩与绝对相位方向的异源影像匹配. *武汉大学学报(信息科学版)*, 46

- (11): 1727-1736 [DOI: 10.13203/j.whugis20200702]
- Yao Y X, Zhang Y J, Wan Y, Liu X Y, Yan X H and Li J Y. 2022. Multi-modal remote sensing image matching considering Co-occurrence filter. *IEEE Transactions on Image Processing*, 31: 2584-2597 [DOI: 10.1109/TIP.2022.3157450]
- Ye F M, Su Y F, Xiao H, Zhao X Q and Min W D. 2018b. Remote sensing image registration using convolutional neural network features. *IEEE Geoscience and Remote Sensing Letters*, 15 (2): 232-236 [DOI: 10.1109/LGRS.2017.2781741]
- Ye Y X, Bruzzone L, Shan J, Bovolo F and Zhu Q. 2019. Fast and robust matching for multimodal remote sensing image registration. *IEEE Transactions on Geoscience and Remote Sensing*, 57 (11): 9059-9070 [DOI: 10.1109/TGRS.2019.2924684]
- Ye Y X, Shan J, Peng J W, Xiong J X and Li W. 2014. Automated multispectral remote sensing image registration using local self-similarity. *Acta Geodaetica et Cartographica Sinica*, 43 (3): 268-275 (叶沅鑫, 单杰, 彭剑威, 熊金鑫, 李维. 2014. 利用局部自相似进行多光谱遥感图像自动配准. *测绘学报*, 43 (3): 268-275) [DOI: 10.13485/j.cnki.11-2089.2014.0039]
- Ye Y X, Shan J, Bruzzone L and Shen L. 2017b. Robust registration of multimodal remote sensing images based on structural similarity. *IEEE Transactions on Geoscience and Remote Sensing*, 55 (5): 2941-2958 [DOI: 10.1109/TGRS.2017.2656380]
- Ye Y X, Shan J, Hao S Y, Bruzzone L and Qin Y. 2018a. A local phase based invariant feature for remote sensing image matching. *ISPRS Journal of Photogrammetry and Remote Sensing*, 142: 205-221 [DOI: 10.1016/j.isprsjprs.2018.06.010]
- Ye Y X, Shen L, Hao M, Wang J C and Xu Z. 2017a. Robust optical-to-SAR image matching based on shape properties. *IEEE Geoscience and Remote Sensing Letters*, 14 (4): 564-568 [DOI: 10.1109/LGRS.2017.2660067]
- Ye Y X, Tan X, Sun M M and Wang M M. 2021. High-resolution remote sensing image classification based on improved DeepLabV3 network. *Bulletin of Surveying and Mapping*, (4): 40-44 (叶沅鑫, 谭鑫, 孙苗苗, 王蒙蒙. 2021. 基于增强DeepLabV3网络的高分辨率遥感影像分类. *测绘通报*, (4): 40-44) [DOI: 10.13474/j.cnki.11-2246.2021.0108]
- Ye Y X, Tang T F, Zhu B, Yang C, Li B and Hao S Y. 2022b. A multi-scale framework with unsupervised learning for remote sensing image registration. *IEEE Transactions on Geoscience and Remote Sensing*, 60: #5622215 [DOI: 10.1109/TGRS.2022.3167644]
- Ye Y X, Zhu B, Tang T F, Yang C, Xu Q Z and Zhang G. 2022a. A robust multimodal remote sensing image registration method and system using steerable filters with first-and second-order gradients. *ISPRS Journal of Photogrammetry and Remote Sensing*, 188: 331-350 [DOI: 10.1016/j.isprsjprs.2022.04.011]
- Yu L, Zhang D R and Holden E J. 2008. A fast and fully automatic registration approach based on point features for multi-source remote sensing images. *Computers and Geosciences*, 34 (7): 838-848 [DOI: 10.1016/j.cageo.2007.10.005]
- Yu X C, Lyu Z H, Hu D, Zhang L B and Xu J D. 2014. Optical remote sensing image registration based on SG-SIFT. *Journal of Beijing University of Posts and Telecommunications*, 37(6): 17-22 (余先川, 吕中华, 胡丹, 张立保, 徐金东. 2014. 基于SG-SIFT的光学遥感影像配准. *北京邮电大学学报*, 37(6): 17-22) [DOI: 10.13190/j.jbupt.2014.06.004]
- Zhang H, Lei L, Ni W P, Tang T, Wu J Z, Xiang D L and Kuang G Y. 2022c. Optical and SAR image matching using pixelwise deep dense features. *IEEE Geoscience and Remote Sensing Letters*, 19: #6000705 [DOI: 10.1109/LGRS.2020.3039473]
- Zhang H, Lei L, Ni W P, Tang T, Wu J Z, Xiang D L and Kuang G Y. 2022a. Explore better network framework for high-resolution optical and SAR image matching. *IEEE Transactions on Geoscience and Remote Sensing*, 60: #4704418 [DOI: 10.1109/TGRS. 2021. 3126939]
- Zhang H, Ni W P, Yan W D, Xiang D L, Wu J Z, Yang X L and Bian H. 2019a. Registration of multimodal remote sensing image based on deep fully convolutional neural network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(8): 3028-3042 [DOI: 10.1109/JSTARS.2019.2916560]
- Zhang J, Ma W P, Wu Y and Jiao L C. 2019b. Multimodal remote sensing image registration based on image transfer and local features. *IEEE Geoscience and Remote Sensing Letters*, 16(8): 1210-1214 [DOI: 10.1109/LGRS.2019.2896341]
- Zhang Y J, Zhang Z X and Gong J Y. 2021. Generalized photogrammetry of spaceborne, airborne and terrestrial multi source remote sensing datasets. *Acta Geodaetica et Cartographica Sinica*, 50 (1): 1-11 (张永军, 张祖勋, 龚健雅. 2021. 天空地多源遥感数据的广义摄影测量学. *测绘学报*, 50 (1): 1-11) [DOI: 10.11947/j. AGCS.2021.20200245]
- Zhang Y X, Liu Y X, Zhang H M and Ma G R. 2022b. Multimodal remote sensing image matching combining learning features and delaunay triangulation. *IEEE Transactions on Geoscience and Remote Sensing*, 60: #5635517 [DOI: 10.1109/TGRS. 2022. 3229366]
- Zhao J W, Yang D F, Li Y F, Xiao P and Yang J L. 2022. Intelligent matching method for heterogeneous remote sensing images based on style transfer. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15: 6723-6731 [DOI: 10.1109/JSTARS.2022.3197748]
- Zheng Z D, Wei Y C and Yang Y. 2020. University-1652: a multi-view multi-source benchmark for drone-based geo-localization//Proceedings of the 28th ACM International Conference on Multimedia. Seattle, USA: ACM: 1395-1403 [DOI: 10.1145/3394171. 3413896]
- Zhou L, Ye Y X, Tang T F, Nan K and Qin Y. 2022. Robust matching

for SAR and optical images using multiscale convolutional gradient features. *IEEE Geoscience and Remote Sensing Letters*, 19: #4017605 [DOI: 10.1109/LGRS.2021.3105567]

Zhu B, Yang C, Dai J K, Fan J W, Qin Y and Ye Y X. 2023. R₂FD₂: fast and robust matching of multimodal remote sensing images via repeatable feature detector and rotation-invariant feature descriptor. *IEEE Transactions on Geoscience and Remote Sensing*, 61: #5606115 [DOI: 10.1109/TGRS.2023.3264610]

Zhu B, Ye Y X, Zhou L, Li Z L and Yin G F. 2021. Robust registration of aerial images and LiDAR data using spatial constraints and

Gabor structural features. *ISPRS Journal of Photogrammetry and Remote Sensing*, 181: 129-147 [DOI: 10.1016/j.isprsjprs.2021.09.010]

作者简介

朱柏,男,博士研究生,主要研究方向为遥感图像处理和多模态图像配准。E-mail:kevin_zhub@163.com

叶沅鑫,通信作者,男,教授,博士生导师,主要研究方向为遥感图像处理、图像配准、变化检测和目标检测。
E-mail:yeyuanxin@home.swjtu.edu.cn