

中图分类号: TP391 文献标识码: A 文章编号: 1006-8961(XXXX)XX-0001-31

论文引用格式: Zhao Zilu, Wang Feng, Jiao Niangang, You Hongjian. From Knowledge-Driven to Data-Driven: A Comprehensive Review of Remote Sensing Image Registration[J/OL]. Journal of Image and Graphics, XXXX:1-31. DOI: 10.11834/jig.260140. (赵子璐, 王峰, 焦念刚, 尤红建. 遥感图像配准技术研究综述: 从知识驱动到数据驱动[J/OL]. 中国图象图形学报, XXXX:1-31. DOI: 10.11834/jig.260140.) [DOI: 10.11834/jig.260140]

## 遥感图像配准技术研究综述: 从知识驱动到数据驱动

赵子璐<sup>1,2,3</sup>, 王峰<sup>1,2</sup>, 焦念刚<sup>1,2</sup>, 尤红建<sup>1,2,3</sup>

1. 中国科学院空天信息创新研究院, 北京 100194; 2. 中国科学院空间信息处理与应用系统重点实验室 北京 100094; 3. 中国科学院大学 电子电气与通信工程学院, 北京 101408

**摘要:** 遥感图像配准作为遥感图像处理中的基础性与关键性任务, 始终是国内外学者的研究热点, 具有重要的理论意义与应用价值。然而, 遥感图像常面临显著几何畸变、非线性辐射差异、复杂地物遮挡及噪声干扰等挑战, 制约了其高精度鲁棒配准的实现。本文系统梳理了遥感图像配准算法: 首先, 概述配准的基本概念与技术框架; 其次, 将现有方法划分为两类进行归纳梳理——基于知识驱动专家设计的方法与基于数据驱动深度学习的方法。前者依托手工设计特征及变换模型, 通过显式特征提取与迭代优化实现传统场景下的可解释性配准, 典型代表为尺度不变特征变换及其改进算法; 后者利用深度学习网络自动学习图像对间的特征或映射关系, 提取高级特征用于配准或端到端预测变换模型, 有效应对显著几何畸变、非线性辐射差异等复杂场景, 典型方法包括基于卷积神经网络的特征匹配网络及其改进网络。最后, 总结现有方法的局限性, 并展望遥感图像配准领域的未来发展趋势与面临的关键挑战。

**关键词:** 遥感图像配准; 知识驱动; 数据驱动; 特征匹配; 深度学习

## From Knowledge-Driven to Data-Driven: A Comprehensive Review of Remote Sensing Image Registration

Zhao Zilu<sup>1,2,3</sup>, Wang Feng<sup>1,2</sup>, Jiao Niangang<sup>1,2</sup>, You Hongjian<sup>1,2,3</sup>

1. Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China; 2. Key Laboratory of Technology in Geo-Spatial Information Processing and Application System, Chinese Academy of Sciences, Beijing 100190, China; 3. School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 101408, China

**Abstract:** As a fundamental and critical task in remote sensing image processing, image registration has consistently been a hot research topic among scholars both domestically and internationally, holding significant theoretical and practical value. This task aims to geometrically align two or more images of the same scene that were acquired at different times, by different sensors, or from different viewpoints, thereby providing a reliable data foundation for subsequent applications such as image mosaicking, image fusion, object detection, change detection, and guidance matching. However, remote sensing images often face significant challenges such as geometric distortion, nonlinear radiometric differences, complex object occlusion, and various types of noise interference. These factors interact with one another, making it difficult to accurately extract and match corresponding features between images. This severely hinders the achievement of high-precision, robust registration and poses greater challenges to the adaptability and robustness of existing registration meth-

收稿日期: 2026-03-19; 修回日期: 2026-06-16

**基金项目:** 的规范中文全称(项目编号: ……)(不同基金之间用分号隔开) Supported by: 基金项目的英文全称(主要基金项目的中英文名称可在学报网站下载中心查找核对)

ods in complex scenes. This paper systematically reviews remote sensing image registration algorithms. First, it outlines the basic concepts and technical framework of remote sensing image registration. Second, existing registration methods are categorized into two major groups for systematic analysis: knowledge-driven expert design methods and data-driven deep learning methods. Knowledge-driven, expert-designed methods rely on manually designed features and transformation models. Through explicit feature extraction and iterative optimization, they achieve interpretable registration in traditional scenarios. These methods can be further subdivided into template-based registration methods, feature-based registration methods, and hybrid registration methods. Template-based registration methods select a template window on the reference image, use its center as the point to be matched, and search for matches on the target image based on similarity criteria to identify corresponding points. These points are then used to solve for geometric transformation parameters, thereby completing remote sensing image registration. Feature-based registration methods identify and effectively describe salient features in remote sensing images. These features are used to establish correspondences between them and, subsequently, to construct geometric transformation relationships between the images. Hybrid registration methods leverage the advantages of multiple techniques to perform remote sensing image registration, thereby compensating for the shortcomings of single-method approaches. Most of these methods integrate various types of image information, such as gradient and phase information, to improve upon classical methods—such as scale-invariant feature transformation—in either the spatial domain or the transform domain, or to enhance registration accuracy by integrating multiple registration strategies. However, knowledge-driven, expert-designed methods suffer from issues such as heavy reliance on specialized knowledge, poor adaptability to complex scenes, and significant limitations in feature selection, making it difficult to meet the registration demands of large-scale, diverse remote sensing images. In recent years, with the continuous advancement of deep learning technology, numerous deep learning network architectures have become increasingly prevalent in the field of image processing, presenting new opportunities for remote sensing image registration methods. Consequently, data-driven deep learning registration methods have emerged. These methods can be further categorized into ensemble learning methods, end-to-end learning methods, style transfer methods, and large-model-driven methods. Ensemble learning methods integrate deep learning networks into template-based or feature-based registration frameworks to generate more robust and efficient feature representations or similarity metrics, thereby facilitating subsequent registration. End-to-end learning methods utilize deep learning networks to perform end-to-end learning from image pair inputs to target outputs, employing joint optimization to accomplish feature extraction, matching, or direct prediction of geometric transformation parameters between images. Style transfer methods use deep learning models, such as generative adversarial networks, to perform style transfer on remote sensing images, significantly reducing radiometric or textural differences between multimodal remote sensing images and creating more consistent image characteristics for subsequent registration operations. Large-model-driven methods leverage the powerful feature representation capabilities and cross-modal generalization abilities of pre-trained large models, providing a novel solution path for addressing challenges such as complex geometric distortions and nonlinear radiometric differences in remote sensing images. By learning directly from the data, these methods can solve image matching and registration problems more efficiently. They utilize deep learning networks to automatically learn feature or mapping relationships between image pairs, extract high-level features for registration, or directly predict transformation models end-to-end, thereby effectively addressing complex scenarios such as significant geometric distortion and nonlinear radiometric differences. Furthermore, this paper systematically reviews and analyzes datasets for multimodal image registration tasks. By summarizing representative datasets in the field of multimodal registration and focusing on their modal composition, data scale, spatial resolution, covered scenarios, and applicability, it provides an effective reference for researchers to select appropriate datasets for registration tasks, thereby facilitating systematic analysis and fair comparison of different methods under unified conditions in future research. Finally, based on the aforementioned review of methods in the field of remote sensing image registration, this paper outlines future development trends in the field.

**Key words:** remote sensing image registration; knowledge-driven; data-driven; feature matching; deep learning

论文引用格式: [DOI: 10. 11834/jig. 260140]

## 0 引言

传感器成像技术的持续创新使得遥感图像在数量与质量上都有了显著提升。遥感图像的普遍获取性极大地推动了遥感图像处理技术的应用和进步。遥感图像配准是遥感数据综合分析的先决条件,也是对遥感图像进行各种深层操作的基石,在图像拼接(Chai 等, 2023)、图像融合(Xu Q 等, 2023)、目标检测(Li Y 等, 2023)、变化检测(Zhong 等, 2015)及制导匹配(滕锡超 等, 2023)等应用中起着重要作用。

遥感图像的类型日趋多样化,它们正在向多模态、多光谱、多分辨率及多时相等多方面发展。遥感图像配准旨在将两幅或多幅不同时相、不同传感器获取、或处于不同视角下的同场景图像进行几何对齐。然而,多类型遥感图像在几何形态和辐射特性上差异显著,加之尺寸庞大、信息复杂且存在噪声干扰,给遥感图像配准带来了显著的挑战(Paul and Pati, 2021)。

为应对这些挑战,国内外学者进行了深入研究,并提出了众多配准技术,推动遥感图像配准技术在过去几十年中取得了显著进展。这一时期内,遥感图像配准方法的持续创新显著提升了配准的精度。从基于模板的配准方法到基于特征的配准方法,学者们已提出了一系列多样化的手工设计配准策略。这些方法大多结合了图像的多类信息,如梯度信息和相位信息等,对众所周知的特征如尺度不变特征变换(scale invariant feature transform, SIFT)在空间域或变换域上进行改进,或通过整合多种配准策略以提高配准精度。这类知识驱动专家设计的配准方法通过显式特征提取与迭代优化,能实现传统场景下的可解释性配准,但存在对专业知识依赖强、对复杂场景适配性不足、特征选择局限性明显等问题,难以满足大规模、多样化遥感图像配准要求。

近年来,基于深度学习的遥感图像配准方法取得了显著进展,利用卷积神经网络(convolutional neural networks, CNN)等深度学习网络进行深度特征采集和非线性表达,具有更好的配准效果和更大的发展潜力(Kuppala 等, 2020)。此类数据驱动深度学习方法利用深度学习网络自动学习图像对间的特征或映射关系,提取高级特征用于配准或端到端

预测变换模型,能有效应对遥感图像配准过程中出现的显著几何畸变以及非线性辐射失真等复杂挑战。

本文将系统梳理遥感图像配准算法。首先概述了遥感图像配准的基本概念;之后将遥感图像配准技术归纳为两大类:基于知识驱动专家设计的配准方法和基于数据驱动深度学习的配准方法(眭海刚 等, 2022)。基于知识驱动专家设计的配准方法核心在于专家对数据处理的经验和专业知识,主要包括基于模板的配准方法、基于特征的配准方法以及混合配准方法等。另一方面,基于数据驱动深度学习的配准方法则基于标注样本数据进行自主学习与优化,主要策略包括集成学习配准方法、端到端学习配准方法、风格迁移配准方法以及大模型驱动的配准方法。这种分类方式有助于更清晰地理解和比较各种遥感图像配准方法的特点和优势。表1是对本文中所提方法的分类总结。

基于上述分类方式,本文将综述近年来不同类型的遥感图像配准方法,对这些方法进行综合回顾,介绍它们的优缺点,并探讨未来遥感图像配准的发展方向 and 面临的挑战。

## 1 遥感图像配准基础

遥感图像配准是将不同时相、不同视角或者不同传感器获取的同一场景的两张或者多张遥感图像进行对齐的一个过程(Zitová and Flusser, 2003)。

在遥感图像配准中,其中一幅图像称为参考图像(reference image, RI),另一幅图像称为感测图像(sensed image, SI)。遥感图像配准的目的就是要将感测图像与参考图像精准对齐。将参考图像和感测图像中的像素灰度值分别记为 $I_r(x_r, y_r)$ 和 $I_s(x_s, y_s)$ ,式中 $(x_r, y_r)$ 和 $(x_s, y_s)$ 分别代表参考图像和感测图像中像素点的坐标。图像配准可定义为下式。

$$I_r(x_r, y_r) = g(I_s(f(x_s, y_s))) \quad (1)$$

式中, $g$ 表示灰度或其他度量值函数(往往不是必须求解)(武越 等, 2023), $f$ 表示将感测图像二维空间坐标映射到参考图像空间坐标的变换函数。上述公式将图像配准问题转化为寻找参考图像

与感测图像之间最佳空间变换模型 $f$ 的问题。

在变换模型 $f$ 估计完成后,将其应用于感测图

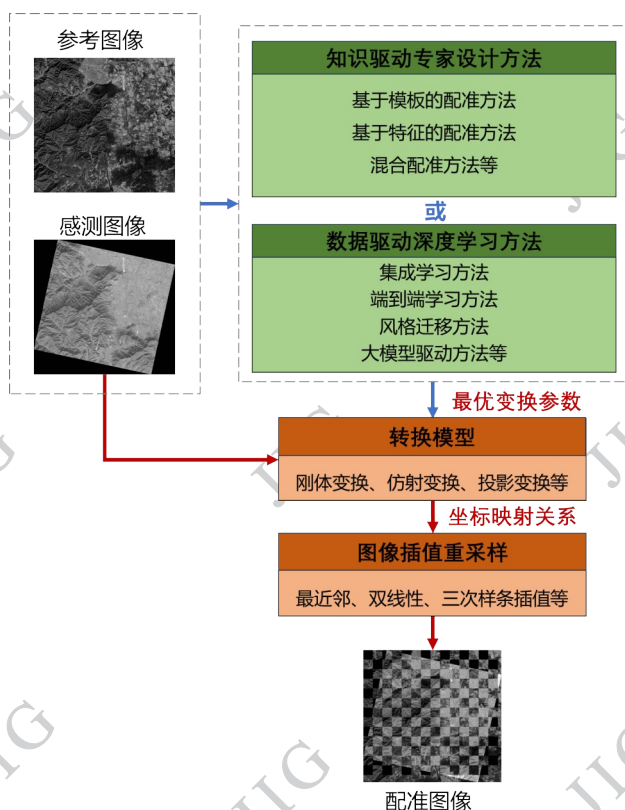


图1 遥感图像配准的主要流程图

Fig. 1 Main flow chart of remote sensing image registration

像,并通过图像插值重采样技术估算非整数位置像素的强度值,从而实现参考图像与感测图像的精确配准。图1为遥感图像配准的主要流程图。

遥感图像配准的核心在于寻求图像对之间的最优空间变换模型,这一过程依赖于图像之间准确的对应关系。需要说明的是,本文中图像匹配特指确定两幅图像之间像素级或特征级对应关系的过程;图像配准则指基于该对应关系估计并应用空间变换模型,实现图像几何对齐的完整流程。匹配作为配准的关键子步骤,贯穿全文;文中所述多数方法,本质上是通过改善匹配结果来提升最终的配准效果。在该领域中,常用的几何变换模型包括刚体变换、仿射变换、投影变换、多项式变换及薄板样条变换等,而常见的图像插值重采样方法主要有最近邻插值、双线性插值和三次样条插值等。

## 2 知识驱动专家设计方法

此类配准方法的设计和建立主要依赖于研究者的专业知识。目前基于知识驱动专家设计的遥感图

像配准方法总体上分为三类,即基于模板的配准方法、基于特征的配准方法与混合配准方法。

### 2.1 基于模板的配准方法

基于模板的配准方法是较早发展起来的图像配准技术,现已发展得非常成熟。该方法基于模板匹配方法获得同名点对应关系:根据遥感图像的固有特性构建相似性度量准则,通过预定义尺寸的模板图像或全局影像检测图像间的匹配关系。图2为模板匹配的示意图。具体操作是在参考图像上选取一个模板窗口,以其中心作为待匹配点,依据相似性准则在目标影像上进行搜索匹配以识别同名点用于求解几何变换参数,完成遥感图像配准。

图3中上半部分为基于模板的配准方法流程图。基于模板的配准方法可以在空间域或变换域中进行。

#### 2.1.1 空间域

常见的基于空间域灰度测度有归一化互相关(normalized cross-correlation, NCC)(Lewis, 1994)、差平方和(sum of squared differences, SSD)(Kybic and Unser, 2003)、互信息(mutual information, MI)(Johnson等, 2001)等。在NCC方面,陆和平和高磊(陆和平和高磊, 2009)提出了一种改进的基于互相关的多源影像匹配算法,针对合成孔径雷达(synthetic aperture radar, SAR)图像和光学图像的特点定义了新的匹配度量方式。对于SSD, Wong和Orchard(Wong and Orchard, 2008)使用穷尽式区域对应搜索技术,利用SSD确定光学与LiDAR影像的对应关系,并利用快速傅里叶变换减少时间成本。

在MI方面,最大互信息法由于能够有效抵抗图像间的灰度差异,近年来被广泛应用于多模态遥感图像的配准。如Ghorbani和Beheshti(Ghorbani and Beheshti, 2007)提出了一种结合了模拟退火(simulated annealing, SA)和ML(Marquardt-Levenberg)优化技术的最大互信息搜索算法,并结合多分辨率金字塔对多模态遥感图像进行配准。Suri和Reinartz(Suri and Reinartz, 2010)使用互信息创建了一种基于直方图的SAR和光学图像配准方法。Dame和Marchand(Dame and Marchand, 2012)提出了一种适用于互信息代价函数的优化过程,可以准确、快速地应用在自然景物和遥感图像等不同类型的图像的配准。Yan等(Yan等, 2020)利用迁移优化最大化互信息的方法来优化变换参数,避免互信息方法陷入

表1 遥感图像配准方法分类与特点

Table 1 Classification and characteristics of remote sensing image registration methods

遥感图像配准方法	特点
基于模板配准方法	基于模板的配准方法在遥感图像中具有较高的精度和简单的实现,但计算复杂度高,实时处理性能差,对旋转、尺度变化和几何畸变等复杂情况的容错率低。
知识驱动专家设计方法 基于特征配准方法	基于特征的配准方法计算成本低、鲁棒性高,能处理复杂背景和多种变换,但依赖特征提取,易受噪声和异常值影响,且在处理质量较低、内容重复、非刚性变形和极端视角变化的图像时易产生歧义和错误。
混合配准方法	混合配准方法通过结合多种配准技术提升配准精度和稳定性,弥补单一方法的不足,但由于计算复杂度高,导致计算速度较慢。
集成学习方法	集成学习方法结合深度学习和传统配准方法,提升特征提取和相似性度量的鲁棒性和有效性,但常受限于手工特征检测器,未充分发挥深度学习的潜力。
端到端学习方法	端到端学习方法联合学习特征的检测和描述或直接学习图像间的变换模型参数,减少了中间步骤和手工干预,缺点是对大量标注数据和计算资源依赖较高,且模型解释性较差。
数据驱动深度学习方法	
风格迁移方法	风格迁移方法的优点是能够减小多模态遥感影像间的强度差异,从而提高配准精度,缺点是转换过程可能引入误差和信息损失,且模型训练复杂度较高。
大模型驱动方法	大模型凭借其强大的特征表征能力与跨模态泛化性能,为突破传统配准方法的局限性提供了新路径。但受限于巨大的计算开销与较高的推理延迟,大模型在端侧部署或实时处理等要求高效率的场景中仍面临较大挑战。

局部最优解而影响配准结果。

### 2.1.2 变换域

基于变换域的配准方法主要包括基于傅里叶变换的相关方法(Kuglin and Hines, 1975)和小波变换法(Le Moigne 等, 2002)等。

基于傅里叶变换的图像配准方法利用快速傅里叶变换(fast Fourier transform, FFT)在频域内实现图像的相似性检测,主要包括频域互相关和相位相关两种方法,它们的主要区别在于是否通过计算相位信息来执行频域中的近似归一化(Tong 等, 2019)。频域互相关技术是指在频域中等效实现相关定理,避免了传统空域相关过程中的迭代搜索,从而实现了相关性的快速计算。相位相关技术则是一种基于

傅里叶变换平移性质的经典方法,仅考虑相位信息,用于校准两幅图像之间的平移错位(Kuglin and Hines, 1975)。由于相位相关方法仅关注相位信息,对图像内容不敏感,因此许多学者将其应用于图像配准。如Reddy

(Reddy and Chatterji, 1996)扩展了相位相关技术,使之能够处理图像间存在平移、旋转和缩放关系的配准问题。Thangavel 和 Kokila (Thangavel and Kokila, 2013)提出了一种扩展的FFT图像配准方案,利用高斯高通滤波器和对数极坐标变换来提升配准的准确性和鲁棒性。此外,蔡潇(蔡潇, 2021)也提出了一种基于相位特征的异源图像匹配算法,利用傅里叶-梅林(Fourier-mellin)算法进行相位相关匹配。

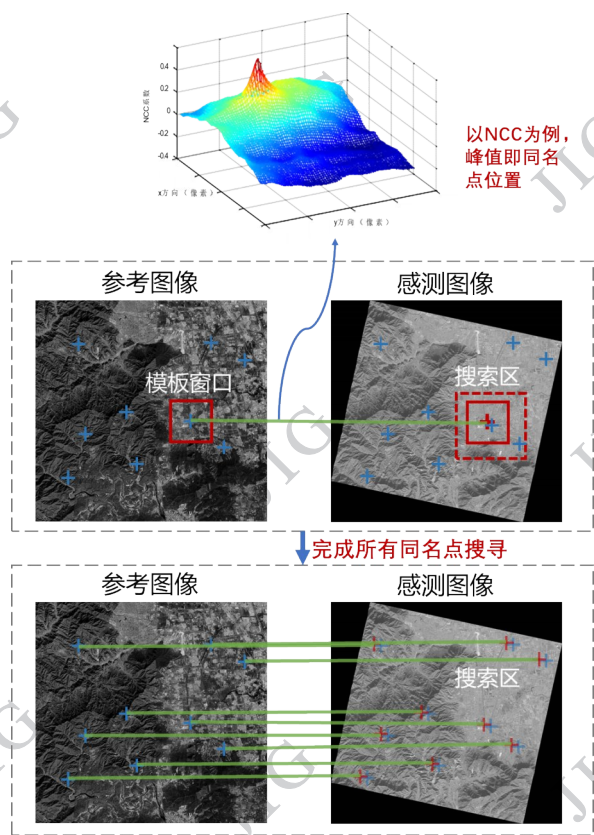


图2 模板匹配示意图

Fig 2 Template matching schematic

这些方法对噪声具有鲁棒性,但对频谱失真则较为敏感。

小波变换具备的多尺度特性能够获得时域和频

域在不同尺度上的信息,这使其非常适用于提取图像的局部特征。伍君(伍君,2005)研究小波变换理论,将小波变换和互信息度结合实现对遥感图像的配准。彭嘉雄和强赞霞(强赞霞和彭嘉雄,2006)在小波变换构建的图像金字塔基础上,结合傅里叶变换与互信息对遥感图像进行分层配准,实现了多种方法的优势互补。

基于模板的配准方法在同源或异源遥感图像的配准中得到了广泛应用。该方法具有较高的精度,且算法简单易实现。然而,基于模板的配准方法逐点进行匹配,计算复杂度高,不利于实时处理。此外,该方法对匹配结果的容错率较低,难以适用于包含旋转、尺度变化和几何畸变等复杂变化的遥感图像配准,多在小幅度的旋转、缩放和局部变形下表现较好(Ma等,2021;王丽娜,2021)。

### 2.2 基于特征的配准方法

基于特征的配准方法首先从图像中识别出显著特征,然后使用特征向量有效地描述这些显著特征。接着,通过比对特征描述符的相似性进行匹配,从而建立特征之间的同名关系,并进一步构建图像间的几何变换关系。

基于特征的配准方法是图像配准领域中的研究热点与重点(樊建伟,2017)。该方法通常包括三个阶段,分别是特征检测、特征描述与特征匹配。图3

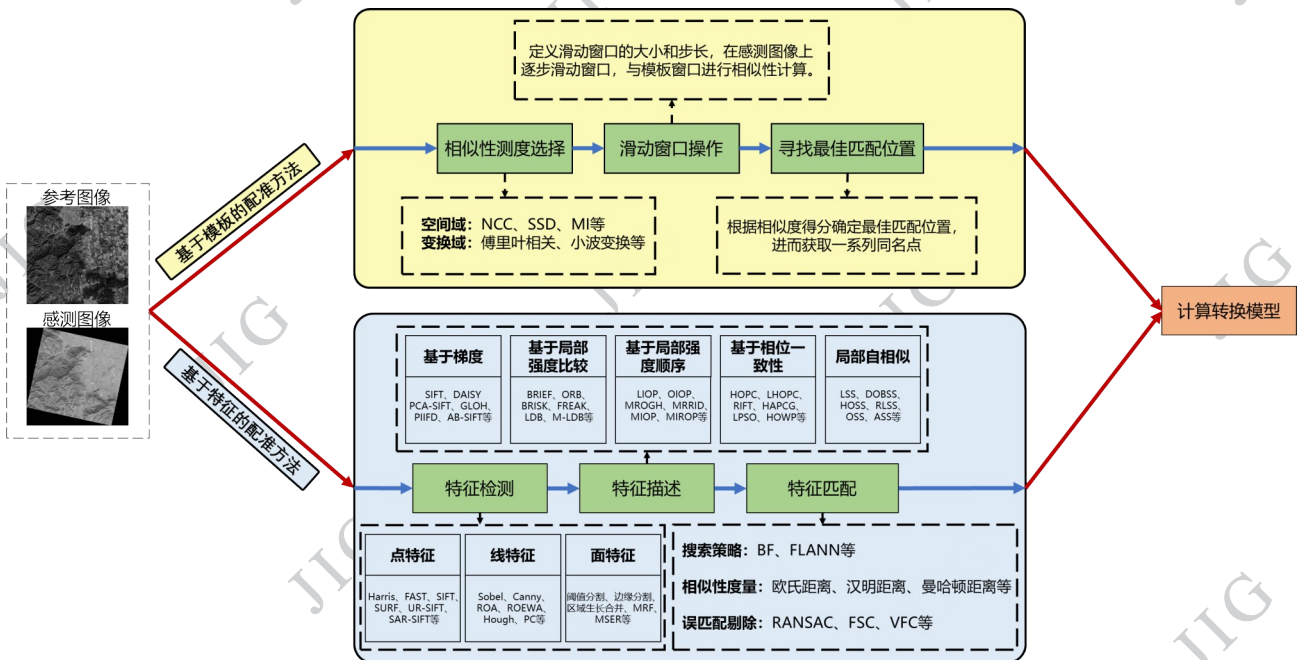


图3 基于模板的配准方法与基于特征的配准方法流程图

Fig. 3 Flowchart of the template-based registration method and the feature-based registration method

的下半部分是基于特征的配准方法流程图。

### 2.2.1 特征检测

基于特征的配准方法首先需要检测图像的显著局部特征。根据特征的物理结构,特征可以被分为点特征(线交叉点、高曲率点、角点等)、线特征(海岸线、道路、区域轮廓等)和面特征(森林、田野、建筑物、湖泊等)(Zitová and Flusser, 2003; 眭海刚等, 2022)。其中,点特征是当前研究较为成熟、应用最为广泛的一类特征。

**1)点特征。**点特征具体可细分为角点特征和斑点特征。角点特征可以定义为两条直线的交点或者等值线的高曲率点。常见的角点特征利用一阶导数或像素强度比较进行特征检测如Moravec(Moravec, 1977)、Harris(Harris and Stephens, 1988)、SUSAN(Smith and Brady, 1997)、FAST(features from accelerated segment test)(Viswanathan, 2011)等。学者们也会针对经典角点特征进行改进,如Mair等(Mair等, 2010)提出了AGAST(adaptive and generic accelerated segment test)特征检测算法,改进了FAST并提升了其检测速度;Fan等(Fan等, 2018)设计了一种基于均匀非线性扩散的Harris特征提取方法,以降低散斑噪声对特征提取的不利影响。随着图像尺度空间理论的出现与完善,常用高斯差分尺度空间(difference of Gaussian, DoG)来近似高斯拉普拉斯尺度空间(Laplacian of Gaussian, LoG)提高金字塔构建效率。许多学者会在图像金字塔中提取Harris或者FAST角点。如Rublee等(Rublee等, 2011)提出的ORB(oriented FAST and rotated BRIEF),Leutenegger等(Leutenegger等, 2011)提出的BRISK(binary robust invariant scalable keypoints)在图像金字塔中利用FAST提取多尺度角点。

斑点特征一般呈现为局部封闭区域,与其周边区域有所区别。学者们常结合尺度空间理论与Hessian矩阵,专门用于检测图像中的斑点特征。典型的特征有Hessian-Affine(Mikolajczyk and Schmid, 2004),SIFT(Lowe, 2004)、SURF(speeded up robust features)(Bay等, 2006)、KAZE(Alcantarilla等, 2012)等,其中不乏SIFT的改进版本。此外,诸多学者进一步提出了基于SIFT的优化策略,以提升实际应用中的性能。如Morel和Yu(Morel and Yu, 2009)提出了一种仿射不变性极高的A-SIFT(affine-SIFT);Schwind等(Schwind等, 2010)提出了SIFT-

OCT方法,通过跳过第一组尺度空间octave来减少噪声对关键点检测的影响;Sedaghat等(Sedaghat等, 2011)提出了一种均匀鲁棒的SIFT方法(uniform robust SIFT, UR-SIFT),通过在每个尺度层分配特征数量实现SIFT特征的均匀分布;Wang等(Wang等, 2012)提出了BF-SIFT(bilateral filter SIFT)方法,利用双边滤波器构建各向异性尺度空间并提取SIFT;Dellinger等(Dellinger等, 2012)提出了SAR-SIFT,根据SAR影像的具体特性引入了新的梯度定义,提高了对SAR影像斑点噪声的鲁棒性;Ma等(Ma等, 2017a)提出了PSO-SIFT,引入新的梯度定义来克服多光谱和多传感器遥感图像之间的强度差异。

基于点特征的特征检测方法以部分像素点表示整幅图像,计算速度快,定位准确,被广泛应用于遥感图像配准。然而,由于不同传感器成像机制的差异和像素表现形式的不同,可能导致图像的灰度和噪声有显著差异,从而降低提取特征重复性,导致配准失败。

**2)线特征。**线特征提取图像的边缘和纹理信息。常见的边缘检测算子有,Roberts、Prewitt、Sobel、Canny(Canny, 1986)、Hough变换等,这些边缘检测算子适用于光学遥感影像,但是对乘性噪声非常敏感。因此有学者提出了考虑边缘方向性的均值比检测器来抵抗乘性噪声带来的影响,如ROA(Touzi等, 1988)和ROEWA(Fjortoft等, 1998)等,被广泛应用于SAR图像的边缘检测。在变换域中还可利用傅里叶变换法或小波变换法进行边缘检测。描述线特征时,常用边界长度、边界直径、曲率等简单描述符,或傅里叶描述符和不变矩等复杂描述符。应用线特征进行遥感图像配准的实例如Habib和Alruzouq(Habib and Alruzouq, 2004)利用改进的迭代Hough变换方法来提取图像中的线特征,实现了多源遥感影像的自动配准;Juan等(Juan等, 2015)通过计算相邻子区域的差异生成一种新型的压缩轮廓特征,并构建多尺度特征图像对SAR场景图像进行由粗到细的匹配;Zhang等(Zhang等, 2017)利用不同线段检测器检测光SAR图像边缘,进行匹配后生成voronoi图,从而获取线交叉点间的点对应关系;Liu和Jiang(Liu and Jiang, 2019)基于线交线结构设计线特征双矩形局部描述符及基于空间关系的异常值去除策略,实现城市卫星遥感图像亚像素精度配准。

由于大多数边缘检测算法依赖于图像本身,无  
© 中国图象图形学报版权所有

法对所有类型的图像都提取出理想的边缘,容易产生不完整、或者虚线特征,因此线特征提取存在局限性。此外,当参考图像与待配准图像的分辨率差异较大时,一张图像中的边缘和轮廓可能在另一张图像中不存在,进一步增加了配准的难度。

随着人们对频域理论研究的不断深入,还出现了许多结合频域边缘特征检测和点特征检测的方法。如 Ma 等(Ma W 等,2018)在相位特征图上获取 SAR-SIFT 特征,并通过位置和方向信息建立空间约束,实现遥感影像对的配准;Li 等(Li 等,2020)提出了 RIFT (radiation-variation insensitive feature transform),该方法利用图像的相位一致性(phase congruency, PC)信息代替图像的梯度与强度信息进行角点特征检测,之后还引入局部归一化滤波器减小多模态影像之间的辐射差异,改进了 FAST 检测器的定向特征,较 RIFT 表现更优(Li J 等,2022);Hu 等(Hu M 等,2023)使用基于 PC 图构建的 DoG 图像金字塔来检测特征点。

**3) 面特征。**面特征是图像中稳定性较强特征,主要通过图像分割的方法得到。面特征通常由它们的重心表示,具备在旋转、缩放和倾斜变换下的不变性,同时对随机噪声和灰度变动也展现出了强大的稳定性。常见的图像分割算法包括阈值分割法、边缘分割法、区域增长与合并法、马尔可夫随机场(Markov Random Field, MRF)分割法、分水岭分割法、多尺度分割法、数学形态学分割法等。这些方法通常会采用不变矩(例如:中心距、Hu 矩、Zernike 矩等)、面积、周长、扁平度和长宽比等特征来描述和匹配图像区域(余先川 等,2013)。应用面特征进行遥感图像配准的实例如 Cheng 等(Cheng 等,2008)通过分层滤波策略去除信息熵较低及空间分布性不佳的 MSER (maximally stable extremal regions) (Matas 等,2004)特征区域并结合 SIFT 描述符,增强特征处理仿射变换时的能力;Zhang 等(Zhang 等,2007)使用受边缘限制的区域生长技术对图像进行分割,提取区域作为不变特征匹配;Wang 等(Wang 等,2014)利用小波变换和光谱残差方法将 SAR 图像划分为显著性区域,通过综合学习粒子群优化对参考图像进行逐层搜索匹配,有效消除了变化区域的干扰和引起的错位;苏娟等(苏娟 等,2016)利用多尺度非线性廓线理论获取 SAR 图像中的封闭均匀区域,并构建具有仿射不变性的区域特征,用于 SAR 图像配

准;Sui 等(Sui 等,2017)通过多尺度分割和凸性模型将图像转换为图像对象,分析这些对象的形状特征以获得共轭区域;李响等(李响 等,2018)采用改进的 MRF 算法分割 SAR 图像获取稳定的显著性区域,随后在显著性区域的边缘附近提取 SIFT 特征点用于图像匹配配准;Zhao 等(Zhao 等,2024a)提出了一种改进的显著 MSER (salinet MSER, SMSER)特征检测方法,在显著图上进行 MSER 提取,如图 4 为不同方法提取的区域特征示意图。SMSER 提取的区域数量虽少但最终成功匹配的区域与其他方法高度一致,表明该方法有效去除冗余特征保留有价值的匹配信息。

面特征相比线特征和点特征包含更多信息,因此可以避免因边缘或点检测不准确而导致的配准结果较差的问题。尽管面特征的提取方法在遥感图像处理中具有重要作用,但其性能受限于分割算法的有效性及面特征自身的尺度。目前少有通用的分割算法能够在多源遥感图像分析中始终取得较佳效果。

### 2.2.2 特征描述

获得遥感图像的显著特征后,需要对这些特征进行描述,以表征检测到的特征并使其具有可区分性和鲁棒性。当前流行的特征描述符通常可归纳为以下几类:基于梯度的描述符、基于局部强度比较的描述符、基于局部强度顺序的描述符、基于相位一致性的描述符以及局部自相似性描述符等(Sedaghat and Mohammadi, 2019)。

**1) 基于梯度的描述符。**SIFT 是一种典型的基于梯度的描述符,自提出以来被广泛借鉴学习,许多梯度描述符都是基于 SIFT 进行扩展的。总体来看,后续的梯度描述符在继承 SIFT 基本框架的基础上,主要沿着三个方向进行了针对性拓展。一是改进直方图统计结构以提升辨别能力和鲁棒性,如 Mikolajczyk 和 Schmid (Mikolajczyk and Schmid, 2005)提出的 GLOH (gradient location and orientation histogram),利用对数极坐标同心圆统计梯度信息直方图;Tola 等(Tola 等,2010)提出的 DAISY,利用不同的各向同性高斯函数高效密集计算梯度直方图,并通过圆形网格提高描述独特性;Sedaghat 等(Sedaghat and Ebadi, 2015a)提出的 AB-SIFT (adaptive binning scale-invariant feature transform),采用自适应分块策略计算特征点局部特征描述符。上述几种描

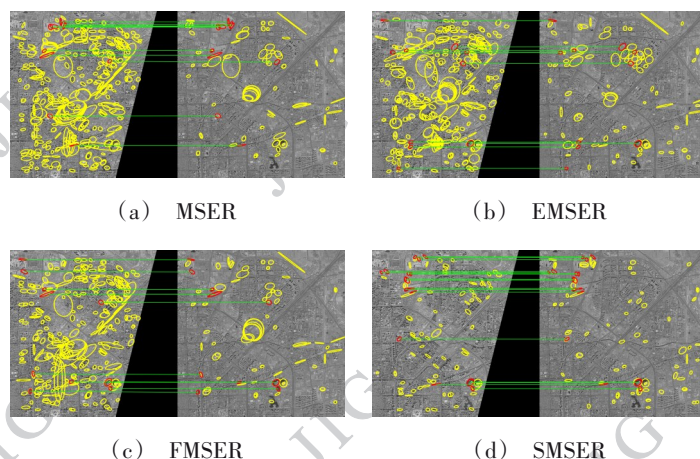


图4 使用不同区域提取方法在同一图像块内的匹配结果示意图(Zhao等,2024a)

Fig 4 Schematic representation of the matching results in the same block using different region extraction methods. (a) MSER. (b) EMSER. (c) fMSER. (d) SMSER. (Zhao等,2024a)

述符的梯度信息直方图统计结构如图5所示。二是改进计算方法或引

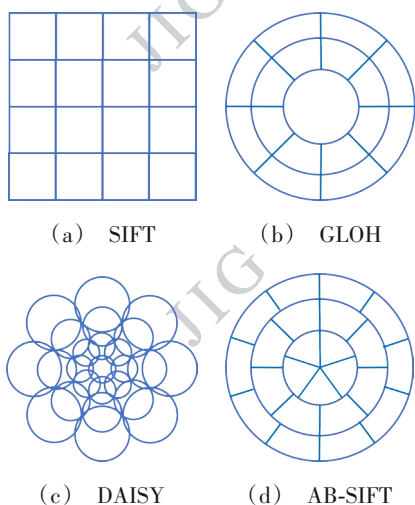


图5 常见类SIFT描述符统计结构

Fig 5 Common SIFT-like descriptor statistical structures. (a) SIFT; (b) GLOH; (c) DAISY; (d) AB-SIFT.

入降维方法提升匹配效率,如 Bay等(Bay等,2006)提出的SURF利用积分图像和Haar小波响应提高计算速度;Ke和Sukthankar(Ke and Sukthankar,2004)提出的PCA-SIFT(principal component analysis SIFT),Duan等(Duan等,2008)提出的ICA-SIFT(independent component analysis SIFT),Liu等(Liu等,2010)提出的KICA-SIFT(kernel independent component analysis SIFT),分别利用主成分分析、独立成分分析和核独立成分分析去除描述符冗余,提升匹

配效率。三是拓展描述符的适用场景,如 Abdel-Hakim和Farag(Abel-Hakim and Farag,2006)提出C-SIFT(colored SIFT)将在灰度空间中的描述符计算转到颜色不变空间中,相较于传统SIFT在颜色和光度变化上更加鲁棒;Chen等(Chen等,2010)提出PIIFD(partial intensity invariant feature descriptor)可以有效地解决多模态图像中的明暗颠倒和非线性强度差异问题。上述改进主要集中于梯度描述符框架内部的优化,然而梯度描述符本身存在计算复杂度较高、对复杂光照变化敏感、对多模态图像匹配适用性不足等局限性。

2)基于局部强度比较的描述符。为了提高特征匹配算法的运算效率,众多研究者基于图像点对间的强度关系来构建二进制描述符(Sedaghat and Mohammadi,2019)。如Calonder等(Calonder等,2010)提出的BRIEF(binary robust independent elementary features)在特征点邻域窗口内随机选取点对进行像素值比较由此生成二进制编码描述。但BRIEF本身不具备方向性,后续研究主要沿着引入旋转尺度不变性和优化采样策略两个方向进行改进,如Ruble等(Ruble等,2011)提出的ORB利用改进的BRIEF创建二进制描述符,相比于原始BRIEF增加了旋转不变性;Leutenegger等(Leutenegger等,2011)提出的BRISK改进了二进制描述符的采样模式,以特征点为中心构建不同半径的同心圆进行均匀等间隔采样;而Alahi等(Alahi等,2012)提出的FREAK(fast retina keypoint)受到人眼视觉系统

的启发,采用视网膜采样模式,这一非均匀采样策略使描述符对特征点中心区域的局部结构更为敏感,同时保持了外围区域的整体信息。Yang 和 Cheng (Yang and Cheng, 2012) 提出的 LDB (local difference binary) 则扩展了特征比较范畴,基于可扩展的多网格策略,通过计算网格对之间的亮度与梯度差异获得二进制描述符;之后 Alcantarilla (Alcantarilla 等, 2013) 等改进的 LDB 描述符 M-LDB (modified LDB), 不仅将二进制集中单独像素的比较改为区域均值之间的比较,还加入了旋转不变性。二进制描述符通过逐像素比较生成,并以汉明距离替代欧氏距离进行相似性度量,虽然能够显著降低存储开销并大幅提升匹配速度,但其在计算效率上的优势常以牺牲一定程度的辨别性与鲁棒性为代价,因此其适用场景往往与梯度描述符形成互补。

**3) 基于局部强度顺序的描述符。**先确定特征主方向再旋转局部区域容易引入方向估计误差,进而导致特征错位与误匹配,因此部分学者提出了基于像素强度相对顺序的旋转不变描述符构造方法 (Sedaghat and Mohammadi, 2019)。如 Wang 等 (Wang 等, 2011) 提出的 LIOP (local intensity order pattern) 利用像素强度顺序信息将特征区域划分为局部子区域,将排序结果作为局部特征编码赋予采样点,最终将所有子区域的特征矢量串联生成 LIOP 特征,由此实现旋转不变性。但该方法对噪声较为敏感。Wang 等 (Wang 等, 2016) 在之前的研究基础上提出 OIOP (overall intensity order pattern), 对局部特征区域上采样点按照灰度次序进行整体量化,使特征描述符对图像噪声更具抵抗力。也有学者基于强度顺序整合多种特征进行匹配,如 Fan 等 (Fan 等, 2012) 提出的 MROGH (multisupport region order-based gradient histogram) 和 MRRID (multisupport region rotation and intensity monotonic invariant descriptor) 分别利用梯度和强度信息在多支持区域上构建特征描述符。Yang 等 (Yang 等, 2018) 提出的 MIOP (multi-neighborhood intensity order pattern) 和 MIROP (multi-neighborhood intensity relative order pattern) 为每个采样点编码多个局部特征提高描述符的鲁棒性。尽管基于强度顺序的描述符在一定程度上降低了对主方向估计精度的依赖,但自身仍存在固有局限:该类描述符依赖于局部区域的像素灰度排序,当图像发生显著的非线性辐射变化时,排序

结果易剧烈变动,因此其在复杂辐射变化场景下的泛化能力仍有待提升。

**4) 基于相位一致性的描述符。**多数基于梯度或像素强度的描述符在存在显著非线性辐射差异的多模态遥感图像对中难以适用 (Ma 等, 2017a)。近年来的研究表明,多模态图像之间共享相似的几何结构和形态特征,通过捕捉这些特征的相似性可实现高精度匹配。如图 6 所示,通过对比梯度图和相位一致性图,可以看出相位特征相对于梯度特征更能适应多模态图像之间的非线性辐射差异,因

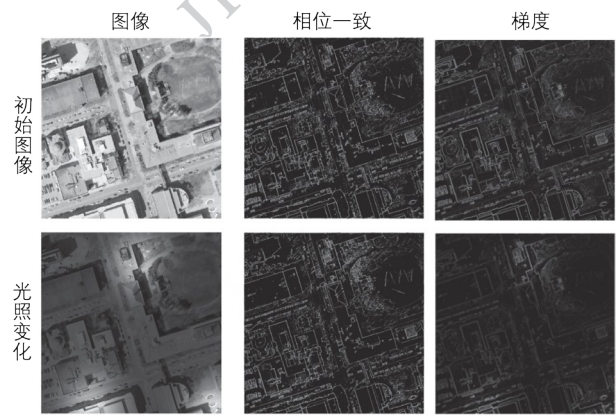


图6 相位一致性图与梯度图的比较 (Ye Y 等, 2018)

Fig 6 Comparison of phase congruency with gradient (Ye Y 等, 2018)

此,一些学者引入 PC 保留图像的几何结构特征。如 Ye 等 (Ye 等, 2017) 借鉴 HOG (histogram of oriented gradient) (Dalal and Triggs, 2005) 框架,基于相位一致性特征值和方向构建 HOPC (histogram of orientated phase congruency) 描述符,能够抵抗非线性灰度差异,之后还基于扩展的 PC 模型和类似 DAISY 的空间结构提出了 LHOPC (local HOPC) (Ye Y 等, 2018)。Fan 等 (Fan 等, 2018) 提出了相位一致性结构描述符 (phase congruency structural descriptor, PCSD), 在 PC 结构图上以分组方式构建 PCSD, 提高了光 SAR 图像配准精度。Li 等 (Li 等, 2020) 提出的 RIFT 则在图像的相位一致性幅度的最大索引图 (maximum index map, MIM) 上利用类 SIFT 统计描述符的方法进行特征描述,如图 7 是 MIM 计算示意图。类似

的描述符还有 HAPCG (histogram of absolute phase consistency gradients) (姚永祥 等, 2021)、LPSO (local phase sharpness orientation) (Yang 等, 2022)、

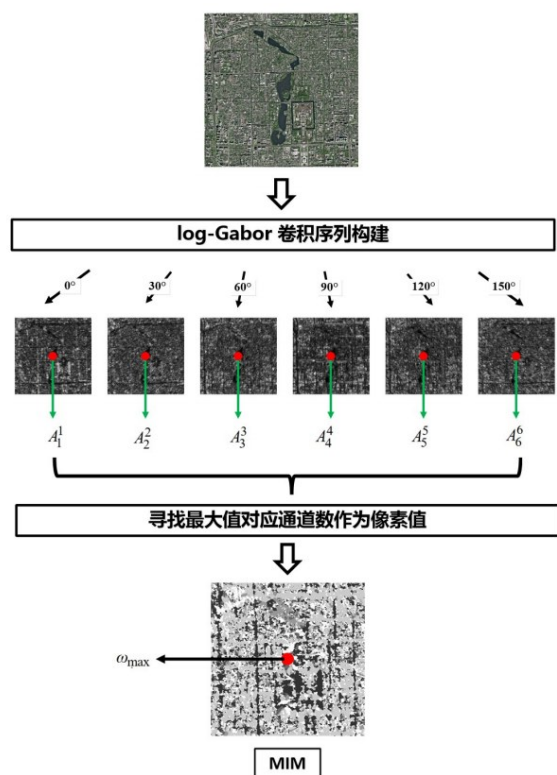


图7 MIM图计算示意图

Fig 7 MIM diagram calculation schematic

HOWP (histogram of the orientation of weighted phase) (Zhang Y 等, 2023) 等。这类相位描述符虽对非线性辐射差异具有较强的适应性, 但PC模型计算复杂度较高, 尤其在大尺寸遥感影像配准中会带来较大的计算压力, 且其性能高度依赖于图像结构。

**5) 局部自相似性描述符。**局部自相似性描述符 (local self-similarity, LSS) (Shechtman and Irani, 2007) 将图像块与周围较大区域进行相关性比较, 为每个像素生成“局部自相似性”描述符, 描述图像斑块内局部自相似性的几何布局, 侧重于提取局部邻域内稳定存在的重复性结构。如图8是具有相似目标的两张图像中的几个对应点以及它们的自相似性描述符。LSS原始形式的描述符不足以应对一些复杂的影像匹配任务。为此, 多项研究从不同角度对LSS进行了改进与扩展。如Sedaghat研究团队将LSS描述符分别扩展为基于差异阶的自相似性 (distinctive order based self-similarity, DOBSS) 描述符 (Sedaghat and Ebadi, 2015b)、定向自相似性直方图 (histogram of oriented self-similarity, HOSS) 描述符 (Sedaghat and Mohammadi, 2019) 以及旋转不变自相

似描述符 (rotation invariant self-similarity, RISS) (Mohammadi 等, 2022), 强化描述符的独特性、光照不敏感性以及旋转不敏感性。Chen等 (Chen 等, 2017) 提出了一种基于自相似性描述符的局部强度顺序模型 (local order pattern based self-similarity, LOPSS)。Xiong

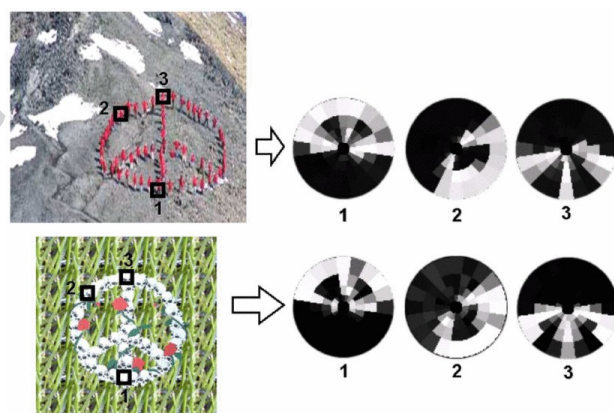


图8 自相似性描述符展示 (Shechtman and Irani, 2007)

Fig 8 LSS descriptor presentation (Shechtman and Irani, 2007)

等 (Xiong 等, 2020) 提出了基于秩的局部自相似性描述符 (rank-based local self-similarity, RLSS) 用于光学与SAR图像的匹配, 通过集成到密集采样网格中得到DRLSS (dense, RLSS) 描述符, 进一步提高描述符的可判别性, 还在之后提出了提出基于多通道自相似图提取定向自相似性描述符 (oriented self-similarity, OSS) (Xiong 等, 2021)。Yang等 (Yang 等, 2024) 利用局部相邻自相似性 (adjacent self-similarity, ASS) 模型构建多维特征, 并最终基于3D卷积进行特征增强和图像特征描述。尽管LSS及其扩展描述符被广泛设计和使用于多模态图像配准任务, 但这些描述符的判别力仍然相对较低 (Fan 等, 2018)。

### 2.2.3 特征匹配

特征匹配的过程本质上是在高维空间中, 通过相似性度量函数执行搜索任务。特征匹配算法可以概括为两个核心步骤: 首先是确定一个有效的相似性度量标准, 其次是设计一个高效的搜索策略。

相似性测度是度量特征描述符间相似性程度的定量指标。特征匹配中常用的相似性测度有欧氏距离、汉明距离、曼哈顿距离、余弦相似度等。针对空间尺度描述符的复杂性, 学者们采用了多种策略进

行特征的匹配搜索,建立初步关系。常见的搜索匹配方法有暴力搜索匹配算法(brute force, BF),快速最近邻搜索库(fast library for approximate nearest neighbors, FLANN)(Muja and Lowe, 2009)等。BF算法无需预处理直接搜索全空间,但其效率较低,增加了计算负担。FLANN方法则专为大数据集和高维特征设计,通过近似匹配加速搜索过程,尽管未必找到最佳匹配结果,但在处理大数据集时效果优于BF算法。常见的建立初步良好对应关系的策略包括固定阈值、最近邻、相互最近邻和最近邻距离比(Ma等, 2021)等方法。

理想的匹配场景中,感测图像的每个点都能在基准图像中找到精确对应的点,理论上这些匹配点对的距离应接近零。然而,实际匹配过程难以达到理想状态,导致初始匹配点对中含有误匹配。这些误匹配点对将对后续的几何变换模型求解产生不利影响。因此,在初步匹配之后,有必要剔除匹配点集中的误匹配点对。

RANSAC(random sample consensus)是最流行的参数误匹配剔除技术(Li等, 2017)。由于RANSAC的良好性能,已经提出了多种用于遥感图像配准的改进版本。Li和Ye(Li and Ye, 2012)在RANSAC中嵌入了鲁棒样本共识判断(robust sample consensus judging, RSCJ)算法,显著提高计算效率。Wu等(Wu等, 2015)提出了快速样本共识(fast sample consensus, FSC),将RANSAC中的数据集分为样本集和共识集,可以减少迭代次数。Wu等(Wu等, 2018)还提出了粒子群优化样本一致性算法,该算法将粒子群优化引入RANSAC中,直接对模态变换参数进行采样,避免了随机选择暂定匹配,有效地在低正确率下实现正确的匹配。

然而,RANSAC在处理复杂场景的多模态匹配时,面对大量外点存在效率和稳健性问题。当误匹配存在比例较高时,RANSAC方法很可能会出错。除了类RANSAC方法外,国内外学者也提出了多种其他误匹配剔除方法。如Ma等(Ma等, 2014)提出的VFC(vector field consensus)通过向量场一致性估计遵循非参数几何变换的内点对应关系,使用高斯混合模型建立图像间变换模型,并通过稀疏逼近降低计算复杂性。在之后还提出了局部线性变换(locally linear transforming, LLT)(Ma等, 2015),和鲁棒引导局部保留匹配(guided locality preserving

matching, GLPM)(Ma J等, 2018)有效处理异常值。Sedaghat和Ebadi(Sedaghat and Ebadi, 2015c)利用K-means聚类来识别和删除不匹配点。Li等(Li等, 2017)采用基于多自适应分箱梯度直方图的支持线描述符进行投票过滤异常值,并引入仿射不变比率细化匹配结果。Jiang等(Jiang等, 2021)提出线性自适应滤波方法,通过先验几何一致性和滤波去噪理论,实现异常值的去除。

相较于基于模板的配准方法,基于特征的配准方法计算成本低,泛化能力强,鲁棒性高。基于特征的配准方法在处理复杂背景和遮挡的图像上具有优势,能稳定处理各种变换,如平移、旋转、尺度和光照,并能应对显著的几何变化(李加元, 2018)。然而,这些方法高度依赖局部特征的提取和描述,易受到噪声和异常值的影响。特别是在处理质量较低、内容重复,以及经历过严重非刚性变形和极端视角变化的图像时,这些问题会更加突出(Ma等, 2021)。而且特征提取方法往往是针对某一特定难点或应用设计的,难以开发出具有高度通用性的特征提取方法(眭海刚等, 2022)。

### 2.3 混合配准方法

国内外学者根据多源影像特性,提出了很多针对性更强的混合配准方法,以适应各种复杂的遥感场景。混合配准方法的主要特点是利用多种技术的优势,弥补单一方法的不足。常见的混合配准方法如结合模板匹配和特征匹配的由粗到细的配准方法;融合多特征或优化多特征提取的配准方法;结合迭代匹配、区域边界约束等策略的配准方法等。图9是混合配准方法的简单分类示意图。

结合模板匹配和特征匹配的由粗到细的配准方法通常利用不同的方法进行粗配准和精配准两个步骤。如Teo和Chen(Teo and Chen, 2011)结合互信息和Canny边缘特征,通过从粗分辨率到细分辨率的迭代过程优化变换参数;张宝尚等(张宝尚等, 2011)利用主动轮廓方法对去噪后的图像进行分割提取区域特征粗匹配,然后利用互信息优化变换参数;Gong等(Gong等, 2014)在利用SIFT对多模态影像进行粗配准后,采用改进的Marquardt-Levenberg搜索策略最大化互信息实现精细配准;Paul和Pati等(Paul and Pati, 2018)首先利用SIFT进行特征匹配,然后使用一阶同步扰动随机近似(simultaneous perturbation stochastic approximation, SPSA)和二阶

SPSA 来实现输入图像之间的互信息精细配准; Xiang 等(Xiang 等, 2018)首先使用 SAR-SIFT 进行粗配准, 然后采用改进的相位一致性模型快速归一化互相关作为新的相似度度量, 实现和参考图像局部搜索区域的精配准; Ye 等利用 FAST 粗配准后通过结合一阶和二阶通道的可控滤波器 (steerable filters of first- and second-Order channels, SFOC) (Ye 等, 2022b) 构建新型结构描述符精细配准多模态遥感图像; Zhang 等(Zhang Yongjun 等, 2024)首先利用自相似性特征 OSS 进行粗匹配, 然后通过构建多维定向自相似特征 (multi-dimensional oriented self-similarity features, MOSS) 进行三维相位相关类模板匹配, 进一步细化粗匹配结果。

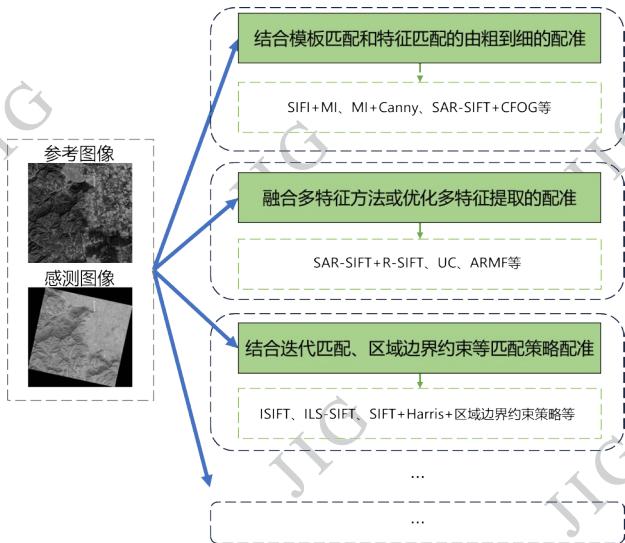


图9 混合配准方法简单分类示意图

Fig. 9 Schematic diagram of a simple classification of hybrid registration methods

许多学者融合多种特征方法或优化多特征提取方法, 综合运用不同类型的特征, 更全面捕捉图像间的对应关系来提高配准的准确性和鲁棒性。如 Zhu 等(Zhu 等, 2016)整合 SAR-SIFT 和 R-SIFT 技术, 用于检测 SAR 图像中的稳定角点和纹理点, 并结合新颖的树状网络进行特征匹配; Sedaghat 和 Mohammadi(Sedaghat and Mohammadi, 2018)提出了一种基于统一能力 (uniform competency UC) 的局部特征提取方法, 可以很容易地应用于各种应用的任何局部特征检测器, 提高匹配性能和改善空间分布的能力, 如图 10 是基于 UC 的局部特征的提取过程; Li 等(Li Z 等, 2022)提出了一种自适应区域多特征 (Adap-

tive regional multiple features ARMF) 匹配方法, 引入自适应区域搜索策略, 利用金字塔放大技术自适应选择匹配特征区域, 提取多种类型特征并自适应选择合适特征描述符。

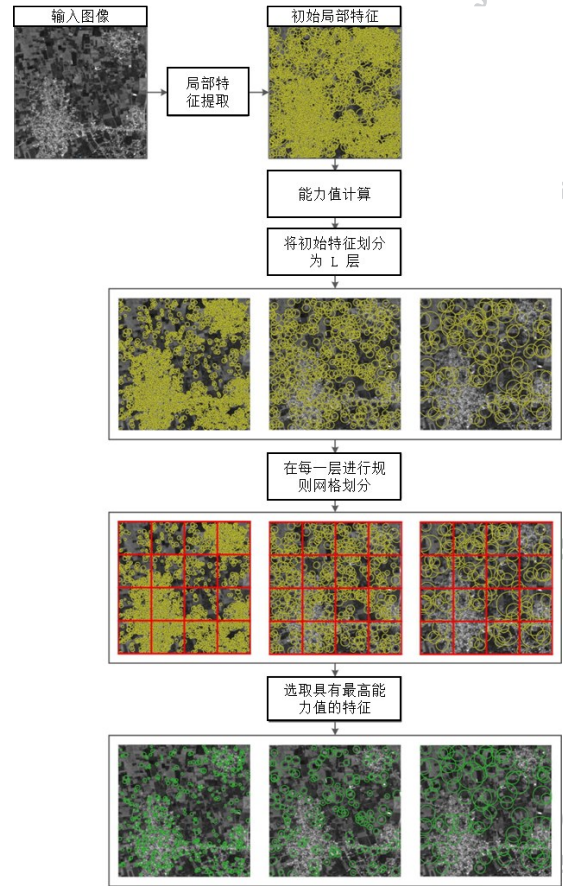


图10 基于 UC 的特征提取过程(Sedaghat and Mohammadi, 2018)

Fig 10 Outline of UC feature extraction process(Sedaghat and Mohammadi, 2018)

近年来还出现了许多结合迭代匹配、区域边界约束等匹配策略的混合配准方法。这些方法通过反复优化匹配过程或对图像进行结构分割来提高配准精度。如 Xu 等(Xu 等, 2015)提出了一种迭代分割集 SIFT (iterative level set and SIFT, ILS-SIFT), 通过提取光学和 SAR 图像中的共轭特征并结合 SIFT 进行配准, 有效避免了因图像分割不良导致的配准失败, 此外, 他们结合区域和线特征等多种特征及多度量方法, 展示了通过迭代策略和细化模型优化配准的潜力(Xu 等, 2016); Chen 等(Chen S 等, 2021)提出了一种名为迭代尺度不变特征变换 (iterative SIFT, ISIFT) 的方法, 该方法包含一个闭环 SIFT 系

统,能够进行返回校正,通过迭代替换当前感知图像提高位置精度;Ma等(Ma等,2017b)利用SIFT和Harris检测特征点,结合边缘检测算法划分图像为结构区域或非结构区域分别匹配,以减少不正确的对应关系。

混合配准方法综合利用多种配准技术的优点,通过相互补充和优化,在处理辐射差异大、几何形变复杂的图像时,表现出更高的匹配精度和稳定性。然而,由于混合配准方法结合了多种技术,其计算复杂度较高,导致计算速度相对较慢。

### 3 数据驱动深度学习方法

多模态遥感影像通常存在复杂的几何和辐射差异,知识驱动专家设计配准方法往往受限于专家知识范围,缺乏高级语义信息,难以确保在多模态遥感图像之间具有高度的可重复性和区分性特征对应关系,从而限制了其在复杂遥感场景中的应用效果(李星华等,2023)。

近年来,随着现代深度学习技术的发展,众多深度学习网络结构如卷积神经网络(CNN)(Lecun等,1998)、孪生神经网络(siamese networks network, SNN)(Chopra等,2005)、生成对抗网络(generative adversarial

network, GAN)(Goodfellow等,2014)、Transformer(Vaswani等,2017)等在图像处理中日益盛行。这些先进的网络结构为遥感图像配准方法带来了新的机遇,基于深度学习的遥感图像配准方法应运而生。通过直接从数据中学习,能够更高效地解决图像匹配和配准问题(Kuppala等,2020)。这些方法由数据驱动,试图从输入数据中抽象出分布结构,不仅能够提取复杂的特征信息,还能够更好地适应多种复杂场景和多模态数据。

目前数据驱动深度学习的配准方法可以进一步细化为以下四类:第一类集成学习配准方法,该类方法以传统配准框架为基础,将深度学习网络与传统算法相结合,用于特征提取、匹配加速或优化模型,从而形成新型混合配准策略,提升传统算法的适应性与处理能力。第二类是端到端学习方法,该类方法通过深度学习网络实现从图像对输入到目标结果的全流程学习,采用联合优化的方式完成特征提取、匹配或直接预测图像间的几何变换参数,相较于集

成学习方法更注重流程自动化。第三类是基于风格迁移的配准方法,通过生成对抗网络等深度学习模型对遥感图像进行风格迁移,显著减小多模态遥感图像间的辐射差异或纹理差异,为后续配准操作创造更加一致的图像特性。第四类是大模型驱动方法,基于预训练大模型强大的特征表征能力和跨模态泛化能力,该类方法为处理遥感图像中复杂非线性形变、多模态成像差异等难题提供了全新的解决路径,展现出重要的发展潜力。

#### 3.1 集成学习方法

集成学习方法通过将深度学习网络融入基于模板或基于特征的配准框架中,生成更加鲁棒且高效的特征表示或相似性度量机制,从而根据不同应用需求构建出更全面、更精准的匹配模型(眭海刚等,2022)。

常规的集成学习方法处理流程是,首先应用点特征提取算法在遥感图像中获取特定数量的特征点,然后以关键点为中心选取相关的图像斑块。接下来,通过描述符学习网络得到深度特征描述符或深度特征图。最后衡量斑块对的相似性(Quan等,2021),以此获得匹配点对用于计算配准变换参数。常见的集成学习方法根据深度学习网络是否参与特征匹配可划分为以下两种:描述符学习方法和度量学习方法。图11是描述符学习网络与度量学习网络简单对比示意图。

对于描述符学习方法,深度学习网络仅用于提取高级特征,其作用局限于特征生成环节,后续匹配及配准步骤由传统算法完成。而在度量学习方法中,深度学习网络不仅用于特征提取,还深度参与到后续的特征匹配及判别步骤,学习相似性度量机制进行后续配准步骤。

##### 1) 描述符学习方法

描述符学习方法基于深度学习网络提取图像的深层特征描述符或者深层特征图,利用特征描述符匹配或者深度特征图模板匹配获得匹配点对用于图像配准。常用的网络是Siamese网络,利用图像斑块的对比学习方式,最小化匹配图像斑块的特征距离,最大化不匹配图像斑块的特征距离。具体实例如Ye等(Ye F等,2018)构建了CNN特征和SIFT的组合特征并将组合特征集成到PSO-SIFT算法中,对遥感影像进行配准;Xu等(Xu W等,2023)提出了一种新的局部特征描述符SODescNet,其网络结构考虑

了SAR图像的特性,利用传统的ORB特征进行特征提取,建立了一个适用于SAR光学匹配的完整特征匹配框架;Zhou等(Zhou等,2022)借助伪孪生网络

提取多尺度卷积梯度特征描述图像结构特性,之后通过快速模板方案实现图像匹配;

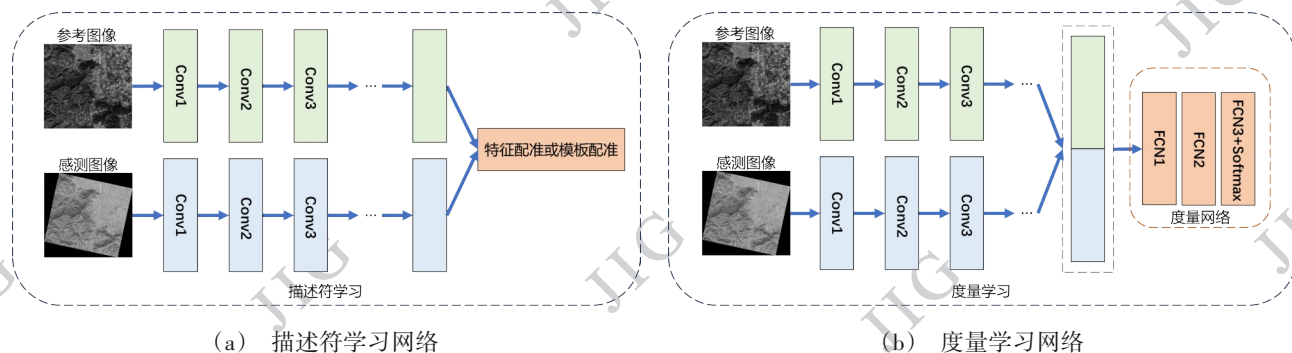


图 11 描述符学习网络与度量学习网络简单对比示意图

Fig. 11 Schematic diagram of a simple comparison between a descriptor learning network and a metric learning network. (a) a descriptor learning network; (b) a metric learning network.

Xiang等(Xiang等,2022)基于特征解耦网络进行光SAR图像复合配准,利用FAST进行关键点检测,然后设计了一个伪孪生全卷积网络和一个残差去噪网络来学习语义和噪声信息的深度特征表示;Zhang等(Zhang H等,2019)训练了一种采用最大正样本和硬负样本特征距离策略的新型损失函数的Siamese全卷积网络学习特征描述符,通过卷积运算计算相似度得分和定位平移偏移量后又提出多级特征融合方法和多频信道注意力模块,构建了光学和SAR图像匹配网络(OSMNet)(Zhang等,2022);Wu(Wu等,2022)等从孪生U-Net网络的两个分支中提取SAR和光学图像的深度特征,并利用互相关层根据这些特征生成影像对的热图,将热图的峰值最大位置作为最佳匹配位置。

描述符学习方法普遍侧重于对遥感影像局部特征进行深层、高维度的语义抽象与表征增强,在提升描述符判别力方面取得了显著进展。然而,在图像配准的整体流程中,其余关键环节仍大量依赖传统方法,缺乏与描述符学习之间的协同优化。

## 2) 度量学习方法

度量学习方法将特征描述和匹配看作一个统一的学习任务,通过学习相似性度量机制来关联图像特征,在此基础上完成后续的配准步骤。(李星华等,2023)。特征匹配质量很大程度上依赖于特征描述符与相似性测度的质量,因此,学习有效的特征相似性度量也很重要。相较于描述符学习方法中的固

定特征距离测量,度量学习方法中自适应相似性学习模式更加灵活(Quan等,2021)。度量学习的核心思想是将图像匹配任务视为二元分类问题,利用权重共享的孪生网络提取输入影像的深度特征,并自动预测输入图像对的匹配标签(权豆,2022)。在计算机视觉领域,常见的方法如MatchNet(Han等,2015)、HardNet(Mishchuk等,2017)等,而在遥感领域,学者们也参考这些方法或提出了专属的配准方案。如Burgmann等(Bürgmann等,2019)以SAR图像中的GCP为中心裁剪图像斑块,并在光学图像上移动搜索多个光学斑块输入到HardNet中计算欧几里得距离分数,以此获取匹配点对;Wang等(Wang等,2018)提出了一种自学习深度学习框架,直接学习遥感图像斑块对其匹配标签之间的端到端映射,应用迁移学习来降低训练成本;Quan等(Quan等,2021)提出了融合多特征关系的多相关注意力学习网络,通过关系注意学习模块自适应地进行图像斑块匹配,之后又提出了自蒸馏特征学习网络(Quan等,2022),利用更多相似性信息进行深度网络优化;Fang等(Fang等,2022)将Siamese U-net、基于FFT的互相关层和softmax2d分类器结合,将图像匹配转换为分类问题,避免了传统图像匹配中通过FFT相关进行耗时的滑动操作,适用于不同尺寸的图像。

除了描述符学习方法和度量学习方法之外,部分研究者也提出了其他与手工设计相结合的方法,

例如在深度特征图上执行特征提取操作。如 Xu 等 (Xu 等, 2022) 在 CNN 生成的深度特征图上识别特征点及其硬描述, 并结合包含邻域梯度信息的软描述符形成联合描述符。或直接结合深度学习与手工设计方法从粗到细的配准遥感图像。如 Ma 等 (Ma 等, 2019) 采用由粗到细的两步策略实现高精度遥感图像配准, 首先利用 VGG-16 提取深度特征对图像对进行粗略对齐, 之后利用局部特征精确调整匹配结果。

集成学习方法特征描述与检测的过程往往相互独立, 通常依赖手工制作的特征检测器来检测关键点。这些方法仍需进行大量的预处理工作并依赖于人工干预, 异常值和不准确的匹配仍然是无法避免的问题 (Hughes 等, 2020)。而且特征提取后丢弃了图像的全局空间信息, 使特征描述仅关注关键点小范围内邻域的局部信息 (Cui 等, 2022)。图像匹配的效果往往受限于手工设计的检测器性能, 未能充分发挥深度学习的潜力。因此, 学者们的研究逐渐转向全流程自动化、端到端的配准模式。

### 3.2 端到端学习方法

当前, 基于深度学习的端到端学习方法主要可以分为两类: 第一类方法是级联特征配准方法, 通过端到端的深度学习网络实现自动、完整的特征提取步骤, 随后完成特征匹配和配准。该方法有效提升了特征提取的自动化与鲁棒性。第二类方法是直接变换配准方法, 利用端到端学习方式直接预测图像的变换模型参数。在此方法中, 无需对图像进行显式的特征描述与匹配操作, 网络能够通过统一的学习框架直接预测图像间的几何变换参数。图 12 是这两种方法的简单示意图。

#### 1) 级联特征配准

近年来, 在计算机视觉领域中, 越来越多的研究者提出专门为关键点检测设计的深度学习方法。此类方法通常利用 CNN 获取图像的特征, 然后建立相应的响应图, 最后从响应图中提取局部响应值最大的点作为特征点。如 TILDE (temporally invariant learned detector) (Verdie 等, 2015)、Key-Net (Barroso-Laguna and Mikolajczyk, 2023) 等是典型算法。而在遥感配准领域少有专门为特征检测设计的网络, 如 Zhang 等 (Zhang S 等, 2023) 提出了一种用于光学和 SAR 图像关键点检测的多级注意力孪生网络, 能够为异构图像匹配提供更多可重复的关键

点。随着深度学习方法在全局特征学习和局部特征学习方面取得显著优势, 使得图像匹配的研究更加倾向于利用联合学习进行特征的检测和描述。本文称为级联特征配准方法, 这种方法旨在增进特征检测与描述之间的紧密联系, 通过端到端网络实现自动化特征提取, 并将提取的高级特征用于后续的特征匹配及配准步骤。

计算机视觉领域中, Kwang 等 (Yi 等, 2016) 提出了一种开创性的学习不变特征变换 (learned invariant feature transform, LIFT) 方法, 该方法是第一个端到端特征检测和描述方法, 结合了三个 CNN 分别进行关键

点检测、方向估计和特征描述, 执行图像匹配; DeTone (DeTone 等, 2018) 提出的 SuperPoint, 输入全尺寸图像并在一次前向传递中联合计算像素级兴趣点位置和相关描述符, 引入了完全卷积模型, 并使用单调自适应模块实现自监督训练, 有效提高检测重复性, 如图 13 是 SuperPoint 的自监督训练过程; Dusmanu 等 (Dusmanu 等, 2019) 提出了一种依赖于描述符中编码的高级特征的关键点检测器 D2-Net, 它在检测和描述网络之间共享所有参数, 并使用可同时对两个任务进行优化的联合公式。遥感领域的学者们也探索了许多特征检测与描述的联合学习方法。如 Cui 等 (Cui 等, 2022) 设计了一种新型的端到端全卷积网络, 利用注意力机制和空间金字塔聚合池化提取和描述适合光 SAR 图像匹配任务的关键特征; Hughes 等 (Hughes 等, 2020) 提出了一个用于光 SAR 图像稀疏匹配的全自动端到端框架, 通过三个深度学习网络替代了全局匹配的阶段, 每个网络都执特定任务包括特征检测、特征匹配和异常值抑制; Li 等 (Li L 等, 2022b) 提出了一种两步“检测+匹配”网络框架, 利用自监督检测网络提取高度可重复的关键点, 并将描述和匹配整合在同一网络中进行优化。

随着 Transformer 架构的发展, 基于该架构的方法通过捕捉图像中的长距离依赖关系, 在遥感图像配准中取得了优异效果。如 Liu 等 (Liu 等, 2023) 提出了 SIFNet, 利用特征金字塔网络提取局部像素多尺度特征, 并通过 Transformer 中的自注意机制实现特征交互, 结合相似性与语义匹配损失函数, 将卫星影像配准转化为回归任务; Zhang 等 (Zhang Yongxian 等, 2024) 提出模态无关一致性匹配方法, 同样利

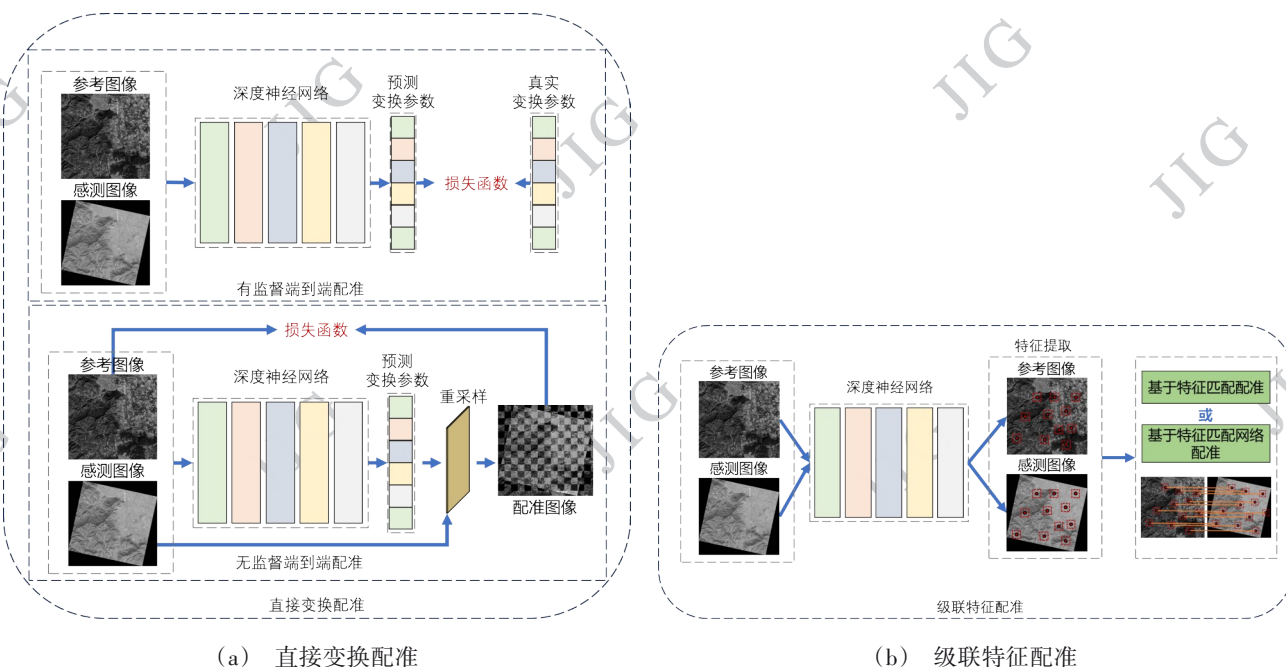


图 12 端到端学习方法示意图

Fig. 12 Schematic diagram of an end-to-end learning approach. (a) direct transformation registration; (b) cascade feature registration.

用 CNN 和 Transformer 注意力机制的能力聚合学习特征,并设计特征一致性关联策略保证正确关联结果;Wang等(Wang等,2025)提出的多层次流场预测网络整合了多尺度特征提取、Swin-Transformer 注意力机制和分层 STN (spatial transformer network) 网络,充分利用图像不同尺度和层级的特征,逐步优化流场预测以对齐局部非刚性畸变。

分割一切模型(segment anything model, SAM) (Kirillov 等, 2023)在边界分割方面取得了重大突破,推动了大规模计算机视觉模型的发展,并已被学

者应用于遥感图像配准领域。如 Gao 等的工作 (Gao 等, 2025)首次将 SAM 应用于多模态遥感图像匹配领域,通过边缘引导位置信息增强局部特征,更好地实现对局部特征的稳健描述,并基于此进行度量学习匹配预测。

2) 直接变换配准

现阶段,深度学习在遥感图像配准领域的主要应用方法是基于深度学习网络提取的深度特征进行配准。然而,这些方法需

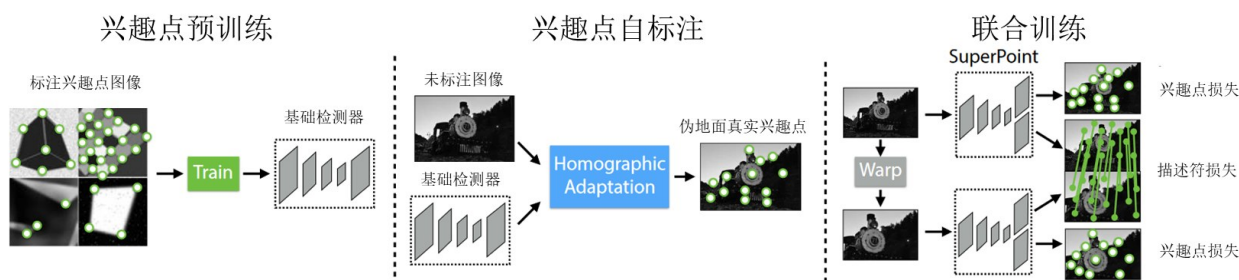


图 13 SuperPoint 自我监督训练概述 (DeTone 等, 2018)

Fig 13 SuperPoint Self-Supervised Training Overview (DeTone 等, 2018)

要处理繁琐的检测-描述-匹配步骤,难以实现全程配准训练。所谓端到端,更多的是指特征提取

过程的端到端。近些年,许多研究利用深度学习网络直接回归图像配准参数进行匹配,无需特征描述

和特征匹配步骤,也就是本节要介绍的直接变换配准方法。直接变换配准方法通过端到端学习框架,利用深度学习的强大计算能力,直接预测图像间的几何变换参数(Li L 等,2022a)。这种直接变换配准方法,根据训练过程中是否依赖真实的变换参数,可划分为有监督方法与无监督方法,有监督方法通过数据驱动直接预测变换参数,并利用预测参数与真实参数之间的损失函数训练网络,以学习真实变换模型;而无监督方法无需真实变换参数,而是将配准后图像与参考图像之间的差异作为损失函数,通过训练隐式地学习合理的空间变换。具体实例如Ye等(Ye 等,2022a)利用无监督深度学习网络(MU-Net)取代了相似性优化过程,直接学习从图像到对其变换参数的端到端映射,并设计了基于结构相似性的损失函数;Li等(Li L 等,2022a)使用D-DenseNet(dual DenseNet)模型对四个角点位移进行回归,然后应用投影成像方程计算定点坐标和动点坐标之间的投影变换矩阵;Li等(Li L 等,2023)将配准定义为回归规定点位移的函数,避免了特征检测-描述-几何约束处理,但是该方法受限与固定图像尺寸;Feng等(Feng 等,2022)开发了一种基于从粗到细的无监督级联卷积网络的自动图像配准算法,首先使用卷积神经网络来学习图像之间的仿射参数,然后使用由具有不同膨胀速率的膨胀卷积组成的编码器-解码器模块进行细化;Xiao等(Xiao 等,2024)提出一种由仿射配准和可变形配准阶段组成的两阶段网络ADNet(affine and deformable registration

networks),对不同视点和变形尺度的遥感图像表现出优异的性能。

端到端学习方法通过联合学习特征的检测和描述,或直接学习图像间的变换模型参数,减少了配准中间步骤和手工干预。这类方法虽然显著简化了配准过程,提高了自动化程度和效率,但是它们对大量高质量标注数据的依赖较高,且需要强大的计算资源支持。此外,端到端学习模型的解释性较差,难以解释其内部工作机制,这些限制因素影响了端到端学习方法在实际场景中的广泛应用。

### 3.3 风格迁移方法

多模态遥感影像间的显著强度差异给图像配准带来了挑战。许多学者利用GAN或其他深度学习网络进行跨模态映射,将多模态图像统一为同一模态、或减小图像差异后,再应用传统手工设计方法或者深度学习方法进行精确配准。如Merkle等(Merkle 等,2018)和Maggiolo等(Maggiolo 等,2022)利用条件对抗网络(conditional adversarial networks, cGANs)将光学斑块转换为伪SAR斑块,之后通过传统的基于模板或基于特征的方法进行配准;Zhang等(Zhang J 等,2019)利用基于CNN的图像传输算法对多模态遥感图像进行预处理,消除差异后采用传统的局部特征方法再进行匹配;Hang等(Huang 等,2019)通过改进的CycleGAN将SAR图像转换为伪光学图像,然后利用SIFT进行配准;Hu等(Hu C 等,2023)通过改进的

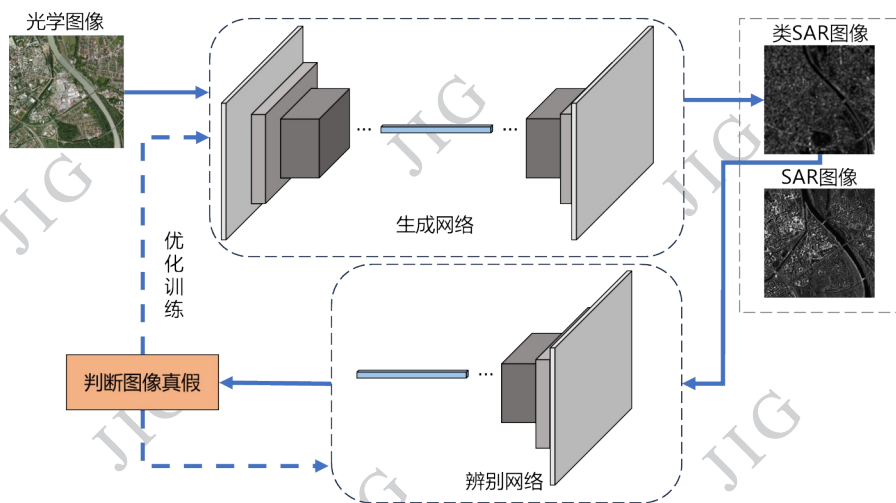


图14 基于GAN的图像风格迁移示意图

Fig. 14 Schematic of GAN based image style migration

Restormer网络(Zamir等,2022)将光学图像转换为伪SAR图像,结合基于ROEWA的多尺度Harris空间和MatchosNet进行特征检测与描述。图14是利用生成对抗网络进行遥感图像风格迁移的示意图。

风格迁移方法在多模态遥感图像配准中亦存在一定局限性,其在改变图像风格的同时,容易损害对配准至关重要的信息;对配准数据依赖强,跨区域泛化能力有限;在风格迁移完成后,后续的配准步骤也会引入额外误差,这些不足制约了风格迁移方法在遥感图像配准中的实用性。

### 3.4 大模型驱动方法

随着大模型的快速发展,学者们已开始将其拓展应用于遥感图像配准领域,探索其在复杂场景下的配准性能与泛化能力。当前,大模型驱动方法主要集中在两个方向:基于大模型的视觉表征增强与语义引导的特征对齐。前者利用视觉基础模型的强大表征能力,从海量数据中学习通用、鲁棒的特征描述子,以提升跨模态、跨场景匹配的泛化性;后者则通过知识蒸馏等技术,将大模型习得的高级语义理解能力注入配准流程,实现由“几何驱动”向“语义引导”的范式转变。

#### 1) 大模型驱动的视觉表征增强模式

大模型具备强大的特征表征与跨模态泛化能力为突破传统配准方法的局限提供了全新路径。以RoMA(Edstedt等,2024,2025)系列为代表的方案,利用视觉基础模型(visual foundation models, VFMs)强大的表征能力,将DINO系列模型作为基础提取多尺度描述符,实现了跨源、跨域影像的高效匹配,其基于DINO v3的更优表现进一步印证了进化后的视觉大模型对极值场景下特征鲁棒匹配的语义锚定能力。基于海量数据驱动的预训练范式,正推动配准模型从任务定制迈向通用适配。GIM(Shen等,2024)框架通过自训练机制从海量无标签视频和遥感序列中挖掘稠密对应关系,使LightGlue(Lindenberg等,2023)等匹配算法也能有效适应未见过的遥感传感器类型,彰显了数据规模驱动的泛化潜力;MatchAnything(He等,2025)通过构建多源数据集混合引擎生成大规模训练数据,并借助DINO模型强化特征的泛化判别力,进一步拓展了通用匹配模型的跨模态适应边界;MINIMA(Ren等,2025)则另辟蹊径,通过在大规模多模态数据上进行对比学习,从

数据分布的底层对齐跨模态语义,显著提升了SAR-光学等异构影像配准的鲁棒性。

随着语义增强路径的深化,研究者意识到仅靠单一模型的语义注入仍难以应对极端场景的挑战,生成模型为配准提供了全新的解决思路。RFDifNet(Diao等,2025)在合成目标影像的同时纠正几何偏差,并利用扩散过程的概率特性对配准不确定性进行显式建模,显著提升了云层覆盖、严重遮挡等复杂场景下的配准稳定性,实现了生成即对齐的范式突破。

#### 2) 语义引导的特征对齐与高效知识蒸馏

在基础特征替代的基础上,研究者开始探索如何让大模型将高级语义理解能力注入配准流程。SAMFeat(Wu J等,2025)首个系统性地将分割一切模型(SAM)(Kirillov等,2023)作为教师模型赋能局部特征学习与匹配,通过知识蒸馏将SAM习得的与类别无关的语义表征注入轻量级学生网络,使模型在推理时无需引入大模型开销即可获得语义增强的局部特征。沿此方向,SemaGlue(Zhang S等,2025)利用预训练分割模型作为语义特征源,通过精巧的对齐机制使几何特征与语义特征在特征层面深度融合,实现边感知、边匹配的新范式,也标志着配准技术从单纯的几何对齐向语义引导的对齐迈进。

尽管相关表现优异,但大模型在处理配准任务时面临运算量巨大与推理延迟高的挑战,导致其难以直接应用于端侧或实时处理场景。DistillMatch(Yang等,2025)则以DINO自监督视觉Transformer为教师,巧妙利用知识蒸馏技术,将大型视觉基础模型强大的通用语义能力传授给轻量级模型,在有效缓解模态差异与数据稀缺难题的同时实现了高效部署,为上述各类先进方法向实际应用转化提供了可行路径。

## 4 遥感图像匹配配准数据集

当前遥感图像配准领域内,大多数方法均在各自独立的数据集和实验设定下进行评测,所报告的精度与鲁棒性指标依赖于不同的场景、数据规模及评价准则,难以直接横向对比。可见,构建统一、公开的基准数据集,对于推动遥感图像配准方法的性能评估与算法设计至关重要。

近年来,随着深度学习在图像配准领域的广泛  
©中国图象图形学报版权所有

应用,丰富多样的基准数据集对模型训练与评测的作用愈发突出。在光学图像方面,计算机视觉领域已有多个成熟的数据集可供沿用,如 MegaDepth (Li and Snavely, 2018)、ScanNet (Dai 等, 2017)、HPatches (Balntas 等, 2017) 等。本节将重点聚焦于针对多模态图像匹配配准任务的数据集进行梳理与

分析。通过系统性总结多模态配准领域中具有代表性的数据集,重点解析其模态构成、数据规模、空间分辨率、覆盖场景及适用性等,为研究者选择针对相关配准任务的数据集提供有效参考,促进后续研究在统一条件下对不同方法进行系统分析与公平比较,汇总参见表2。

表2 遥感图像匹配数据集

Table 2 Remote sensing image matching dataset

数据集名称	影像类型	影像对数量	地理覆盖与场景说明	数据集链接
SEN1-2	光-SAR	282384对	全球陆地区域,覆盖四季,具有地形和季节多样性	<a href="https://mediatum.ub.tum.de/1436631">https://mediatum.ub.tum.de/1436631</a>
WHU-OPT-SAR	光-SAR	100对(大幅面)	覆盖湖北省约5万平方公里,地形、植被类型多样	<a href="https://github.com/AmberHen/WHU-OPT-SAR-dataset.git">https://github.com/AmberHen/WHU-OPT-SAR-dataset.git</a>
QXS-SAROPT	光-SAR	20000对	覆盖圣迭戈、上海、青岛三地;涵盖港口、集装箱堆场、道路、建筑、海面、植被、船舶等	<a href="https://github.com/yaoxu008/QXS-SAROPT">https://github.com/yaoxu008/QXS-SAROPT</a>
OSDataset2.0	光-SAR	6,476对(局部)+1对(全幅)	覆盖14个国家6类典型地物;全幅场景覆盖丰富地物	<a href="https://radars.ac.cn/web/data/get-Data?dataType=DatasetinthePaper">https://radars.ac.cn/web/data/get-Data?dataType=DatasetinthePaper</a>
MultiResSAR	光-SAR	10850对	全球多个国家和城市;城市、乡村、平原、丘陵、山区、水域六类场景	<a href="https://github.com/betterlll/Multi-Resolution-SAR-dataset-">https://github.com/betterlll/Multi-Resolution-SAR-dataset-</a>
SOMA-1M	光-SAR	1300000+对	全球多尺度覆盖,包含12种典型土地覆盖类别	<a href="https://github.com/PeihaoWu/SOMA-1M">https://github.com/PeihaoWu/SOMA-1M</a>
MapData	光-电子地图	121781对	全球范围;城市建成区、乡村聚落、山地、平原盆地、荒漠等	<a href="https://github.com/PeihaoWu/Map-Glue">https://github.com/PeihaoWu/Map-Glue</a>
GLS-MIFT	多时相光学、光学-SAR、光学-红外、光学-地图等	1110对	跨领域场景(计算机视觉、医学、遥感),遥感部分主要覆盖城市及郊区	<a href="https://github.com/Zhongli-Fan/GLS-MIFT">https://github.com/Zhongli-Fan/GLS-MIFT</a>

同时,本节以 MultiResSAR 数据集为例,展示其原文基于该数据集构建的可参照定量对比框架 (Zhang W 等, 2025), 参见表3。建议后续研究尽量基于统一公开基准进行方法对比,以更清晰地揭示不同方法在相同条件下的相对优劣及适用范围。

1) SEN1-2 数据集 (Schmitt 等, 2018) 是目前应用较为广泛的光学-SAR 配准基准数据集。该数据集由哨兵1号 (Sentinel-1) SAR 影像与哨兵2号 (Sentinel-2) 光学影像配对生成,共包含 282384 对图像块,每块大小为 256×256 像素,空间分辨率均为 10m。其数据采集自全球陆地,覆盖四个季节,具有较好的地形多样性和季节多样性。但 SEN1-2 的数据分辨率较低且图像块较小,难以支撑高分辨率大尺寸影像配准任务。

2) WHU-OPT-SAR 数据集 (Li X 等, 2022) 包含 100 对大幅光学-SAR 图像对,图像尺寸为 5556×3704 像素,覆盖湖北省约 5 万平方公里的区域,地形类型多样。其中,光学影像来自高分一号 (GF-1), SAR 影像来自高分三号 (GF-3),影像重采样后分辨率为 5m。该数据集涵盖多种地形与植被类型,具有丰富的场景多样性与地形复杂性。该数据集最初面向土地利用分类任务构建,提供了成对的光学与 SAR 影像,亦可尝试用于图像配准任务。

3) QXS-SAROPT 数据集 (Huang 等, 2021) 包含 20,000 对 256×256 像素的光学-SAR 图像对,空间分辨率为 1 米, SAR 数据来自 GF-3; 光学数据来自 GoogleEarth。覆盖圣迭戈、上海与青岛三地,覆盖港口、集装箱堆场、道路与建筑等典型目标,同时保留

海面与植被等自然要素以及船舶目标(向俞明等, 2025), 该数据集能够有效支持光-SAR 图像匹配。然而, 该数据集的地理覆盖范围仅限于三个港口城市, 且图像对尺寸固定, 难以满足高分辨率大尺度影像配准的需求。

4) OSDataset2.0 数据集(向俞明等, 2025)是大规模 SAR-光学影像匹配基准数据集。其局部训练集包含覆盖 14 个国家、6 类典型地物的 6,476 对 512×512 像素配准图像块, 重采样至 1 m 分辨率; 全幅测试集提供一对 Google Earth 光学影像与 GF-3 SAR 影像, 0.43 m 分辨率, 涵盖丰富地物, 并收录 72 对高精度同名点。该数据集兼顾训练与整景测试需求, 规模较大且标注精细。但基于该数据集的评测显示, 各算法在全幅场景下的均方根误差 (root

mean square error, RMSE) 均高于 4 个像素, 表明高精度 SAR-光学匹配仍面临一定挑战。

5) MultiResSAR 数据集(Zhang W 等, 2025) 针对多分辨率 SAR 与光学遥感图像配准而创建, 包含 10850 对多源、多分辨率、多场景的 SAR 与光学图像对。SAR 数据来源涵盖 Sentinel-1、GF-3、HT1-A 及 Umbra, 光学影像来自 Google Earth, 分辨率范围为 0.16m 至 10m。数据集覆盖全球多个国家和城市, 包含城市、乡村、平原、丘陵、山区及水域等六类典型场景。其突出特点在于系统性地覆盖了不同分辨率组合下的配准场景, 为评估算法在分辨率差异显著条件下的鲁棒性提供了重要的基准数据。数据集虽然覆盖了 0.16m-10m 多分辨率, 但仅有 1.1 万对样本, 无法支撑大模型预训练。

表 3 MultiResSAR 数据集的平均配准性能评估结果(Zhang W 等, 2025)

Table 3 Average registration performance evaluation results for the MultiResSAR dataset(Zhang W 等, 2025)

算法	成功率(%)	均方根误差 (像素)	正确匹配 数	时间(秒)	
RIFT(Li 等, 2020)	66.51%	3.58	108.40	5.283	
HOWP(Zhang Y 等, 2023)	52.63%	4.11	115.58	12.420	
ASS(Xiong 等, 2022)	39.34%	3.83	86.31	5.717	
知识驱动专家设计方法	MOSS(Zhang Yongjun 等, 2024)	17.93%	3.98	114.36	9.666
SRIF(Li J 等, 2023)	13.28%	5.35	230.06	7.373	
SAR-SIFT(Dellinger 等, 2012)	1.68%	4.47	38.59	3.569	
LNIFT(Li J 等, 2022)	0.41%	4.71	61.53	1.602	
XoFTR(Tuzcuoglu 等, 2024)	40.58%	3.03	244.26	0.032	
RoMa(Edstedt 等, 2024)	35.26%	3.15	589.70	1.065	
XFeat(Potje 等, 2024)	36.29%	4.94	73.24	0.017	
DKM(Edstedt 等, 2023)	29.85%	3.41	395.96	0.443	
数据驱动深度学习方法	SuperGlue(Sarlin 等, 2020)	29.02%	3.39	74.03	1.926
LightGlue(Lindenberger 等, 2023)	26.41%	3.34	122.85	0.042	
Efficient LoFTR(Wang 等, 2024)	15.71%	3.70	100.17	0.040	
LoFTR(Sun 等, 2021)	9.50%	3.68	80.77	0.037	
SGM-NET(Chen H 等, 2021)	2.05%	3.16	122.15	0.550	

6) SOMA-1M 数据集(Wu 等, 2026)是于 2026 年发布的超大规模遥感图像配准数据集, 包含超过 130 万对像素级精确对齐的 512×512 像素图像对。该数据集整合了 Sentinel-1、PIESAT-1、Capella Space 及 Google Earth 等多源卫星影像, 实现 0.5 m 至 10 m

的全球多尺度覆盖, 涵盖 12 种典型土地覆盖类别, 显著提升了场景的多样性与复杂性。SOMA-1M 采用严格的由粗到精图像匹配框架生成高精度对齐数据, 为多模态基础模型的相关研究提供了重要的数据基础设施支撑。

7) MapData数据集(Wu P等,2025)是大规模多模态电子地图与可见光图像数据集,提供121781对对齐的电子地图与可见图像,图像来源于Google Earth,涵盖城市建成区、乡村聚落、山地地形、平原盆地及荒漠区域等多种典型场景,覆盖全球范围。

8) GLS-MIFT数据集(Fan等,2024)涵盖计算机视觉、医学以及遥感领域的12种类型,共计1110组多模态影像对,囊括了以上领域的典型应用场景。其中,遥感数据部分包含多时相-光学、光学-SAR、光学-红外以及光学-地图等数据类型,主要覆盖城市及郊区区域。该数据集的优势在于其跨领域的多模态特性和广泛的应用场景,然而,数据量相对较小,可能难以满足大规模深度学习模型的训练需求。

## 5 发展与展望

在过去的数十年中,遥感图像配准作为遥感图像处理的重要环节,得到了广泛的关注和应用,并在技术上取得了显著的进步。然而,尽管已经取得了一定的成果,遥感图像配准仍然面临着许多挑战和问题。遥感图像配准领域的研究难点和焦点主要集中在以下几个方面。

1) 关注遥感图像特性的深度学习配准:目前,多数深度学习配准模型及训练数据集源自计算机视觉领域,忽视了遥感影像特有的成像物理模型、大幅面多模态特性及地学知识,且缺乏高质量标注数据,导致复杂环境下的多源遥感影像配准效果不佳。如Du等(Du等,2025)已指出现有深度学习方法局限于视觉表现,忽视了SAR的内在特性,进而提出了面向SAR图像理解量身定制的模型。SAR独特的成像机制与物理特性亦不能简单利用计算机视觉领域训练集,迄今仍缺少像计算机视觉通用任务中那样广泛认可的大规模基准数据集(向俞明等,2025)。因此,未来研究应集中于整合多模态遥感数据的成像探测机理与遥感地理学的知识,构建专门针对遥感图像配准的深度学习模型与数据集,深入挖掘遥感图像的丰富语义层次,实现高效精准的多模态配准,从而推动遥感图像分析的智能化发展。

2) 遥感图像在轨实时配准:工程应用中对遥感图像的配准速度往往有较高的要求。随着遥感影像获取技术的日益成熟,遥感影像数据量呈爆发式增长。传统地面处理模式难以满足实时性与高效性需

求,亟需将算法迁移至星载计算平台,以降低传输带宽、提升数据时效,服务于应急减灾、交通运输、目标检测等领域(Zhao等,2024b)。近年来,已有研究尝试探索面向星上环境的轻量化配准方案,如轻量化模型(Priyasad等,2026)与轻量化参考数据等(Zhao等,2024b)。然而,现有方法多面向特定传感器或有限场景,尚缺乏兼顾高精度、强鲁棒与低资源消耗的通用型星上配准方案。如何在保证配准精度的前提下,提高遥感图像配准的效率,仍是一个亟待解决的问题。未来的研究需要进一步在算法轻量化与硬件加速方面进行协同创新,以实现快速且精确的图像配准,满足实际在轨应用中对实时响应与低资源开销的双重要求。

3) 大尺寸遥感图像深度学习配准:目前,利用深度学习进行的遥感图像配准实验大多基于分块数据或特定的数据集进行。而大尺寸遥感图像覆盖范围广、分辨率高、地物信息丰富,对算法的计算效率和场景适应能力提出了更高要求。将通用图像匹配算法直接应用于大幅卫星数据时,其兼容性问题尤为突出,具体表现为计算资源需求急剧增加,以及对复杂地理环境与显著地表变化处理的不足(Geng等,2025)。这使得大幅遥感图像的配准成为一项具有挑战性的任务。尽管深度学习在图像处理领域已经取得了显著的进展,但专门面向大幅遥感图像的配准框架仍然相对匮乏。为此,亟需研发能够高效处理大幅遥感图像的深度学习方法,如设计适配大幅面特性的新型网络架构、优化现有模型的训练与推理策略等方法。

## 6 结语

遥感图像配准是遥感图像处理领域的关键环节,其结果直接关系到后续图像融合、图像拼接、目标检测和变化感知等多个重要应用。在过去几十年中,众多研究人员提出了多样的配准方法,在该领域取得重大进展。本文对现有的遥感图像配准方法进行了全面的讨论,将其归纳为两大类,一类是知识驱动专家设计方法,一类是数据驱动深度学习方法,并分别进行了详细的归纳与分析。最后总结了当前面临的挑战与未来发展方向。

致谢:。

## 参考文献(References)

- Abdel-Hakim A E and Farag A A. 2006. CSIFT: a SIFT descriptor with color invariant characteristics//2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06): Vol. 2. 1978-1983 [DOI: 10.1109/CVPR.2006.95].
- Alahi A, Ortiz R and Vanderghyest P. 2012. FREAK: fast retina key-point//2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, RI: IEEE: 510-517 [DOI: 10.1109/CVPR.2012.6247715].
- Alcantarilla P F, Bartoli A and Davison A J. 2012. KAZE features//Fitzgibbon A, Lazebnik S, Perona P, Sato Y, Schmid C. Computer Vision - ECCV 2012. Berlin, Heidelberg: Springer: 214-227 [DOI: 10.1007/978-3-642-33783-3\_16].
- Alcantarilla P, Nuevo J and Bartoli A. 2013. Fast explicit diffusion for accelerated features in nonlinear scale spaces//Proceedings of the British Machine Vision Conference 2013. Bristol: British Machine Vision Association: 13.1-13.11 [DOI: 10.5244/C.27.13].
- Baltas V, Lenc K, Vedaldi A and Mikolajczyk K. 2017. HPatches: a benchmark and evaluation of handcrafted and learned local descriptors//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 3852-3861 [DOI: 10.1109/CVPR.2017.410].
- Barroso-Laguna A and Mikolajczyk K. 2023. Key.net: keypoint detection by handcrafted and learned CNN filters revisited. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(1): 698-711 [DOI: 10.1109/TPAMI.2022.3145820].
- Bay H, Tuytelaars T and Van Gool L. 2006. SURF: speeded up robust features//Leonardis A, Bischof H, Pinz A. Computer Vision - ECCV 2006. Berlin, Heidelberg: Springer: 404-417 [DOI: 10.1007/11744023\_32].
- Bürgmann T, Koppe W and Schmitt M. 2019. Matching of TerraSAR-X derived ground control points to optical image patches using deep learning. ISPRS Journal of Photogrammetry and Remote Sensing, 158: 241-248 [DOI: 10.1016/j.isprsjprs.2019.09.010].
- Cai X. 2021. Heterogeneous image matching algorithm based on phase feature. Optics & Optoelectronic Technology, 19(1): 48-53 (蔡潇, 2021. 基于相位特征的异源图像匹配算法. 光学与光电技术, 19(1): 48-53) [DOI: 10.19519/j.cnki.1672-3392.2021.01.007].
- Calonder M, Lepetit V, Strecha C and Fua P. 2010. BRIEF: binary robust independent elementary features//Daniilidis K, Maragos P, Paragios N. Computer Vision - ECCV 2010. Berlin, Heidelberg: Springer: 778-792 [DOI: 10.1007/978-3-642-15561-1\_56].
- Canny J. 1986. A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-8(6): 679-698 [DOI: 10.1109/TPAMI.1986.4767851].
- Chai X, Chen J, Mao Z and Zhu Q. 2023. An upscaling - downscaling optimal seamline detection algorithm for very large remote sensing image mosaicking. Remote Sensing, 15(1): 89 [DOI: 10.3390/rs15010089].
- Chen H, Luo Z, Zhang J, Zhou L, Bai X, Hu Z, et al. 2021. Learning to match features with seeded graph matching network//2021 IEEE/CVF International Conference on Computer Vision (ICCV). 6281-6290 [DOI: 10.1109/ICCV48922.2021.00624].
- Chen J, Tian J, Lee N, Zheng J, Smith R T and Laine A F. 2010. A partial intensity invariant feature descriptor for multimodal retinal image registration. IEEE Transactions on Bio-Medical Engineering, 57(7): 1707-1718 [DOI: 10.1109/TBME.2010.2042169].
- Chen S, Li X and Zhao L. 2017. A novel local pattern based self-similarity descriptor for multisource remote sensing image registration//2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). 5153-5156 [DOI: 10.1109/IGARSS.2017.8128162].
- Chen S, Zhong S, Xue B, Li X, Zhao L and Chang C I. 2021. Iterative scale-invariant feature transform for remote sensing image registration. IEEE Transactions on Geoscience and Remote Sensing, 59(4): 3244-3265 [DOI: 10.1109/TGRS.2020.3008609].
- Cheng L, Gong J, Yang X, Fan C and Han P. 2008. Robust affine invariant feature extraction for image matching. IEEE Geoscience and Remote Sensing Letters, 5(2): 246-250 [DOI: 10.1109/LGRS.2008.915599].
- Chopra S, Hadsell R and LeCun Y. 2005. Learning a similarity metric discriminatively, with application to face verification//2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05): Vol. 1. 539-546 vol. 1 [DOI: 10.1109/CVPR.2005.202].
- Cui S, Ma A, Zhang L, Xu M and Zhong Y. 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: 1-13 [DOI: 10.1109/TGRS.2021.3066432].
- Dai A, Chang A X, Savva M, Halber M, Funkhouser T and Nießner M. 2017. ScanNet: richly-annotated 3D reconstructions of indoor scenes//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2432-2443 [DOI: 10.1109/CVPR.2017.261].
- Dalal N and Triggs B. 2005. Histograms of oriented gradients for human detection//2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05): Vol. 1. San Diego, CA, USA: IEEE: 886-893 [DOI: 10.1109/CVPR.2005.177].
- Dame A and Marchand E. 2012. Second-order optimization of mutual information for real-time image registration. IEEE Transactions on Image Processing, 21(9): 4190-4203 [DOI: 10.1109/TIP.2012.2199124].
- Dellinger F, Delon J, Gousseau Y, Michel J and Tupin F. 2012. SAR-

- SIFT: a SIFT-like algorithm for applications on SAR images//2012 IEEE International Geoscience and Remote Sensing Symposium. 3478-3481 [DOI: 10.1109/IGARSS.2012.6350671].
- DeTone D, Malisiewicz T and Rabinovich A. 2018. SuperPoint: self-supervised interest point detection and description//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Salt Lake City, UT, USA: IEEE: 337-33712 [DOI: 10.1109/CVPRW.2018.00060].
- Diao W, Hu L, Zhang K and Xiao L. 2025. Dual-conditionally guided diffusion models for fusion of unregistered multisource remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 63: 1-17 [DOI: 10.1109/TGRS.2025.3577073].
- Du Y, Chen Y, Huang L, Yang Y, Ghamisi P and Du Q. 2025. SUMMIT: a SAR foundation model with multiple auxiliary tasks enhanced intrinsic characteristics. *International Journal of Applied Earth Observation and Geoinformation*, 141: 104624 [DOI: 10.1016/j.jag.2025.104624].
- Duan C, Meng X, Tu C and Yang C. 2008. How to make local image features more efficient and distinctive. *IET Computer Vision*, 2(3): 178 [DOI: 10.1049/iet-cvi:20070049].
- Dusmanu M, Rocco I, Pajdla T, Pollefeys M, Sivic J, Torii A, et al. 2019. D2-net: a trainable CNN for joint description and detection of local features//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 8084-8093 [DOI: 10.1109/CVPR.2019.00828].
- Edstedt J, Athanasiadis I, Wadenbäck M and Felsberg M. 2023. DKM: dense kernelized feature matching for geometry estimation//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 17765-17775 [DOI: 10.1109/CVPR52729.2023.01704].
- Edstedt J, Nordström D, Zhang Y, Bökman G, Astermark J, Larsson V, et al. 2025. RoMa v2: harder better faster denser feature matching. *arXiv* [DOI: 10.48550/arXiv.2511.15706].
- Edstedt J, Sun Q, Bökman G, Wadenbäck M and Felsberg M. 2024. RoMa: robust dense feature matching//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 19790-19800 [DOI: 10.1109/CVPR52733.2024.01871].
- Fan B, Wu F and Hu Z. 2012. Rotationally invariant descriptors using intensity order pooling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(10): 2031-2045 [DOI: 10.1109/TPAMI.2011.277].
- Fan J W. 2017. Research on SAR image registration based on feature point. Xian: Xidian University (樊建伟. 2017. 基于特征点的SAR图像配准算法研究. 西安: 西安电子科技大学)
- Fan J, Wu Y, Li M, Liang W and Cao Y. 2018. SAR and optical image registration using nonlinear diffusion and phase congruency structural descriptor. *IEEE Transactions on Geoscience and Remote Sensing*, 56(9): 5368-5379 [DOI: 10.1109/TGRS. 2018. 2815523].
- Fan Z, Pi Y, Wang M, Kang Y and Tan K. 2024. GLS - MIFT: a modality invariant feature transform with global-to-local searching. *Information Fusion*, 105: 102252 [DOI: 10.1016/j.inffus. 2024. 102252].
- Fang Y, Hu J, Du C, Liu Z and Zhang L. 2022. SAR-optical image matching by integrating siamese U-net with FFT correlation. *IEEE Geoscience and Remote Sensing Letters*, 19: 1-5 [DOI: 10.1109/LGRS.2021.3100531].
- Feng R, Li X, Bai J and Ye Y. 2022. MID: a novel mountainous remote sensing imagery registration dataset assessed by a coarse-to-fine unsupervised cascading network. *Remote Sensing*, 14(17): 4178 [DOI: 10.3390/rs14174178].
- Fjortoft R, Lopes A, Marthon P and Cubero-Castan E. 1998. An optimal multiedge detector for SAR image segmentation. *IEEE Transactions on Geoscience and Remote Sensing*, 36(3): 793-802 [DOI: 10.1109/36.673672].
- Gao T, Lan C, Huang W and Wang S. 2025. SFA-net: a SAM-guided focused attention network for multimodal remote sensing image matching. *ISPRS Journal of Photogrammetry and Remote Sensing*, 223: 188-206 [DOI: 10.1016/j.isprsjprs.2025.02.032].
- Geng Z, Yang B, Pi Y, Yang Y and Zhao Y. 2025. FRIM: a framework for satellite-borne remote sensing image matching. *IEEE Transactions on Geoscience and Remote Sensing*, 63: 1-18 [DOI: 10.1109/TGRS.2025.3599378].
- Ghorbani H and Beheshti A A. 2007. Multiresolution registration of multitemporal remote sensing images by optimization of mutual information using a simulated annealing based marquardt-levenberg technique//2007 International Conference on Intelligent and Advanced Systems. 685-690 [DOI: 10.1109/ICIAS.2007.4658475].
- Gong M, Zhao S, Jiao L, Tian D and Wang S. 2014. A novel coarse-to-fine scheme for automatic image registration based on SIFT and mutual information. *IEEE Transactions on Geoscience and Remote Sensing*, 52(7): 4328-4338 [DOI: 10.1109/TGRS. 2013. 2281391].
- Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. 2014. Generative adversarial nets//Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2: Vol. 2. Cambridge, MA, USA: MIT Press: 2672-2680.
- Habib A F and Alruzouq R I. 2004. Line-based modified iterated hough transform for automatic registration of multi-source imagery. *Photogrammetric Record*, 19(105): 5-21 [DOI: 10.1111/j.0031-868X. 2003.00254.x].
- Han X, Leung T, Jia Y, Sukthankar R and Berg A C. 2015. MatchNet: unifying feature and metric learning for patch-based matching// 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 3279-3286 [DOI: 10.1109/CVPR.2015.7298948].
- Harris C G and Stephens M J. 1988. A combined corner and edge detector//Alvey vision conference [DOI: 10.5244/C.2.23].

- He X, Yu H, Peng S, Tan D, Shen Z, Bao H, et al. 2025. MatchAnything: universal cross-modality image matching with large-scale pre-training. arXiv [DOI: 10.48550/arXiv.2501.07556].
- Hu C, Zhu R, Sun X, Li X and Xiang D. 2023. Optical and SAR image registration based on pseudo-SAR image generation strategy. *Remote Sensing*, 15(14): 3528 [DOI: 10.3390/rs15143528].
- Hu M, Sun B, Kang X and Li S. 2023. Multiscale structural feature transform for multi-modal image matching. *Information Fusion*, 95: 341-354 [DOI: 10.1016/j.inffus.2023.02.026].
- Huang M, Xu Y, Qian L, Shi W, Zhang Y, Bao W, et al. 2021. The QXS-SAROPT dataset for deep learning in SAR-optical data fusion. arXiv [DOI: 10.48550/arXiv.2103.08259].
- Huang X, Wen L and Ding J. 2019. SAR and optical image registration method based on improved CycleGAN//2019 6th Asia-Pacific Conference on Synthetic Aperture Radar (APSAR). 1-6 [DOI: 10.1109/APSAR46974.2019.9048448].
- 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.isprsjprs.2020.09.012].
- Jiang X, Ma J, Fan A, Xu H, Lin G, Lu T, et al. 2021. Robust feature matching for remote sensing image registration via linear adaptive filtering. *IEEE Transactions on Geoscience and Remote Sensing*, 59(2): 1577-1591 [DOI: 10.1109/TGRS.2020.3001089].
- Johnson K, Cole-Rhodes A, Zavorin I and Moigne J L. 2001. Mutual information as a similarity measure for remote sensing image registration//Geo-Spatial Image and Data Exploitation II: Vol. 4383. SPIE: 51-61 [DOI: 10.1117/12.428251].
- Juan S, Wei C and Yang-yang Z. 2015. A multi-scale SAR scene matching algorithm based on compressed contour feature. *Zhang Y J. Image and graphics: Vol. 9217. Cham: Springer International Publishing: 188-197* [DOI: 10.1007/978-3-319-21978-3\_17].
- Ke Y and Sukthankar R. 2004. PCA-SIFT: a more distinctive representation for local image descriptors//Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.: Vol. 2. II-II [DOI: 10.1109/CVPR.2004.1315206].
- Kirillov A, Mintun E, Ravi N, Mao H, Rolland C, Gustafson L, et al. 2023. Segment anything//2023 IEEE/CVF International Conference on Computer Vision (ICCV). 3992-4003 [DOI: 10.1109/ICCV51070.2023.00371].
- Kuglin C D and Hines D C. 1975. The phase correlation image alignment method. *IEEE International Conference on Cybernetics and Society*, 6: 163-165.
- Kuppala K, Banda S and Barige T R. 2020. An overview of deep learning methods for image registration with focus on feature-based approaches. *International Journal of Image and Data Fusion*, 11(2): 113-135 [DOI: 10.1080/19479832.2019.1707720].
- Kybic J and Unser M. 2003. Fast parametric elastic image registration. *IEEE Transactions on Image Processing*, 12(11): 1427-1442 [DOI: 10.1109/TIP.2003.813139].
- Le Moigne J, Campbell W J and Crompt R F. 2002. An automated parallel image registration technique based on the correlation of wavelet features. *IEEE Transactions on Geoscience and Remote Sensing*, 40(8): 1849-1864 [DOI: 10.1109/TGRS.2002.802501].
- Lecun Y, Bottou L, Bengio Y and Haffner P. 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11): 2278-2324 [DOI: 10.1109/5.726791].
- Leutenegger S, Chli M and Siegwart R Y. 2011. BRISK: binary robust invariant scalable keypoints//2011 International Conference on Computer Vision. 2548-2555 [DOI: 10.1109/ICCV. 2011.6126542].
- Lewis J P. 1994. Fast template matching. *Vis. Interface*, 95: 1-5.
- Li B and Ye H. 2012. RSCJ: robust sample consensus judging algorithm for remote sensing image registration. *IEEE Geoscience and Remote Sensing Letters*, 9(4): 574-578 [DOI: 10.1109/LGRS. 2011.2175434].
- Li J, Hu Q and Ai M. 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 J, Hu Q, Ai M and Zhong R. 2017. Robust feature matching via support-line voting and affine-invariant ratios. *ISPRS Journal of Photogrammetry and Remote Sensing*, 132: 61-76 [DOI: 10.1016/j.isprsjprs.2017.08.009].
- Li J, Hu Q and Zhang Y. 2023. Multimodal image matching: a scale-invariant algorithm and an open dataset. *ISPRS Journal of Photogrammetry and Remote Sensing*, 204: 77-88 [DOI: 10.1016/j.isprsjprs.2023.08.010].
- Li J, Xu W, Shi P, Zhang Y and Hu Q. 2022. LNIFT: locally normalized image for rotation invariant multimodal feature matching. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1-14 [DOI: 10.1109/TGRS.2022.3165940].
- Li J Y. 2018. Research on key techniques of robust remote sensing image feature matching. Wuhan: Wuhan University (李加元. 2018. 鲁棒性遥感影像特征匹配关键问题研究. 武汉: 武汉大学).
- Li L, Han L, Ding M and Cao H. 2023. Multimodal image fusion framework for end-to-end remote sensing image registration. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 1-14 [DOI: 10.1109/TGRS.2023.3247642].
- Li L, Han L, Ding M, Liu Z and Cao H. 2022a. Remote sensing image registration based on deep learning regression model. *IEEE Geoscience and Remote Sensing Letters*, 19: 1-5 [DOI: 10.1109/LGRS. 2020.3032439].
- Li L, Han L and Ye Y. 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 X H, Ai W H, Feng R T and Luo S J. 2023. Survey of remote sensing image registration based on deep learning. *National Remote Sensing Bulletin*, 27(2): 267-284 (李星华, 艾文浩, 冯蕊涛, 罗少杰. 2023. 遥感影像深度学习配准方法综述. *遥感学报*, 27(2): 267-284)
- Li X, Su J and Wu W. 2018. A SAR image registration algorithm based on salient region segmentation. *Journal of Yangzhou University (Natural Science Edition)*, 21(3): 69-73 (李响, 苏娟, 伍薇. 2018. 一种基于显著性区域分割的SAR图像配准算法. *扬州大学学报(自然科学版)*, 21(3): 69-73 [DOI: 10.19411/j.1007-824x.2018.03.015]
- Li X, Zhang G, Cui H, Hou S, Wang S, Li Xin, et al. 2022. MCANet: a joint semantic segmentation framework of optical and SAR images for land use classification. *International Journal of Applied Earth Observation and Geoinformation*, 106: 102638 [DOI: 10.1016/j.jag.2021.102638].
- Li Y, Xu Q, He Z and Li W. 2023. Progressive task-based universal network for raw infrared remote sensing imagery ship detection. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 1-13 [DOI: 10.1109/TGRS.2023.3275619].
- Li Z and Snavely N. 2018. MegaDepth: learning single-view depth prediction from internet photos//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA: IEEE: 2041-2050 [DOI: 10.1109/CVPR.2018.00218].
- Li Z, Yue J and Fang L. 2022. Adaptive regional multiple features for large-scale high-resolution remote sensing image registration. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1-13 [DOI: 10.1109/TGRS.2022.3141101].
- Lindenberg P, Sarlin P E and Pollefeys M. 2023. LightGlue: local feature matching at light speed//2023 IEEE/CVF International Conference on Computer Vision (ICCV). 17581-17592 [DOI: 10.1109/ICCV51070.2023.01616].
- Liu M, Zhou G, Ma L, Li L and Mei Q. 2023. SIFNet: a self-attention interaction fusion network for multisource satellite imagery template matching. *International Journal of Applied Earth Observation and Geoinformation*, 118: 103247 [DOI: 10.1016/j.jag.2023.103247].
- Liu S and Jiang J. 2019. Registration algorithm based on line-intersection-line for satellite remote sensing images of urban areas. *Remote Sensing*, 11(12): 1400 [DOI: 10.3390/rs11121400].
- Liu X, Tian Z, Leng C and Duan X. 2010. Remote sensing image registration based on KICA-SIFT descriptors//2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery: Vol. 1. 278-282 [DOI: 10.1109/FSKD.2010.5569671].
- 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, Jiang J, Zhou H, Zhao J and Guo X. 2018. Guided locality preserving feature matching for remote sensing image registration. *IEEE Transactions on Geoscience and Remote Sensing*, 56(8): 4435-4447 [DOI: 10.1109/TGRS.2018.2820040].
- Ma J, Jiang X, Fan A, Jiang J and Yan J. 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 J, Zhao J, Tian J, Yuille A L and Tu Z. 2014. Robust point matching via vector field consensus. *IEEE Transactions on Image Processing*, 23(4): 1706-1721 [DOI: 10.1109/TIP.2014.2307478].
- Ma J, Zhou H, Zhao J, Gao Y, Jiang J and Tian J. 2015. Robust feature matching for remote sensing image registration via locally linear transforming. *IEEE Transactions on Geoscience and Remote Sensing*, 53(12): 6469-6481 [DOI: 10.1109/TGRS.2015.2441954].
- Ma W, Wen Z, Wu Y, Jiao L, Gong M, Zheng Y, et al. 2017a. 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, Wu Y, Liu S, Su Q and Zhong Y. 2018. Remote sensing image registration based on phase congruency feature detection and spatial constraint matching. *IEEE Access*, 6: 77554-77567 [DOI: 10.1109/ACCESS.2018.2883410].
- Ma W, Wu Y, Zheng Y, Wen Z and Liu L. 2017b. Remote sensing image registration based on multifeature and region division. *IEEE Geoscience and Remote Sensing Letters*, 14(10): 1680-1684 [DOI: 10.1109/LGRS.2017.2728604].
- Ma W, Zhang J, Wu Y, Jiao L, 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].
- Maggiolo L, Solarna D, Moser G and Sarpico S B. 2022. Registration of multisensor images through a conditional generative adversarial network and a correlation-type similarity measure. *Remote Sensing*, 14(12): 2811 [DOI: 10.3390/rs14122811].
- Mair E, Hager G D, Burschka D, Suppa M and Hirzinger G. 2010. Adaptive and generic corner detection based on the accelerated segment test//Daniilidis K, Maragos P, Paragios N. *Computer Vision - ECCV 2010*. Berlin, Heidelberg: Springer: 183-196 [DOI: 10.1007/978-3-642-15552-9\_14].
- Matas J, Chum O, Urban M and Pajdla T. 2004. Robust wide-baseline stereo from maximally stable extremal regions. *Image and Vision Computing*, 22(10): 761-767 [DOI: 10.1016/j.imavis.2004.02.006].
- Merkle N, Auer S, Müller R and Reinartz P. 2018. Exploring the poten-

- tial 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].
- Mikolajczyk K and Schmid C. 2004. Scale & affine invariant interest point detectors. *International Journal of Computer Vision*, 60(1): 63-86 [DOI: 10.1023/B:VISI.0000027790.02288.f2].
- 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].
- Mishchuk A, Mishkin D, Radenović F and Matas J. 2017. Working hard to know your neighbor's margins: local descriptor learning loss// *Proceedings of the 31st International Conference on Neural Information Processing Systems*. Red Hook, NY, USA: Curran Associates Inc.: 4829-4840.
- Mohammadi N, Sedaghat A and Jodeiri Rad M. 2022. Rotation-invariant self-similarity descriptor for multi-temporal remote sensing image registration. *Photogrammetric Record*, 37(177): 6-34 [DOI: 10.1111/phor.12402].
- Moravec H P. 1977. Towards automatic visual obstacle avoidance. Morgan Kaufmann Publishers Inc. [DOI: http://dx.doi.org/].
- Morel J M and Yu G. 2009. ASIFT: a new framework for fully affine invariant image comparison. *SIAM Journal on Imaging Sciences*, 2(2): 438-469 [DOI: 10.1137/080732730].
- Muja M and Lowe D G. 2009. Fast approximate nearest neighbors with automatic algorithm configuration://*Proceedings of the Fourth International Conference on Computer Vision Theory and Applications*. Lisboa, Portugal: SciTePress - Science and Technology Publications: 331-340 [DOI: 10.5220/0001787803310340].
- Paul S and Pati U C. 2018. Coarse-to-fine registration of remote sensing optical images using SIFT and SPSA optimization. *Soft computing: theories and applications*. Springer, Singapore: 771-781 [DOI: 10.1007/978-981-10-5687-1\_69].
- Paul S and Pati U C. 2021. A comprehensive review on remote sensing image registration. *International Journal of Remote Sensing*, 42(14): 5396-5432 [DOI: 10.1080/01431161.2021.1906985].
- Potje G, Cadar F, Araujo A, Martins R and Nascimento E R. 2024. XFeat: accelerated features for lightweight image matching//*2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2682-2691 [DOI: 10.1109/CVPR52733.2024.00259].
- Priyasad D, Fernando T, Haghghat M, Gammulle H, Del Prete R and Fookes C. 2026. Two-stage offline knowledge distillation for onboard registration of multispectral satellite images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 232: 319-335 [DOI: 10.1016/j.isprsjprs.2025.12.009].
- Qie Z X and Peng J X. 2006. Pyramid approach for remote sensing image registration based on FFT and mutual information. *Computer Engineering and Applications*(36): 32-34), 54 (强赞霞, 彭嘉雄. 2006. 基于 FFT 与互信息准则的分层遥感图像配准. *计算机工程与应用*(36): 32-34), 54
- Quan D. 2022. Deep feature representation learning for image matching. Xian: Xidian University (权豆. 2022. 基于深度特征表示学习的图像匹配研究. 西安: 西安电子科技大学) [DOI: 10.27389/d.cnki.gxadu.2022.003130]
- Quan D, Wang S, Li Y, Yang B, Huyan N, Chanussot J, et al. 2021. Multi-relation attention network for image patch matching. *IEEE Transactions on Image Processing*, 30: 7127-7142 [DOI: 10.1109/TIP.2021.3101414].
- Quan D, Wei H, Wang S, Lei R, Duan B, Li Y, et al. 2022. Self-distillation feature learning network for optical and SAR image registration. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1-18 [DOI: 10.1109/TGRS.2022.3173476].
- 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].
- Ren J, Jiang X, Li Z, Liang D, Zhou X and Bai X. 2025. MINIMA: modality invariant image matching//*2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 23059-23068 [DOI: 10.1109/CVPR52734.2025.02147].
- Rublee E, Rabaud V, Konolige K and Bradski G. 2011. ORB: an efficient alternative to SIFT or SURF//*2011 International Conference on Computer Vision*. 2564-2571 [DOI: 10.1109/ICCV.2011.6126544].
- Sarlin P E, DeTone D, Malisiewicz T and Rabinovich A. 2020. SuperGlue: learning feature matching with graph neural networks//*2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 4937-4946 [DOI: 10.1109/CVPR42600.2020.00499].
- 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].
- Schwind P, Suri S, Reinartz P and Siebert A. 2010. Applicability of the SIFT operator to geometric SAR image registration. *International Journal of Remote Sensing*, 31(8): 1959-1980 [DOI: 10.1080/01431160902927622].
- 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 Ebadi H. 2015c. Very high resolution image matching

- based on local features and k-means clustering. *Photogrammetric Record*, 30(150): 166-186 [DOI: 10.1111/phor.12101].
- Sedaghat A and Mohammadi N. 2018. Uniform competency-based local feature extraction for remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 135: 142-157 [DOI: 10.1016/j.isprsjprs.2017.11.019].
- 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].
- Shechtman E and Irani M. 2007. Matching local self-similarities across images and videos//2007 IEEE Conference on Computer Vision and Pattern Recognition. 1-8 [DOI: 10.1109/CVPR.2007.383198].
- Shen X, Cai Z, Yin W, Müller M, Li Z, Wang K, et al. 2024. GIM: learning generalizable image matcher from internet videos. *arXiv* [DOI: 10.48550/arXiv.2402.41095].
- Smith S M and Brady J M. 1997. SUSAN—a new approach to low level image processing. *International Journal of Computer Vision*, 23(1): 45-78 [DOI: 10.1023/A:1007963824710].
- Su J, Li B 和 Wang Y Z. 2016. SAR image registration algorithm based on closed uniform regions. *Journal of Electronics & Information Technology*, 38(12): 3282-3288 (苏娟, 李彬, 王延钊. 2016. 一种基于封闭均匀区域的SAR图像配准方法. *电子与信息学报*, 38(12): 3282-3288)
- Sui H G, Liu C, Gan Z, Jiang Z J 和 Xu C. 2022. Overview of multi-modal remote sensing image matching methods. *Acta Geodetica 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)
- Sui H, Song Z, Gao D and Hua L. 2017. Automatic image registration based on shape features and multi-scale image segmentation//2017 2nd International Conference on Multimedia and Image Processing (ICMIP). 118-122 [DOI: 10.1109/ICMIP.2017.17].
- Sun J, Shen Z, Wang Y, Bao H and Zhou X. 2021. LoFTR: detector-free local feature matching with transformers//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 8918-8927 [DOI: 10.1109/CVPR46437.2021.00881].
- 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].
- Teng X C, Ye Y B 和 Liu X C. 2023. Multi-modal multi-view scene matching and its application in precision guidance systems. *National Defense Technology*, 44(5): 26-34 (滕锡超, 叶熠彬, 刘学聪. 2023. 多模态多视角景象匹配及精确制导应用. *国防科技*, 44(5): 26-34) [DOI: 10.13943/j.issn1671-4547.2023.05.04]
- Teo T A and Chen S Y. 2011. Feature-based image registration of ALOS PALSAR and AVNIR-2 images//2011 IEEE International Geoscience and Remote Sensing Symposium. 566-569 [DOI: 10.1109/IGARSS.2011.6049191].
- Thangavel P and Kokila R. 2013. An extension of FFT based image registration//Meghanathan N, Nagamalai D, Chaki N. *Advances in Computing and Information Technology*. Berlin, Heidelberg: Springer: 729-737 [DOI: 10.1007/978-3-642-31552-7\_74].
- 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, Ye Z, Xu Y, Gao S, Xie H, Du Q, et al. 2019. Image registration with Fourier-based image correlation: a comprehensive review of developments and applications. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(10): 4062-4081 [DOI: 10.1109/JSTARS.2019.2937690].
- Touzi R, Lopes A and Bousquet P. 1988. A statistical and geometrical edge detector for SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 26(6): 764-773 [DOI: 10.1109/36.7708].
- Tuzcuoğlu Ö, Köksal A, Sofu B, Kalkan S and Alatan A A. 2024. XoFTR: cross-modal feature matching transformer//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). 4275-4286 [DOI: 10.1109/CVPRW63382.2024.00431].
- Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, et al. 2017. Attention is all you need//Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc.: 6000-6010.
- Verdie Y, Yi K M, Fua P and Lepetit V. 2015. TILDE: a temporally invariant learned DEtector//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 5279-5288 [DOI: 10.1109/CVPR.2015.7299165].
- Viswanathan D. 2011. Features from accelerated segment test (FAST).
- Wang G, Liu X, Jiao L and Wang S. 2014. A multiscale region-based approach to automatic SAR image registration using CLPSO//2014 IEEE Geoscience and Remote Sensing Symposium. 5226-5229 [DOI: 10.1109/IGARSS.2014.6947677].
- Wang L M. 2021. Research on the registration technology of optical and SAR remote sensing images. Changchun: Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences (王丽娜. 2021. 可见光与SAR遥感图像配准技术研究. 长春: 中国科学院大学(中国科学院长春光学精密机械与物理研究所)) [DOI: 10.27522/d.cnki.gkcegs.2021.000015]
- Wang S, Quan D, Liang X, Ning M, Guo Y and Jiao L. 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 S, You H and Fu K. 2012. BFSIFT: a novel method to find feature matches for SAR image registration. *IEEE Geoscience and Remote Sensing Letters*, 9(4): 649-653 [DOI: 10.1109/LGRS.2011.2177437].
- Wang W, Mu K and Liu H. 2025. A multihierarchy flow field prediction network for multimodal remote sensing image registration. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 18: 5232-5243 [DOI: 10.1109/JSTARS.2025.3532939].
- Wang Y, He X, Peng S, Tan D and Zhou X. 2024. Efficient LoFTR: semi-dense local feature matching with sparse-like speed//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 21666-21675 [DOI: 10.1109/CVPR52733.2024.02047].
- Wang Z, Fan B, Wang G and Wu F. 2016. Exploring local and overall ordinal information for robust feature description. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(11): 2198-2211 [DOI: 10.1109/TPAMI.2015.2513396].
- Wang Z, Fan B and Wu F. 2011. Local intensity order pattern for feature description//2011 International Conference on Computer Vision, Barcelona, Spain: IEEE: 603-610 [DOI: 10.1109/ICCV.2011.6126294].
- Wong A and Orchard J. 2008. Efficient FFT-accelerated approach to invariant optical - LIDAR registration. *IEEE Transactions on Geoscience and Remote Sensing*, 46(11): 3917-3925 [DOI: 10.1109/TGRS.2008.2001685].
- Wu J. 2005. The research of image registration methods based on wavelet transformation. Changsha: Hunan University (伍君. 2005. 基于小波变换的图像配准方法研究. 长沙: 湖南大学)
- Wu J, Xu R, Wood-Doughty Z, Wang C, Xu S and Lam E Y. 2025. Segment anything model is a good teacher for local feature learning. *IEEE Transactions on Image Processing*, 34: 2097-2111 [DOI: 10.1109/TIP.2025.3554033].
- Wu P, Yao Y, Wan Y, Zhang W, Zhao R, Li J, et al. 2026. SOMA-1M: a large-scale SAR-optical multi-resolution alignment dataset for multi-task remote sensing. *arXiv* [DOI: 10.48550/arXiv.2602.05480].
- Wu P, Yao Y, Zhang W, Wei D, Wan Y, Li Y, et al. 2025. MapGlue: multimodal remote sensing image matching <https://arxiv.org/abs/2503.16185v1>.
- Wu W, Xian Y, Su J and Ren L. 2022. A siamese template matching method for SAR and optical image. *IEEE Geoscience and Remote Sensing Letters*, 19: 1-5 [DOI: 10.1109/LGRS.2021.3108579].
- Wu Y, Bai Z F, Gong M G, Qu B Y T, Li H, Zhang M Y, et al. 2023. Application of swarm intelligence and a bioinspired computing algorithm in remote sensing image registration. *Scientia Sinica Technologica*, 53(2): 147-166 (武越, 白壮飞, 公茂果, 曲博艺婷, 李豪, 张明阳, 等. 2023. 遥感图像配准中的群智汇聚方法. *中国科学: 技术科学*), 53(2): 147-166
- Wu Y, Ma W, Gong M, Su L and Jiao L. 2015. A novel point-matching algorithm based on fast sample consensus for image registration. *IEEE Geoscience and Remote Sensing Letters*, 12(1): 43-47 [DOI: 10.1109/LGRS.2014.2325970].
- Wu Y, Miao Q, Ma W, Gong M and Wang S. 2018. PSOSAC: particle swarm optimization sample consensus algorithm for remote sensing image registration. *IEEE Geoscience and Remote Sensing Letters*, 15(2): 242-246 [DOI: 10.1109/LGRS.2017.2783879].
- Xiang D, Xie Y, Cheng J, Xu Y, Zhang H and Zheng Y. 2022. Optical and SAR image registration based on feature decoupling network. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1-13 [DOI: 10.1109/TGRS.2022.3211858].
- Xiang Y M, Chen J Y, Hong Z H, Jiao N G, Wang F, You H J, et al. 2025. OSDataset2.0: SAR-optical image matching dataset and evaluation benchmark. *Journal of Radars*: 1-14 (向俞明, 陈锦杨, 洪中华, 焦念刚, 王峰, 尤红建, 等. 2025. OSDataset2.0: SAR-光学影像匹配数据集及评估基准. *雷达学报(中英文)*): 1-14
- Xiang Y, Wang F and You H. 2018. An automatic and novel SAR image registration algorithm: a case study of the Chinese GF-3 satellite. *Sensors*, 18(2): 672 [DOI: 10.3390/s18020672].
- Xiao Y, Zhang C, Chen Y, Jiang B and Tang J. 2024. ADRNet: affine and deformable registration networks for multimodal remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 1-13 [DOI: 10.1109/TGRS.2024.3373568].
- Xiong X, Jin G, Xu Q and Zhang H. 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, Xu Q, Zhang H, Wang L 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: 1-17 [DOI: 10.1109/TGRS.2022.3197357].
- Xu C, Liu C, Li H, Ye Z, Sui H and Yang W. 2022. Multiview image matching of optical satellite and UAV based on a joint description neural network. *Remote Sensing*, 14(4): 838 [DOI: 10.3390/rs14040838].
- Xu C, Sui H G, Li D R, Sun K M and Liu J Y. 2016. An automatic optical and SAR image registration method using iterative multi-level and refinement model. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B7: 593-600 [DOI: 10.5194/isprs-archives-XLI-B7-593-2016].
- Xu C, Sui H, Li H and Liu J. 2015. An automatic optical and SAR image registration method with iterative level set segmentation and SIFT. *International Journal of Remote Sensing*, 36(15): 3997-4017 [DOI: 10.1080/01431161.2015.1070321].
- Xu Q, Li Y, Nie J, Liu Q and Guo M. 2023. UPanGAN: unsupervised pansharpening based on the spectral and spatial loss constrained generative adversarial network. *Information Fusion*, 91: 31-46 [DOI: 10.1016/j.inffus.2022.10.001].

- Xu W, Yuan X, Hu Q and Li J. 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].
- Yan X, Zhang Y, Zhang D, Hou N and Zhang B. 2020. Registration of multimodal remote sensing images using transfer optimization. *IEEE Geoscience and Remote Sensing Letters*, 17 (12): 2060-2064 [DOI: 10.1109/LGRS.2019.2963477].
- Yang M, Fan F, Li Z, Deng S, Ma Y and Ma J. 2025. DistillMatch: leveraging knowledge distillation from vision foundation model for multimodal image matching. *arXiv* [DOI: 10.48550/arXiv.2509.16017].
- Yang W, Mei L, Ye Z, Wang Y, Hu X, Zhang Y, et al. 2024. Adjacent self-similarity 3-D convolution for multimodal image registration. *IEEE Geoscience and Remote Sensing Letters*, 21: 1-5 [DOI: 10.1109/LGRS.2024.3351158].
- Yang W, Xu C, Mei L, Yao Y and Liu C. 2022. LPSO: multi-source image matching considering the description of local phase sharpness orientation. *IEEE Photonics Journal*, 14(1): 1-9 [DOI: 10.1109/JPHOT.2022.3144227].
- Yang X and Cheng K T. 2012. LDB: an ultra-fast feature for scalable augmented reality on mobile devices//2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). 49-57 [DOI: 10.1109/ISMAR.2012.6402537].
- Yang Y, Duan F, Ma L and Jiang J. 2018. A robust method for constructing rotational invariant descriptors. *Signal Processing: Image Communication*, 60: 224-236 [DOI: 10.1016/j.image.2017.10.010].
- 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].
- Ye F, Su Y, Xiao H, Zhao X and Min W. 2018. 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, Shan J, Bruzzone L and Shen L. 2017. 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, Shan J, Hao S, Bruzzone L and Qin Y. 2018. 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, Tang T, Zhu B, Yang C, Li B and Hao S. 2022a. A multiscale framework with unsupervised learning for remote sensing image registration. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1-15 [DOI: 10.1109/TGRS.2022.3167644].
- Ye Y, Zhu B, Tang T, Yang C, Xu Q and Zhang G. 2022b. 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].
- Yi K M, Trulls E, Lepetit V and Fua P. 2016. LIFT: learned invariant feature transform. *Leibe B, Matas J, Sebe N, Welling M. Computer vision - ECCV 2016: Vol. 9910. Cham: Springer International Publishing: 467-483* [DOI: 10.1007/978-3-319-46466-4\_28].
- Yu X C, Lv Z H and Hu D. 2013. Review of remote sensing image registration techniques. *Optics and Precision Engineering*, 21(11): 2960-2972 (余先川, 吕中华, 胡丹. 2013. 遥感图像配准技术综述. *光学精密工程*, 21(11): 2960-2972).
- Zamir S W, Arora A, Khan S, Hayat M, Khan F S and Yang M H. 2022. Restormer: efficient transformer for high-resolution image restoration//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, LA, USA: IEEE: 5718-5729 [DOI: 10.1109/CVPR52688.2022.00564].
- Zhang B S, Tian Z and Yan W D. 2011. An SAR image registration algorithm based on segmentation-derived regions. *Chinese Journal of Engineering Mathematics*, 28(1): 7-14 (张宝尚, 田铮, 延伟东. 2011. 基于分割区域的SAR图像配准方法研究. *工程数学学报*, 28(1): 7-14).
- Zhang D, Yu L and Cai Z. 2007. Automatic registration for ASAR and TM images based on region features//*Geoinformatics 2007: Remotely Sensed Data and Information: Vol. 6752. SPIE: 1406-1414* [DOI: 10.1117/12.761309].
- Zhang G, Sui H, Song Z, Hua F and Hua L. 2017. Automatic registration method of SAR and optical image based on line features and spectral graph theory//2017 2nd International Conference on Multimedia and Image Processing (ICMIP). 64-67 [DOI: 10.1109/ICMIP.2017.49].
- Zhang H, Lei L, Ni W, Tang T, Wu J, Xiang D, et al. 2022. Explore better network framework for high-resolution optical and SAR image matching. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1-18 [DOI: 10.1109/TGRS.2021.3126939].
- Zhang H, Ni W, Yan W, Xiang D, Wu J, Yang X, et al. 2019. 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, Wu Y and Jiao L. 2019. 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 S, Fu Z, Liu J, Su X, Luo B, Nie H, et al. 2023. Multilevel

- attention siamese network for keypoint detection in optical and SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 1-17 [DOI: 10.1109/TGRS.2023.3293143].
- Zhang S, Zhu Z, Li Z, Lu T and Ma J. 2025. Matching while perceiving: enhance image feature matching with applicable semantic amalgamation. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(10): 10094-10102 [DOI: 10.1609/aaai.v39i10.33095].
- Zhang W, Zhao R, Yao Y, Wan Y, Wu P, Li J, et al. 2025. Multi-resolution SAR and optical remote sensing image registration methods: a review, datasets, and future perspectives. *arXiv* [DOI: 10.48550/arXiv.2502.01002].
- Zhang Yongxian, Lan C, Zhang H, Ma G and Li H. 2024. Multimodal remote sensing image matching via learning features and attention mechanism. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 1-20 [DOI: 10.1109/TGRS.2023.3348980].
- Zhang Y, Yao Y, Wan Y, Liu W, Yang W, Zheng Z, et al. 2023. Histogram of the orientation of the weighted phase descriptor for multimodal remote sensing image matching. *ISPRS Journal of Photogrammetry and Remote Sensing*, 196: 1-15 [DOI: 10.1016/j.isprsjprs.2022.12.018].
- Zhang Yongjun, Zhang W, Yao Y, Zheng Z, Wan Y and Xiong M. 2024. Robust registration of multi-modal remote sensing images based on multi-dimensional oriented self-similarity features. *International Journal of Applied Earth Observation and Geoinformation*, 127: 103639 [DOI: 10.1016/j.jag.2023.103639].
- Zhao Z, Wang F and You H. 2024a. Robust region feature extraction with salient MSER and segment distance-weighted GLOH for remote sensing image registration. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17: 2475-2488 [DOI: 10.1109/JSTARS.2023.3344474].
- Zhao Z, Wang F and You H. 2024b. Lightweight and stable multi-feature databases for efficient geometric localization of remote sensing images. *Remote Sensing*, 16(7): 1237 [DOI: 10.3390/rs16071237].
- Zhong Y, Liu W, Zhao J and Zhang L. 2015. Change detection based on pulse-coupled neural networks and the NMI feature for high spatial resolution remote sensing imagery. *IEEE Geoscience and Remote Sensing Letters*, 12(3): 537-541 [DOI: 10.1109/LGRS.2014.2349937].
- Zhou L, Ye Y, Tang T, 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: 1-5 [DOI: 10.1109/LGRS.2021.3105567].
- Zhu H, Ma W, Hou B and Jiao L. 2016. SAR image registration based on multifeature detection and arborescence network matching. *IEEE Geoscience and Remote Sensing Letters*, 13(5): 706-710 [DOI: 10.1109/LGRS.2016.2539207].
- Zitová B and Flusser J. 2003. Image registration methods: a survey. *Image and Vision Computing*, 21(11): 977-1000 [DOI: 10.1016/S0262-8856(03)00137-9].

## 作者简介

赵子璐,女,博士,研究方向为大数据分析 with 遥感信息提取。

E-mail: zhaozilu16@mails.ucas.ac.cn

王峰,通信作者,男,副研究员,主要研究方向为遥感图像配准、大区域镶嵌拼接、三维重建、遥感图像增强、时序目标解译等。E-mail: wangfeng003020@aircas.ac.cn

焦念刚,男,助理研究员,主要研究方向为遥感影像配准、大区域镶嵌拼接、三维重建等。E-mail: jiaong@aircas.ac.cn

尤红建,男,研究员,主要研究方向为遥感影像配准、遥感影像分割、遥感信息应用等。E-mail: hjyou@mail.ie.ac.cn