

引文格式:樊仲黎,张力,王庆栋,等.SAR 影像和光学影像梯度方向加权的快速匹配方法[J].测绘学报,2021,50(10):1390-1403. DOI: 10.11947/j.AGCS.2021.20200587.
FAN Zhongli, ZHANG Li, WANG Qingdong, et al. A fast matching method of SAR and optical images using angular weighted orientated gradients[J]. Acta Geodaetica et Cartographica Sinica, 2021, 50(10): 1390-1403. DOI: 10.11947/j.AGCS.2021.20200587.

SAR 影像和光学影像梯度方向加权的快速匹配方法

樊仲黎¹, 张 力¹, 王庆栋¹, 刘思婷^{1,2}, 叶沅鑫³

1. 中国测绘科学研究院, 北京 100039; 2. 兰州交通大学测绘与地理信息学院, 甘肃 兰州 730070; 3. 西南交通大学地球科学与环境工程学院, 四川 成都 611756

A fast matching method of SAR and optical images using angular weighted orientated gradients

FAN Zhongli¹, ZHANG Li¹, WANG Qingdong¹, LIU Siting^{1,2}, YE Yuanxin³

1. Chinese Academy of Surveying and Mapping, Beijing 100039, China; 2. Faculty of Geomatics, Lanzhou Jiaotong University, Lanzhou 730070, China; 3. Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu 611756, China

Abstract: To solve the problem of matching difficulty caused by the significant nonlinear grayscale differences between SAR and optical images, this paper proposes a fast matching algorithm based on image structural properties named SOFM(SAR-to-optical fast matching algorithm). The traditional methods based on image grayscale are generally difficult to resist the nonlinear grayscale differences between SAR and optical images, but the geometric constructs and shape features can exist stably among different types of images, so in our the proposed method both the magnitude and orientation information of image gradient are used to build a geometric structural feature descriptor named AWOG (angular weighted orientated gradients), then based on the template matching strategy, the sum of squared difference of the descriptors is used to define the similarity metric for matching and then the image matching function expressed in the frequency domain is given. A complete set of image matching process is established based on SOFM, and has been validated using multiple pairs of SAR and optical images, the results show that the proposed method can effectively resist the nonlinear grayscale differences between SAR and optical images, and outperforms the traditional classical image grayscale-based methods and existing image structural-based methods in matching performance and precision.

Key words: SAR images; optical images; structural properties; image gradient; template matching

Foundation support: The National Key Research and Development Project of China (Nos. 2019YFB1405602; 2017YFB0503000)

摘 要:针对 SAR 影像与光学影像之间存在显著的非线性灰度差异导致影像匹配困难的问题,本文提出了一种基于影像结构特性的快速匹配方法(SAR-to-optical fast matching algorithm, SOFM)。传统基于影像灰度的匹配方法一般难以抵抗影像间的非线性灰度差异,而影像中的几何结构和形状特征在不同类型的影像之间较为稳定,因此本文综合利用影像的梯度幅值和梯度方向信息构建出一种能够有效表达影像结构的特征描述符—角度加权方向梯度(angular weighted orientated gradients, AWOG),随后基于模板匹配的策略,选择描述符之间的差值的平方和(sum of squared difference, SSD)建立用于匹配的相似性测度,并给出了在频率域中表达的影像匹配函数。基于 SOFM 方法建立了一套完整的影像匹配流程,随后选择多组影像进行匹配试验,结果表明,本文方法能够有效抵抗 SAR 影像与光学影像之间的非线性灰度差异,并且在匹配性能和匹配精度等方面都优于经典的基于影像灰度的匹配方法以及其他基于影像结构特性的匹配方法。

关键词: SAR 影像;光学影像;结构特性;影像梯度;模板匹配

中图分类号: P237

文献标识码: A

文章编号: 1001-1595(2021)10-1390-14

基金项目: 国家重点研发计划(2019YFB1405602;2017YFB0503000)

近年来,随着 SAR 影像数据在地理国情监测^[1]、地表沉降变化检测^[2]等领域发挥的作用越来越显著,其与光学影像截然不同的成像机理与特性引起学术界和工业界的广泛关注。结合 SAR 影像与光学影像各自优势,实现对大量地表特性信息的互补,能够更好地解决变化检测、目标识别、地物分类等任务。目前 SAR 影像和光学影像联合应用的关键前提在于影像匹配,影像匹配的本质是在多幅影像之间寻找同名点的过程^[3],但是由于 SAR 影像独特的成像机理,使其与光学影像之间存在显著的非线性灰度差异,同时 SAR 影像还会受到严重的相干斑噪声干扰^[4],这些影响因素使光学影像上的显著特征点在 SAR 影像上极不稳定,甚至失效,导致现有的影像匹配技术难以应用于 SAR 影像和光学影像间的匹配,因此 SAR 影像和光学影像的匹配技术是目前遥感领域的难点,也是热点研究问题之一^[5-7]。

影像匹配方法主要可分为基于特征的方法、基于区域的方法^[8]和基于深度学习的方法^[9-10]。基于特征的方法通过在输入的影像之间提取共有特征,并检测特征间的相似性来实现影像匹配。这些影像特征可以是点特征^[11]、线特征^[12]、面特征^[13]或是局部不变性特征(SIFT^[14]、SURF^[15]等)。为了得到较好的匹配效果,检测到的影像特征应具有良好的稳定性、可区分性和可重复性,但是在具有显著非线性灰度差异的影像之间很难提取到具有高重复率的共有特征,从而导致匹配不到足够数量的同名点,使得此类方法在 SAR 影像和光学影像匹配任务上失败。

基于区域的方法也被称为模板匹配方法,通过在影像间设置一定尺寸大小的影像窗口,并选择某种相似性测度作为检测影像相似性的准则来实现影像匹配^[16]。传统的模板匹配方法通常是利用影像的灰度信息进行匹配,代表性的方法有:互信息(mutual information, MI)^[17]、归一化互相关(normalized cross correlation, NCC)^[18]等。由于 NCC 方法能够对影像间的线性灰度变化保持稳健,因此被广泛应用于光学遥感影像匹配,但其对影像间的非线性灰度差异较为敏感。对于 MI 方法,其本质是依赖于影像灰度的统计信息进行

匹配,与影像间灰度变化的方式相关性较弱,故理论上其对影像间的非线性灰度差异有一定的适应性,然而 MI 方法忽略了邻域像素的影响,使得计算结果容易陷入局部极值而导致误匹配,并且其对模板窗口大小比较敏感,匹配过程计算量较大,因此 MI 方法在 SAR 影像和光学影像的匹配时效果不佳。

受到深度学习技术的影响,一些基于深度神经网络进行 SAR 影像和光学影像匹配的方法被提出。文献[9]提出了一种基于图像风格迁移的匹配方法,通过改进的 VGG-19 网络模型将 SAR 影像和光学影像进行风格互换,再利用 SIFT 或 SURF 等算法对互换风格后的影像进行匹配,该方法能够获取较多的正确匹配点。文献[10]改进了一种用于检测图像相似性的 Siamese 网络模型^[19-20],通过抽取特征以及去除池化层对原网络进行优化,该方法取得了较高的正确匹配率和匹配精度。现阶段基于深度学习的方法存在一定的局限性,首先需要花费大量的时间和精力来制作训练数据集以及训练模型,并且由于地表形态构造的多样性,使得训练数据集难以具有良好的通用性,导致对于不同类型的地表形态,训练模型缺乏良好的泛化能力;其次,此类方法的执行效率较低,由于需要将遥感影像输入到神经网络模型中进行训练,这个过程需要对海量样本数据进行迭代计算,对计算环境的软硬件配置提出挑战,并且非常耗时,因此这类方法难以达到实际应用的要求,缺乏时效性。

除上述的各种方法之外,近年来,一些研究者发现影像的几何结构和形状特征能够有效抵抗影像间的非线性灰度差异^[21-22],因而将其引入到 SAR 影像和光学影像的匹配当中,并取得了较好的匹配效果。这类方法不直接利用影像的灰度信息进行匹配,而是首先利用如影像的梯度、局部自相似性(local self-similarity, LSS)、相位一致性(phase congruency, PC)等能有效地表达影像结构的信息来构建特征描述符,随后建立描述符之间的相似性测度,并采用模板匹配的策略对影像进行匹配。文献[23]利用具有对比度和光照不变性的相位一致性特征对方向梯度直方图

(histograms of oriented gradients, HOG)^[24]进行了改进,提出了相位一致性方向直方图(histogram of orientated phase congruency, HOPC),在光学和 SAR 影像的匹配任务上取得了较好的效果。但是, HOG 和 HOPC 是在一个稀疏的采样格网(非逐像素)内进行特征构建,是一种相对稀疏的特征表达方式,难以精确地捕获影像的细节结构信息,另外其计算效率较低。文献[25]利用方向梯度信息构建了一种稠密结构特征描述符(channel features of orientated gradients, CFOG),通过逐像素地提取邻域结构特征,增强了描述符对于结构信息的细节表达能力,同时在频率域中完成影像匹配,显著提高了匹配性能及计算效率。不过, CFOG 是利用水平和垂直方向的梯度内插出的方向梯度,其计算精度还可进一步提高。针对这一问题,本文提出一种更精确的方向梯度计算方法,并以此构建特征描述符进行 SAR 影像和光学影像的匹配。

本文提出的 SOFM 方法使用影像的梯度信息构建出一种表达影像结构的特征描述符,即角度加权方向梯度 AWOG,随后选择描述符之间的差值的平方和(SSD)作为用于匹配的相似性测度,并利用快速傅里叶变换将匹配过程转换到频

率域进行。设计了一套基于本文方法的自动匹配流程,并通过设计多组试验来验证本文方法的有效性和匹配性能。

1 结构特性表达的影像相似性

本节将给出利用影像梯度信息构建 AWOG 描述符的详细步骤,以及给出在频率域中表达的影像匹配函数的详细推导过程。

1.1 特征描述符

AWOG 描述符的提出受到 HOG^[23] 和 CFOG^[25] 的启发。HOG 描述符最初被用于处理行人检测任务,由于对影像结构具有良好的表达能力,其应用范围被拓展到目标识别^[26]、影像分类^[27] 和影像检索^[28] 等领域。CFOG 是在 HOG 的基础上发展的一种逐像素的结构特征描述符,其使用方向梯度信息来表达影像结构特性,能够有效抵抗影像间的光照和对比度变化,以及非线性灰度差异。因此,本文同样选择利用影像梯度幅值和方向信息来构建 AWOG 特征描述符。

如图 1 所示,对于一幅给定的影像,首先使用 x 方向的滤波器 $d_x = [-1, 0, 1]$ 和 y 方向的滤波器 $d_y = [-1, 0, 1]^T$ 对影像进行滤波,得到影像在 x 轴和 y 轴方向的方向梯度 g_x 和 g_y 。

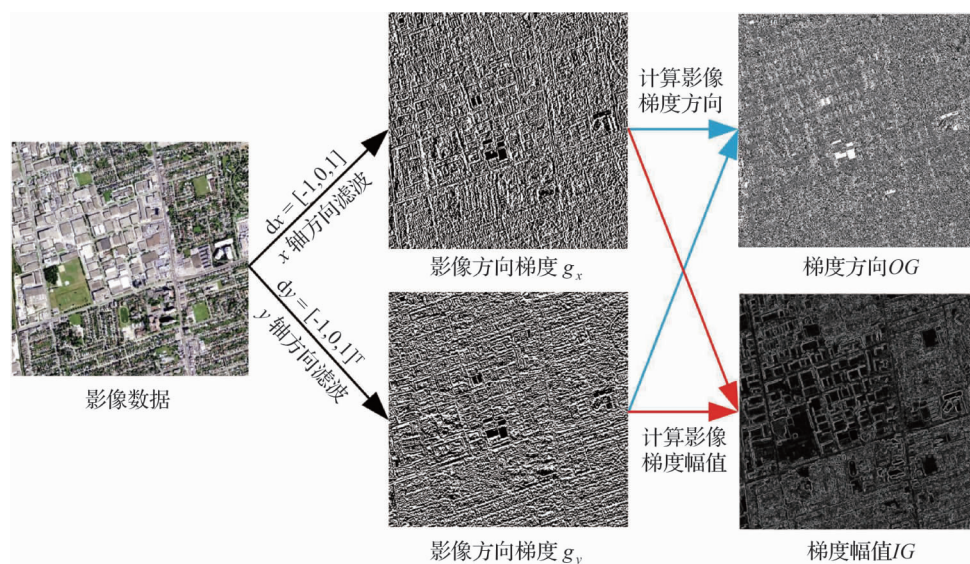


图 1 影像梯度方向 OG 和梯度幅值 IG 的计算

Fig.1 Computational process of orientation and magnitude of image gradient

随后根据式(1)和式(2)计算影像的梯度幅值 IG 和梯度方向 OG。由于不同类型的传感器所获取的影像之间存在辐射强度反转的现象^[29],因

此为削弱 SAR 影像和光学影像间因辐射强度反转所造成的影响,以提高描述符的稳健性,对梯度方向 OG 进行一步归化操作^[25,30],根据式(2)计

算得到的梯度方向 OG 其值域为 $[0^\circ, 360^\circ)$, 将属于区间 $[180^\circ, 360^\circ)$ 的值统一减去 180° , OG 的值域被归化到 $[0^\circ, 180^\circ)$

$$IG(x, y) = \sqrt{g_x(x, y)^2 + g_y(x, y)^2} \quad (1)$$

$$OG(x, y) = \arctan^{-1} \left(\frac{g_y(x, y)}{g_x(x, y)} \right) \quad (2)$$

如图 2 所示,按角度间隔为 22.5° 将 OG 的值域 8 等分,等分的区间被 9 个角度方向所隔开,由于特征描述符将在这 9 个角度方向上统计得到,故将它们称为 9 个特征方向,按角度从小到大为 9 个特征方向赋予索引号 0~8。

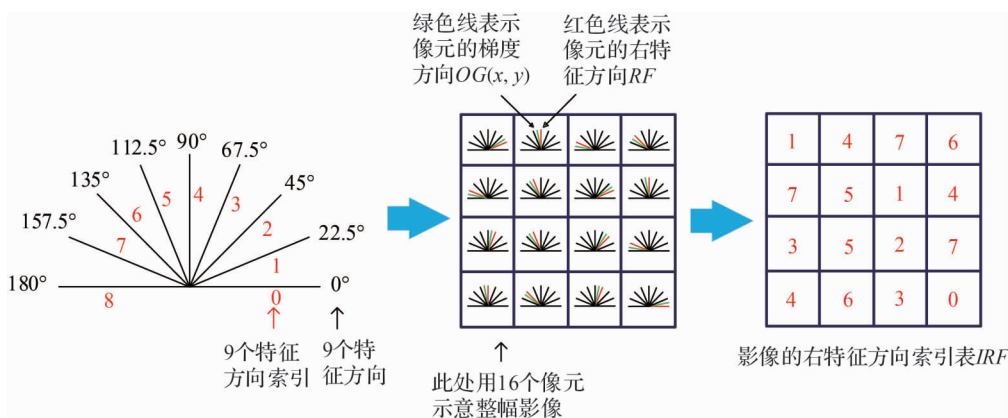


图 2 右特征方向索引表 IRF 的构建
Fig.2 Construction of index table IRF

对于影像中任意像元 $p(x, y)$,其梯度方向值 $OG(x, y)$ 存在于某两个特征方向所构成的角度区间内(极少数恰好位于某特征方向上,但不影响后续计算),将该角度区间中角度值较大的特征方向称为像元的左特征方向 LF ,角度值较小的特征方向称为像元的右特征方向 RF ,根据式(3)进行计算

$$IRF(x, y) = \left\lfloor \frac{OG(x, y)}{22.5} \right\rfloor \quad (3)$$

式中, $\lfloor \cdot \rfloor$ 符号表示取不大于符号内数值的整数。

可得到像元右特征方向的索引号 $IRF(x, y)$,由此对整幅影像进行计算,得到影像的右特征方向索引表 IRF ,该索引表将用于加速后续的计算过程。

如图 3 所示,对于任意像元 $p(x, y)$,其梯度方向值 $OG(x, y)$ 存在于其左特征方向 LF 和右特征方向 RF 所构成的角度区间内,因此可根据梯度方向值 $OG(x, y)$ 与左右特征方向的角度差,将该像元的梯度幅值 $IG(x, y)$ 加权分配到其左右特征方向上, $OG(x, y)$ 与其左右特征方向中某一个的角度差越小,说明 $OG(x, y)$ 距离该特征方向越近,从而 $IG(x, y)$ 向该特征方向分配的值越大;反之, $OG(x, y)$ 与该特征方向的角度差越大,说明 $OG(x, y)$ 距离该特征方向越远, $IG(x, y)$

向该特征方向分配的值越小。

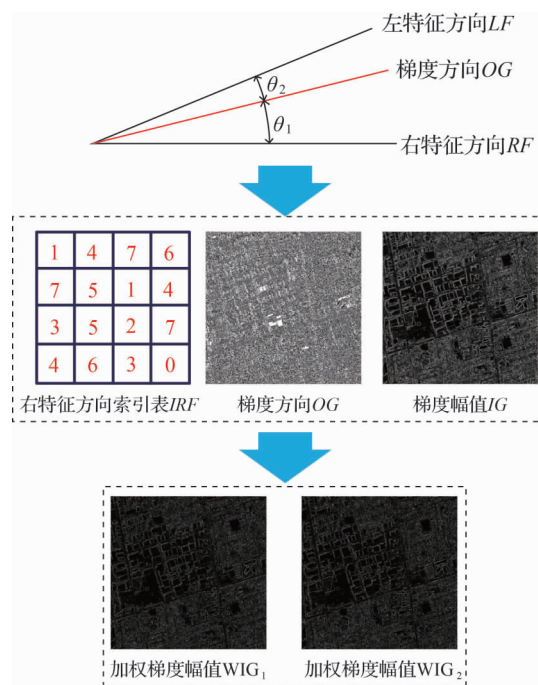


图 3 加权梯度幅值 WIG_1 、 WIG_2 的计算过程
Fig.3 Computational process of weighted magnitude of image gradient

基于上述加权原则,根据式(4)计算像素的梯度方向值 $OG(x, y)$ 与右特征方向 RF 的角度差

$\theta_1(x, y)$ 以及与左特征方向 LF 的角度差 $\theta_2(x, y)$

$$\left. \begin{aligned} \theta_1(x, y) &= OG(x, y) - IRF(x, y) \cdot 22.5 \\ \theta_2(x, y) &= 22.5 - \theta_1(x, y) \\ WIG_1(x, y) &= \frac{\theta_2(x, y)}{22.5} \cdot IG(x, y) \\ WIG_2(x, y) &= \frac{\theta_1(x, y)}{22.5} \cdot IG(x, y) \end{aligned} \right\} (4)$$

随后根据 $\theta_1(x, y)$ 和 $\theta_2(x, y)$ 对梯度幅值 $IG(x, y)$ 进行加权分配, 由此对整幅影像进行处理, 得到影像的加权梯度幅值 WIG_1 和 WIG_2 。

逐像元为中心在 WIG_1 和 WIG_2 上同时开辟尺寸为 3×3 像素的统计窗口, 对两个窗口内 9 个特征方向上的加权梯度幅值进行统计, 利用右特征方向索引表 IRF 加速统计过程, 联合两个统计窗口, 将统计结果以向量的形式输出, 该向量即为

统计窗口中心像元的特征向量。使用索引表进行统计相比于不使用索引表的统计方式, 在统计窗口内的每个像元处均可减少两次乘法、两次除法和一次减法运算而只保留两次加法运算, 因此, 使用索引表进行统计将极大地提高特征描述符的提取速度。

将每个像元处的特征向量沿垂直于影像平面的 z 轴方向排列便形成初步的影像描述符。随后使用 z 轴方向的滤波器 $dz = [1, 3, 1]^T$ 对初步描述符进行滤波, z 轴上的滤波操作对不同特征方向的特征值起到平滑作用, 该操作能够减小影像间局部几何和强度变形引起的畸变所造成的影响^[25]。对 z 轴滤波后的描述符进行归一化便得到最终的 $AWOG$ 特征描述符。

综上, $AWOG$ 特征描述符的构建过程如图 4 所示。

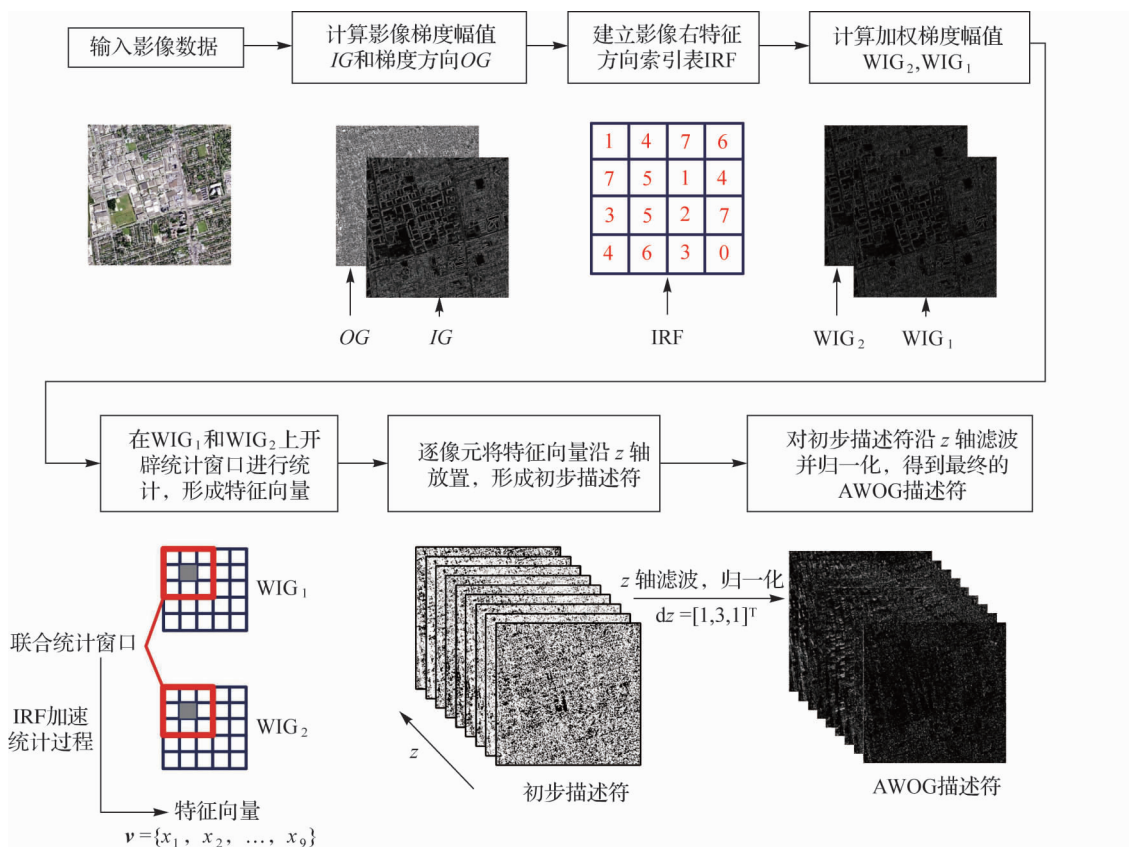


图 4 AWOG 特征描述符的构建
Fig.4 Construction of AWOG

1.2 相似性测度

由于 $AWOG$ 描述符具有多层级结构, 相比于传统的单层描述符来说其数据量较大, 在空间

域中进行匹配较为耗时, 因此选择在频率域中进行匹配^[31], 在频率域中能够显著提高匹配时的计算速度^[25], 选择描述符之间差值的平方和作为同

名点匹配时的相似性准则,并利用快速傅里叶变换推导出在频率域中表达的影像匹配函数。

差值的平方与表达的 AWOG 描述符之间的相似性测度表示为

$$d(v) = \sum_x [A_1(x) - A_2(x - v)]^2 \quad (5)$$

式中, x 表示某像元的像素坐标; $A_1(x)$ 和 $A_2(x - v)$ 分别表示参考影像描述符和输入影像描述符被模板窗口覆盖的区块; 变量 v 表示输入影像描述符的搜索窗口中某像元相对于搜索窗口中心像元的偏移向量; $d(v)$ 为两个描述符区块之间的差值的平方和函数。

为了得到参考影像中的某像元在输入影像上的最佳匹配位置,需对式(5)求最小值,由此,同名点的搜索问题转化为计算使式(5)达到最小值的偏移向量 v_i 。进而给出如下影像匹配函数

$$v_i = \operatorname{argmin}_v \left\{ \sum_x [A_1(x) - A_2(x - v)]^2 \right\} \quad (6)$$

对于模板匹配的过程,需要在输入影像搜索窗口中的每个像元位置上,计算与模板窗口间的相似性值,该计算过程非常耗时,因此,为了提高计算效率,采用快速傅里叶变换将该过程转换到频率域进行加速。

将式(6)中的平方项展开

$$v_i = \operatorname{argmin}_v \left\{ \sum_x A_1^2(x) + \sum_x A_2^2(x - v) - 2 \sum_x A_1(x) \cdot A_2(x - v) \right\} \quad (7)$$

对于式(7),等式右侧第 1 项的计算值是一个常数,因此计算时可将其忽略。后两项可以视为描述符之间的相关计算,由于空间域中的相关对应频率域中的相乘,因此可利用快速傅里叶变换(二维或三维)将后两项转换到频率域中计算。影像的匹配函数转化为

$$v_i = \operatorname{argmin}_v \left\{ F^{-1} [F(A_2(x - v))F^*(A_2(x - v))] - 2F^{-1} [F(A_2(x - v))F^*(A_1(x))] \right\} \quad (8)$$

式中, F 和 F^{-1} 表示快速傅里叶变换的正变换和逆变换; F^* 表示对 F 求复共轭。

对于式(8),等式右侧第 1 项表达输入影像描述符与自身的相关计算,对于进行了归一化的特征描述符如 AWOG、FLSS^[32]、FHOG^[32] 等,右侧第 1 项的计算值几乎为一个常数,同时其在影像不同位置处的计算值的变化率对于式(8)右侧第 2 项的计算结果几乎没有影响,因此参照文献[32],选择忽略式(8)右侧第 1 项而只保留第 2 项。因此同名点匹配被转化为计算对式(8)右侧

第 2 项反符号求得最大值时对应的偏移向量 v_i 。最终得到如下影像匹配函数

$$v_i = \operatorname{argmax}_v \left\{ F^{-1} [F(A_2(x - v))F^*(A_1(x))] \right\} \quad (9)$$

式(9)表达了两幅影像的描述符在频率域中的互相关,该计算会得到一幅相关图,因此通过确定最大反应值所在位置,便可以得到同名匹配点。

为验证上述影像匹配函数的有效性,利用 x 方向的相似性曲线将其与 MI^[17] 和 NCC^[18] 进行比较,选择尺寸为 41×41 像素的模板窗口在一组高分辨率城区影像上进行测试,测试影像如图 5 所示。

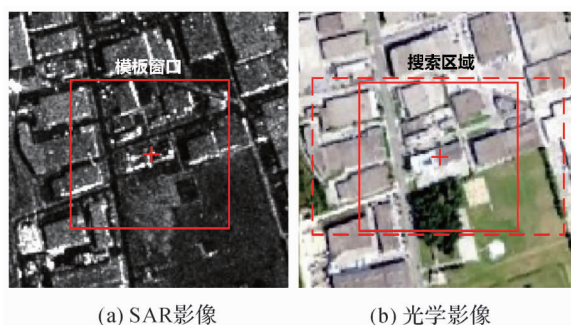


图 5 相似性曲线测试影像
Fig.5 Images for testing similarity curve

图 6 展示 SOFM、MI^[17] 和 NCC^[18] 等 3 种方法在 X 方向上的相似性曲线,通过对比可以发现,SOFM 方法的相似性曲线比较光滑并成功检测到了正确匹配位置,相比之下,MI 和 NCC 方法的相似性曲线则不够光滑,且峰值位置与正确匹配位置存在一定偏差。通过相似性曲线的对比,初步证明了本文方法的有效性,关于本文方法匹配性能的分析见第 3 节。

2 SAR 影像和光学影像模板匹配流程

(1) 将 SAR 影像作为参考影像,光学影像作为输入影像,为两影像构建影像金字塔。若影像间分辨率差异不超过 2 倍,一般为影像金字塔设置 4 个层级,层级间的缩放比为 2;若影像间分辨率差异较大,则根据分辨率差异倍数为分辨率较高的影像增加一定的金字塔影像层数,并通过建立查找表的方式在分辨率差异较小的层级间进行匹配。由低层级影像构建高层级影像时,先对低层级影像进行高斯滤波,后按照双线性内插的采样方式生成高层级影像。随后由顶层金字塔影像进入后续处理步骤。

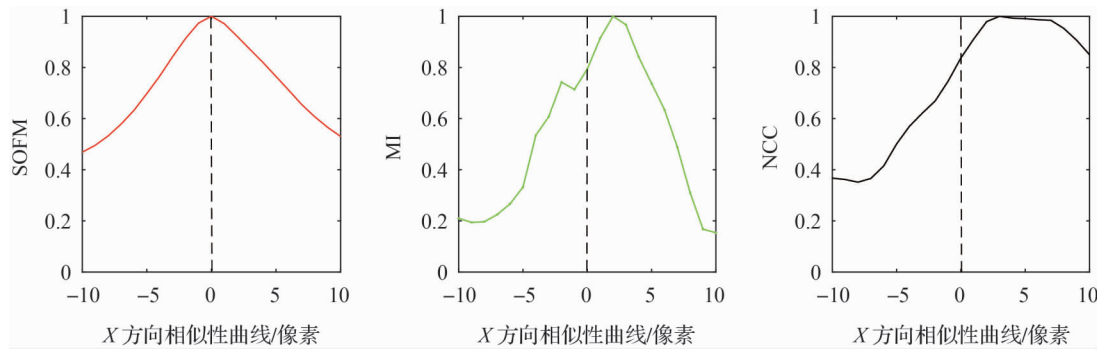


图6 SOFM、MI和NCC方法的相似性曲线
Fig.6 Similarity curves of SOFM, MI and NCC

(2) 为保证所提取影像特征点的均匀性,选择使用分块 Harris 算子,计算参考影像中每个像元处的 Harris 特征值,随后将影像划分为 $n \times n$ 个互不重叠的格网区域,提取每个格网内特征值较大的 k 个点作为特征点,整幅影像将得到 $k \times n \times n$ 个特征点。关于划分的格网数 n 和格网内特征点数 k 的选择,一般经验性的将 k 设置为 1~5 个点,再根据期望得到的总点数计算出格网数 n 。

(3) 考虑到影像之间存在旋转与尺度差异,而 AWOG 描述符不具备旋转与尺度不变性,因此在匹配之前对影像进行 image-reshaping 操作(图 7)来消除影像间的旋转和尺度差异,目前所允许的影像间的缩放倍数约为 0.2~5。处理步骤为:①以特征点为中心在参考影像上确定一定尺寸的搜索区域;②给定搜索区域的近似高程面(区域平均高程面或已有概略 DEM),利用参考影像的定向参数将搜索区域投影到近似高程面上,得到搜索区域对应的物方面元;③利用影像成像模型(如 RFM 模型)将物方面元反投影到输入影像上得到相应的搜索区域范围;④建立参考影像的搜索区域和输入影像的搜索区域之间的透视变换或仿射变换模型,并根据所建立的变换模型对输入影像的搜索区域进行重采样,基本上消除其与参考影像之间的旋转和尺度差异。

(4) 根据本文提出的 SOFM 算法,对经过步骤(3)处理的参考影像和输入影像的影像块提取 AWOG 描述符,并根据给出的影像匹配函数完成同名点匹配。

(5) 重复步骤(3)一步骤(4),直至遍历所有特征点,得到初始的同名点集合,由于匹配过程在提取的影像块上进行,此时同名点坐标由局部像

素坐标系所表达,因此通过坐标转换将同名点坐标表达在原始输入影像的像素坐标系上,并输出初始同名点集合。

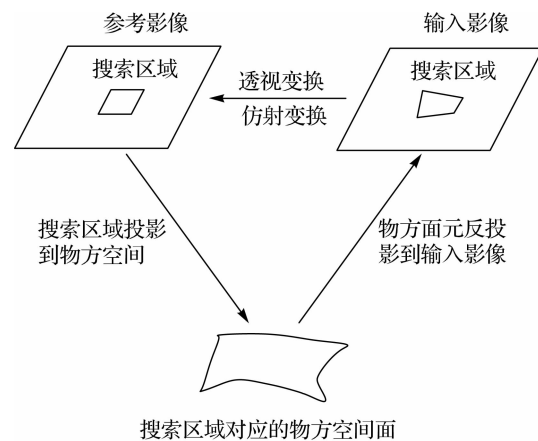


图7 影像匹配时的 image-reshaping 操作
Fig.7 Image-reshaping processing used for image matching

(6) 构建二维透视变换表达的 RANSAC 框架,设置一定的粗差阈值及迭代次数,将步骤(5)输出的初始同名点集合输入该框架,迭代剔除残差大于阈值的点对,输出内点率最高的同名点集合。

(7) 构建二维透视变换表达的整体最小二乘模型,将步骤(6)剔除粗差后的同名点集合作为输入,计算单个点的残差和整体的标准差,迭代删除残差较大的同名点对,直至标准差小于给定的阈值,将剩余的同名点集合作为匹配结果,并由此计算当前层级金字塔影像的仿射变换参数。

(8) 重复步骤(2)一步骤(7),并将计算出的仿射变换参数向低层级影像传递,从而实现由粗到精的影像匹配策略,直至完成金字塔底层影像

的匹配,并记录最终输出结果。

3 试验与分析

为验证所提出方法的有效性,本文试验将仅考虑影像间的平移差异,对于影像间的旋转和尺度差异可通过第 2 节中的 image-resaping 操作进行消除,使得影像间只剩余平移参数未知,而平移参数与影像自身的无控定位精度有关。因此,为简化试验步骤,试验中所使用的参考影像均为

经过多视处理后的 SAR 影像的数字正射影像图(以下简称 DOM),输入影像均为经过正射纠正后的光学影像的 DOM。试验过程中使用的计算机配置为 Intel(R)Core(TM)i9-10885H CPU @ 2.4 GHz,64 GB。

3.1 试验数据

选择 5 组不同分辨率且覆盖不同地面场景的 SAR 影像和光学影像的 DOM 作为试验数据,如图 8 所示。具体的数据描述见表 1。

表 1 试验数据描述

Tab.1 Description of datasets used in the matching experiments

数据编号	参考影像 DOM		输入影像 DOM		影像特点
	传感器(数据源)获取时间	大小/像素分辨率/m	传感器(数据源)获取时间	大小/像素分辨率/m	
1	TerraSAR-X 2007-12	534×5243	Google Earth 2007-11	528×5243	影像位于城市地区,房屋明显,道路清晰
2	TerraSAR-X 2013-07	600×60 030	TM 波段 3 2014-03	600×60 030	影像位于郊外地区,地块分明,水系丰富
3	天绘二号 2019-08	777×73 710	Google Earth 2019-04	777×73 710	影像位于城市近郊,地物类型丰富,包含河流、道路、厂房
4	高分三号 2017-02	1000×100 010	Google Earth 2017-08	1001×100 110	影像位于山区,山脊线明显,地势较缓区域存在农田
5	天绘二号 2019-12	3624×396 310	Google Earth 2019-06	3625×396 410	影像覆盖我国某城市及其周边地区

图 8 为试验影像的缩略图,可以看出每组影像间均存在显著的非线性灰度差异。

3.2 SOFM 方法对模板窗口大小的适应性分析

使用表 1 中的第 5 组试验数据,按照第 2 节的匹配流程对影像进行匹配,固定搜索窗口为 31×31 像素,改变模板窗口的尺寸(41×41 像素~101×101 像素),记录最终输出的点数以及匹配到的总点数,试验结果见表 2。

表 2 不同模板窗口下的影像匹配结果

Tab.2 The results of image matching with different template sizes

模板窗口尺寸/像素	输出点数	总点数	比率/(%)
41×41	865	1209	71.55
51×51	1007	1199	83.99
61×61	1121	1211	92.57
71×71	1065	1110	95.95
81×81	1080	1100	98.18
91×91	1075	1085	99.08
101×101	1075	1081	99.44

分析表 2 中的结果可知,随着设置的模板窗口尺寸的增大,最终输出的点数逐渐增多且趋于

稳定,输出点数占匹配总点数的比率逐渐上升,当模板窗口尺寸设置为 61×61 像素时,输出点的数量以其占总点数的比率均已达到较高水平,因此在应用中将模板窗口的尺寸固定为 61×61 像素即可。

3.3 SOFM 方法对搜索窗口大小的适应性分析

使用表 1 中的第 5 组试验数据,对试验影像建立影像金字塔(4 级),相邻层级间的缩放倍数为 2,选择 61×61 像素的模板窗口,对顶层金字塔影像进行整个影像范围内的搜索,计算影像间的仿射变换参数并向低层级影像传递,并对低层级影像重新提取特征点,在低层级影像中设置 21×21 像素的搜索窗口,记录最终输出的点数以及匹配到的总点数,试验结果见表 3。

表 3 影像金字塔匹配结果

Tab.3 The results of image matching using image pyramid

金字塔层级	搜索窗口尺寸	输出点数	总点数	比率/(%)
1 : 8	整幅影像	143	143	100
1 : 4	21×21	290	299	96.99
1 : 2	21×21	610	639	95.46
1 : 1	21×21	1117	1211	92.23

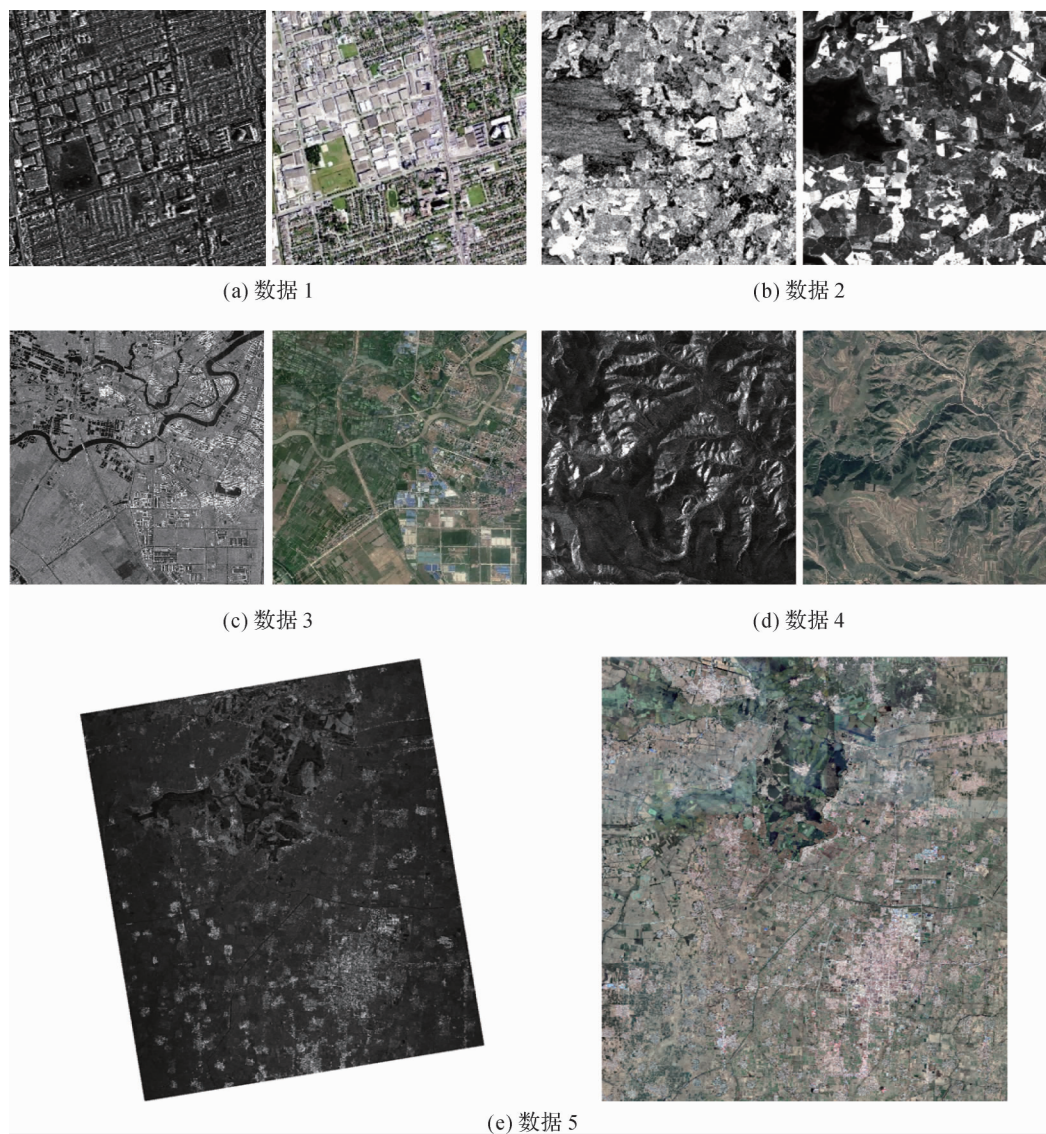


图 8 试验影像

Fig.8 The images used in the experiments

对表 3 进行分析可知,在影像匹配时,通过使用对影像建立影像金字塔的方式,既能加快影像匹配的速度,又能稳定地获取数量足够多的同名点。

3.4 匹配性能对比分析

为验证 SOFM 方法的匹配性能,使用表 1 中的第 1~4 组试验数据,选择 CFOG^[25]、HOPC^[23]、DLSC^[33]、MI^[17]、NCC^[18] 等方法进行对比分析与评价,对比的内容包括:正确匹配率(正确匹配点数/总匹配点数)、计算效率和均方根误差(RMSE)等 3 个方面。试验步骤为:①在参考影像上提取 200 个 Harris 特征点;②利用影像金字塔为每个特征点在输入影像上确定一个 21×21 像素的搜索窗口,采用不同尺

寸的模板窗口(31×31 像素~ 91×91 像素)对影像进行匹配并记录不同方法的性能表现。

图 9 展示了参与对比的 6 种匹配方法在 4 组试验影像上的正确匹配率(容差为 1.5 像素)。

通过对图 9 进行分析可以得到以下结论。

(1) NCC 方法总体上在各组试验中的正确匹配率最低,甚至在第 3 组试验中出现了近乎匹配失败的情况,究其原因,NCC 方法直接利用影像的灰度信息进行匹配,其能够对影像间的线性灰度变化具有良好的适应性,但对影像间的非线性灰度差异非常敏感,因此 NCC 方法在进行光学影像间的匹配时表现良好,但在匹配 SAR 影像和光学影像时效果较差。

(2) MI 方法在多组试验中的表现虽然优于

NCC 方法,但其总体上的正确匹配率仍然较低,MI 方法本质是依赖影像灰度的统计信息进行匹配,故理论上讲其匹配效果应与影像间灰度变化的方式相关性较弱,但 MI 方法在匹配过程中

不考虑邻域像素的影响,导致计算结果容易陷入局部极值而产生误匹配,因此 MI 方法不适用于 SAR 影像和光学影像的匹配。

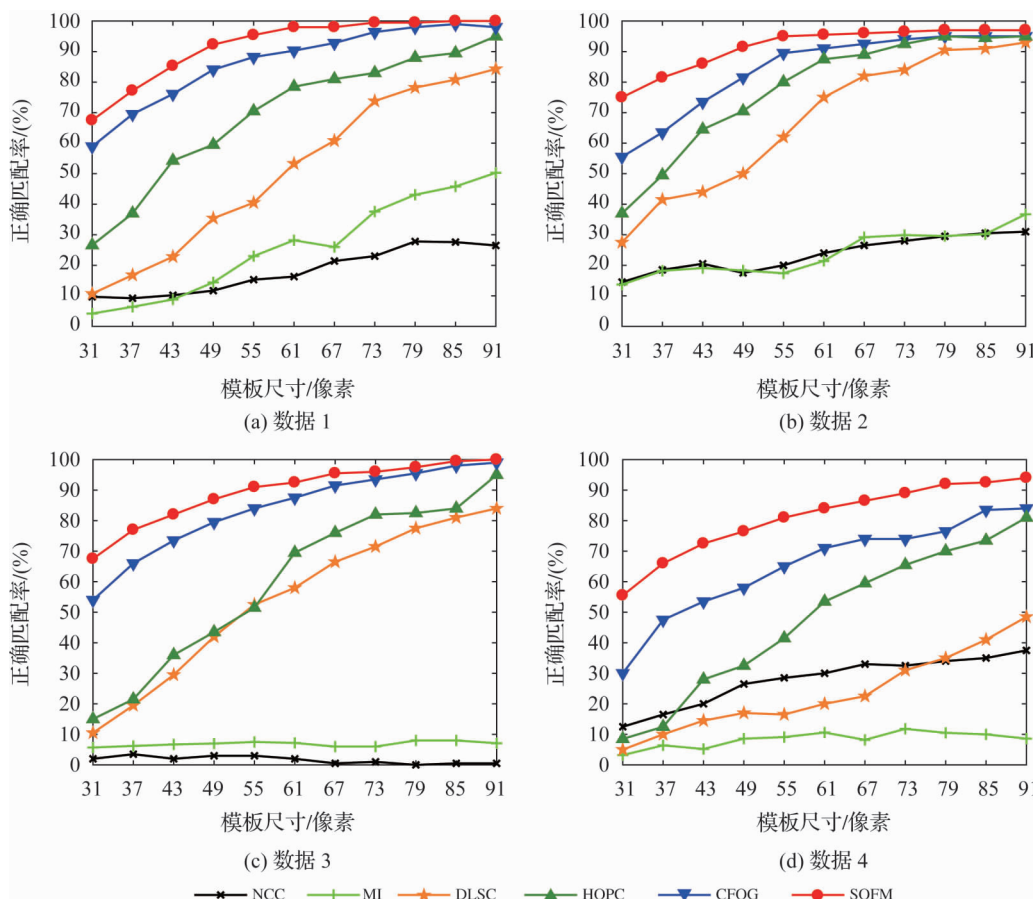


图 9 SOFM、CFOG、HOPC、DLSC、MI 和 NCC 在不同模板尺寸下的正确匹配率

Fig.9 The correct matching ratios of SOFM, CFOG, HOPC, DLSC, MI and NCC with different template sizes

(3) HOPC 和 DLSC 方法的表现较好,总体来看,HOPC 在各组试验影像上的正确匹配率均高于 DLSC。这两种方法属于基于影像结构特性的匹配方法,HOPC 和 DLSC 方法分别利用了影像的相位一致性(PC)和局部自相似性(LSS)信息来表达影像的几何构造,但是由图 9 可知,相比于 SOFM 这两种方法的匹配性能受模板尺寸变化的影响较大,在模板尺寸较大时,能够得到较高的正确匹配率;但当设置的模板尺寸逐渐减小时,这两种方法的正确匹配率会出现较明显的降低。

(4) 相比于上述的几种方法,SOFM 和 CFOG 在各试验像对上的正确匹配率更高,但是整体而言,SOFM 方法在每组试验中的表现均优于 CFOG。SOFM 的正确匹配率受模板尺寸变化的影响更小,性能更加稳定,在模板尺寸较小时

仍能保证较高的正确匹配率。当设置模板尺寸为 31×31 像素时,SOFM 方法在 4 组试验上的正确匹配率平均要比 CFOG 方法高 17.8%。分析其原因,可能是由于 CFOG 方法是利用水平和垂直方向的梯度插值得到多个固定方向的方向梯度,并对方向梯度进行高斯卷积来得到目标像素的特征向量。而这种计算方式会引入一定程度的模糊,降低特征向量之间的区分性,因此当设置模板尺寸较小或影像质量较差时,会导致其匹配效果出现一定程度的下降。而 SOFM 方法则是以方向间的角度差为权重,将像素的梯度值加权分配到与梯度方向相关性最强的两个固定方向上,同时结合索引表进行邻域统计得到目标像素的特征向量,增强了特征向量的独特性,因此其稳健性更好,匹配效果更优。

图 10 展示了各方法在 4 组试验中的平均运行时间。

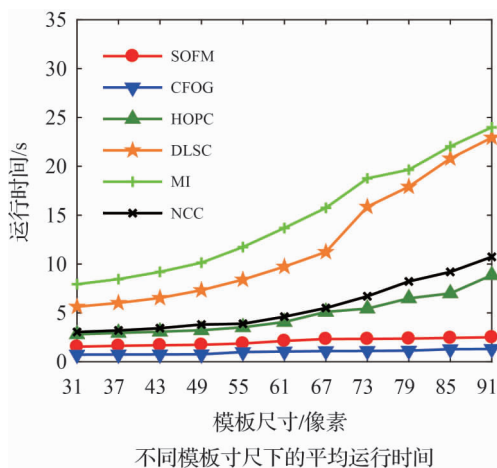


图 10 SOFM、CFOG、HOPC、DLSC、MI 和 NCC 在不同模板尺寸下的平均运行时间

Fig.10 The average running times of SOFM, CFOG, HOPC, DLSC, MI and NCC with different template sizes

分析图 10 可得如下结论。

(1) 总体上 MI 方法的计算效率最低,其次较低的是 DLSC 方法,NCC 方法的计算效率虽然较高,但是其匹配效果较差,HOPC 方法的计算效率略高于 NCC。CFOG 方法的计算效率最高,SOFM 方法的计算效率略低于 CFOG 方法。

(2) SOFM 和 CFOG 方法的时间消耗随着模板尺寸的增大只有微小增加,而其他 4 种方法的时间消耗会随着模板尺寸的增大而迅速增加。当模板尺寸较大时(47×47 像素及以上),HOPC 和 DLSC 方法能够得到较高的正确匹配率,但同时也会有更长的时间消耗。

(3) 当设置模板尺寸 91×91 像素时,SOFM、CFOG、HOPC、DLSC、MI、NCC 等方法的时间消耗分别为 2.51、1.31、11.18、30.17、29.44 和 14.99 s,通过数据对比可以看出,除 CFOG 比 SOFM 方法快 1.2 s 之外,NCC 方法的时间消耗约是 SOFM 方法的 6 倍,最慢的 DLSC 方法的时间消耗约是 SOFM 方法的 12 倍。因此,相较之下,SOFM 方法在保证较高的正确匹配率的前提下,达到了快速匹配的效果。

图 11 展示了当模板尺寸为 91×91 像素时各方法在 4 组数据上的 RMSE。由图 11 可知 SOFM 方法在各组试验影像上均具有最小的

RMSE,这说明相较于其余几种方法,SOFM 方法具有最高的匹配精度,能够在进行 SAR 影像和光学影像的匹配时更具优势。

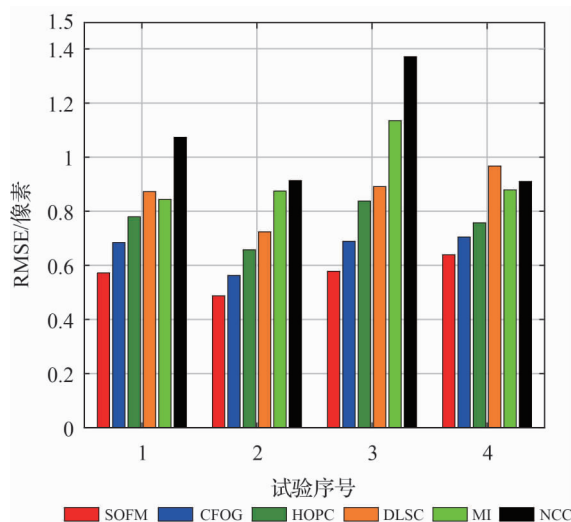


图 11 SOFM、CFOG、DLSC、HOPC、MI 和 NCC 在各试验影像上的 RMSE

Fig.11 The RMSEs of correct matches of SOFM, CFOG, HOPC, DLSC, MI and NCC

图 12 展示了 SOFM 方法在各组试验影像上的匹配结果(模板尺寸 61×61 像素),可以看出匹配到的同名点数量充足且分布均匀,观察局部放大图能够看到同名点的对应位置关系是正确的。

通过以上的各组试验和对比分析,证明了本文提出的 SOFM 方法在正确匹配率、匹配稳定性、计算效率和匹配精度等方面均更具优势,能够快速、高效地完成 SAR 影像与光学影像之间的自动匹配任务。

4 讨论

针对当前遥感领域较为困难的 SAR 影像和光学影像间的匹配问题,本文提出了一种基于影像结构特性的快速匹配方法 SOFM。该方法利用影像的梯度信息构建了一种特征描述符—角度加权方向梯度 AWOG,然后利用描述符之间的差值的平方和建立了用于模板匹配的相似性测度,并推导出了利用快速傅里叶变换在频率域中表达的影像匹配函数,在频率域中进行影像匹配能够大幅提高计算效率,随后给出了使用本文方法的模板匹配流程。通过与现有的影像匹配方法 CFOG、HOPC、DLSC、MI 和 NCC 在多个方面进行对比分析,结果显示:SOFM 方法能够较好地

抵抗 SAR 影像与光学影像之间的非线性灰度差异,其匹配性能良好且更具稳定性,同时在计算效率和匹配精度方面也有着优异的表现。SOFM

方法能够快速地完成 SAR 影像与光学影像之间的自动匹配任务,获得精度高且分布均匀的同名匹配点,为后续的相关应用提供了可靠保障。

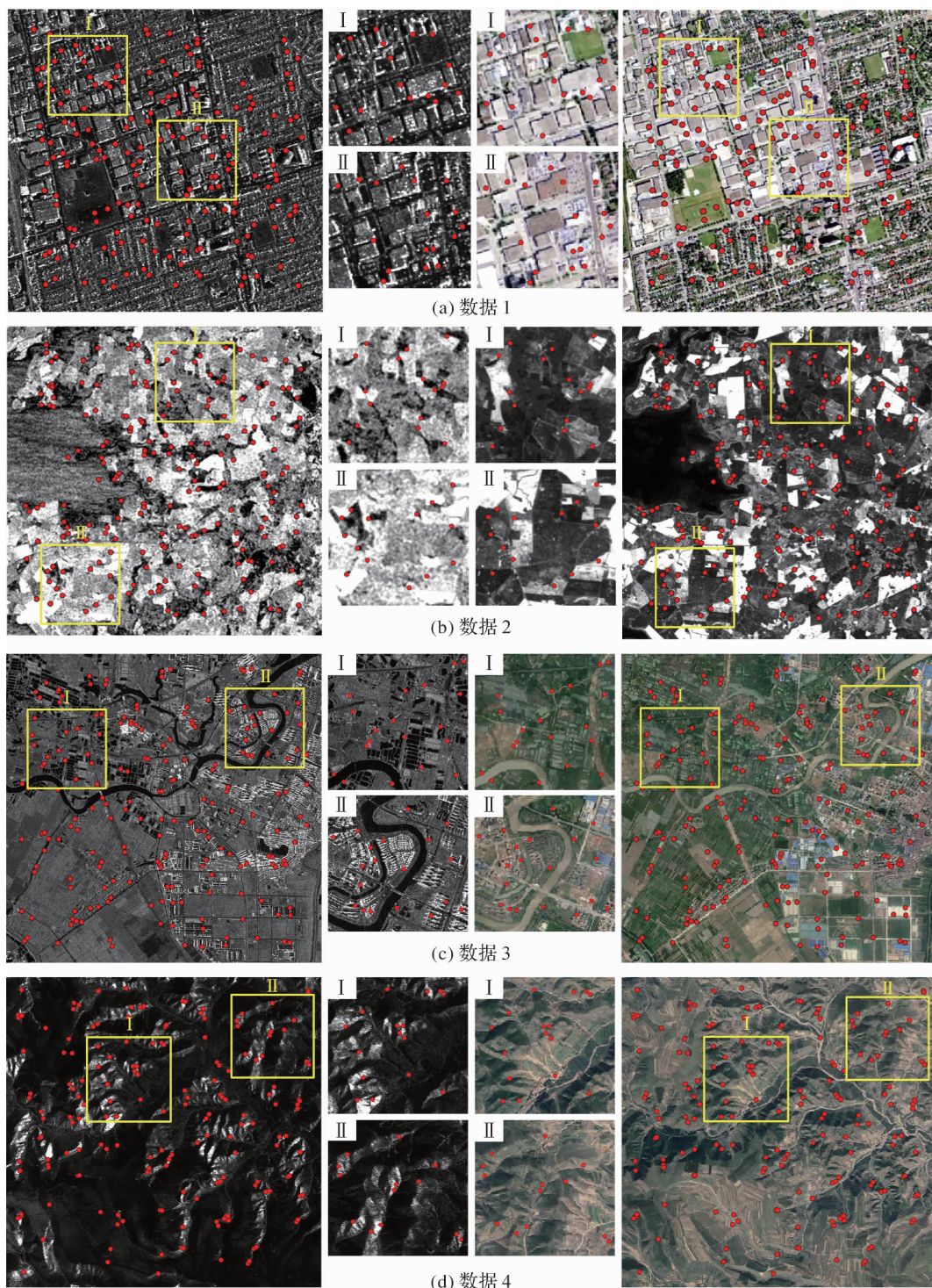


图 12 SOFM 的匹配结果

Fig.12 The image matching results of SOFM

考虑到 SAR 影像生成的正射影像图(DOM)的平面精度基本只与传感器的定轨精度有关,与

瞬时成像姿态角几乎没有相关性,而目前卫星定轨精度已经达到足够高的水平,因此 SAR 影像的

DOM能够保证较高的平面精度^[34]。笔者下一步的工作将使用由天绘二号卫星数据生成的正射影像图(DOM)及数字高程模型(DEM)作为控制信息,将光学原始影像与之进行匹配,匹配得到的同名点将作为控制点来精化光学原始影像的RPC参数,从而可形成一套完全基于国产SAR卫星数据和光学卫星数据的无控测图方法流程,以期在无法实地测得控制点的境外地区实现高精度的无控测图。

试验过程中也发现本文方法目前存在一定不足:①SOFM方法中的AWOG描述符不具有旋转和尺度不变性,但可利用影像的定向参数基本消除影像间的旋转和尺度差异(见第2节)。试验结果表明,对于地形起伏较小的区域如平地、丘陵地甚至是山体高度变化不大的山区,本文方法都能得到较好的匹配结果,但是对于高程变化剧烈的山区如我国横断山区,方法便有可能失效,因此尚需要进一步研究能够抵抗局部影像间旋转和尺度差异的方法,以更好地解决复杂地形区域的SAR影像与光学影像之间的匹配。②SOFM方法依赖于影像中蕴含的结构特征进行匹配,如果影像的结构特征不够丰富,如影像中含有大面积的森林区域,则方法得到的有效特征点数量会有一定程度的减少,同时其分布情况也会在一定程度上变差,因此,在进一步的研究中将考虑引入纹理特征来辅助进行匹配。

参考文献:

- [1] WANG Yafei, ZHAO Feifei, CHENG Liang, et al. Framework for monitoring the conversion of cultivated land to construction land using SAR image time series[J]. *Remote Sensing Letters*, 2015, 6(10): 794-803.
- [2] STROZZI T, U WEGMÜLLER, TOSI L, et al. Land subsidence monitoring with differential SAR interferometry [J]. *Photogrammetric Engineering & Remote Sensing*, 2001, 67(11): 1261-1270.
- [3] 叶沅鑫, 慎利, 陈敏, 等. 局部相位特征描述的多源遥感影像自动匹配[J]. *武汉大学学报(信息科学版)*, 2017, 42(9): 1278-1284.
YE Yuanxin, SHEN Li, CHEN Min, et al. An automatic matching method based on local phase feature descriptor for multi-source remote sensing images[J]. *Geomatics and Information Science of Wuhan University*, 2017, 42(9): 1278-1284.
- [4] 陈敏, 朱庆, 朱军, 等. SAR影像与光学影像的高斯伽玛型边缘强度特征匹配法[J]. *测绘学报*, 2016, 45(3): 318-325.
CHEN Min, ZHU Qing, ZHU Jun, et al. Feature matching for SAR and optical images based on Gaussian-gamma-shaped edge strength map[J]. *Acta Geodaetica et Cartographica Sinica*, 2016, 45(3): 318-325.
- [5] 闫利, 王紫琦, 叶志云. 顾及灰度和梯度信息的多模态影像配准算法[J]. *测绘学报*, 2018, 47(1): 71-81.
YAN Li, WANG Ziqi, YE Zhiyun. Multimodal image registration algorithm considering grayscale and gradient information[J]. *Acta Geodaetica et Cartographica Sinica*, 2018, 47(1): 71-81.
- [6] 岳春宇, 江万寿. 几何约束和改进SIFT的SAR影像和光学影像自动配准方法[J]. *测绘学报*, 2012, 41(4): 570-576.
YUE Chunyu, JIANG Wanshou. An automatic registration algorithm for SAR and optical images based on geometry constraint and improved SIFT[J]. *Acta Geodaetica et Cartographica Sinica*, 2012, 41(4): 570-576.
- [7] 李雨谦, 皮亦鸣, 王金峰. 基于水平集的SAR图像与光学图像的配准[J]. *测绘学报*, 2010, 39(3): 276-282.
LI Yuqian, PI Yiming, WANG Jinfeng. The registration between SAR and optical image based on level set[J]. *Acta Geodaetica et Cartographica Sinica*, 2010, 39(3): 276-282.
- [8] ZITOVÁ B, FLUSSER J. Image registration methods: a survey[J]. *Image and Vision Computing*, 2003, 21(11): 977-1000.
- [9] ZHANG Jun, MA Wenping, WU Yue, et al. Multimodal remote sensing image registration based on image transfer and local features [J]. *IEEE Geoscience and Remote Sensing Letters*, 2019, 16(8): 1210-1214.
- [10] 南轲, 齐华, 叶沅鑫. 深度卷积特征表达的多模态遥感影像模板匹配方法[J]. *测绘学报*, 2019, 48(6): 727-736.
NAN Ke, QI Hua, YE Yuanxin. A template matching method of multimodal remote sensing images based on deep convolutional feature representation[J]. *Acta Geodaetica et Cartographica Sinica*, 2019, 48(6): 727-736.
- [11] YU Le, ZHANG Dengrong, HOLDEN E J. A fast and fully automatic registration approach based on point features for multi-source remote-sensing images[J]. *Computers & Geosciences*, 2008, 34(7): 838-848.
- [12] CHEN Min, SHAO Zhenfeng. Robust affine-invariant line matching for high resolution remote sensing images[J]. *Photogrammetric Engineering & Remote Sensing*, 2013, 79(8): 753-760.
- [13] GONCALVES H, CORTE-REAL L, GONCALVES J A. Automatic image registration through image segmentation and SIFT[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2011, 49(7): 2589-2600.
- [14] LOWE D G. Object recognition from local scale-invariant features[C]// *Proceedings of the 7th IEEE International Conference on Computer Vision*. Kerkyra, Greece: IEEE, 1999: 1150-1157.
- [15] BAY H, TUYTELAARS T, VAN GOOL L. SURF: speeded

- up robust features[C]//Proceedings of the European conference on computer vision. Berlin, Heidelberg:Springer, 2006: 404-417.
- [16] 叶沉鑫, 郝思媛, 曹云刚. 基于几何结构属性的光学和 SAR 影像自动配准[J]. 红外与毫米波学报, 2017, 36(6): 720-726.
YE Yuanxin, HAO Siyuan, CAO Yungang. Automatic registration of optical and SAR image using geometric structural properties [J]. Journal of Infrared and Millimeter Waves, 2017, 36(6): 720-726.
- [17] COLE-RHODES A A, JOHNSON K L, LEMOIGNE J, et al. Multiresolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient[J]. IEEE Transactions on Image Processing, 2003, 12(12): 1495-1511.
- [18] MARTINEZ A, GARCIA-CONSUEGRA J, ABAD F. A correlation-symbolic approach to automatic remotely sensed image rectification[C]//Proceedings of 1999 IEEE International Geoscience and Remote Sensing Symposium. Hamburg, Germany:IEEE, 1999: 336-338.
- [19] BROMLEY J, BENTZ J W, BOTTOU L, et al. Signature verification using a “Siamese” time delay neural network [J]. Advances in neural information processing systems, 1994(1): 737-737.
- [20] BERTINETTO L, VALMADRE J, HENRIQUES J F, et al. Fully-convolutional siamese networks for object tracking[C]//Proceedings of the European Conference on Computer Vision. Cham:Springer, 2016.
- [21] SHECHTMAN E, IRANI M. Matching local self-similarities across images and videos[C]// Proceedings of 2007 IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, MN, USA:IEEE, 2007: 1-8.
- [22] 王峰, 尤红建. 基于级联变换的光学和 SAR 图像配准算法[J]. 红外与毫米波学报, 2015, 34(4): 486-492.
WANG Feng, YOU Hongjian. Robust registration method of SAR and optical remote sensing Images based on cascade transforms[J]. Journal of Infrared and Millimeter Waves, 2015, 34(4): 486-492.
- [23] YE Yuanxin, SHEN Li. Hopc: a novel similarity metric based on geometric structural properties for multi-modal remote sensing image matching[J]. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2016, III-1: 9-16.
- [24] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]//Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, CA, USA:IEEE, 2005: 886-893.
- [25] YE Yuanxin, BRUZZONE L, SHAN Jie, et al. Fast and robust matching for multimodal remote sensing image registration [J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(11): 9059-9070.
- [26] FELZENSZWALB P F, GIRSHICK R B, MCALLESTER D, et al. Object detection with discriminatively trained part-based models[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(9): 1627-1645.
- [27] HARZALLAH H, JURIE F, SCHMID C. Combining efficient object localization and image classification [C] // Proceedings of the 12th IEEE International Conference on Computer Vision. Kyoto, Japan:IEEE, 2009: 237-244.
- [28] SHRIVASTAVA A, MALISIEWICZ T, GUPTA A, et al. Data-driven visual similarity for cross-domain image matching[C]//Proceedings of the 2011 SIGGRAPH Asia Conference. Hong Kong, China: ACM Press, 2011: 1-10.
- [29] SEDAGHAT A, EBADI H. Distinctive order based self-similarity descriptor for multi-sensor remote sensing image matching[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2015, 108: 62-71.
- [30] LU Jiazhen, SUN Fangde, DONG Jing. A novel multi-sensor image matching algorithm based on adaptive multi-scale structure orientation [J]. IEEE Access, 2019, 7: 177474-177483.
- [31] REDDY B S, CHATTERJI B N. An FFT-based technique for translation, rotation, and scale-invariant image registration[J]. IEEE Trans Image Process, 1996, 5(8): 1266-1271.
- [32] YE Yuanxin, BRUZZONE L, SHAN Jie, et al. Fast and robust structure-based multimodal geospatial image matching[C]//Proceedings of 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Fort Worth, TX, USA:IEEE, 2017: 5141-5144.
- [33] YE Yuanxin, SHEN Li, HAO Ming, et al. Robust optical-to-SAR image matching based on shape properties [J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14(4): 564-568.
- [34] 楼良盛, 刘志铭, 张昊, 等. 天绘二号卫星工程设计与实现[J]. 测绘学报, 2020, 49(10): 1252-1264.
LOU Liangsheng, LIU Zhiming, ZHANG Hao, et al. TH-2 satellite engineering design and implementation[J]. Acta Geodaetica et Cartographica Sinica, 2020, 49(10): 1252-1264.

(责任编辑:张艳玲)

收稿日期: 2020-12-08

修回日期: 2021-07-21

第一作者简介: 樊仲黎(1997—),男,硕士生,研究方向为星载多源传感器数据高精度协同处理。

First author: FAN Zhongli (1997—), male, postgraduate, majors in high-precision collaborative processing of satellite-borne multi-sensor data.

E-mail: fzl10466@163.com

通信作者: 王庆栋

Corresponding author: WANG Qingdong

E-mail: wangqd@casm.ac.cn