

一种基于深度和局部特征的分布式光电图像配准方法

陈晓露, 刘奕

(中国电子科技集团公司第十一研究所, 北京 100015)

摘要:针对分布式光电成像系统采集的红外和可见光图像在配准时易受噪声影响,配准精度不高问题,提出一种基于卷积神经网络深度特征和 RIFT 局部特征的图像配准算法。首先基于改进的 AVIRnet 提取待配准红外和可见光图像的卷积深度特征,利用深度特征进行初匹配,得到初步的空间关系;然后在重叠图像区域内提取 RIFT 特征点;最后对局部特征点进行修正,得到最终的匹配点对,估算出精确的变换矩阵。实验结果表明:本文方法通过深度特征和局部特征两次匹配,对非线性辐射差异具有不变性,满足了分布式光电红外和可见光图像配准的精度要求。

关键词:分布式光电;红外和可见光图像;图像配准;特征匹配

中图分类号:TP391;TN29 **文献标识码:**A **DOI:**10.3969/j.issn.1001-5078.2023.07.021

A distributed electro-optical image registration method based on deep and local features

CHEN Xiao-lu, LIU Yi

(The 11th Research Institute of CETC, Beijing 100015, China)

Abstract: To address the problems that the infrared and visible images collected by distributed electro-optical image system are susceptible to noise and poor alignment accuracy, an image registration algorithm based on convolutional neural network features and RIFT local features is proposed in this paper. Firstly, the convolution deep features of infrared and visible images to be registered are extracted based on improved AVIRnet, and the initial feature point pairs are obtained by deep features rough match. Secondly, the RIFT feature points are abstracted from overlap image area. Then, the local feature points are revised to obtain the final matching points pairs, and the accurate transformation matrix is estimated. The experiment result indicated that by the two-step registrations on deep and local features, the proposed method is invariant to nonlinear radiation difference and essentially satisfied the precise requirement of distributed electro-optical infrared and visible image registration.

Keywords: distributed electro-optical; infrared and visible image; image registration; feature matching

1 引言

随着“分布式杀伤”概念的提出,单平台之间的离散式对抗转变为体系间的协同对抗^[1],在此背景下,提出分布式光电成像系统,即以多种异质的小型、低成本光电载荷实现或超越高性能、高价值的单个光电载荷系统能力。分布式光电发挥作

用的关键之一是图像信息能够有效协同利用,如多帧图像超分辨率重建提高目标空间分辨率、多视点图像拼接扩大侦察范围、多波段融合增加目标有效信息量。图像配准在这里作为前置步骤,精度高、鲁棒性好的配准算法更有利于分布式光电图像的协同应用。与上述应用对应的是不同时

间、不同视角、不同模式的图像配准。尽管应用不同,但配准算法并不独立,如多视点图像拼接的图像数据通常是不同时间采集的,多模式图像融合的图像数据也存在视角差异。现阶段分布式光电采集的主要是红外图像和可见光图像,因此本文针对红外和可见光的多视点图像配准算法研究。

由于成像原理不同,红外和可见光图像的纹理和灰度分布差异巨大,一直是多模图像配准中的难点。图像配准步骤^[2]一般包括:寻找参考图像和模板图像之间的对应关系、几何变换参数估计、图像插值变换。其中参考图像和模板图像之间的相似性度量是图像配准的核心。相似性度量的方法大致可分为基于特征和基于区域信息,前者提取兴趣点,为每个兴趣点计算一个描述符,相似度为相应描述符之间的距离;后者是利用两图像相关性、互信息等直接测量相似性。由于遮挡、热量不均匀、杂波、对比度反转等影响,过去红外和可见光图像的配准针对特定场景或特定目标,匹配点数少、配准精度往往较低。根据多模图像非线性异构的特点,Li 等^[3]提出相位一致性模型,检测的特征点具有非线性辐射差异不变性;孙兴龙等^[4]基于图像轮廓特征,克服了红外可见光图像灰度和纹理差异的影响;职玉等^[5-6]提出一种基于特征轮廓四边形的热红外图和可见光图的配准方法,解决已有算法特征轮廓应用不足导致的精度较差问题;赵洪山等^[7]将显著性梯度信息和归一化互信息相结合作为配准的测度函数,解决红外图像模糊导致的配准精度下降问题。近年来,随着深度学习的发展,卷积神经网络(CNN)方法被应用于多模图像配准中,Zhao 等^[8]利用超列组合 CNN 中不同层的特征匹配红外和可见光图像中的特征点;毛远宏等^[9]提出由特征提取和特征匹配两部分组成的图像块匹配网络;刘洋等^[10]利用人工免疫网络算法配准红外和可见光图像。

研究表明 CNN 能够提取图像的特征,浅层网络可检测边缘、纹理等特征,深层网络可检测更为复杂稳定的深度特征,异构图像更容易通过深度特征进行匹配。分布式光电成像系统实际采集的红外和可见光图像存在显著的非线性辐射差异,且易受到噪声影响,传统的基于特征和区域的方法很难保证配准的准确性,本文提出一种基于深度特征和局部特征新的两步配准方法以解决分布式光电红外和可见

光图像配准精度问题。

2 分布式光电图像配准总体方案

待配准的红外和可见光图像由装载于无人机平台的多个分布式光电载荷采集得到,每套光电载荷均由光电传感器、伺服稳定平台、电源组件等部分组成,其中光电传感器包括中/长波红外传感器、可见光 CCD、激光测距机等。在实际应用中,光电载荷可对搭载的光电传感器进行选配,整个分布式光电系统可实现光电图像可见光—红外波段的覆盖。

本文针对分布式光电载荷采集的中波红外和可见光图像存在显著的非线性辐射、视角差异,采用深度特征进行图像初匹配,初匹配可在图像特征层中进行,得到基于深度特征的近似变换矩阵,再结合局部特征提高配准的精度。具体的方法流程如图 1 所示。

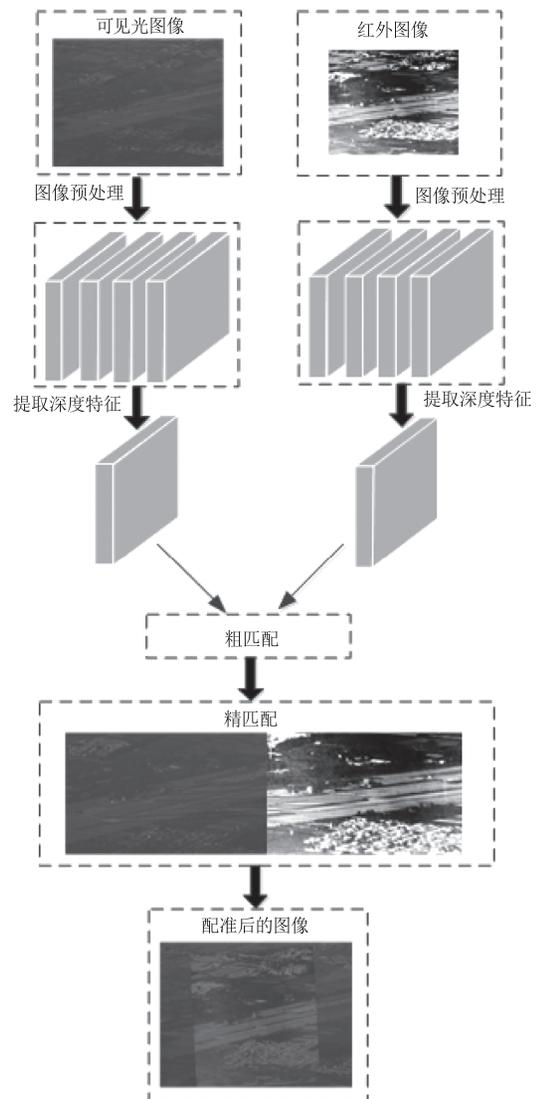


图 1 本文方法结构图

Fig. 1 Structure of the method presented in this paper

3 基于深度和局部特征的配准方法

3.1 深度特征提取

图像配准可看作是视觉相似位置匹配的过程,由于成像机理的不同,经典算法 SIFT 很难对外观明显变化的红外和可见光图像直接匹配,而 CNN 模型可挖掘图像本质内容特征。VGG16 是使用较多的深度神经网络,当学习样本不足时,常用 ImageNet 预训练网络进行微调的方式提高训练效率,但分布式光电图像与 ImageNet 数据集图像特征差距较大,学习的效果不理想。为提高实际航拍红外和可见光图像的匹配率,本文引入可见光和红外深度学习匹配网络 AVIRNet^[11],并对 AVIR 块进行改进,以更好的提取图像深度特征。

改进后的整个网络结构与 AVIRNet 相同,包括 5 个卷积块、5 个 AVIR 块、5 个最大池化层、1 个全局平均池化层与 1 个全连接层,如下图所示,其中卷积块包括 3×3 滤波器、批标准化 BN 和激活函数 ReLU。

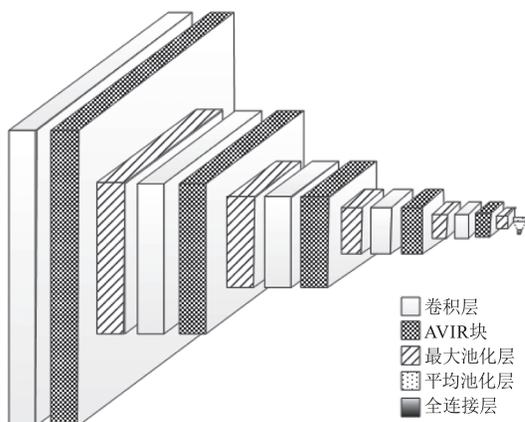


图2 AVIRnet 网络结构

Fig. 2 Structure of AVIRnet

人眼观察异构图像的匹配效果,首先关注的是边缘,其次是纹理,基于这一思想,原 AVIRNet 设计了 AVIR 块,卷积后的特征通过水平滤波器和垂直滤波器提取边缘。线性滤波器对噪声比较敏感,无法抑制噪声的影响,直接应用于分布式光电图像边缘检测效果差,因此本文首先使用高斯滤波器平滑和去除噪声,然后使用 Sobel 算子计算水平和垂直方向的一阶导数检测边缘。用改进的 AVIRNet 作为特征提取器实现分布式图像的特征提取,由于只使用卷积层提取特征,所以输入图像可是任意大小,但较大的输入图像需要较大的计算量,同时对特征

描述需要适当大小的感受野,因此充分考虑航拍图像特性,输入图像大小为 128×128 ,卷积层输出分别为 $128 \times 128 \times 64$ 、 $64 \times 64 \times 128$ 、 $32 \times 32 \times 256$ 、 $16 \times 16 \times 512$ 和 $8 \times 8 \times 512$ 。更深的层有更多的抽象特征,但是考虑到第五层特征点数量太少,因此不采用该层,从而获得了特征图。

3.2 空间块匹配

受最大池化的影响,高层特征图尺寸更小,可降低匹配的复杂度,同时具有更深层次的抽象特征,因此初匹配从第四层开始。我们将从改进的 AVIRNet 提取的红外和可见光图像特征记为 $\{M_I^L(F), M_V^L(F)\}$,其中 L 表示特征层数, $L = 1, 2, 3, 4$ 。设 a_{ij}^L 表示 $M_I^L(F)$ 中某一位置特征向量, b_{ij}^L 表示 $M_V^L(F)$ 中某一位置特征向量,其中 ij 为特征图尺寸。则第四卷积层特征图 $\{M_I^4(F), M_V^4(F)\}$,第四层输出通道为 512,于是特征向量大小为 512,特征图大小为 16×16 ,则第四层特征图在原始图像中是一个 8×8 的块,原图像被分成 16×16 个块。 a_{ij}^4 和 b_{ij}^4 的匹配是寻找图像中相似的块,每个块的中心表示块的位置,以特征向量之间的欧式距离互为最小的匹配对作为候选匹配对。

3.3 局部特征提取

得到近似变换矩阵后,进一步提取图像的局部特征,得到最终的对应关系以计算得到准确的变换矩阵。经典的 SIFT 算子是以图像的梯度信息检测特征点,对红外和可见光非线性强度差异较大的图像,差分梯度不具有不变性。因此本文采用 RIFT^[3] 提取局部特征,用相位一致性代替差分梯度进行特征点检测。

相位一致性模型表示为:

$$PC(x,y) = \frac{\sum_s \sum_o \omega_o(x,y) [A_{so}(x,y) \Delta \Phi_{so}(x,y) - T]}{\sum_s \sum_o A_{so}(x,y) + \zeta} \quad (1)$$

式中, $\omega_o(x,y)$ 是权重函数; ζ 是一个极小的常量; $[\cdot]$ 为防止值为负的运算符,当值为负时取 0; $\Delta \Phi_{so}(x,y)$ 为相位偏差函数。

根据矩分析算法,最小矩对应的轴为主轴,表示特征的方向信息,可用于检测角点;最大矩对应的轴与主轴垂直,一般反映特征的显著性,可用于检测边缘点,因此局部特征提取结合了角点和边缘点特征。

由于分布式光电图像数据量较大,直接使用 RIFT 算法将提取大量的特征点,可见光图像覆盖目标范围大于红外图像,检测出的特征点大部分不在红外图像上,利用这些特征点搜索匹配点不仅耗时而且影响配准效率。因此在特征点提取前,用下式先确定红外和可见光图像的重叠区域范围:

$$S_i(v) = \sum_x [D_I(x) - D_V(x - v)]^2 \quad (2)$$

式中, D_I 和 D_V 是红外和可见光图像待匹配图像模板区域; x 是模板区域像素点位置; v 为 D_I 和 D_V 的偏移量,相似度 S_i 最小时求取 v 。

并且 RIFT 是基于相位一致性进行特征点检测和描述,不能像 SIFT 构造尺度空间,不具有尺度不变性,而分布式光电采集的红外和可见光图像之间存在分辨率差异。因此为获得尺度不变性,对已提取的四层红外和可见光图像特征图,逐层计算相位一致性模型构造最小矩图和 MIM 图,从而得到多个尺度下的特征表达,最后将每一层图像上特征点按各层降采样倍数转换到原图像坐标系上。

3.4 求取全局配准矩阵

实际采集的分布式光电图像由于噪声和强度差异,红外图像和可见光图像在局部特征匹配时的最佳匹配点不总是最邻近的,为提高配准精度,需去除明显的错误点对。在块匹配中对挑选 n 对最近邻距离比值倒数最高的点对,然后随机挑选出 4 对,建立归一化重心坐标系,通过下式计算:

$$\sqrt{\sum_{i=1}^4 (x_i - x'_i)^2} < \delta \quad (3)$$

式中, x_i 和 x'_i 为匹配点对; δ 为阈值。

直到选出满足上式的 4 对特征点对,用以估计仿射变换模型,并用得到的仿射变换模型对初始匹配点对进行变换,计算残差,误差小于阈值 ε 的点对作为内点。迭代一定次数后,选出具有最大内点数所对应的仿射变换模型为变换矩阵。

4 实验结果与分析

4.1 数据集构建

根据分布式光电试飞拍摄的环境,包括机场、道路、建筑和农田,分为 3 组数据训练,1 组数据测试,每组数据至少包含两种场景。实验图像数据由分布式光电成像系统采集,红外图像选用中波红外,工作波段为 $3.7 \mu\text{m} \sim 4.8 \mu\text{m}$,原始图像分辨率为 640×512 像素,可见光图像原始图像分辨率为 $1920 \times$

1080 像素。为提高训练效率,将图像尺寸缩放为 512×512 像素。

实验数据集经过筛选,剔除高重叠度的图像(重叠度 50% 以上),样本图像对包含了显著的外观变化、视点变化和轻微的错位和旋转。从中选取三个不同场景的图像对进行实验比较,如图 3 所示,上排为红外图像,下排为可见光图像,它们的分辨率、噪声分布和灰度值都有很大差异,且由于试飞时间在傍晚,可见光图像整体较暗,噪点较多。从图 3(a) ~ (c) 分别对应着建筑、农田和机场不同的景物特征。

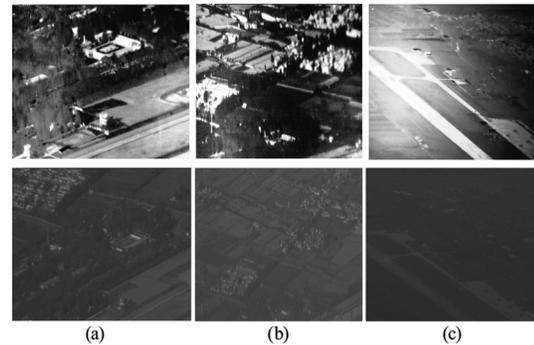


图 3 红外和可见光实验图像对

Fig. 3 Test data sets

4.2 图像配准实验

对数据集中不同场景、不同视点的图像对进行配准实验,图 4 显示了特征匹配的结果,其中图 4(a) 和图 4(b) 是本文由粗到精两步匹配和传统 SIFT 匹配精度结果对比,正确匹配用黄线表示,错误匹配用蓝色表示。结果表明,与传统 SIFT 方法相比本文方法生成的对应关系更加准确,特征点在图像上的分布集中在红外和可见光图像重叠区域,降低了无效特征点检出率。图 4(c) ~ 4(e) 是图 3(a) ~ 3(c) 对应图像对的配准结果,采用全局配准矩阵对红外图像进行插值变换,并将其叠加到可见光图像上,以红外图像镶嵌显示最终的配准结果。当图像对外观差异较大,使用改进 AVIR 网络提取的卷积特征对外观变化的鲁棒性更好,并且局部特征描述子利用了高级卷积多尺度特征,然后在尺度空间上进行特征点检测,相比原 RIFT 提出的描述子具有尺度不变性,得到较为理想的配准效果。

用均方根误差 RMSE 和正确匹配率 CMR 来衡量配准算法的性能,实验选择用 SIFT 和 RIFT 配准算法与所提方法进行比较。RMSE 定义为:

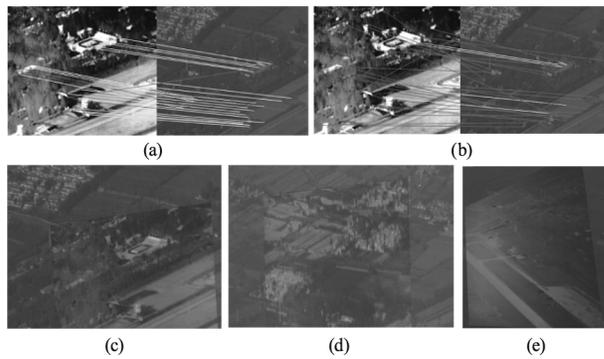


图4 红外和可见光图像特征匹配及配准结果

Fig. 4 Infrared and visible images feature matching and registration results

$$RMSE = \sqrt{\frac{1}{n} \left(\sum_{i=1}^n (x_i - \tilde{x}'_i)^2 + (y_i - \tilde{y}'_i)^2 \right)} \quad (4)$$

式中, (x_i, y_i) 为红外图像匹配点的坐标; $(\tilde{x}'_i, \tilde{y}'_i)$ 为对应的可见光图像经变换后的理论匹配点坐标; n 为匹配特征点对数。

$$CMR = \frac{n_c}{n} \quad (5)$$

式中, n_c 为正确匹配对数。

表1 不同实验场景算法性能对比

Tab. 1 Performance comparison of algorithm in different experiments

实验场景	配准方法	CMR/%	RMSE
1	SIFT	53	10.883
	RIFT	83	4.236
	本文方法	94	2.816
2	SIFT	49	12.846
	RIFT	81	3.872
	本文方法	88	3.034
3	SIFT	62	8.957
	RIFT	78	5.187
	本文方法	89	2.941

如表1所示,配准结果的RMSE越小表示配准精度越高,CMR越大匹配性能越好,手动选择20对均匀分布的特征点对估计仿真变换模型来计算比较各方法的RMSE和CMR。图像上获取足够数量的正确匹配点对是配准成功的关键,由于红外

和可见光图像之间的非线性辐射差异,用SIFT提取的正确匹配点对太少,很难直接生成变换矩阵。相比之下,RIFT的相位一致性表现出较好的抗红外和可见光图像间非线性辐射差异的能力,正确的匹配点对数明显增加,但对于分辨率差异较大的图像,难以获得稳定的配准效果。从表中可以看出本文用深度特征块匹配,然后在粗配准结果上使用改进的RIFT算法进行精匹配,在三组实验数据中均得到较好的匹配结果。

5 结论

应用分布式光电成像系统对实际场景成像,对采集的红外和可见光图像实施配准实验时发现,直接用传统的SIFT算法提取特征点,由于异质图像间较大的非线性辐射差异,提取的匹配点对不足且不准确,直接影响了分布式光电图像的有效协同利用。本文提出的由粗到精两步配准策略,基于真实红外可见光图像的实验效果验证了方法的有效性。但同时由于使用了深度特征和局部特征,算法复杂度也增加,后续可针对两种特征存在的冗余,进一步提高配准的计算速度。

参考文献:

- [1] Tang Shengjing, Si Songwei, Zhang Yao, et al. Review on the development of intelligence-based distributed cooperative operational system[J]. Air & Space Defense, 2019, 2(1):6-13. (in Chinese)
唐胜景, 史松伟, 张尧, 等. 智能化分布式协同作战体系发展综述[J]. 空天防御, 2019, 2(1):6-13.
- [2] Mikhail L Uss, Benoit Vozel, Sergey K Abramov, et al. Selection of a similarity measure combination for a wide range of multimodal image registration cases[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(1):60-75.
- [3] Jiayuan Li, Qingwu Hu, Mingyao Ai. RIFT: multi-modal image matching based on radiation-variation insensitive feature transform[J]. IEEE Transactions on Image Processing, 2020, 29:3296-3310.
- [4] Sun Xinglong, Han Guangliang, Guo Lihong, et al. Infrared-visible video automatic registration with contour feature matching[J]. Optics and Precision Engineering, 2020, 28(5):1140-1151. (in Chinese)
孙兴龙, 韩广良, 郭立红, 等. 采用轮廓特征匹配的红外-可见光视频自动配准[J]. 光学精密工程, 2020,

- 28(5):1140-1151.
- [5] Zhi Yu, Zhu Juanjuan, Liu Rui, et al. Registration of thermal infrared image and visible image based on featured contour quadrilateral[J]. *Infrared and Laser Engineering*, 2021, 50(S2):20200520. (in Chinese)
职玉, 朱娟娟, 刘锐, 等. 应用特征轮廓四边形的热红外图与可见光图配准[J]. *红外与激光工程*, 2021, 50(S2):20200520.
- [6] Jiang Qian, Liu Yadong, Fang Jian, et al. Registration method for power equipment infrared and visible images based on contour feature[J]. *Chinese Journal of Scientific Instrument*, 2020, 41(11):252-260. (in Chinese)
姜骞, 刘亚东, 方健, 等. 基于轮廓特征的电力设备红外和可见光图像配准方法[J]. *仪器仪表学报*, 2020, 41(11):252-260.
- [7] Zhao Hongshan, Zhang Zeyan. Power equipment infrared and visible images registration based on cultural wolf pack algorithm [J]. *Acta Optica Sinica*, 2020, 40(16):1610003. (in Chinese)
赵洪山, 张则言. 基于文化狼群算法的电力设备红外和可见光图像配准[J]. *光学学报*, 2020, 40(16):1610003.
- [8] Zhenbing Zhao, Lingling Zhao, Yincheng Qi, et al. Infrared and visible image registration based on hypercolumns [C]//in CCF Chinese Conference on Computer Vision. Springer, 2017:529-539.
- [9] Mao Yuanhong, Ma Zhong, He Zhanzhuang. Infrared-visible image patches matching via convolutional neural networks[J]. *Infrared and Laser Engineering*, 2021, 50(5):20200364. (in Chinese)
毛远宏, 马钟, 贺占庄. 采用卷积神经网络的红外和可见光图像块匹配[J]. *红外与激光工程*, 2021, 50(5):20200364.
- [10] Liu Yang, Yang Chen. Infrared and visible image registration based on deeplearning theory [J]. *Laser Journal*, 2020, 41(11):81-85. (in Chinese)
刘洋, 杨晨. 基于深度学习理论的红外和可见光图像配准[J]. *激光杂志*, 2020, 41(11):81-85.
- [11] Inhopark, Jongminjeong, Sunghokim. Accent for visible and infrared registration (AVIR): Attention Block for increasing patch matching rate through edge emphasis[J]. *IEEE Access*, 2021, 9:160661-160674.