

基于相位一致性增强的相邻自相似性多模态遥感影像匹配^①



曹徐鹏^{1,2}, 喻国荣^{1,2}, 吴俊卿^{1,2}

¹(武汉科技大学 计算机科学与技术学院, 武汉 430070)

²(武汉科技大学 智能信息处理与实时工业系统湖北省重点实验室, 武汉 430070)

通信作者: 喻国荣, E-mail: ygrong190@163.com

摘要: 由于多模态图像之间存在着显著的辐射差异和几何差异, 给多模态遥感图像之间的高精度配准带来了巨大挑战. 为了解决这些问题, 本文提出了一种基于相位一致性增强的相邻自相似性多源遥感影像匹配方法——PC-ASS. 首先, 利用非线性扩散滤波构建多尺度影像表示, 在抑制噪声的同时保留共有的边缘与结构, 为后续特征检测提供基础. 随后, 基于多尺度多方向 Log-Gabor 滤波器计算得到相位一致性幅度图, 用以衡量影像中结构显著区域. 本文将相位一致性幅度值作为加权因子引入到相邻自相似性响应计算中, 增强了影像中的结构特征: 高相位一致性区域获得更强的响应, 从而提升了边缘与角点等稳健特征的检测数量与质量. 进一步地, 在描述符构造阶段, 本文在极坐标统计直方图框架中引入相位一致性加权机制, 对每个像素的相邻自相似性值按照其相位一致性幅值进行加权, 使结构显著区域在描述符中占据更大比例, 从而增强了特征对噪声、纹理干扰及跨模态辐射差异的鲁棒性. 最后, 通过最近邻距离比匹配策略和快速抽样一致性算法 (FSC) 剔除错误匹配, 实现高精度配准. 在 3 个公开的多模态遥感影像数据集上, 本文与 PSO-SIFT、OSS、HAPCG、RIFT 和 ASS 这 5 种匹配方法进行对比实验. 实验结果表明, PC-ASS 在平均正确匹配点数量、均方根误差和正确匹配率指标上均优于现有方法, 具有较强的鲁棒性和适用性.

关键词: 多模态遥感影像; 非线性扩散; 相位一致性; 相邻自相似性

引用格式: 曹徐鹏, 喻国荣, 吴俊卿. 基于相位一致性增强的相邻自相似性多模态遥感影像匹配. 计算机系统应用. <http://www.c-s-a.org.cn/1003-3254/10112.html>

Multimodal Remote Sensing Image Matching Using Phase-congruency-enhanced Adjacent Self-similarity

CAO Xu-Peng^{1,2}, YU Guo-Rong^{1,2}, WU Jun-Qing^{1,2}

¹(School of Computer Science and Technology, Wuhan University of Science and Technology, Wuhan 430070, China)

²(Hubei Province Key Laboratory of Intelligent Information Processing and Real-time Industrial System, Wuhan University of Science and Technology, Wuhan 430070, China)

Abstract: Significant radiometric and geometric discrepancies among multimodal remote sensing images present substantial challenges for achieving high-precision registration. To address these issues, this study proposes a phase-congruency-enhanced adjacent self-similarity matching method, termed as PC-ASS, for multimodal remote sensing imagery. First, a multi-scale image representation is constructed via nonlinear diffusion filtering to suppress noise while preserving common edges and structural information, thus providing a reliable foundation for subsequent feature detection. Next, phase congruency amplitude maps are computed using multi-scale, multi-orientation Log-Gabor filters to characterize structurally salient regions in the images. The phase congruency amplitudes are then used as weighting

① 基金项目: 国家自然科学基金 (42501559); 湖北省自然科学基金 (2025AFB544)

收稿时间: 2025-09-16; 修改时间: 2025-10-10; 采用时间: 2025-10-29; csa 在线出版时间: 2026-01-16

factors in computing adjacent self-similarity responses, thus enhancing structural features: regions with higher phase congruency yield stronger responses, increasing both the number and the quality of robust features such as edges and corners. Furthermore, during descriptor construction, a phase-congruency-weighted mechanism is incorporated into the polar statistical histogram framework, weighting each pixel's adjacent self-similarity value by its phase congruency amplitude. This ensures that structurally salient regions contribute more prominently to the descriptor, thereby improving robustness against noise, texture interference, and cross-modal radiometric differences. Finally, incorrect matches are eliminated through a nearest-neighbor distance ratio strategy combined with the fast sample consensus (FSC) algorithm, enabling high-precision registration. Comparative experiments on three publicly available multimodal remote sensing datasets against five representative methods (PSO-SIFT, OSS, HAPCG, RIFT, and ASS) demonstrate that PC-ASS outperforms existing approaches in average correct matches, mean root-mean-square error, and correct matching rate, highlighting its robustness and broad applicability.

Key words: multimodal remote sensing image; nonlinear diffusion; phase-congruency (*PC*); adjacent self-similarity

1 引言

随着对地观测技术的发展,传感器的种类日渐丰富,多传感器的综合观测可为地表监测提供更加稳定的结果。然而,不同传感器的成像机理导致了影像间存在着显著的非线性辐射差异和复杂的几何畸变。这为传统的、依赖于影像灰度或梯度信息的匹配算法带来了严峻挑战,难以保持稳定的鲁棒性与高精度,从而严重制约了多源数据的融合与应用。因此,深入研究高效、鲁棒的多模态遥感影像配准方法,是实现多源数据融合、地表变化监测、高精度三维重建等应用的必要前提和基础保障^[1]。

目前,多源遥感影像的匹配方法通常可按照匹配策略与信息依赖的不同分为3大类:基于区域的方法、基于特征的方法以及基于深度学习的方法。

第1类是基于区域的方法。该类方法主要通过直接比较影像的灰度或统计分布特征来衡量相似性。在遥感影像配准研究中,Viola等人^[2]首次提出将互信息作为图像配准的相似性度量,突破了传统归一化互相关等方法仅能处理同源影像的局限,使区域法能够应用于多模态遥感影像的配准问题,并在跨模态条件下表现出较强的鲁棒性。但该方法计算开销较大,且主要针对刚性配准场景,对大范围地形起伏或非刚性畸变的适应性不足。Zitová等人^[3]在其综述中系统总结了区域法的代表性方法,将互信息与互相关、相位相关等并列归纳,指出区域法在多模态遥感影像中的作用,同时强调其在运算效率和对噪声、几何畸变敏感性方面的局限。进一步地,Tuytelaars等人^[4]在局部不变特征检

测的研究中将区域法作为对照,明确指出其仅在光照差异较小、几何变换有限的遥感场景下有效,而在大视角、多时相或强非刚性变化条件下往往失效,由此凸显出特征法在遥感应用中的优势。以互信息为代表的区域法为多源遥感影像配准奠定了基础,但研究重心逐渐转向特征法,以满足复杂遥感成像条件下的更高鲁棒性需求。

第2类是基于特征的方法。影像之间相似的结构特征尤其是轮廓信息,为提取稳健特征提供了基础。Shechtman等人^[5]提出了一种名为局部自相似性的特征描述方法,通过量化局部窗口中灰度像素的统计同现特性,成功地捕获了图像中的几何图层结构,对辐射变化具有一定的鲁棒性。Sedaghat等人^[6]使用一种改进的局部自相似性描述符,通过自相似性计算中产生的相关性曲面来代替梯度信息并为特征点分配主方向,但显式分配方向的精度仍有限。陈聪鹏等人^[7]提出结合相位对称与排序局部自相似性的RPCLSS方法,有效提升了抗非线性辐射差异的能力,在多模态影像匹配精度上优于现有方法,但计算开销较大且在复杂纹理下表现有限。Chen等人^[8]将局部三值模式引入局部自相似性框架,构建的描述符在模拟和少量真实多源遥感影像数据上取得了较好效果。Xiong等人^[9]提出一种名为ASS算法,使用相邻自相似性的特征来精细且快速地捕捉图像结构。与传统的自相似性不同,ASS仅在单位像素半径的极小区域内进行计算,并引入局部统计加权差分,有效抑制SAR影像的乘性噪声并提高效率,在此基础上,利用ASS生成的最小自相似图和索

引图提出了一种新的特征检测器和描述符,使其对SAR影像的斑点噪声具有较强鲁棒性.此外,频域信息也可用于构建对非线性辐射差异不敏感的稳健特征描述符.傅里叶变换是一种全局分析工具,能够将整幅图像分解为整体的幅度和相位信息,但要捕捉图像中特定位置的局部特征即局部相位和局部幅值还需其他方法.Nunes等人^[10]提出基于Log-Gabor滤波器的局部特征描述符,对光照和对比度变化具有较强的鲁棒性,但其计算成本较高且旋转与尺度不变性仍是难点.Zhu等人^[11]结合Log-Gabor滤波器与多通道自相关策略提取稳定特征点,同时在Log-Gabor响应序列基础上构建RMLG特征描述符,具备较强的抗辐射差异与旋转畸变能力,但其不足在于缺乏尺度不变性.Kovesi等人^[12]提出的相位一致性(PC)分析模型来检测图像特征为后续的频域匹配方法提供了理论基础.姚永祥等人^[13]提出了绝对相位方向的特征点主方向分配策略,并利用各向异性加权力矩来分析特征点的局部频域信息,构建了更具区分度的特征描述符,取得了稳健的旋转不变性成果.Zhang等人^[14]将相位方向与加权策略相结合,提高了频域描述符的稳定性和判别力.Ye等人^[15]将相位一致性和方向直方图策略结合,根据图像的几何结构属性来描述特征,提升了多模态影像匹配的准确性.Li等人^[16]提出了一种建立在相位一致性上的多模态图像匹配方法,利用PC的最大和最小力矩进行特征检测,并根据最大索引图来构建特征描述符,实现了对剧烈非线性辐射差异的稳健性,在各类多源遥感影像上均取得了较好的匹配结果.

第3类是基于深度学习的方法.Quan等人^[17]提出一种生成式匹配网络,用于无监督地对齐光学与SAR图像,通过深度学习实现跨模态信息嵌入与配准.该方法无需人工配对准关键点,能自动学习匹配映射;在一定程度上改善了形变和跨模态差异带来的配准困难.南轲等人^[18]使用了一种结合了传统特征点检测和孪生网络(Siamese network)深度特征表达的模板匹配方法,匹配效果优于传统模板匹配方法.蓝朝桢等人^[19]提出一种名为CMM-Net(cross-modality matching network)的深度学习特征匹配方法,使用卷积神经网络提取异源影像的高维特征图,然后使用传统特征匹配方法进行特征匹配和粗差剔除.与多种传统异源影像匹配算法相比,CMM-Net提取的特征具有更高的匹配精度和稳定性,取得了稳健的匹配结果.Zeng等人^[20]使用一

种结合深度学习语义分割和互信息的区域匹配方法,将影像从原始的像素域转换到语义标签域,有效地消除了不同传感器带来的非线性辐射差异,使得原本不相似的影像变得可比较.基于深度学习的方法在多模态遥感影像匹配领域的研究仍然面临挑战,主要是因为缺乏足量且合适的多模态遥感影像数据用于模型训练与测试,并且由于高分辨率遥感影像的数据量庞大,模型的训练和推理对计算资源的需求量较高,尤其是处理大规模数据集时,计算成本和时间开销成为限制因素.

2 基础知识

2.1 非线性扩散滤波

非线性扩散(nonlinear diffusion, ND)滤波是一种自适应图像平滑技术,能够在保留边缘和结构特征的同时有效抑制噪声.该方法基于热传导偏微分方程,通过控制扩散系数来实现选择性平滑.Perona-Malik非线性扩散滤波器的迭代更新公式如式(1):

$$I_{t+1} = I_t + \lambda \cdot \sum_{d \in \{N, S, E, W\}} c_d \cdot \nabla_d I_t \quad (1)$$

其中, I_t 为第 t 次迭代的图像, λ 为时间步长参数,其控制扩散过程的速度和稳定性,参考Fan等人^[21]工作中对时间步长参数的定义,本文方法中将时间步长参数设置为0.15. $\nabla_d I_t$ 表示图像 I_t 在特定方向 d 上的梯度, c_d 为扩散系数.

扩散系数 c_d 是梯度的函数,其计算公式如式(2):

$$c_d = \exp\left(-\frac{\left(\frac{\nabla_d I_t}{k}\right)^2}{2}\right) \quad (2)$$

其中, k 为扩散阈值参数,控制梯度大小与扩散强度的关系.当梯度 $\nabla_d I_t$ 远小于 k 时, c_d 接近1,实现强扩散;当梯度远大于 k 时, c_d 接近0,保留边缘.

2.2 Log-Gabor 滤波器

Log-Gabor 滤波器是实现相位一致性分析的核心工具,其频域表达式如式(3):

$$LG(\omega) = \exp\left(-\frac{\log\left(\frac{\omega}{\omega_0}\right)^2}{2\left(\log\left(\frac{k}{\omega_0}\right)\right)^2}\right) \quad (3)$$

其中, ω 为当前频率, ω_0 是滤波器的中心频率, k/ω_0 为常量不变量,控制滤波器的带宽.本文方法中滤波器的中心频率 ω_0 按几何级数确定,其表达式如式(4):

$$\omega_0^{(n)} = \frac{1}{\lambda_{\min} \cdot mult^{(n-1)}} \quad (4)$$

其中, 最小波长 λ_{\min} 设置为 3 像素, 波长倍长因子 $mult$ 设置为 2.0, n 为尺度索引设置范围为 {1, 2, 3, 4}.

2.3 ASS 算法的描述符

ASS 描述符的构建在局部自相似思想的基础上, 引入单位像素邻域的相邻自相似性^[9]. 对于中心像素 P 与其单位像素距离处的邻域像素 q , 定义 ASS 特征值为加权绝对差的加和, 其计算公式如式 (5), 示意图如图 1 所示.

$$ASS(p, q) = \sum_{\omega_1(i, j, r_1)} |k(I_p(i, j) - I_q(i, j))| \quad (5)$$

其中, $\omega_1(i, j, r_1)$ 为半径 $r_1 = 2$ 的圆形窗口, I_p 与 I_q 分别表示以 p 与 q 为中心的图像块. 接着采用偏移均值滤波计算多通道自相似图 R_o , 得到最小自相似图及其对应的方向索引图. 在索引图上建立对数极坐标网格, 在每个网格单元内统计方向索引值的分布直方图, 并将整个直方图按网格顺序进行拼接, 最终形成一个最终的高维描述向量符.

3 研究方法

3.1 总体结构

本文的研究方法如图 2 所示, 主要包括以下几个步骤: (1) 非线性扩散预处理: 非线性扩散实现参考图

像和待匹配图像的多尺度图像构建, 抑制噪声的同时保持边缘结构信息. (2) 构造相位一致性幅度图: 基于多尺度多方向 Log-Gabor 滤波器组的频域分析原理, 计算参考影像和待匹配影像的相位一致性幅度图. (3) 进行特征点的检测与提取: 将相位一致性幅值加权到最小自相似性图中, 实现特征点的检测和提取. (4) 生成特征描述符: 在相邻自相似性的径向-角度划分基础上, 使用相位一致性幅值对每个像素进行加权统计, 生成新的特征描述符. (5) 进行匹配与粗差剔除: 采用最近邻距离比策略和快速抽样一致性算法 (FSC) 算法剔除错误匹配实现高精度配准.

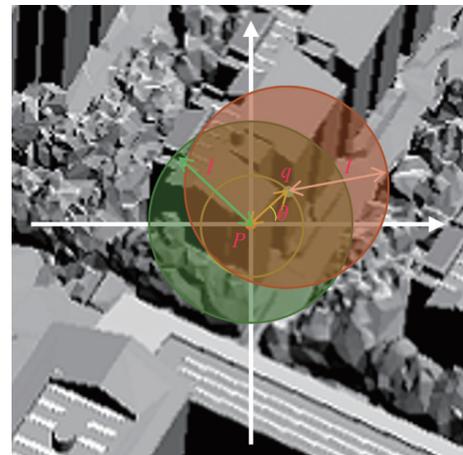


图 1 ASS 特征值计算示意图

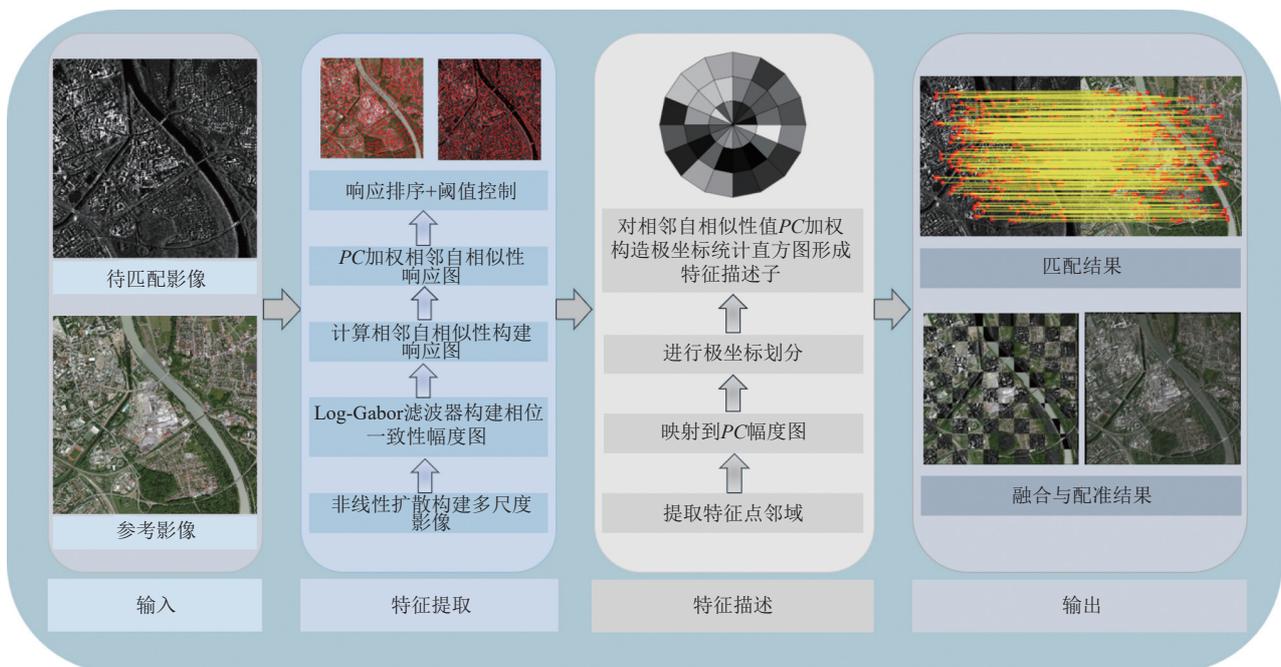


图 2 本文方法流程图

3.2 多模态特征提取

在多模态遥感影像匹配中,参考图像和待匹配图像在成像机理和辐射特性上存在显著差异,传统特征提取方法往往难以有效识别跨模态间的共同结构特征.本文提出的基于PC增强的相邻自相似性算法通过融合非线性扩散和相位一致性技术,构建了一种对比度不变的多模态特征提取方法.

首先使用非线性扩散实现多尺度图像表示.非线性扩散基于Perona-Malik模型如式(6):

$$\frac{\partial I}{\partial t} = \text{div}[c(|\nabla I|)\nabla I] \quad (6)$$

其中, div 为散度算子, ∇I 表示图像 I 的梯度,扩散系数采用指数形式如式(7):

$$c(|\nabla I|) = \exp\left(-\frac{|\nabla I|^2}{\kappa^2}\right) \quad (7)$$

其中, κ 为梯度阈值,该扩散系数具有边缘保持特性:在边缘区域即 $|\nabla I|$ 远大于 κ 时,扩散系数接近0,抑制扩散;在平滑区域即 $|\nabla I|$ 远小于 κ 时,扩散系数接近1,促进扩散.

多尺度空间采用几何级数增长的尺度序列设计,本文方法中尺度因子 F 设置为1.26,确保相邻尺度层之间的平滑过渡,尺度层数根据图像的高度和宽度自适应确定.本文对于每个尺度层 i ,扩散阈值 k 根据当前尺度因子进行自适应调整,如式(8):

$$k(\sigma_i) = k_0 \times \sigma_i \quad (8)$$

其中, k_0 为基准扩散阈值, σ_i 为当前尺度因子.

在进行多尺度图像构建后,通过多尺度多方向Log-Gabor滤波器计算PC幅度图.相位一致性的计算基于复数滤波响应,其中每个尺度和方向的滤波响应通过频域卷积获得,如式(9)和式(10):

$$G_n = \text{FFT}(I(x)) \cdot \text{LG}(n) \cdot \text{AngleFilter}_n \quad (9)$$

$$[en(x), on(x)] = \text{IFFT}[G_n] \quad (10)$$

其中, $\text{FFT}(I(x))$ 是图像的傅里叶变换. $\text{LG}(n)$ 表示第 n 个尺度的径向Log-Gabor滤波器,计算如式(3)所示. AngleFilter_n 表示对应方向的角度滤波器,是一个高斯函数. $\text{IFFT}[G_n]$ 是逆傅里叶变换.相位一致性度量的计算公式如式(11):

$$PC = E_{PC\theta}(x) = \frac{\sqrt{\left(\sum_n en_\theta(x)\right)^2 + \left(\sum_n on_\theta(x)\right)^2}}{\sum_n A_{n,\theta}(x) + \varepsilon} \quad (11)$$

其中, $en_\theta(x)$ 和 $on_\theta(x)$ 分别表示在位置 x 、方向 θ 和尺度 n 上的偶响应和奇响应. $A_{n,\theta}(x) = \sqrt{en_\theta(x)^2 + on_\theta(x)^2}$ 为尺度 n 和方向 θ 上滤波响应的幅值, ε 为防止除0的小常数.

由于相位一致性高的区域对应图像中结构信息丰富的位置,通过相位一致性(PC)幅值能够有效突出这些结构显著区域的特征响应,同时抑制噪声和纹理区域的干扰.故在响应图生成阶段,每个方向的相邻自相似性响应值都通过相位一致性幅值进行增强,如式(12):

$$R'_\theta(x, y) = R'_\theta(x, y) \cdot [1 + E_{PC}(x, y)] \quad (12)$$

在特征点筛选阶段,进一步结合相位一致性幅值对特征响应进行二次增强,如式(13):

$$\text{ComValue} = \text{ResValue} \cdot [1 + E_{PC}(x, y)^2] \quad (13)$$

本文方法采用 $[1 + E_{PC}(x, y)]$ 和 $[1 + E_{PC}(x, y)^2]$ 两种方式来增强基于以下考虑:首先,这种形式保证了增强的单调性和非负性,避免了其他可能的增强方式如直接相乘或指数增强可能导致的数值不稳定或过度增强问题;其次,平方运算的增强策略符合特征检测从粗到精的处理流程,第1层线性增强保持整体动态范围平衡,第2层二次增强实现选择性突出,相比于统一使用单一增强函数更加精细化;最后,相位一致性幅值的物理意义决定了增强函数的形式,PC幅值在 $[0, 1]$ 范围内反映结构显著程度,使用 $(1 + PC)$ 和 $(1 + PC^2)$ 的组合既保证了 $PC=0$ 时不产生负面影响,又确保了 $PC=1$ 时获得合理的增强效果,这种适度增强避免了对数增强 $\log(1 + PC)$ 增强不足或幂函数增强 PC^n ($n > 2$)过度增强的问题,经实验验证在多模态图像匹配中取得了最佳的性能平衡.

最终的特征点筛选采用响应值排序和阈值控制的方式,首先根据增强后的响应值对所有候选特征点进行降序排列,然后根据预设的特征点比例系数进行筛选.特征点保留数量阈值通过公式 $npts = \text{round}(H \times W \times K)$ 计算,其中 H 和 W 分别为图像的高度和宽度, K 为特征点筛选比例系数,通常设置为0.05.本文通过非线性扩散保持结构完整性,结合相位一致性进行特征增强,降低了影像变化对特征点检测的影响,提高了特征点的数量和质量,图3是ASS算法^[9]与本文方法对比示意图.

3.3 描述符的构造

相邻自相似性描述符采用极坐标统计直方图的方法

式构建特征描述符,对图像块内所有像素采用等权重统计,无法区分结构性的重要性.在多模态图像中,边缘和角点等结构显著区域应该比平坦区域具有更高的权重.此外,由于缺乏结构显著性判断机制,相邻自相似性描述符容易受到噪声和复杂纹理的干扰,尤其是在 SAR-图像的相干斑噪声环境下,描述符的稳定性较差.针对上述问题,本文引入了相位一致性加权机制,提升相邻自相似性描述符在多模态遥感影像中的判别能力和鲁棒性.

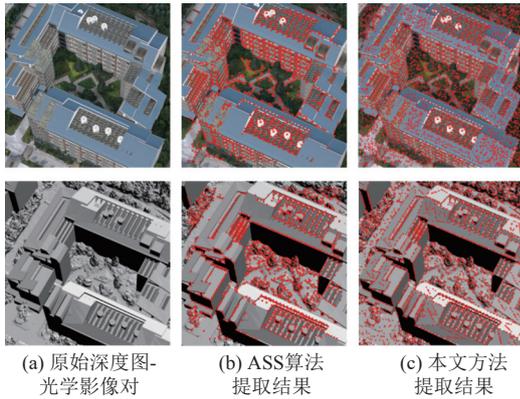


图3 ASS算法^[9]与本文方法对比结果展示

首先,将提取到的特征点及其周围的图像块坐标映射到原始图像的 PC 幅度图上,建立特征点图像块与原始图像 PC 幅度图之间的精确映射关系,从而获得对应区域的相位一致性幅值.由于相邻自相似性描述符采用多尺度金字塔处理,特征点坐标需要从当前尺度空间映射到原始图像空间.假设当前尺度图像尺寸为 (H_s, W_s) ,原图像尺寸为 (H_o, W_o) ,则映射关系如式 (14):

$$\begin{cases} x_{PC} = \text{round}\left(\frac{x_s}{W_s} \times W_o\right) \\ y_{PC} = \text{round}\left(\frac{y_s}{H_s} \times H_o\right) \end{cases} \quad (14)$$

其中, x_s/W_s 和 y_s/H_s 为特征点在当前尺度图像中的相对位置, $\text{round}(\cdot)$ 将浮点数坐标转换为整数像素坐标,用于确定 PC 幅度图的具体像素值. x_{PC} 和 y_{PC} 是特征点映射在原始图像坐标系中的位置.通过双线性插值,将 PC 幅度图精确映射到与相邻自相似性响应图像块相同的尺寸,确保每个像素都能获得对应的 PC 幅值.

在完成 PC 幅度图的映射后,结合极坐标划分结构,需要将每个像素的 PC 幅值分配到其所在的子区域中,作为后续加权统计的依据. PC 加权函数采用线性加权形式如式 (15):

$$w_{PC}(x, y) = 1 + \alpha \times E_{PC}(x, y)^\gamma \quad (15)$$

其中, α 为权重增强系数,控制 PC 增强的强度, γ 为非线性指数,调节权重分布的非线性特性.本文方法中权重增强系数 α 设置为 1.1,非线性指数 γ 设置为 1.2.该设计确保了当 PC 幅值为 0 时 PC 权重为 1. PC 权重可由式 (15) 计算所得,确保基础权重.当 PC 幅值增大时权重非线性增长,突出结构显著区域. PC 权重值关于 PC 幅值单调递增,结构越显著,权重越大.这种自适应非线性权重函数既保证数值稳定性,又能有效区分结构重要性,使得边缘、角点等高 PC 幅值区域在描述符构建中获得更大的权重.在此基础上,将传统的等权重直方图统计公式,如式 (16):

$$H_{i,j}(b) = \frac{|\{p \in \Omega_{i,j} \mid ASS(p) = b\}|}{|\Omega_{i,j}|} \quad (16)$$

替换为 PC 加权统计公式,如式 (17):

$$H'_{i,j}(b) = \frac{\sum_{p \in \Omega_{i,j}, ASS(p)=b} w_{PC}(p)}{\sum_{p \in \Omega_{i,j}} w_{PC}(p)} \quad (17)$$

其中,分子表示所有相邻自相似性响应值为 b 的像素的加权总和,分母表示区域内所有像素的权重总和. $ASS(p)$ 的计算如式 (5).

在每个子区域内,相邻自相似性响应值将像素分配到不同的直方图中,但每个像素的贡献不再是简单的计数,而是根据其 PC 权重进行加权累加,最后使用权重总和而非像素总数进行归一化,归一化保证描述符在不同区域、不同图像间具有可比性.权重归一化计算公式如式 (18):

$$Normalization = \frac{WeightedSum}{\sum_{p \in \Omega_{i,j}} w_{PC}(p)} \quad (18)$$

其中,分子为某个直方图的加权计数值,分母为该区域所有像素权重的总和.此外,为了进一步提升描述符的判别能力,本文在构建过程中还考虑了多尺度信息融合策略.通过在不同金字塔层级上计算 PC 增强的 ASS 特征描述符,并采用加权融合的方式将多尺度信息整合到最终的描述符中,增强了描述符对尺度变化的鲁棒性.同时,针对旋转变换,在极坐标划分时采用主方向对齐策略,确保描述符具有旋转不变性.本文方法设计的描述符和 ASS 算法形成的描述符如图 4 所示.本文方法的描述符在对数极坐标上的结果更加相似,能够更好地实现多模态影像之间的配准.

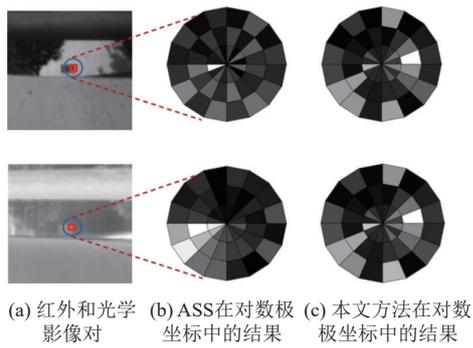


图4 在对数极坐标中的对比结果图

ASS 描述符在红外和光学影像对上有着明显的不相似情况, 而本文提出的描述符有明显改进, 不仅提升了对边缘、角点等结构显著区域的辨识能力, 还有效抑制了噪声和弱结构区域的干扰, 增强了描述符在复杂遥感场景下的鲁棒性。

3.4 匹配和粗差别除

在匹配和粗差别除阶段, 首先采用最近邻距离比策略实现初匹配。随后为了剔除初始匹配的错误匹配,

采用 FSC 算法从中提取正确匹配点对^[22]。

4 实验和结果

4.1 数据集

本文实验采用的数据集包含 Skyearth 团队公开的多模态遥感影像数据集 CoFSM 和 HOWP 数据集, 以及 Brown 等人^[23]公开的场景识别数据集。CoFSM 数据集涵盖 6 种典型数据类型, 包括光学-光学、红外-光学、深度图-光学、MAP-光学、SAR-光学和昼-夜, 每个类型包含 10 幅影像对, 共 60 对影像。HOWP 数据集则涵盖 SAR-光学、红外-光学、深度图-光学、MAP-光学和昼-夜这 5 种数据类型, 每个类型包含 10 幅影像对, 共 50 对影像。文献^[23]公开的场景识别数据集包含 477 对经过配准的 RGB 和 NIR 影像。数据集分为 9 个类别, 包括乡村、田野、森林、山脉、老建筑、街道、城市和水域和室内。在本文实验中, 我们从该数据集的 9 个类别中, 从每个类别中随机挑选 12 对影像, 共计 108 对影像进行实验。部分实验数据如图 5 所示。

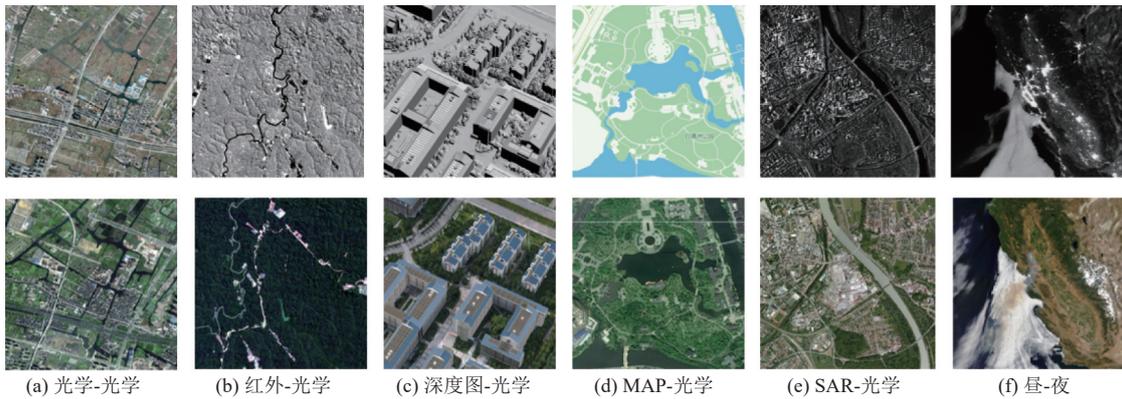


图5 部分实验数据展示

4.2 评价指标

为了验证本文所提出算法的效果, 本文将其与 5 种当前先进的多模态遥感影像匹配算法 ASS^[9]、PSO-SIFT^[24]、RIFT^[16]、OSS^[25]和 HAPCG^[13]进行对比分析。为客观评估匹配算法的性能, 本文选取平均正确匹配点数量 (average number of correct matches, NCM)、均方根误差 (root mean square error, RMSE) 以及正确匹配率 (correct matches ratio, CMR) 作为综合评估指标。CMR 计算公式如式 (19), RMSE 计算公式如式 (20):

$$CMR = \frac{CM}{CM + FM} \quad (19)$$

$$RMSE = \sqrt{\frac{1}{M} \sum_{i=1}^M [(x_i - x'_i)^2 + (y_i - y'_i)^2]} \quad (20)$$

其中, CM 为正确匹配点数, FM 为错误匹配点数。式 (20) 中 M 表示总匹配点对数量, (x_i, y_i) 为参考影像的特征点坐标, (x'_i, y'_i) 则是待匹配影像中对应点经仿射变换后的坐标。

正确匹配点对依据数据集提供的 Ground truth 文件进行判定: 首先在待匹配影像上对特征点坐标进行仿射变换, 然后计算其与参考影像对应特征点的欧氏距离, 如果距离误差不超过 3 个像素, 就判定为正确匹配。本文依据摄影测量与遥感的通用标准将阈值设置

为3个像素,3个像素以内的均方根误差可满足多源数据融合、地表变化监测、高精度三维重建等实际应用的要求。

4.3 定性分析

在数据集 CoFSM 上,将本文方法 PC-ASS 与 PSO-SIFT、OSS、HAPCG、RIFT 和 ASS 这5种匹配方法在6种多模态遥感影像对上进行匹配,获得的可视化匹配结果如图6所示。

4.4 定量分析

除了基于视觉效果定性评估外,本文还对各类匹配方法进行了系统的定量分析。匹配性能主要通过正确匹配点数 (NCM) 和匹配点的均方根误差 (*RMSE*) 来表示。为验证本文方法的有效性,在3个数据集上进行对比实验,并统计各方法的平均正确匹配点数量 (NCM)、均方根误差 (*RMSE*) 和正确匹配率 (*CMR*) 进行量化评价。

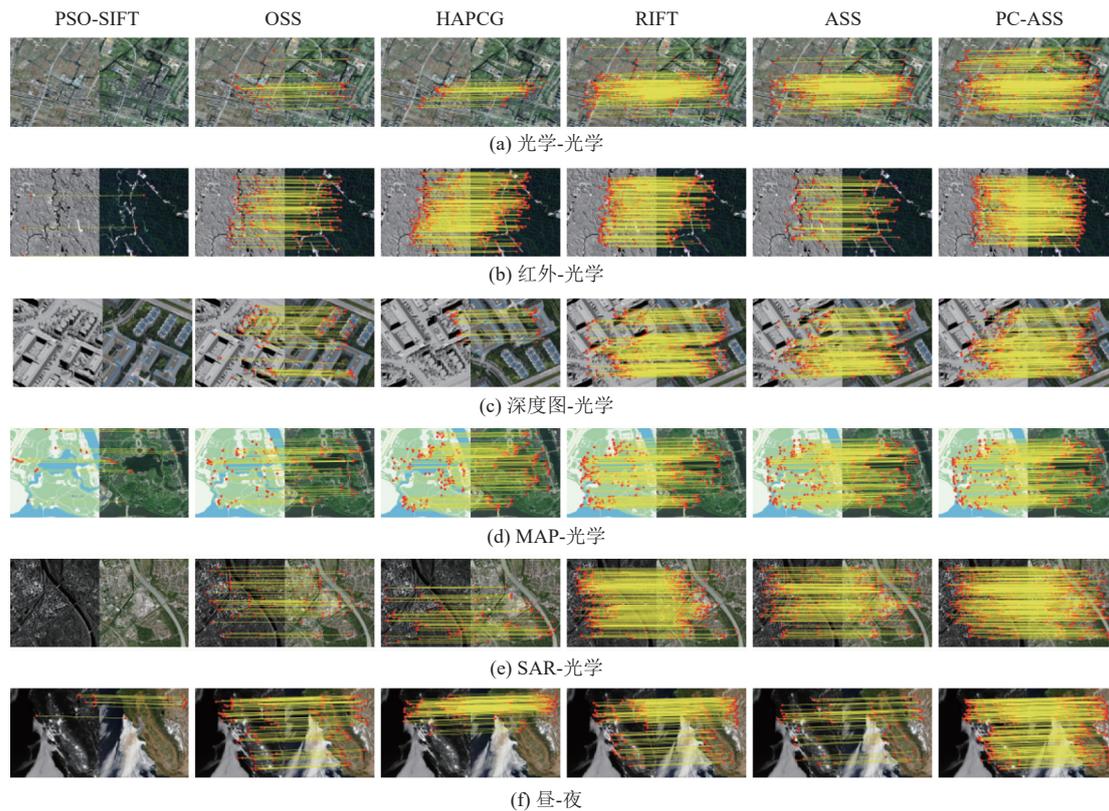


图6 各方法在数据集 CoFSM 的匹配结果

各方法在 CoFSM 数据集上的平均正确匹配点数量和均方根误差如表1所示。

从表1可以看出,在全部6种数据类型中,PC-ASS在平均正确匹配点数量和均方根误差两项关键指标上均取得了最佳性能。PSO-SIFT方法效果最差,其在大多数影像类型上的 NCM 极低,例如在深度图-光学和 SAR-光学数据集上分别仅有0.4和0.9个正确匹配点,均方根误差值也因匹配点过少而无法计算。其因为这类依赖梯度方向直方图的描述符,其梯度信息对多模态影像间的非线性辐射畸变非常敏感,导致难以在辐射差异显著的区域形成稳定有效的描述。OSS方法

在昼-夜和 SAR-光学类型影像上表现不及其他类型影像上的表现,平均正确匹配点数量不超过50。OSS是基于自相似性特征进行匹配,能够在一定程度上捕获局部图像结构的重复模式,但当不同成像机理下相同物体的局部模式差异巨大时判别性就不足。HAPCG方法在6类影像上的平均 *RMSE* 指标都优于 OSS 算法,但在深度图-光学和 SAR-光学类型影像上的平均正确匹配点数不如 OSS 算法,只有79.1和12.3。这是因为 HAPCG 算法依赖相位一致性来识别结构显著区域,但深度图像通常具有平滑的深度渐变和简单的几何结构,缺乏丰富的纹理信息,导致相位一致性响应较弱;而

SAR 图像由于相干散射机制产生的斑点噪声会干扰相位一致性计算,使得算法难以准确识别真正的结构特征. RIFT 算法在 6 类影像匹配效果良好,在光学-光学、红外-光学、深度图-光学、MAP-光学、SAR-光

学类型上的平均 NCM 都超过了 100,在昼-夜类型上的匹配效果略差,平均 NCM 只有 93.1. ASS 算法的平均 NCM 和平均 RMSE 指标都优于 RIFT 算法,但跟 PC-ASS 算法相比略差.

表 1 6 种方法在数据集 CoFSM 上的定量比较结果

数据类别	评价指标	PSO-SIFT	OSS	HAPCG	RIFT	ASS	PC-ASS
光学-光学	NCM	14.2	85.6	128.1	216.9	223.7	322.4
	RMSE	—	2.2155	1.8206	1.8637	1.7960	1.7530
红外-光学	NCM	33.2	178.4	221.7	298.1	272.3	334.9
	RMSE	—	2.0802	1.8507	1.8777	1.7990	1.7286
深度图-光学	NCM	0.4	81.9	79.1	178.6	191.1	235.4
	RMSE	—	2.4059	1.8583	1.9482	1.9140	1.8435
MAP-光学	NCM	2.7	72.1	104.6	148.6	190.3	197.9
	RMSE	—	2.2372	1.8161	2.0039	1.8486	1.7751
SAR-光学	NCM	0.9	37.7	12.3	120.1	116.4	153.0
	RMSE	—	2.4124	1.9079	1.9949	1.9577	1.8961
昼-夜	NCM	7.6	45.2	53.1	93.1	94.2	114.8
	RMSE	—	2.1073	1.7161	1.9430	1.7957	1.5484

注:“—”表示RMSE>4,误差过大,本文不再讨论;加粗表示最优结果

此外,通过计算 PC-ASS 方法与其余 5 种方法在影像上的正确匹配率进行分析,结果如图 7 所示.由图可知,PSO-SIFT 方法在深度图-光学、MAP-光学和 SAR-光学类型影像上正确匹配率较低在 20% 以上,在另外 3 种类型影像上,正确匹配率超过 40%. OSS 方法的整体表现尚可,但在处理 SAR-光学类型影像时正确匹配率只有 34%. HAPCG 方法整体表现偏弱,在所有 6 种数据类型上,其正确匹配率均未超过 43%. RIFT 方法在光学-光学、红外-光学、深度图-光学数据上表现良好,正确匹配率分别为 54.6%、63.6%、58.0%,在其他 3 种类型上表现一般. ASS 方法在 6 种类型影像上正确匹配率均高于前面 4 种算法,在 SAR-光学类型上的正确匹配率最低,但也达到 47.7%. PC-ASS 方法则表现优于 ASS 方法,在光学-光学、红外-光学、深度图-光学、MAP-光学类型影像上的正确匹配率都达到了 60% 以上,在 SAR-光学和昼-夜类型影像上的正确匹配率也在 50% 左右,达到较好的匹配效果.

在 HOWP 数据集上,本文方法同样取得了很好的效果,如表 2 所示.在深度图-光学和 MAP-光学类型影像上 PC-ASS 方法平均正确匹配点数量都在 150 左右,并且在匹配质量上效果也很好,其均方根误差为所有算法中最低的 1.7255 和 1.6830,且平均正确匹配率分别高达 0.714 和 0.801,表现优于其他 5 种方法. PC-ASS 方法在红外-光学和 SAR-光学类型影像上正确匹配点数量低于深度图-光学和 MAP-光学类型影像的数

量,平均正确匹配点数量为 80.8 和 85.8,平均正确匹配率超过 0.6.在昼-夜类型影像上正确匹配点数量仅有 51.2,平均 RMSE 表现较好仅有 1.4638,平均 CMR 为 0.426.整体而言,在 HOWP 数据集上本文方法均优于其他 5 种方法,能够取得较好的匹配性能.

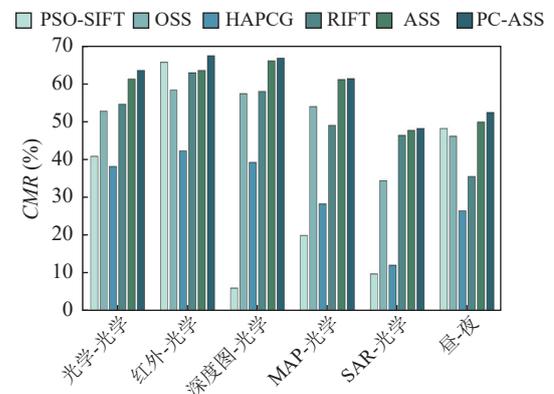


图 7 6 种方法在数据集 CoFSM 上的正确匹配率展示

在文献[23]公开的红外-可见光类型影像数据集中,包含水域影像、植被影像、城市影像和室内影像.如表 3 所示,本文方法的平均 NCM 达到了 1210.2,平均 RMSE 也只有 1.3656,并且正确匹配率达到了 0.8018.本文方法在这个影像数据集中与其他方法相比较展现出较好的匹配性能.

表 4 展示了本文方法基于 CoFSM 数据集的消融实验结果.本文在非线性扩散 (ND) 和相位一致性 (PC)

增强这两个改进点进行了消融实验. 结果显示, 虽然使用传统相邻自相似性算法表现出一定的匹配性能, 但仍有改进空间. 单独使用非线性扩散处理或相位一致性增强的匹配方法在性能上各有贡献, 其中相位一致性增强使正确匹配率 (CMR) 提升至 0.604, 均方根误差 (RMSE) 降低至 1.7854, 而非线性扩散处理在匹配点数量上表现突出, 平均正确匹配点数量 (NCM) 达到

224.1. 但相比基线方法, 各项指标均有所提升, 表明这些改进能够进一步提高本文方法的特征匹配性能. 最终结合非线性扩散处理和相位一致性增强的完整方法取得了最佳性能, NCM 达到 226.4, CMR 提升至 0.617, RMSE 降低至 1.7574, 相比基线方法分别提升了 24.9%、5.8% 和降低了 5.1%, 验证了两个改进的有效性和互补性.

表 2 6 种方法在数据集 HOWP 上的定量比较结果

数据类别	指标	PSO-SIFT	OSS	HAPCG	RIFT	ASS	PC-ASS
红外-光学	NCM	4.2	72.1	10.6	36.4	39.8	80.8
	RMSE	—	2.1151	1.8502	1.8854	1.7565	1.6895
	CMR	0.374	0.571	0.179	0.298	0.501	0.638
深度图-光学	NCM	0	66.1	1.6	65.1	102.3	141.9
	RMSE	—	2.4053	1.7589	2.0018	1.9283	1.7255
	CMR	0	0.528	0.57	0.405	0.546	0.714
MAP-光学	NCM	7.5	96.4	90.4	103.5	128.2	153.5
	RMSE	—	2.2281	1.8113	1.9835	1.828	1.683
	CMR	0.513	0.573	0.24	0.435	0.584	0.801
SAR-光学	NCM	1.4	36.9	4.8	30.1	62.4	85.8
	RMSE	—	2.3907	1.5963	1.9403	1.9316	1.5772
	CMR	0.219	0.423	0.045	0.178	0.551	0.669
昼-夜	NCM	2.1	27	10.1	14	37.4	51.2
	RMSE	—	1.9309	1.5694	1.6573	1.7284	1.4638
	CMR	0.24	0.294	0.063	0.103	0.321	0.426

注: “—”表示RMSE>4, 误差过大, 本文不再讨论; 加粗表示最优结果

表 3 6 种方法在红外-可见光类型影像数据集上的定量比较结果

指标	PSO-SIFT	OSS	HAPCG	RIFT	ASS	PC-ASS
NCM	139.3	536.1	970.9	717.3	806.6	1210.2
RMSE	—	1.5643	1.5091	1.4841	1.4356	1.3656
CMR	0.7566	0.7691	0.735	0.7231	0.774	0.8018

注: “—”表示RMSE>4, 误差过大, 本文不再讨论; 加粗表示最优结果

表 4 消融实验

PC	ND	NCM	RMSE	CMR
—	—	181.3	1.8518	0.583
√	—	188.7	1.7854	0.604
—	√	224.1	1.8228	0.597
√	√	226.4	1.7575	0.617

一方面, 非线性扩散的处理过程利用图像的梯度信息引导特征提取过程, 通过平滑图像噪声同时保持边缘结构, 有助于算法更精确地定位和描述关键特征点. 另一方面, 相位一致性增强不仅能够帮助算法感知图像中更丰富的相位信息, 还可以更好地保留图像中微小的结构变化和边缘特征. 从而增强了匹配方法对全局纹理信息和局部几何特征的感知能力. 结合非线性扩散处理和相位一致性增强的匹配方法能够同时利用两者的优势提升图像匹配的性能. 非线性扩散处理和相位一致性增强在 NCM、RMSE 和 CMR 指标上的

出色表现证明了本文方法在多模态遥感影像匹配应用中的优越性和广泛适用性.

5 结论与展望

本文方法 PC-ASS 通过非线性扩散滤波实现多尺度影像构建, 并结合相位一致性与相邻自相似性特征, 有效提高了特征点的检测质量和描述能力, 展现出较强的鲁棒性与广泛的适用性. 尽管本文方法在实验中取得了较好的效果, 但仍有进一步研究的空间. 本文方法在特征点检测与匹配阶段仍存在一定的计算开销, 未来可结合并行计算或轻量化策略以提升整体效率. 再次, 对于纹理复杂或存在强噪声干扰的遥感场景, 还需探索更具判别力的特征增强机制, 以进一步提升算法的稳定性和匹配精度. 综上所述, 本文提出的方法在多模态遥感影像匹配中表现出良好的性能, 为多源数据的融合与应用提供了有力支持, 同时也为后续研究奠定了基础.

参考文献

- 1 朱柏, 叶沅鑫. 多模态遥感图像配准方法研究综述. 中国图象图形学报, 2024, 29(8): 2137-2161.

- 2 Viola P, Wells III WM. Alignment by maximization of mutual information. *International Journal of Computer Vision*, 1997, 24(2): 137–154. [doi: [10.1023/A:1007958904918](https://doi.org/10.1023/A:1007958904918)]
- 3 Zitová B, Flusser J. Image registration methods: A survey. *Image and Vision Computing*, 2003, 21(11): 977–1000. [doi: [10.1016/S0262-8856\(03\)00137-9](https://doi.org/10.1016/S0262-8856(03)00137-9)]
- 4 Tuytelaars T, Mikolajczyk K. Local invariant feature detectors: A survey. *Foundations and Trends® in Computer Graphics and Vision*, 2008, 3(3): 177–280.
- 5 Shechtman E, Irani M. Matching local self-similarities across images and videos. *Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition*. Minneapolis: IEEE, 2007. 1–8.
- 6 Sedaghat A, Ebadi H. Distinctive order based self-similarity descriptor for multi-sensor remote sensing image matching. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2015, 108: 62–71. [doi: [10.1016/j.isprsjprs.2015.06.003](https://doi.org/10.1016/j.isprsjprs.2015.06.003)]
- 7 陈聪鹏, 喻国荣, 鲍海洲, 等. 结合相位对称与基于排序 LSS 的多模态遥感影像匹配. *计算机系统应用*, 2024, 33(10): 56–65. [doi: [10.15888/j.cnki.csa.009668](https://doi.org/10.15888/j.cnki.csa.009668)]
- 8 Chen SH, Li XR, Yang H, *et al.* Robust local feature descriptor for multisource remote sensing image registration. *Journal of Applied Remote Sensing*, 2018, 12(1): 015005.
- 9 Xiong X, Jin GW, Xu Q, *et al.* Robust registration algorithm for optical and SAR images based on adjacent self-similarity feature. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5233117.
- 10 Nunes CFG, Pádua FLC. A local feature descriptor based on Log-Gabor filters for keypoint matching in multispectral images. *IEEE Geoscience and Remote Sensing Letters*, 2017, 14(10): 1850–1854. [doi: [10.1109/LGRS.2017.2738632](https://doi.org/10.1109/LGRS.2017.2738632)]
- 11 Zhu B, Yang C, Dai JK, *et al.* 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*, 2023, 61: 5606115.
- 12 Kovese P. Image features from phase congruency. *Videre: Journal of Computer Vision Research*, 1999, 1(3): 1–26.
- 13 姚永祥, 张永军, 万一, 等. 顾及各向异性加权力矩与绝对相位方向的异源影像匹配. *武汉大学学报·信息科学版*, 2021, 46(11): 1727–1736.
- 14 Zhang YJ, Yao YX, Wan Y, *et al.* Histogram of the orientation of the weighted phase descriptor for multi-modal remote sensing image matching. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2023, 196: 1–15. [doi: [10.1016/j.isprsjprs.2022.12.018](https://doi.org/10.1016/j.isprsjprs.2022.12.018)]
- 15 Ye YX, Shen L. HOPC: A novel similarity metric based on geometric structural properties for multi-modal remote sensing image matching. *Proceedings of the 2016 ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Prague: ISPRS, 2016. 9–16.
- 16 Li JY, Hu QW, Ai MY. RIFT: Multi-modal image matching based on radiation-variation insensitive feature transform. *IEEE Transactions on Image Processing*, 2020, 29: 3296–3310. [doi: [10.1109/TIP.2019.2959244](https://doi.org/10.1109/TIP.2019.2959244)]
- 17 Quan D, Wang S, Liang XF, *et al.* Deep generative matching network for optical and SAR image registration. *Proceedings of the 2018 IEEE International Geoscience and Remote Sensing Symposium*. Valencia: IEEE, 2018. 6215–6218.
- 18 南轲, 齐华, 叶沅鑫. 深度卷积特征表达的多模态遥感影像模板匹配方法. *测绘学报*, 2019, 48(6): 727–736.
- 19 蓝朝楨, 卢万杰, 于君明, 等. 异源遥感影像特征匹配的深度学习算法. *测绘学报*, 2021, 50(2): 189–202.
- 20 Zeng L, Du YL, Lin HP, *et al.* A novel region-based image registration method for multisource remote sensing images via CNN. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 1821–1831. [doi: [10.1109/JSTARS.2020.3047656](https://doi.org/10.1109/JSTARS.2020.3047656)]
- 21 Fan JW, Wu Y, Wang F, *et al.* SAR image registration using phase congruency and nonlinear diffusion-based SIFT. *IEEE Geoscience and Remote Sensing Letters*, 2015, 12(3): 562–566. [doi: [10.1109/LGRS.2014.2351396](https://doi.org/10.1109/LGRS.2014.2351396)]
- 22 Wu Y, Ma WP, Gong MG, *et al.* A novel point-matching algorithm based on fast sample consensus for image registration. *IEEE Geoscience and Remote Sensing Letters*, 2015, 12(1): 43–47. [doi: [10.1109/LGRS.2014.2325970](https://doi.org/10.1109/LGRS.2014.2325970)]
- 23 Brown M, Süssstrunk S. Multi-spectral SIFT for scene category recognition. *Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition*. Colorado Springs: IEEE, 2011. 177–184.
- 24 Ma WP, Wen ZL, Wu Y, *et al.* Remote sensing image registration with modified SIFT and enhanced feature matching. *IEEE Geoscience and Remote Sensing Letters*, 2017, 14(1): 3–7. [doi: [10.1109/LGRS.2016.2600858](https://doi.org/10.1109/LGRS.2016.2600858)]
- 25 Fan JW, Xiong Q, Ye YX, *et al.* Combining phase congruency and self-similarity features for multimodal remote sensing image matching. *IEEE Geoscience and Remote Sensing Letters*, 2023, 20: 4001105.

(校对责编: 张重毅)