

引文格式: 蓝朝桢, 卢万杰, 于君明, 等. 异源遥感影像特征匹配的深度学习算法[J]. 测绘学报, 2021, 50(2): 189-202. DOI: 10.11947/j. AGCS.2021.20200048.  
LAN Chaozhen, LU Wanjie, YU Junming, et al. Deep learning algorithm for feature matching of cross modality remote sensing images[J]. Acta Geodaetica et Cartographica Sinica, 2021, 50(2): 189-202. DOI: 10.11947/j. AGCS.2021.20200048.

## 异源遥感影像特征匹配的深度学习算法

蓝朝桢<sup>1</sup>, 卢万杰<sup>1</sup>, 于君明<sup>2</sup>, 徐 青<sup>1</sup>

1. 信息工程大学地理空间信息学院, 河南 郑州 450001; 2. 中国电子科技集团公司第二十七研究所, 河南 郑州 450047

## Deep learning algorithm for feature matching of cross modality remote sensing images

LAN Chaozhen<sup>1</sup>, LU Wanjie<sup>1</sup>, YU Junming<sup>2</sup>, XU Qing<sup>1</sup>

1. Institute of Geospatial Information, Information Engineering University, Zhengzhou 450001, China; 2. China Electronic Technology Group Corporation 27th Research Institute, Zhengzhou 450047, China

**Abstract:** Focusing on the problem of difficulty in matching due to the differences in imaging modality, time phases, and resolutions of cross modality remote sensing images, a new deep learning feature matching method named CMM-Net is proposed. First, a convolutional neural network is used to extract high-dimensional feature maps of the cross modality remote sensing images. The key points are selected according to the conditions that both the channel maximum and local maximum are met, and the 512-dimensional descriptors in corresponding location are extracted on the feature map to complete the feature extraction. In the matching stage, after completing the fast-nearest neighbor searching, in order to solve the problem of lots of mismatched points, a purification algorithm with dynamic adaptive Euclidean distance and RANSAC constraints is proposed to ensure that the mismatches are effectively eliminated while retaining the correct matches. The algorithm was tested using multiple sets of cross modality remote sensing images and compared with other algorithms. The results show that the proposed algorithm has the ability to extract similar scale invariant features in cross modality images, and has strong adaptability and robustness.

**Key words:** deep learning; image matching; cross modality image; convolution neural network; satellite image

**Foundation support:** The National Key Research and Development Project (No. 2017YFC1200305); The National Natural Science Foundation of China (Nos. 41876105; 41371436)

**摘 要:** 针对异源遥感影像的成像模式、时相、分辨率等不同导致匹配困难的问题, 提出了一种基于深度学习特征的匹配方法 CMM-Net。首先, 利用卷积神经网络提取异源遥感影像的高维特征图, 根据同时满足通道最大和局部最大两种条件选取关键点, 并在特征图上提取相应位置的 512 维描述符。在匹配阶段, 完成快速最近邻搜索特征匹配后, 为解决误匹配点多的问题, 提出了动态自适应欧氏距离阈值和 RANSAC 共同约束的提纯算法, 保证误匹配有效剔除的同时, 最大限度保留正确匹配点。利用多组异源遥感影像对算法进行了测试, 并与多种异源影像匹配算法进行了比较, 结果表明本文算法能够提取出异源影像的尺度不变相似特征, 具有较强的适应性和稳健性。

**关键词:** 深度学习; 影像匹配; 异源影像; 卷积神经网络; 卫星影像

**中图分类号:** P236      **文献标识码:** A      **文章编号:** 1001-1595(2021)02-0189-14

**基金项目:** 国家重点研发计划(2017YFC1200305); 国家自然科学基金(41876105; 41371436)

随着遥感技术的飞速发展, 可见光、红外、合成孔径雷达(synthetic aperture radar, SAR)等多

类型传感器对地观测影像<sup>[1]</sup>日趋丰富。不同平台和传感器获取的异源影像之间具有一定互补性,为遥感信息的深入挖掘、大数据分析提供了海量的数据来源<sup>[2]</sup>。影像之间的匹配是异源影像进一步处理及分析的核心问题。由于成像机理、波段、时相等不同,异源影像在辐射特征、几何特征上都存在巨大差异,它们之间的匹配一直都是影像匹配的研究难点。

国内外学者针对这一问题提出了多种匹配方法。主要包括两大类,一类是基于特征的匹配方法,另一类是基于模板匹配的方法。图像特征匹配通常提取关键点周围一定邻域内的局部特征信息描述符,通过比较描述符来确定匹配点,其中最著名的是尺度不变特征转换(scale invariant feature transformation, SIFT)描述符<sup>[3]</sup>。SIFT描述符能够很好地抵抗影像间的旋转和尺度差异,但由于基于影像局部邻域的梯度分布描述关键点,对异源影像的匹配效果较差。因此,众多学者试图通过改进 SIFT 算法或者结合其他约束信息开展异源影像匹配研究<sup>[4-8]</sup>。但是当异源影像间的辐射和几何差异较大时,特别是存在较大的非线性误差时,由于较小的邻域强度和梯度信息不能提供稳定的特征<sup>[9]</sup>,因此很难获得良好的效果。

与 SIFT 及其改进算法相比,基于模板的匹配可以获得更大影像范围内的不变量。其中,在异源影像匹配方面比较成功的相似性测度包括互信息(mutual information, MI)以及相位一致性<sup>[9-13]</sup>。此外,文献[14]提出了一种顾及灰度和梯度信息的多模态影像配准算法,文献[15]提出了基于相位一致性和最大索引图的辐射不变特征匹配方法 RIFT,都取得很好的匹配效果。

以上算法都依赖人工设计的描述符,当面对的遥感影像在来源、模式和波段等方面变化越来越多的情况下,其扩展和表达能力有限,基于模板的匹配效率也相对较低。近几年,深度学习特别是卷积神经网络(convolutional neural network, CNN)在计算机视觉任务如图像分类、目标检测和分割方面取得了巨大的进步和性能提升。自 2014 年文献[16]首次引入以来,学者们就开始将 CNN 应用于图像特征提取过程,并逐渐从 SIFT 特征向 CNN 特征转变<sup>[16-22]</sup>。传统的人工设计的描述子只能提取和表示影像相对低层的特征,而 CNN 通常被认为能够提取更高层、更抽

象的语义特征。利用高层的语义信息进行匹配,有着很强的泛化性,更接近人类视觉观察原理,理论上应该更能够抵抗由于波段、成像模式、季节变换等带来的干扰,有望在匹配适应性方面获得较大提升。

早期提出的 CNN 特征提取方法大多使用全连接层展平提取整幅图像或者图像块的特征向量,通常用于图像分类或者识别,不进行关键点特征的匹配。2016 年以后,学习型关键点特征检测和描述算法得到了迅速发展,相继提出了 LIFT<sup>[23]</sup>、SuperPoint<sup>[24]</sup>、DELFT<sup>[25]</sup>、D2-Net<sup>[26]</sup>等。其中,D2-Net 利用 30 多万个预匹配的立体像对进行训练,在解决变化场景下影像匹配方面取得了重要进展,且表现出巨大的潜力。但这些算法模型提出的主要目的是用于光照和视角变化的地面近景可见光影像,多用于地面车辆视觉导航等。文献[27]专门针对跨模态异源的影像匹配问题,提出了用跨模态上下文增强的局部描述子,取得了积极进展。

本文试图在引入 D2-Net 特征提取基本思想的基础上,提出一种基于深度学习特征的异源遥感影像匹配算法(cross modality matching net, CMM-Net),希望能够通过 CNN 提取的高层语义局部特征,自动学习和寻找异源影像同名点之间几何和辐射不变量,形成一种稳健的端对端的异源遥感影像匹配方法。

## 1 异源遥感影像不变局部特征提取与匹配原理

### 1.1 算法基本思想

要实现异源遥感影像稳健的特征匹配,核心问题在于如何减小异源影像辐射和几何差异带来的影响,找到不变特征表示方法。为达到这一目的,本文提出的 CMM-Net 着重在以下 3 个方面进行考虑。①构建一个适合特征关键点提取和描述的 CNN 网络和提取算法。来自 CNN 较深层的特征图可以看作是更高层的语义信息,高层抽象的语义信息要比低层梯度信息更能够适应辐射和几何上的变化;适当扩大提取特征对应的原始输入影像范围(感受域),有利于异源影像不变特征的提取。②利用已经配对好的光照和拍摄角度都存在较大差异的数据对 CNN 网络进行训练,让 CNN 特征提取器能学习到光照、几何等变化影像的不变性特征。③采取“以多求可靠”的策

略。异源影像差异大,即使考虑到以上两点,对于模式不同的影像(如 SAR 和光学影像)进行匹配还是存在很大困难。因此,本文考虑“以多求可靠”的策略,让提取的候选特征具有较多的数量,通过提升匹配过程的筛选机制进行有效约束,以获得更多可靠、更均匀的匹配对。

此外,特征定位也是特征匹配过程中不可忽视的问题。通常 CNN 经过卷积抽象后,输出特征图分辨率会成倍下降。CNN 深度越深,特征的不变性表达能力越强,但定位精度越差,特征的抽象和精确定位在 CNN 中是一对矛盾。本文在特征提取过程中,选择抽取 CNN 网络的中间层,并通过改变池化步长方法来保证网络最后一次池化输出分辨率不降低,同时,与 SIFT 等传统算法一样,采用关键点内插的方式,获得更高的定位精度。

### 1.2 异源遥感影像深度学习不变特征提取

#### 1.2.1 网络模型

为使 CNN 网络适合特征提取,CMM-Net 选取经典 VGG16 网络模型<sup>[28]</sup>并对其进行适应性改造。经典的 VGG16 模型共 5 个卷积网络,主要用于分类任务。通常,网络的前几层感受域很小,得到的特征是相对底层的边缘、角点等局部特征,但定位精度较高;网络层数越深,提取的特征越抽象,信息越全局,越能抵抗异源影像带来的干扰,但定位精度越差。因此,为了能够使特征点既有足够抽象性、也可以获得较高的定位精度,本文算法丢弃 VGG16 的最后一个卷积层,选取中间第 4 层中的最后一个(第 3 个)卷积层(Conv4\_3)输出作为关键点提取的特征图,网络结构如图 1 所示。特征图是原始影像经过 CNN 网络多层卷积、池化后得到的输出,其维度和同一层卷积核数量相同。选取 Conv4\_3 的输出作为特征图,既具有较深的 CNN 不变性表达能力,又能保留一定的分辨率。

卷积神经网络每一层经过池化后,分辨率一般都会下降,为了保持特征图的分辨率,本文将最后一个(第 3 个)池化层窗口滑动步长从 2 像素,替换为 1 像素,并且池化方法也由最大池化,替换为平均池化。第 4 层 3 个卷积(Conv4\_1 至 Conv4\_3)采用空洞卷积率(dilation)为 2 的空洞卷积,这样可以扩大感受域,提高特征表达泛化能力,有利于异源图像特征的不变性表达。经过改造后,相对经典的 VGG16 网络,输出的特征图从

原始图像的 1/8 扩大到 1/4,定位精度提高一倍<sup>[26]</sup>。

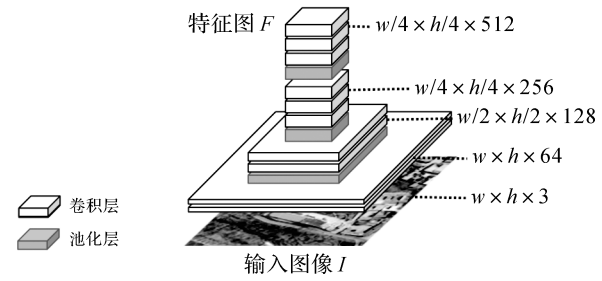


图 1 网络结构

Fig.1 The neural network structure

#### 1.2.2 深度和平面局部极值点选择与特征描述

利用上节设计的网络结构,如果直接使用 Conv4\_3 输出的特征图全部像素作为特征,特征将过于密集,且大部分特征不够显著。因此需要从特征图中选择一些特征明显的关键点用于异源影像匹配,这一步称为特征筛选。

假设输入原始影像为  $I$ , 尺寸为  $w \times h$ , 设网络输出特征图为 3D 张量  $F = \mathcal{F}(I)$ ,  $F \in \mathbb{R}^{w \times h \times n}$ , 其通道数  $n = 512$ 。特征关键点筛选和描述符提取即在  $F$  中进行。

为了在  $\mathbb{R}^{w \times h \times n}$  特征空间筛选出较为显著的特征点,CMM-Net 采用了在高维特征图的通道方向和局部平面内同时最大的筛选策略<sup>[26]</sup>, 公式为

$$k = \arg \max_t D_{ij}^k \quad (D_{ij}^k \text{ 为 } D^k \text{ 中局部最大}) \quad (1)$$

式中,  $D^k$  为第  $k$  层特征值, 且  $D^k \in \mathbb{R}^{w \times h}$ ;  $D_{ij}^k$  为特征图上点位  $(i, j)$  处的特征值。对一个待选点  $P(i, j)$ , 首先进行通道选择, 即从  $n$  个通道特征图中选出响应值最大的通道  $k$ , 得到相应通道上的特征图为  $D^k$ ; 接着验证  $D_{ij}^k$  是否为局部最大。如果满足两个条件, 说明该待选点  $P$  为显著点, 即选为特征点。为了得到足够数量的特征点, 平面极值在  $3 \times 3$  邻域内选取。

同时, 在特征图  $F$  上提取  $(i, j)$  处的 512 维通道向量, 并进行 L2 范式归一化, 得到特征描述符  $\hat{d}_{ij}$ <sup>[26]</sup>

$$\hat{d}_{ij} = \mathbf{d}_{ij} / \|\mathbf{d}_{ij}\|_2 \quad (2)$$

式中,  $\mathbf{d}_{ij} = F_{ij}$ ,  $\mathbf{d} \in \mathbb{R}^n$ 。但离散空间的极值点并不是真正的极值点, 为获得更精确的关键点的位置, 借鉴 SIFT 算法, 采用特征图局部插值加密的方法获得子像素级定位精度, 同时描述子也在邻

域内用双线性插值获得。最终插值得到的  $\hat{d}_{ij}$  是一个  $n$  维向量,利用该向量,即可根据欧氏空间距离进行匹配。

### 1.2.3 特征离散金字塔多尺度模型

虽然 CNN 模型使用尺度上有差异的训练样本进行训练,特征描述符能够学习一定程度的尺度不变性,但也难以应对尺度变化较大的情况。因此,CMM-Net 采取了离散影像金字塔模型来应对尺度的较大变化。对输入图像  $I$ ,金字塔影像  $I^\rho$ ,本文采用 0.25、0.5、1.0、2.0 倍分辨率(即  $\rho=0.25,0.5,1,2$ )4 个离散尺度层来适应两幅图像分辨率的剧烈变化。与通过有一定尺度差异的训练数据学习到的 CNN 模型一起构成算法对输入影像连续且较大尺度变化的适应性。

金字塔的每一层分别提取 CNN 特征  $F^\rho$ ,然后利用式(3)进行累加融合<sup>[26]</sup>。关键点特征描述将通过累加得到的融合特征图  $\tilde{F}^\rho$  进行提取

$$\tilde{F}^\rho = F^\rho + \sum_{\gamma < \rho} F^\gamma \quad (3)$$

由于金字塔分辨率不同,低分辨率特征图需要线性内插成与高分辨率特征图相同大小才能进行累加。此外,为了防止不同层级检测特征的重复,CMM-Net 从最粗的尺度开始,标记检测到的位置;这些位置将被上采样到高一个尺度的特征图中作为模板,如果高一级分辨率特征图提取的关键点落入该模板中就会被舍弃。

## 1.3 损失函数与模型训练

ImageNet 预训练的 VGG16 网络模型不具备对异源影像相似性的表达能力。为让 CNN 网络模型能够在辐射与几何差异较大的异源遥感影像上学习共同特征的表达,需要重新设计损失函数并利用合理的训练数据进行模型微调。

### 1.3.1 三元组距离排序损失函数

CMM-Net 采用了三元组距离排序函数(triplet margin ranking loss, TMRL)<sup>[29]</sup>作为损失函数。因为在特征检测过程中,希望特征点具有一定通用性,以适应不同环境光照辐射和几何差异的影响;但同时,在特征描述过程中,又希望特征向量尽可能具有独特性,以便寻找同名像点。针对这个问题,三元组距离排序损失函数通过惩罚导致错误匹配的任何不相关描述符来增强相关描述符的独特性。此外,为寻求检测特征的可重复性,将检测项添加到损失函数中<sup>[26]</sup>,公式为

$$\mathcal{L}(I_1, I_2) = \sum_{c \in C} \frac{s_c^{(1)} s_c^{(2)}}{\sum_{q \in C} s_q^{(1)} s_q^{(2)}} m(p(c), n(c)) \quad (4)$$

式中, $p(c)$ 和 $n(c)$ 分别代表同名像点的正对距离和负对距离; $s_c^{(1)}$ 和 $s_c^{(2)}$ 为在影像 $I_1$ 和 $I_2$ 上 $A$ 、 $B$ 两个点处得到的特征检测得分,具体计算方法可参照文献[26]; $C$ 是影像 $I_1$ 和 $I_2$ 上所有同名像点的集合。

上述损失函数会基于所有匹配项的检测得分来生成距离因子 $m$ 的加权平均值。因此,为了使损失最小化,具有较低的距离因子的最相关的对应关系,将获得较高的相对得分,并让具有较高相对得分的对应关系获得与其余特征不同的相似描述符,提高匹配的稳健性。

### 1.3.2 训练数据

利用式(4),为了让 CNN 模型能够学习到辐射和几何差异下像素级特征相似性的表达,训练数据除了数量要足够外,还必须同时具备以下两个条件:①训练影像具备较大的辐射和几何差异;②训练影像具备像素级的对应关系。文献[30]构建了一种 MegaDepth 数据集。该数据集提出的初衷是进行单张影像的深度恢复,由 100 多万幅从互联网获得的地标影像组成,这些影像拍摄光照、尺度差异都比较大,还包含了大量昼夜影像对,如图 2 所示。此外,该数据集还从这些影像中筛选出约 10 万幅优质影像,使用开源运动恢复结构软件 COLMAP<sup>[31]</sup>构建了近 200 个不同全球地标的三维场景,包括了相机内/外部参数、场景结构信息及深度图。从这些三维场景中,可以获得立体影像对,并利用场景三维信息和相机参数,第 2 幅影像上的像点可以投影到第 1 幅中,建立像对像素级的对应关系。由此可见,MegaDepth 数据集可以满足上述两个条件。因此,本文算法选用该数据集进行模型训练。



图 2 训练数据示例

Fig.2 Examples of training data

### 1.3.3 模型训练方法

CMM-Net 骨干网络采用了基于海量 ImageNet 数据集进行预训练的 VGG16 模型,对网络模型中最后一个密集特征提取器 Conv4\_3 采用迁移学习微调训练的方法训练网络模型。初始学习率设置为  $10^{-3}$ ,然后每 10 个 Epoch 减小一半,对于每对同名像点,分别选择以同名像点为中心的  $256 \times 256$  像素的随机影像区域送入到网络进行训练。通过训练可以使网络能从较大辐射和几何差异的异源遥感影像上获得相似特征的良好性能。

## 1.4 异源遥感影像深度学习不变特征匹配方法

### 1.4.1 基本思想与流程

CMM-Net 匹配算法主要思想为:将 1.2 节算法作为特征提取器,在原始影像上提取深度特征。特征匹配方法采用快速最近邻搜索(FLANN)方法。由于异源影像差异较大,提取的候选特征点数量较多,不可避免地会存在大量误匹配的情况,本文提出了动态自适应距离约束条件和随机采样一致性(random sample consensus,RANSAC)约束相结合的误匹配点剔除方法,实现“以多求可靠”策略中有效筛选这关键一环。算法流程如图 3 所示。

### 1.4.2 动态自适应阈值匹配点提纯

采用 RANSAC 或者类似几何约束算法是行之有效的匹配点提纯方法,但如果待提纯的匹配点中含有大量差异较大的误匹配,RANSAC 的随机取值迭代方法受这些错误匹配点的影响,算法变得极不稳定。因此,在几何约束之前,需要进行匹配点对的粗提纯。通过 FLANN 算法搜索出来待筛选的匹配对中,包含欧氏距离最近的第 1 匹配点和次近的第 2 匹配点。通常认为,对于待筛选的匹配对中的第  $j$  个匹配对,第 1 匹配点距离  $dis_j$  比第 2 匹配点距离  $dis'_j$  越小,说明匹配质量越好。传统的算法都采取固定比例因子  $t$  作为阈值,即当满足  $dis_j < t \cdot dis'_j$  时,选入候选匹配对。但由于不同传感器和时相的影像差距较大,CNN 特征进行欧氏空间距离搜索时,欧氏距离差分布范围很难预料。因此,每一对影像通常需要不断手动调整阈值  $t$  到合适数值,才能筛选出优质的匹配对。为解决这一问题,提高算法的适应性,本文设计了动态自适应欧氏距离约束方法。该方法针对待提纯的匹配对数据进行统计,根据数据特点自动的配置相应参数。首先,从

FLANN 搜索出来的  $N$  个包含大量误匹配的匹配对中统计第 1 匹配点和第 2 匹配点距离差的均值

$$avgdis = \left( \sum_{j=1}^N dis'_j - dis_j \right) / N \quad (5)$$

对于每一个待筛选匹配对,剔除的条件为第 1 距离小于第 2 距离与距离差均值 avgdis 之差,公式为

$$dis_j < dis'_j - avgdis \quad (6)$$

算法通过从数据中统计得到的距离差均值,作为判别比较标准,可以很好地适应不同来源的像对之间的差异,能够很好地进行第 1 轮筛选,保留优质匹配点,提高 RANSAC 输出的稳定性。

RANSAC 中几何约束关系要根据待匹配影像的成像几何关系选择合适的模型。在实际工程应用中,应尽可能采用严格的约束模型。如画幅式面阵影像通常可用单应矩阵、本质矩阵等约束模型;线阵卫星影像多采用基于 PRC 模型或者多项式模型的核心线约束。本文试验中,由于选取的试验遥感影像摄影距离较远、地面相对高差不大、选取的影像区域也较小,故采用仿射变换模型,以适应不同成像模型的像对之间的缩放、平移、旋转和错切等变换。

## 2 试验与分析

本节首先对 CMM-Net 算法提取的异源影像同名特征的相似性进行验证,并分别针对尺度和方向上的性能进行试验;然后,为评价 CMM-Net 的性能,与多个异源影像匹配算法进行对比试验;最后,利用实际异源遥感影像进行配准试验,考查在影像配准上的实际应用效果。试验过程中,CMM-Net 深度学习模型在 PyTorch 框架下实现,试验用计算机为华硕 ROG 笔记本,CPU 为 i7-9750H,显卡为 GeForce RTX 2060 (6 GB 显存),内存为 32 GB;实现语言为 Python,操作系统为 Ubuntu 18.04。

### 2.1 特征相似性试验

提取的多源影像上的同名特征是否具有相似性,是异源影像匹配的关键。本文选取了高分三号 SAR 卫星、资源三号光学卫星、谷歌地球以及无人机等几种来源的异源影像,见图 4。测试数据中成像模式不同、时相不同、分辨率也不同,具有代表性。几幅图像的中心位置经过人工配准,试验通过该中心提取的特征向量的相似度进

行评估。

试验分别利用 SIFT 算法与本文方法提取了 4 幅试验图像中心同名像点的特征向量,统一进行 L2 范式归一化,并以曲线的形式绘制,横坐标为特征向量维度,纵坐标为归一化的特征值。这种特征曲线走向可以反映异源影像特征之间的相似程度。为显示清晰,随机截取了其中一段范围,如图 5 所示。

对比图 5 直观展现出的 4 个异源图像同名像点的特征向量的走向。很明显,本文算法相对于 SIFT 算法明显具有比较高的相似性,在很多维度坐标(横轴)上,曲线的波峰波谷变化规律接近。特别需要指出的是,差异最大的 SAR 图像(红色实线)和其他几个特征曲线走向相似度也很高。为了量化特征向量直接的差异,统计了特征向量之间的余弦距离如表 1 所示。由于 SIFT 与本文算法提取的特征维度和量化标准不同,相对于欧氏距离,代表向量之间角度差异的余弦距离更能反映不同维度向量之间的相似程度。

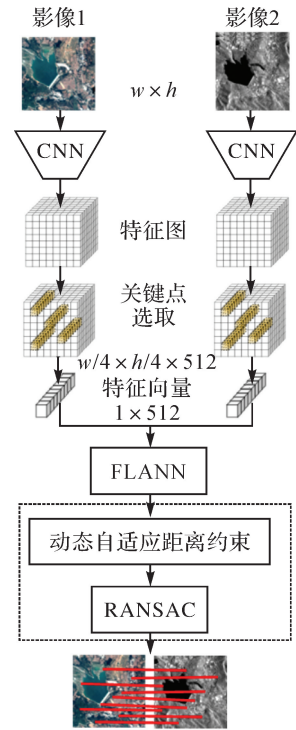


图 3 异源遥感影像匹配算法流程

Fig.3 Overview of the proposed matching method pipeline

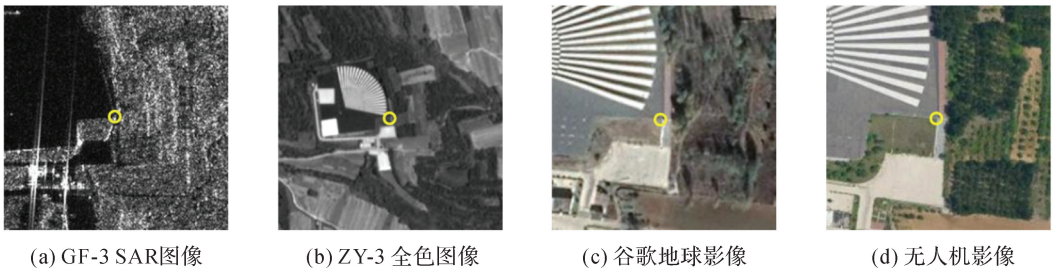
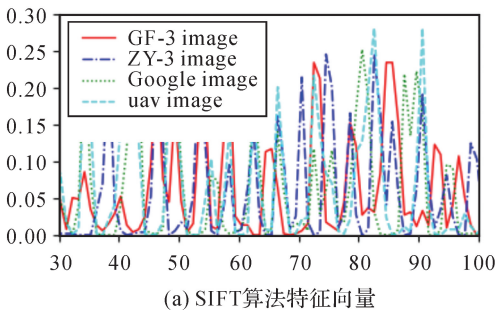
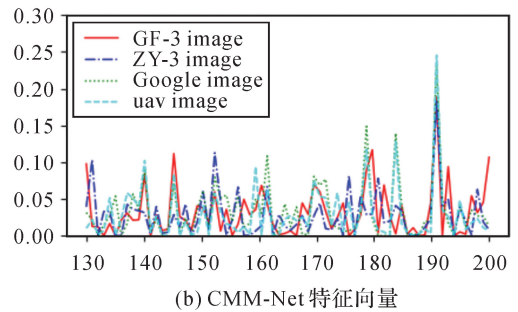


图 4 特征不变性测试数据

Fig.4 Image data for feature invariance test



(a) SIFT算法特征向量



(b) CMM-Net 特征向量

图 5 特征向量不变性测试结果

Fig.5 Vector invariance test result

前期通过大量正确匹配点和错误匹配点的特征向量的余弦距离统计得出一个经验,相似特征的余弦距离通常小于 0.4。根据这个原则,表 1 的

结果可以发现,试验中 SIFT 算法无法描述差异较大的异源图像的不变特征。而本文算法提出的特征表现出很好的相似性。同时,不同的成像模

式和高噪音的 SAR 图像和光学影像的特征相似度虽然要弱于光学影像之间的相似度,但也在一个比较高的相似性水平。通过本节的试验,说明利用本文的 CNN 模型和训练方法,特征提取器可以为异源影像提出不变特征,能够适应由于成像模式和环境造成的辐射非线性畸变的影响。

表 1 特征向量的余弦距离

Tab.1 Cosine distance of feature vectors

算法	影像对	GF-3	ZY-3	Google
SIFT	GF-3	—	—	—
	ZY-3	0.52	—	—
	Google	0.68	0.73	—
CMM-Net	GF-3	—	—	—
	ZY-3	0.37	—	—
	Google	0.31	0.28	—
算法	UAV	0.58	0.61	0.31
	UAV	0.34	0.28	0.15

2.2 特征尺度和方向适应性试验

为验证 CMM-Net 的离散金字塔多尺度模型的可行性,并测试特征在方向上的性能,本文设计了两个试验分别验证。试验以图 4(d)为基准,人工缩小或绕中心旋转图 4(c),得到在尺度和旋转上连续变化的图像,再通过比较两幅图像中心点提取的特征向量之间的余弦距离来评估算法在尺度和方向上的性能。试验结果如图 6 和图 7 所示。

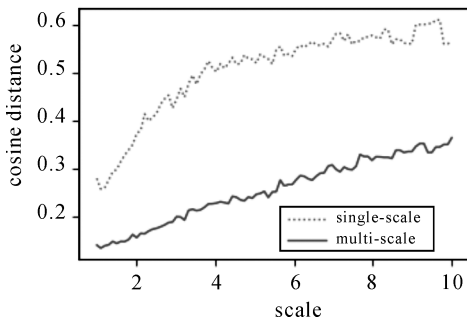


图 6 尺度和旋转不变性评估

Fig.6 Scale and rotate invariance test result

图 6 为尺度不变性评估结果。横轴表示影像在尺度上缩小的倍数,试验按照每隔 0.1 倍取样;纵轴为尺度变化后提取的特征与基准影像上同名特征之间的余弦距离。图中可以反映尺度变化,所提取的特征描述符相似性的变化情况。其中,实线为应用本文离散金字塔多尺度模型后特征相似性的变化。虚线是没有采用多尺度模型的结果。

果。从图 6 可以看出,采用离散金字塔多尺度模型后,两幅图像上同名特征提取的相似度水平提升显著。虽然会随着尺度差异变大而有所下降,但下降相对缓慢,且在缩小近 10 倍的情况下,特征余弦距离仍然保持在 0.3 左右;相反,没有采用多尺度模型情况下,相似性随尺度改变变化剧烈,尺度变化 2~3 倍后,特征相似性基本消失(余弦距离大于 0.4)。通过这个试验还可以看出,虽然本文采取的是 4 层离散(0.25,0.5,1,2 倍)金字塔模型,但整个相似性曲线变化是连续的,并没有出现阶跃现象。这说明,本文算法具备良好的连续多尺度适应能力。分析其原因,主要在于 CNN 模型本身具备一定范围的尺度适应能力,离散金字塔模型扩大了算法尺度适应的范围。

算法的旋转适应性试验结果如图 7 所示。试验对第 2 幅影像以 1°为步长,旋转 360°,计算旋转后影像提取的特征和基准图上特征的余弦距离作为相似度指标。



(a) 0~360°



(b) 0~90°

图 7 旋转适应性评估结果

Fig.7 Rotate invariance test result

图 7(b)为图 7(a)曲线的局部放大,从中可以看出,总体上算法不具备旋转不变性,但是在一定角度范围(约 15°)内,特征相似性可以保持一个较好的水平。原因在于,本文算法并没有对 CNN 特征旋转不变性进行专门的设计,小角度下的不变性是 CNN 网络从具备一定角度变化的训练图像上学习得到的。因此,本文算法只适用于较小

角度(约  $15^\circ$ )变换影像的匹配。说明本文算法还有很大的提升空间,下一步对训练数据进行旋转增广,有可能改善这一性能。

另一方面,在工程应用中,遥感影像通常具有相对准确的轨道位姿或者 PRC 等先验信息,预先可在一定的精度范围内计算影像之间在尺度和旋转角度上的几何差异,并进行消除。因此,匹配算法能够适应一定角度旋转和尺度变化即可满足实际要求。

### 2.3 关键点筛选试验

由于异源影像特别是模式不同的影像(例如 SAR 和光学)特征相似性相对较弱,因此,筛选较多待选特征点的“以多求可靠”的策略是有必要的。试验选择  $250 \times 250$  像素大小的 SAR 影像和无人机(UAV)光学影像进行特征点筛选,并可视化提取的特征图和关键点筛选结果,如图 8 所示。

从特征图可视化结果图 8(b)、图 8(e)可以看出,CNN 通过多层卷积、池化后,提取出了更为抽象的局部特征。利用深度和局部极大值的原则,CMM-Net 算法在无人机影像块上筛选出了 436 个特征,如图 8(c)所示;在 SAR 影像上筛选了 358 个特征,如图 8(f)。都提取出了较多的待选特征点,这为下一步特征匹配奠定了基础。

## 2.4 匹配性能对比试验

### 2.4.1 试验数据

试验数据情况如表 2 所示。测试数据源涵盖了星载传感器、无人机传感器、谷歌地球影像,波段和模式上包括了可见光、SAR、热红外、夜光遥感图像、栅格地图图像和深度图等。分辨率不同,影像模式多,时间与季节跨度大,对测试算法适应性具有很好的代表性。

表 2 测试数据基本情况

Tab.2 Summarization of testing datasets

组别 ID	1	2	3	4	5	6
参考影像类型	谷歌光学影像	ZY-3 PAN 光学影像	光学影像	光学影像	谷歌光学影像	WorldView-2 光学影像
分辨率/m	0.5	2.5	—	—	15	2.0
影像尺寸/pix	$1500 \times 1500$	$1000 \times 1000$	$800 \times 800$	$1000 \times 1000$	$500 \times 500$	$500 \times 500$
待匹配影像类型	谷歌光学影像	GF-3 SL SAR 影像	红外影像	SNPP/VHIS 夜光影像	OpenStreetMap 栅格地图	Lidar 深度图
分辨率/m	0.5	1.0	—	—	15	2.0
影像尺寸/像素	$1500 \times 1500$	$2500 \times 2500$	$800 \times 800$	$1000 \times 1000$	$500 \times 500$	$500 \times 500$
特点	冬季与夏季、 地物变化大	光学-SAR, 尺度差异 2.5 倍	可见光-红 外,不同波 段,有一定 旋转	昼夜不同	可见光-栅格地 图不同模式	深度图-可见 光不同模式

图 9 为测试影像的缩略图。其中,第 1 组均为光学影像,但时间与季节跨度大、地物变化明显。第 2 组为资源 3 号全色影像和高分 3 号 SAR 影像, SAR 影像分辨率为光学影像的 2.5 倍,并且 SAR 影像上存在较大噪音。第 3 组为气象卫星获取的可见光和红外波段影像,地物反射率明显不同,还存在一定旋转关系。第 4 组为光学影像与夜光遥感影像,在辐射上存在明显非线性畸变。第 5 组和第 6 组分别为光学影像与栅格地图影像、光学图像与 LiDAR 构建的深度晕渲图,具有完全不同的模式。

### 2.4.2 评价指标

试验采用正确匹配点数量(NCM)、匹配准

确率(SR)、匹配点均方根误差(RMSE)和匹配消耗时间(MT)来评价算法性能。指标说明如下。

#### 2.4.2.1 正确匹配点数量(NCM)

正确匹配点判断如式(7)所示

$$\text{Corr}(x) : \|x_i - \hat{x}_j\| \leq \epsilon \quad (7)$$

参考影像上通过算法匹配出的一个特征点位置  $x_i$ ,和理论点位  $\hat{x}_j$  之间像素距离如果小于给定的精度阈值  $\epsilon$ ,则认为正确匹配点。

正确匹配点的数量为整幅影像上,满足上述条件的所有匹配点的数量。影像配准等其他应用对正确匹配点的数量都有一定要求,该指标可以反映匹配算法的性能。



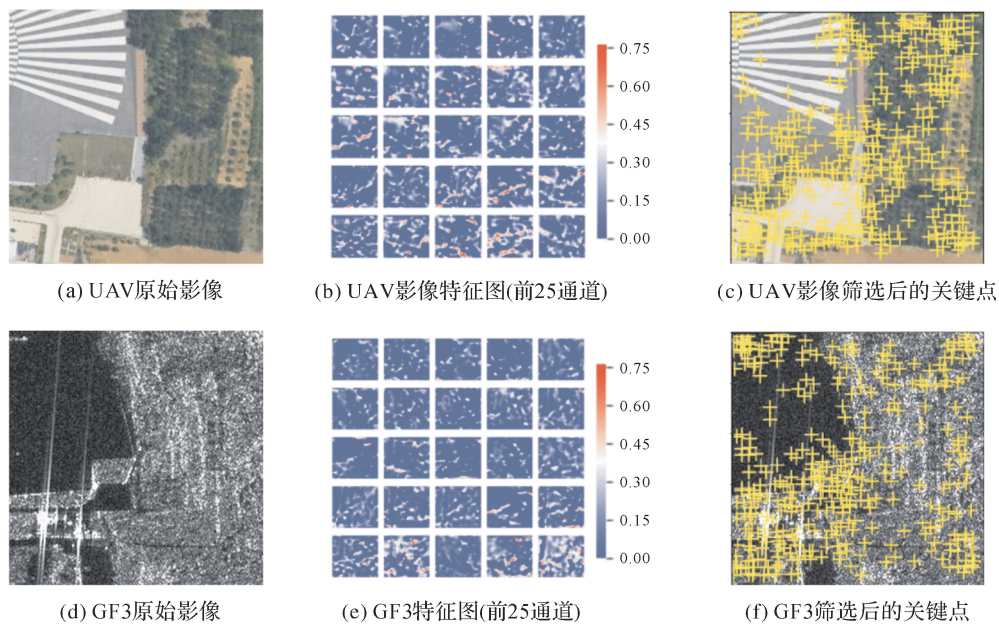


图 8 关键点选择试验结果  
Fig.8 Results of key points selection test

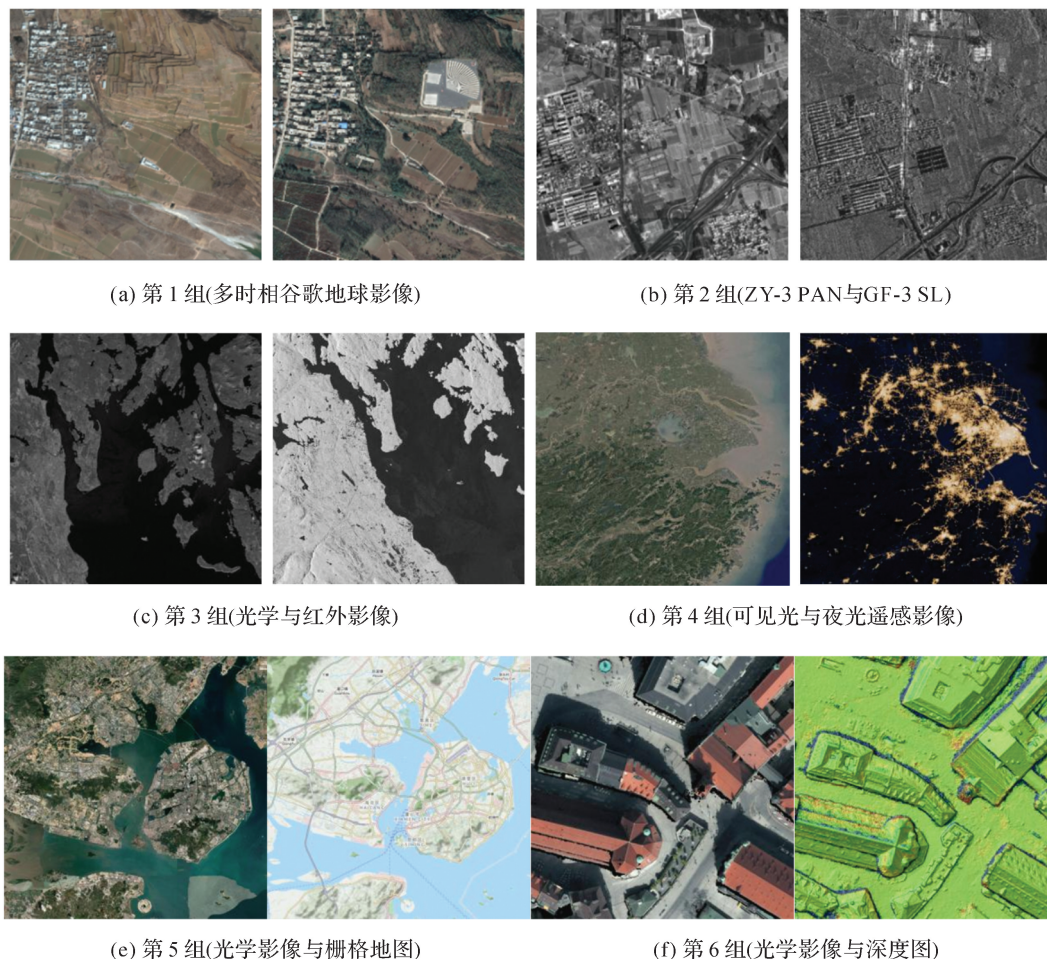


图 9 试验影像对数据  
Fig.9 Image pairs in the test

### 2.4.2.2 匹配准确率(SR)

准确率表达为满足式(7)的正确匹配点数占算法给出的所有匹配点数(NTP)的百分比,这个指标可以反映出算法得到的匹配点对成功率。

### 2.4.2.3 匹配点均方根误差(RMSE)

匹配点均方根误差 RMSE 可以反映匹配点的准确程度,计算公式为

$$RMSE = \frac{1}{NCM} \sum_i \| \mathbf{H}(\mathbf{x}_i) - \hat{\mathbf{x}}_i \| \quad (8)$$

式中,NCM 为正确匹配点的数量; $\mathbf{H}$  代表两张影像的真实变换模型,本文采用人工检查点拟合的仿射变换模型代替,该指标可以反映匹配点在像素上的位置偏移误差。

### 2.4.3 试验结果与分析

试验比较了多种在多源影像匹配中比较有代表性的匹配算法,包括基于传统的影像互信息(MI)的匹配算法(商业软件 ENVI v5.3 实现)、基于深度学习特征的 DELF 与 ContextDesc 算法。其中,ContextDesc 是专门为多模态影像设计的深度学习匹配算法。试验统计了各算法的正确匹配点数 NCM、总匹配点数 NTP 以及正确匹配点的均方根误差 RMSE,并计算了匹配点的准确率。精度阈值  $\epsilon$  取 3 个像素,结果如表 3 所示。

通过对比分析表 3 数据可知,在试验采用的 6 种不同模式的像对匹配中,本文提出的 CMM-Net 算法对所有像对均能给出一定数量的正确匹配点,匹配点均方根误差约 2 个像素。而 ENVI-MI、RIFT、ContextDesc 3 种算法对差异较大的 SAR 图像和可见光影像都没能匹配出正确的点。RIFT 算法在尺度相同的多模式影像上,都取得了非常好的匹配效果。DELF 算法虽然也有一定适应性,但匹配点数量明显偏少。ENVI-MI 算法对红外和可见光的试验数据也未能正确匹配,ContextDesc 算法对昼夜影像、地图与卫星影像的试验数据匹配未能成功。由此说明,在以上 5 种算法中,CMM-Net 在多模态影像匹配适应性方面性能最为稳定。

从正确匹配点数量和精度上看,采用传统相位一致性匹配算法的 RIFT 总体上要优于 CMM-Net。这是由于基于窗口模板匹配的定位精度要高于 CNN 特征,而基于滑动窗口的算法,可以提取更加密集的待匹配点。但从匹配准确率(SR)上看,CMM-Net 相对高一些。此外,与适应性较

强的 RIFT 和 DELF 算法相比,CMM-Net 算法耗时最短。

表 3 匹配结果

Tab.3 Matching results

组别 ID	1	2	3	4	5	6	
CMM-Net	NCM	97	31	255	68	208	197
	NTP	413	105	479	184	467	492
	SR/(%)	23	30	53	37	41	40
	RMSE	2.07	2.02	1.92	1.97	1.91	1.87
	耗时/s	6.8	8.4	5.6	3.3	3.1	3.4
	ENVI-MI	NCM	19	0	0	37	163
NTP		75	38	26	129	213	98
SR/(%)		25	0	0	29	77	69
RMSE		1.75	—	—	1.86	1.39	1.64
耗时/s		35	38	30	26	24	25
RIFT		NCM	127	0	141	121	336
	NTP	539	336	436	385	1232	1094
	SR/(%)	23	0	32	31	27	31
	RMSE	1.737	—	3.24	1.091	1.856	1.120
	耗时/s	7.9	8.6	6.7	4.8	4.3	4.2
	DELF	NCM	75	5	15	2	27
NTP		188	32	143	16	33	27
SR/(%)		40	16	10	12	82	19
RMSE		1.94	2.12	2.08	2.45	1.01	1.87
耗时/s		13.6	18.4	5.6	7.5	3.2	3.2
Context-Desc		NCM	3	0	111	0	0
	NTP	61	42	158	22	40	82
	SR/(%)	5	0	70	0	0	7
	RMSE	3.64	—	1.38	—	—	2.13
	耗时/s	3.9	4.3	3.4	3.8	3.7	3.9

### 2.4.4 匹配效果对比

图 10、图 11 和图 12 直观显示了几种匹配算法的匹配效果。从中可以看出,CMM-Net 除了可以匹配出数量较多的匹配点外,其分布也较为均匀。

图 10 的结果表明,对于有一定旋转的影像,CMM-Net 获得了比 RIFT 更多的匹配点,主要原因在于,训练数据具有一定的视角变化,使得 CMM-Net 具有一定范围内的旋转不变性。从图 11 也可以看出,CMM-Net 和 RIFT 对夜光影像都取得了较多的匹配数量,DELF 只匹配了很少数量的点,而 ContextDesc 算法则完全失效。图 12 为尺度差异 2.5 倍的 SAR 影像和可见光影像,由于模式不同且噪音较大,参与比较的算法中,只有 DELF 和 CMM-Net 有效,并且 CMM-Net 在匹配数量和点位分布均匀性上都优于 DELF,说明 CMM-Net 算法对噪音、模式变化适应性较好。

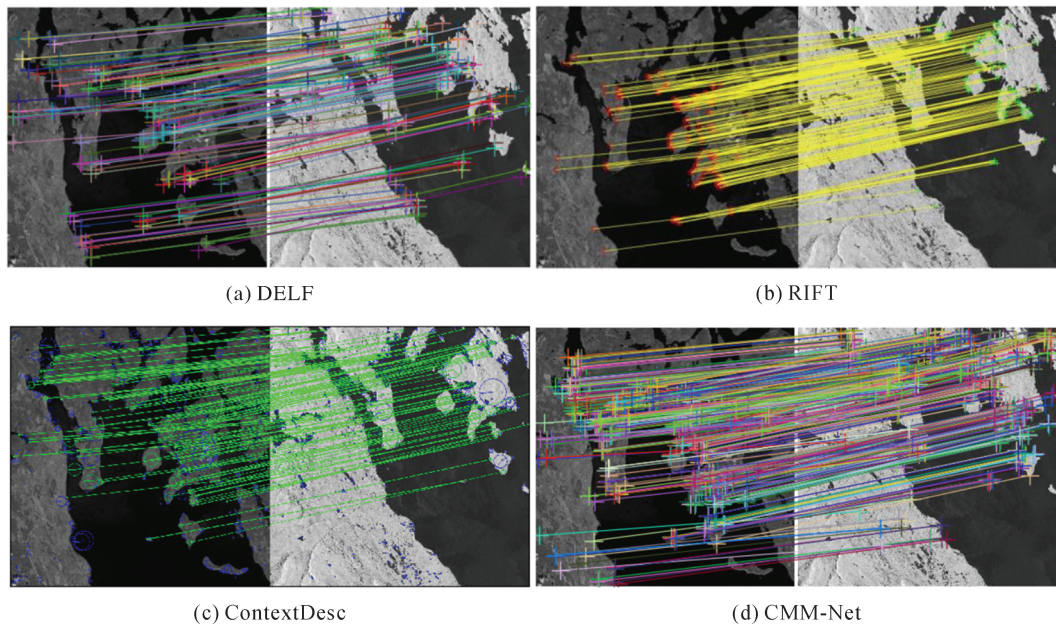


图 10 可见光与红外影像(第 3 组)匹配结果  
 Fig.10 Matching results of visible light and infrared images(Group 3)

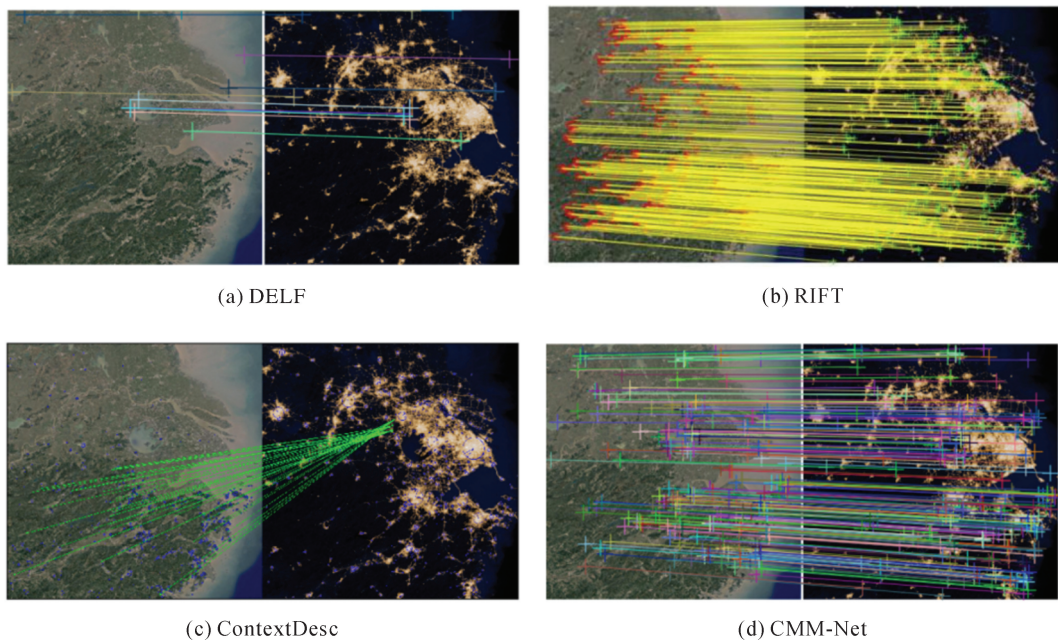


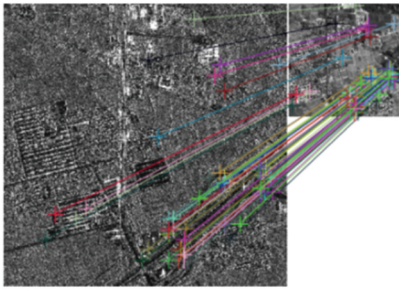
图 11 可见光与夜光遥感影像(第 4 组)匹配结果  
 Fig.11 Matching results of visible light and luminous remote sensing images(Group 4)

### 2.5 影像配准效果

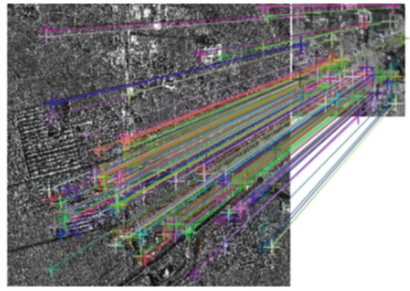
异源影像匹配的主要目的之一是影像配准,为了更好地说明 CMM-Net 算法的有效性,本文还在匹配算法的基础上,进行了影像配准试验。首先,根据 CMM-Net 匹配给出的匹配点组,计算仿射变换参数,再利用该参数对第一幅影像进行纠正,得到与第 2 幅影像配准的图像。配准的效

果可以很好地反映匹配算法的精度,为便于检查,试验采用叠加开窗的显示方式。图 13 为 4 组影像的配准效果图,配准图右侧为局部放大图。

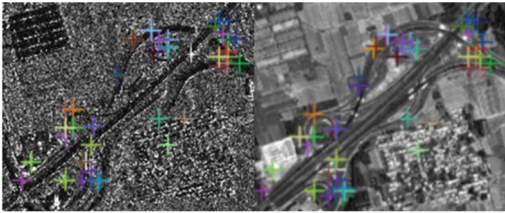
从配准结果可以看出,算法对 SAR、红外、夜光以及栅格地图图像与可见光的配准图像上各区域的配准误差基本都控制在 2 像素以内,肉眼很难看出明显的错位,说明 CMM-Net 匹配效果良好。



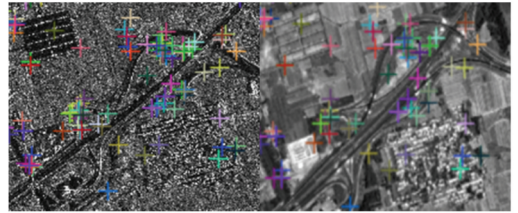
(a) DELF



(b) CMM-Net



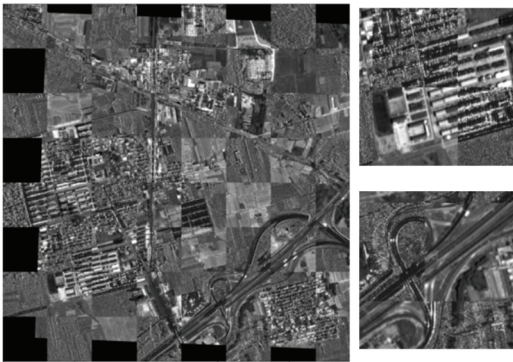
(c) DELF匹配结果放大图



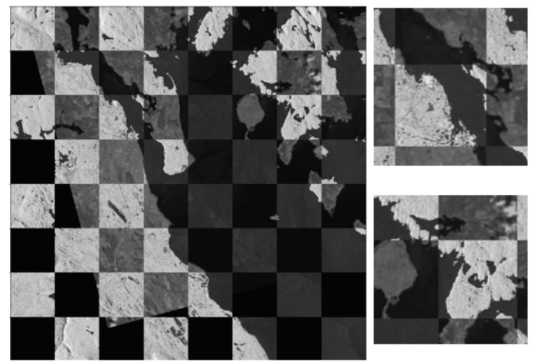
(d) CMM-Net匹配结果放大图

图 12 高分 3 号 SAR 影像与资源三号全色影像(第 2 组)匹配结果

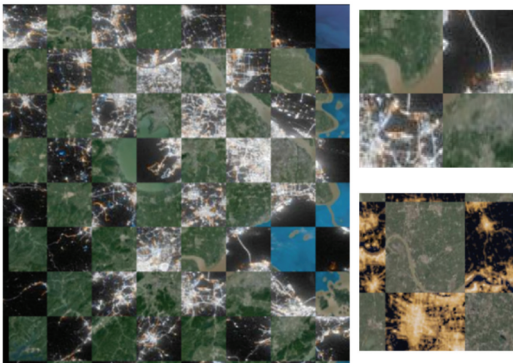
Fig.12 Matching results of GF-3 SAR image & ZY-3 image(Group 2)



(a) 第 2 组影像配准结果与放大图



(b) 第 3 组影像配准结果与放大图



(c) 第 4 组影像配准结果与放大图



(d) 第 5 组影像配准结果与放大图

图 13 影像配准结果

Fig.13 Image registration results

### 3 结论与展望

异源遥感影像之间由于成像模式、时相、分辨率等存在差异，匹配工作面临巨大的挑战。本文

提出了一种基于卷积神经网络的匹配方法。试验结果表明，该异源影像匹配算法具有较强的稳健性，在适应性方面要强于其他几种算法，且在匹配点数量、分布、效率等方面都有一定优势。虽然该

算法目前还无法完全超越基于人工设计的相位一致性匹配算法,但也为异源遥感影像稳健匹配提供了一种思路。但是,本文仅仅在利用卷积神经网络特征进行异源影像不变特征的提取与搜索匹配方面进行了探索和试验,还存在如下局限性:①受限于GPU内存,单次无法处理大像幅的影像;②从试验结果看,本文提出的CNN特征还很难做到子像素级精度的匹配;③算法流程中考虑先验几何约束信息较少。

因此,在进一步研究中,可以采用分块特征提取方法,突破GPU内存的限制;改进特征提取网络,做到子像素级的特征检测;在匹配算法流程中,还可以综合考虑先验的RPC或位姿参数、由粗到精的金字塔匹配策略以及物方约束等多种条件,提高匹配速度,并最大程度减小误匹配率,以增强工程实用性。

本文匹配算法源代码和测试数据开源在Github上(<https://github.com/lan-cz/cnn-matching>),供感兴趣的同行做进一步研究。

## 参考文献:

[1] ZHAO Y, JIANG M, MA Z F. Integration of SAR polarimetric features and multi-spectral data for object-based land cover classification[J]. *Journal of Geodesy and Geoinformation Science*, 2019, 2(4): 64-72.

[2] 李德仁. 展望大数据时代的地球空间信息学[J]. *测绘学报*, 2016, 45(4): 379-384. DOI: 10.11947/j. AGCS. 2016. 20160057.

LI Deren. Towards geo-spatial information science in big data era[J]. *Acta Geodaetica et Cartographica Sinica*, 2016, 45(4): 379-384. DOI: 10.11947/j. AGCS. 2016. 20160057.

[3] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. *International Journal of Computer Vision*, 2004, 60(2): 91-110.

[4] 戴激光, 宋伟东, 李玉. 渐进式异源光学卫星影像SIFT匹配方法[J]. *测绘学报*, 2014, 43(7): 746-752. DOI: 10.13485/j.cnki.11-2089.2014.00.

DAI Jiguang, SONG Weidong, LI Yu. Progressive SIFT matching algorithm for multi-source optical satellite images[J]. *Acta Geodaetica et Cartographica Sinica*, 2014, 43(7): 746-752. DOI: 10.13485/j.cnki.11-2089.2014.00.

[5] 王峰, 尤红建, 傅兴玉, 等. 应用于多源SAR图像匹配的级联SIFT算法[J]. *电子学报*, 2016, 44(3): 548-554.

WANG Feng, YOU Hongjian, FU Xingyu, et al. Cascade SIFT matching method for multi-source SAR images[J]. *Acta Electronica Sinica*, 2016, 44(3): 548-554.

[6] 叶沅鑫, 单杰, 熊金鑫, 等. 一种结合SIFT和边缘信息的

多源遥感影像匹配方法[J]. *武汉大学学报(信息科学版)*, 2013, 38(10): 1148-1151, 1260.

YE Yuanxin, SHAN Jie, XIONG Jinxin, et al. A node localization method in wireless sensor network based on *k*-means cluster[J]. *Geomatics and Information Science of Wuhan University*, 2013, 38(10): 1148-1151, 1260.

[7] 王瑞瑞, 马建文, 陈雪. 多传感器影像配准中基于虚拟匹配窗口的SIFT算法[J]. *武汉大学学报(信息科学版)*, 2011, 36(2): 163-166.

WANG Ruirui, MA Jianwen, CHEN Xue. SIFT algorithm based on visual matching window for registration between multi-sensor imagery[J]. *Geomatics and Information Science of Wuhan University*, 2011, 36(2): 163-166.

[8] DELLINGER F, DELON Julie, GOUSSEAU Y, et al. SAR-SIFT: a SIFT-like algorithm for SAR images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2015, 53(1): 453-466.

[9] 凌志刚, 梁彦, 程咏梅, 等. 一种稳健的多源遥感图像特征配准方法[J]. *电子学报*, 2010, 38(12): 2892-2897.

LING Zhigang, LIANG Yan, CHENG Yongmei, et al. A robust multi-source remote-sensing image registration method based on feature matching[J]. *Acta Electronica Sinica*, 2010, 38(12): 2892-2897.

[10] KOVESI P. Image features from phase congruency[J]. *Journal of Computer Vision Research*, 1999, 1(3): 1-26.

[11] WONG A, CLAUSI D A. ARRSI: automatic registration of remote-sensing images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2007, 45(5): 1483-1493.

[12] 陈敏, 朱庆, 朱军, 等. 多光谱遥感影像亮度空间相位一致性特征点检测[J]. *测绘学报*, 2016, 45(2): 178-185. DOI: 10.11947/j. AGCS. 2016. 20150030.

CHEN Min, ZHU Qing, ZHU Jun, et al. Interest point detection for multispectral remote sensing image using phase congruency in illumination space [J]. *Acta Geodaetica et Cartographica Sinica*, 2016, 45(2): 178-185. DOI: 10.11947/j. AGCS. 2016. 20150030.

[13] 叶沅鑫, 单杰, 彭剑威, 等. 利用局部自相似进行多光谱遥感图像自动配准[J]. *测绘学报*, 2014, 43(3): 268-275. DOI: 10.13485/j.cnki.11-2089.2014.0039.

YE Yuanxin, SHAN Jie, PENG Jianwei, et al. Automated multispectral remote sensing image registration using local self-similarity [J]. *Acta Geodaetica et Cartographica Sinica*, 2014, 43(3): 268-275. DOI: 10.13485/j.cnki.11-2089.2014.0039.

[14] 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. DOI: 10.11947/j. AGCS. 2018. 20170368.

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. DOI: 10.11947/j. AGCS. 2018. 20170368.

[15] LI Jiayuan, HU Qingwu, AI Mingyao. RIFT: multi-modal

- image matching based on radiation-variation insensitive feature transform[J]. *IEEE Transactions on Image Processing*, 2020, 29: 3296-3310.
- [16] DOSOVITSKIY A, FISCHER P, SPRINGENBERG J T, et al. Discriminative unsupervised feature learning with exemplar convolutional neural networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 38(9): 1734-1747.
- [17] 龚健雅, 季顺平. 摄影测量与深度学习[J]. *测绘学报*, 2018, 47(6): 693-704. DOI: 10.11947/j. AGCS. 2018.20170640.
- GONG Jianya, JI Shunping. Photogrammetry and deep learning [J]. *Acta Geodaetica et Cartographica Sinica*, 2018, 47(6): 693-704. DOI: 10.11947/j. AGCS. 2018.20170640.
- [18] ZAGORUYKO S, KOMODAKIS N. Learning to compare image patches via convolutional neural networks[C]// 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE, 2015: 4353-4361.
- [19] 范大昭, 董杨, 张永生. 卫星影像匹配的深度学习神经网络方法[J]. *测绘学报*, 2018, 47(6): 844-853. DOI: 10.11947/j. AGCS. 2018.20170627.
- FAN Dazhao, DONG Yang, ZHANG Yongsheng. Satellite image matching method based on deep convolution neural network [J]. *Acta Geodaetica et Cartographica Sinica*, 2018, 47(6): 844-853. DOI: 10.11947/j. AGCS. 2018.20170627.
- [20] YANG Zhuoqian, DAN Tingting, YANG Yang. Multi-temporal remote sensing image registration using deep convolutional features [J]. *IEEE Access*, 2018, 6: 38544-38555.
- [21] YE Famao, SU Yanfei, XIAO Hui, et al. Remote sensing image registration using convolutional neural network features[J]. *IEEE Geoscience and Remote Sensing Letters*, 2018, 15(2): 232-236.
- [22] GONG J Y, JI S P. Photogrammetry and deep learning [J]. *Journal of Geodesy and Geoinformation Science*, 2018, 1(1): 1-15.
- [23] YI K M, TRULLS E, LEPETIT V, et al. LIFT: learned invariant feature transform[C]// Proceedings of the 14th European Conference on Computer Vision. Amsterdam, Netherlands: Springer, 2016: 467-483.
- [24] DETONE D, MALISIEWICZ T, RABINOVICH A. SuperPoint: self-supervised interest point detection and description[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Salt Lake City, UT: IEEE, 2018: 224-236.
- [25] NOH H, ARAUJO A, SIM J, et al. Large-scale image retrieval with attentive deep local features [C] // 2017 IEEE International Conference on Computer Vision (ICCV). Venice: IEEE, 2017: 3456-3465.
- [26] DUSMANU M, ROCCO I, PAJDLA T, et al. D2-Net: a trainable CNN for joint description and detection of local features[C]// Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, CA: IEEE, 2019.
- [27] LUO Zixin, SHEN Tianwei, ZHOU Lei, et al. Context-Desc: local descriptor augmentation with cross-modality context[C]// The Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA: IEEE, 2019: 2522-2531.
- [28] KAREN Simonyan, REW Zisserman. Very deep convolutional networks for large-scale image recognition[C]// Proceedings of the International Conference on Learning Representation. San-diego, California, USA: [s.n.], 2014.
- [29] MISHCHUK A, MISHKIN D, RADENOVIC F, et al. Working hard to know your neighbor's margins: local descriptor learning loss[C]// Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017.
- [30] LI Zhengqi, SNAVELY N. MegaDepth: learning single-view depth prediction from internet photos [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT: IEEE, 2018.
- [31] SCHÖNBERGER J L, FRAHM J M. Structure-from-motion revisited[C]// The Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV: IEEE, 2016: 4104-4113.

(责任编辑:张艳玲)

收稿日期: 2020-02-17

修回日期: 2020-12-14

第一作者简介: 蓝朝桢(1979—),男,博士,副教授,研究方向为摄影测量与遥感。

First author: LAN Chaozhen(1979—), male, PhD, associate professor, majors in photogrammetry and remote sensing.

E-mail: lan\_cz@163.com

通信作者: 徐青

Corresponding author: XU Qing

E-mail: 13937169139@139.com