

引文格式: 王晨霖, 金新, 韩林森, 等. 无人机视觉定位的稀疏特征匹配算法[J]. 航空制造技术, 2025, 68(19): 96–102.

WANG Chenji, JIN Xin, HAN Linsen, et al. Sparse feature matching algorithm for UAV visual localization[J]. Aeronautical Manufacturing Technology, 2025, 68(19): 96–102.

无人机视觉定位的稀疏特征匹配算法

王晨霖¹, 金新², 韩林森², 黄文君¹, 罗远哉¹

(1. 西北大学, 西安 710127;

2. 中国航空制造技术研究院, 北京 100024)

[摘要] 目前, 无人机(Unmanned aerial vehicle, UAV) 导航定位主要使用全球卫星导航系统(Global navigation satellite system, GNSS)。然而, 当卫星信号不佳或受到干扰时, 会严重影响无人机任务的完成, 甚至危及无人机安全飞行。针对这一问题, 本文提出了一种适用于 GNSS 拒止环境下的视觉定位算法, 可确保无人机实现安全、长期的稳定飞行, 该算法通过无人机航拍图像与带有地理标签的卫星地图匹配解算无人机地理坐标。首先, 设计了一种卫星地图预处理方法, 降低无人机飞行过程中的计算量; 其次, 采用学习感知图像块相似度(Learned perceptual image patch similarity, LPIPS) 指标进行初始检索; 最后, 结合基于深度学习的 SuperPoint 稀疏特征提取算法与 LightGlue 特征匹配算法进行图像匹配与偏移量估计, 实现无人机的视觉定位。提出的方法在 ALTO 数据集上进行测试验证, 在 R@1 指标上超过目前最先进方法 17.2%, 证明了该方法的可行性与先进性。

关键词: 无人机(Unmanned aerial vehicle, UAV); 视觉定位; 全球卫星导航系统拒止; 特征提取; 特征匹配

中图分类号: TP751

文献标志码: A

DOI: 10.16080/j.issn1671-833x.2025.19.096

Sparse Feature Matching Algorithm for UAV Visual Localization

WANG Chenji¹, JIN Xin², HAN Linsen², HUANG Wenjun¹, LUO Hangzai¹

(1. Northwest University, Xi'an 710127, China;

2. AVIC Manufacturing Technology Institute, Beijing 100024, China)

[ABSTRACT] Currently, unmanned aerial vehicles (UAVs) primarily rely on the global navigation satellite system (GNSS) for navigation and localization. However, in scenarios where satellite signals are weak or interfered with, the completion of UAV missions is severely affected, and the safe flight of UAVs could even be jeopardized. To address this issue, this paper proposes a visual localization algorithm to ensure the safe and long-term flight of UAVs in GNSS-denied environments. The algorithm calculates the UAV's geographic coordinates by matching aerial images captured by the UAV with geotagged satellite maps. Firstly, a satellite map preprocessing method is designed to reduce the computational load during UAV flights. Secondly, the learned perceptual image patch similarity (LPIPS) metric is used for initial retrieval. Finally, image matching and offset estimation are performed by combining the deep learning-based SuperPoint sparse feature extraction algorithm and the LightGlue feature matching algorithm, thus finally achieving UAV visual localization. The proposed method is tested on the ALTO dataset, achieving a 17.2% improvement over the current state-of-the-art methods in terms of the R@1 metric, which demonstrates its feasibility and advancement.

Keywords: Unmanned aerial vehicle (UAV); Visual localization; GNSS-denied; Feature extraction; Feature matching

近年来,无人机因其研制周期短、成本低,以及使用风险较低等独特优势,广泛应用于战场侦察、灾后救援、田间农药喷洒、商业航拍等诸多民用及军用领域。随着军/民用需求对无人机性能的要求越来越高,无人机技术已不再是将人机简单地无人化,而是需要解决一系列独特的关键技术^[1-2]。目前无人机主要依赖卫星导航定位,但在战时等特殊环境下,卫星信号容易被干扰或欺骗^[3]。为解决这一问题,本文设计了一种新型无人机视觉定位方案,在全球卫星导航系统(Global navigation satellite system, GNSS)完全拒止的情况下也能保证无人机的长时间稳定飞行。

无人机视觉地理定位可以视为图像检索任务^[4],通过将无人机航拍图像与携带地理位置信息的卫星瓦片地图进行匹配来估计无人机的地理位置。近年来,随着深度学习技术的快速发展,基于深度学习方法的图像检索算法成为国内外的研究重点。2016年,NetVLAD算法通过引入卷积神经网络(Convolutional neural network, CNN),在大规模视觉地点识别任务上取得显著成果^[5-6]。2020年,DELG模型在两个公开数据集上实现最先进的图像检索^[7]。2021年提出的Patch-NetVLAD算法,在NetVLAD的基础上通过结合局部和全局描述符,进一步提升了识别准确性^[8],尽管性能略逊于DELG模型,但其运行速度更快,更适用于移动端计算平台。2023年,TransVLAD模型通过融合卷积神经网络与注意力机制,在3个数据集上均取得了当前最先进的结果^[4,9]。然而,以上方法大多集中在自然图像领域的视觉定位任务中,并未针对无人机视觉定位任务进行设计和优化。同年,西北工业大学Hu等^[10]提出了多阶段细化方法CurriculumLoc,通过分阶段语义场景识别和几何结构测量,在无人机视觉位置识别公开数据集ALTO上的测试结果显著优于其他算法^[11]。然而,CurriculumLoc方法的特征提取模块采用了复杂的Swin Transformer架构^[12],其LSA(移位局部自注意力)的计算方式实现难度较大,且受限于独特的计算单元设计,难以在移动端计算平台上高效部署。

本文为无人机视觉定位任务提出了一种基于稀疏特征匹配的解决方案。首先针对无人机视觉定位任务的延迟敏感特性设计了一种卫星地图预处理方法,通过降低无人机飞行过程中的计算量来提升计算速度。同时采用轻量级Alexnet架构学习感知图像块相似度(Learned perceptual image patch similarity, LPIPS)指标快速筛选出候选卫星图像^[13-14],并结合基于深度学习的SuperPoint稀疏特征提取算法与LightGlue特征匹配算法实现航拍图像与卫星图像的高精度配准^[15-16]。值得注意的是,本文采用的特征提取模块仅采用CNN架构,具备在移动端

高效部署的特点,进一步增强了其在实际应用中的可行性与可用性。

1 基于稀疏特征匹配的无人机视觉定位方法

本文提出的基于稀疏特征匹配的无人机视觉定位方法的整体流程如图1所示。首先,在飞行前对卫星地图进行预处理,得到离线地图数据库,在执行视觉定位任务中,利用LPIPS指标对航拍图像与卫星图像的相似度进行评估,筛选出前50张卫星图像作为候选图像;随后,将航拍图像与候选图像通过稀疏特征提取算法与特征匹配算法计算匹配点对,并根据平均像素距离进行重排序;最后,对匹配效果最好的卫星图像进行单应性估计,解算出无人机地理坐标。

1.1 卫星地图预处理

为满足不同飞行高度下的视觉定位任务需求,从谷歌卫星地图服务下载合适层级的瓦片地图。在谷歌卫星地图中,瓦片地图金字塔式的层级结构表示不同的缩放级别,其层级范围通常为0~18,且每个层级的地图分辨率均为固定值,因此可以根据瓦片地图的层级与编号得到卫星图像上任意一点的地理坐标。

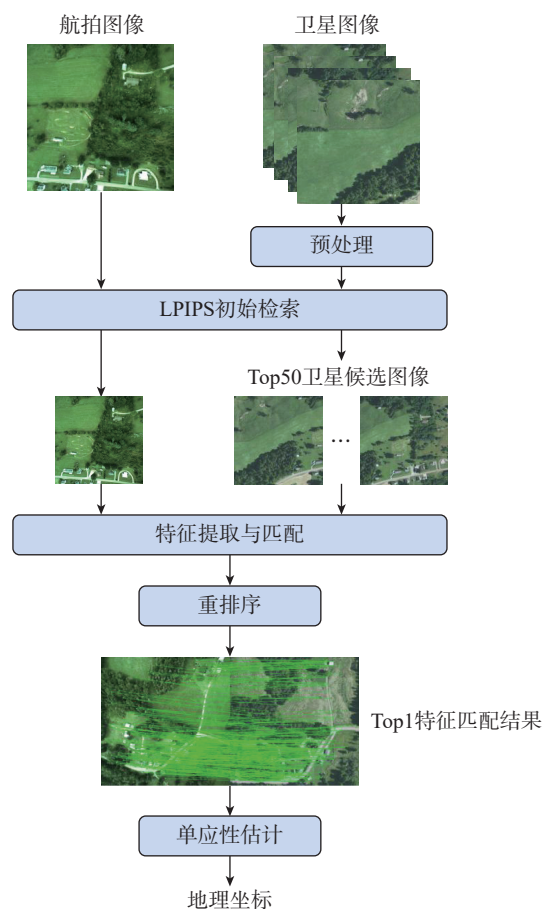


图1 算法整体流程图

Fig.1 Overall flowchart of algorithm

对于执行长距离飞行任务的无人机,在将航拍图像与大范围卫星图像库匹配的过程中,需要对卫星地图进行多次特征提取,这会占用大量的计算资源。因此,为节省计算资源、降低视觉定位的延迟,在完成卫星地图下载后,本文采用 SuperPoint 稀疏特征提取模型对卫星地图进行关键点提取和描述符生成,并将提取出的关键点坐标及特征描述符保存到无人机的机载硬盘中作为离线地图数据库。这意味着在每次进行图像匹配时,只需读取预先提取并保存在硬盘中的特征参数,无须每次都对所有瓦片地图进行特征提取。如图 2 所示, SuperPoint 模型对卫星地图进行稀疏特征提取(图 2 中白色圆点为得到的关键点),随后将关键点对应的坐标与描述符存入离线地图数据库。

1.2 基于 LPIPS 的初始检索

在无人机坐标丢失的情况下,直接将航拍图像与大量的卫星图像匹配需要耗费大量时间和计算资源。因此在进行稀疏特征匹配之前,可通过计算航拍图像与卫星图像之间的相似性来确定航拍图像对应的卫星图像范围,从而过滤掉差异较大的卫星图像。

在计算机视觉的研究中,图像相似度感知是一个关键问题。传统的度量标准,如 L2 距离、峰值信噪比(Peak signal-to-noise ratio, PSNR)和结构相似性指数(Structural similarity index measure, SSIM)等,其计算主要基于像素或特征的统计性质^[17],因此在航拍图像与遥感图像之间的跨域图像相似度感知任务中表现不佳。为了解决这

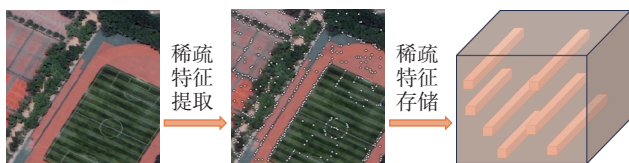


图 2 卫星地图预处理示意图

Fig.2 Schematic diagram of satellite map preprocessing

个问题,本文采用 AlexNet 架构的 LPIPS 指标对航拍图像与卫星图像的相似度进行评估,筛选出最相似的 50 张卫星图像作为初始检索结果,送入后续的稀疏特征匹配模块。LPIPS 是一种基于深度学习方法的图像质量评估指标,与传统基于像素差异或统计差异的评估方法不同, LPIPS 通过在大型图像数据集上训练的神经网络,来学习图像感知差异的度量方式。常见的 LPIPS 指标包括 AlexNet 架构和 VGG 架构两种^[18],其中 AlexNet 架构比 VGG 架构的运行速度快,但精度低。AlexNet 架构在 500 × 500 分辨率的 ALTO Offset_0_None 验证集上的测试精度与 VGG 架构相比,仅相差 1.6%,运行速度相差约 3 倍。因此,为了满足无人机视觉定位任务的实时性要求,本文采用了更加轻量级的 AlexNet 架构 LPIPS 指标,通过卷积神经网络提取航拍图像与卫星图像的高维特征表征,并计算特征向量距离得到相似度分数,其中数值越低,表征图像间相似度越高。图 3 (卫星图像左上角标注的红色数值为 LPIPS 相似度分数)展示了基于 LPIPS 指标的前 10 个检索结果,按相似度分数由左至右、由上至下的顺序排列,其中卫星图像边框色度反映该卫星图像与航拍图像的实际地理定位偏差。

1.3 基于 SuperPoint+Lightglue 的图像匹配

经过初始检索后,本文将航拍图像与 50 张候选卫星图像进行匹配。考虑到无人机视觉定位任务计算资源受限与实时性的要求,本文将图像匹配分为稀疏特征提取与特征匹配两个模块。相较于密集特征方法,稀疏特征方法计算量更低、运行速度更快。此外,得益于稀疏特征方法分阶段计算的特性,本文能够对卫星地图进行特征预处理,从而进一步降低无人机飞行过程的计算负载。

在稀疏特征提取模块中,本文采用基于全卷积神经网络(Fully convolutional network, FCN)架构的 SuperPoint 网络作为图像关键点检测与特征描述子生成的核心算法。SuperPoint 是一种常用的基于深度学习的

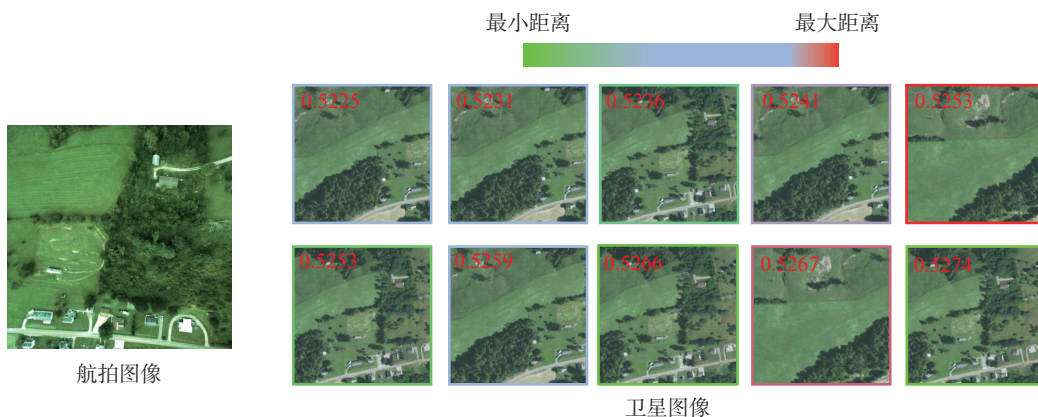


图 3 LPIPS 初始检索结果(前 10 张)

Fig.3 LPIPS initial retrieval results (top 10 images)

特征提取模型,能够通过神经网络自主学习图像中的关键点和对应描述符,因此具有更强的泛化能力。相较于 SIFT、ORB 等传统的特征提取算法^[19-20], SuperPoint 算法可以生成更密集、更正确的匹配。同时由于其 FCN 架构的特性,相较于 CurriculumLoc 方法中基于 Swin Transformer 架构的特征提取模块, SuperPoint 更适合部署在无人机机载计算平台。图 4 为稀疏特征提取可视化结果,白色实心圆点代表被提取出的特征点,可以看出, SuperPoint 模型在两幅图像上提取的特征点分布相似,均主要集中在房屋、道路、树木等具有明显特征的区域。

在稀疏特征匹配模块中,本文选用 LightGlue 网络作为图像特征匹配的核心算法,通过特征点的位置信息与描述符在两个图像之间建立准确的特征点对应关系。作为稀疏匹配领域的最新技术, LightGlue 在内存和计算方面更高效、更准确。在运算速度上, LightGlue 远快于密集特征匹配器,更符合无人机视觉导航任务的实时性要求。同时在稀疏匹配器中, LightGlue 在预测精度与推理时间上都极具竞争力。图 5 展示了稀疏特征匹配可视化结果,绿色实心圆点为成功匹配的特征点,绿色线段表示其对应关系。

在完成稀疏特征匹配后,本文采用 Hu 等^[10]研究中

的重排序策略,根据两张图像匹配点的平均像素距离对 50 张候选图像进行重排序,将平均距离最小的图像视为匹配结果。图 6 展示了稀疏特征匹配重排序后的前 10 个检索结果,按匹配点平均像素距离由左至右、由上至下依次递增排列,可以看出,经稀疏特征匹配后排名靠前的卫星图像更接近正确结果。

在完成重排序后,通过 RANSAC 算法计算航拍图像与排名第 1 位的卫星图像之间的单应性变换矩阵,解算出航拍图像中心点在卫星图像中对应位置^[21],如图 7 所示(绿色实心圆点为图像中心点,红色实心圆点为经航拍图像中心点单应性变换后估算的无人机坐标点)。最后根据卫星地图地理坐标信息及空间分辨率转换为最终的无人机地理坐标。

2 试验与结果分析

2.1 测试数据集

为验证本文方法的有效性,本文分别在公开数据集 ALTO 和自采航拍影像数据集上进行定量试验和定性试验。ALTO 数据集由直升机航拍数据和卫星地图数据构成,包括两次大型飞行(150 km 和 260 km),覆盖多样地形,如森林、城市、乡村、河流、湖泊和平原,如图



(a) 航拍图像 (b) 卫星图像

图 4 稀疏特征提取可视化结果

Fig.4 Visualization results of sparse feature extraction

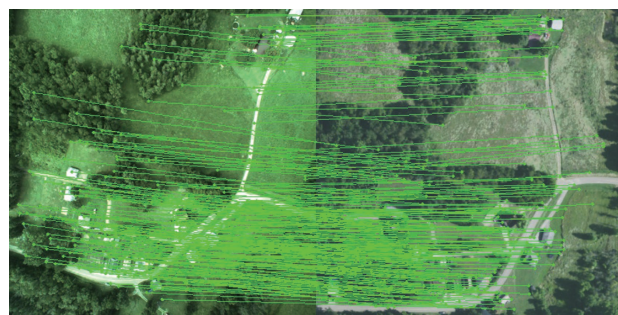


图 5 稀疏特征匹配可视化结果

Fig.5 Visualization results of sparse feature matching

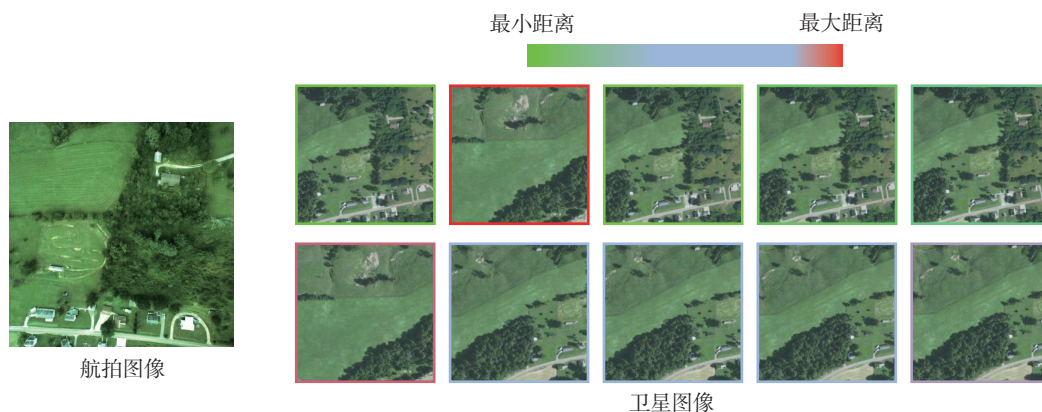


图 6 稀疏特征匹配重排序结果(前 10 张)

Fig.6 Visualization results of sparse feature matching reranking (top 10 images)

8 所示。该数据集旨在为无人机视觉位置识别和定位方法的研发以及基准测试提供丰富、多样化的试验场景和数据支持。其中航拍数据由直升机机载传感器收集,包括高精度 GPS-INS 地面真实位置数据、高精度加速度计读数、激光高度计读数和垂直向下的 RGB 图像。

自航拍影像数据集由 714 对航拍图像与卫星参考图像构成,航拍数据由大疆无人机记录,包括城市、道桥、植被、水域、农田 5 种常见场景,以及光照变换、视点变换、纹理变换 3 种具有挑战性的变换条件,如图 9 所示。该数据集旨在模拟无人机真实飞行环境,以充分验证算法在实际应用场景的有效性与其鲁棒性。

2.2 评价指标

在图像检索任务中, Recall@K (R@K) 是最常用的评价指标^[22-23],因此本文选择 R@K 作为 ALTO 数据集上试验结果的评价指标。R@K 表示检索结果的前 K 个候选图像中,包含正确检索图像的平均准确率,即

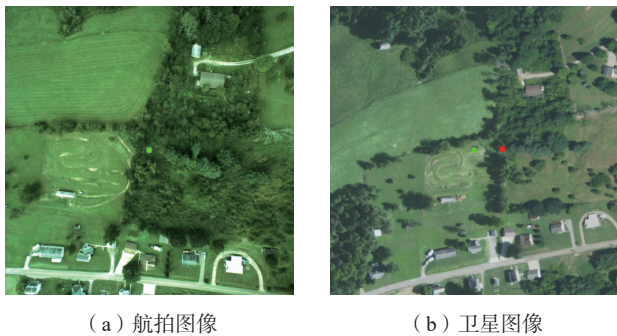


图 7 无人机坐标解算可视化结果

Fig.7 Visualization results of UAV coordinate calculation



图 8 ALTO 数据集图像示例

Fig.8 Examples of ALTO dataset images

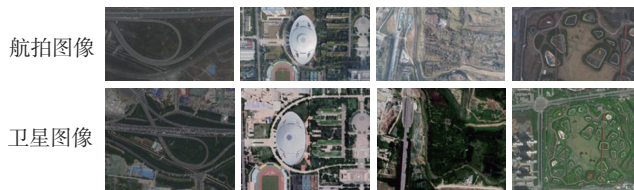


图 9 实际航拍影像示例

Fig.9 Examples of real-world aerial images

$$R@K = \frac{1}{N} \sum_{i=1}^N \min \left(\sum_{j=1}^K S(q_i, r_j), 1 \right) \quad (1)$$

其中

$$S(q_i, r_j) = \begin{cases} 1 & \text{if } d(q_i, r_j) < \text{dis} \\ 0 & \text{if } d(q_i, r_j) \geq \text{dis} \end{cases} \quad (2)$$

式中, q_i 表示第 i 张航拍图像; r_j 表示第 j 张卫星图像; N 为航拍图像数量; K 为卫星图像数量; d 表示两张图像的真实地理距离; dis 表示阈值距离,在试验中取值 20 m。

在无人机视觉定位任务中,通常将候选卫星图像中的第 1 位作为后续图像配准的绝对参考,因此 R@1 值是试验的重要评价指标。

2.3 测试结果

将本文方法与国内外先进方法进行对比,在 ALTO 验证集上进行测试验证,本文方法使用的网络模型均采用公开的预训练模型,将单应性变换后得到的地理坐标作为候选图像地理坐标。试验结果如表 1 所示,在阈值设定为 20 m 时,本文方法在 R@1 这一项无人机定位任务的重要指标上达到 79.8%,超过 CurriculumLoc 方法 17.2%,同时与其他指标的差距在 6% 左右,证明了本文方法的先进性。

鉴于无人机视觉定位任务的延迟敏感特性,算法的运算速度也是重要的评价指标,因此在 ALTO 验证集上将本文方法与国内外先进方法的运算速度进行对比,试验内容为单张航拍图像与 2295 张卫星图像匹配的运算时间,如图 10 所示。为保证测试结果的公平性,运算速度测算的所有步骤均在 Intel (R) Core (TM) i5-12600KF CPU@3.70 GHz, NVIDIA RTX 3060Ti GPU, 32 GB 内存的硬件平台上进行。从图 10 可以看出,本文方法在运算速度上有显著优势。

表 1 与其他先进视觉定位方法的对比结果 (DIST=20 m)

Table 1 Comparison results with other state-of-the-art visual localization methods (DIST=20 m)

方法	ALTO			
	R@1/%	R@5/%	R@10/%	R@20/%
NetVLAD	20.4	58.3	77.1	88.4
Patch-NetVLAD-2048	33.9	78.3	91.2	95.6
Patch-NetVLAD-4096	31.1	74.8	90.4	94.9
DELG	35.9	79.9	92.3	95.6
TransVLAD	34.7	78.8	91.9	95.3
CurriculumLoc	62.6	92.6	95.6	96.1
本文方法	79.8	87.5	89.6	90.7

为验证本文方法中各模块的有效性,从ALTO验证集随机抽取200张航拍图像,在此基础上开展消融试验,结果如表2所示。SPLG、SMP分别为稀疏特征匹配模块与卫星图像预处理模块,统计单张航拍图像与2295张卫星图像匹配的计算时长。仅使用SPLG模块,R@10与R@20的两项指标达到91.5%与95.0%,但在R@1这一重要指标上仅为37.0%,同时所需计算时间也较长,这一结果表明,SPLG模块可以很好地将正确图像对匹配,但也会对相似纹理区域产生大量误匹配,进而影响R@1的精度。在引入AlexNet架构的LPIPS模块进行初始检索后,与SPLG方法相比,R@1指标提升40.5%,这一结果表明,初始检索模块可以有效滤除大量错误卫星图像,减少稀疏特征匹配模块的误匹配率,在精度提升的同时大大缩短了计算时间。在此基础上,本文提出的SMP模块在不影响精度的情况下,将计算时长压缩至22.5 s。

为验证本文方法的有效性与鲁棒性,在自采的实际航拍数据上对无人机视觉定位任务中常见的5种场景与3种具有挑战性的变换条件进行了定性测试,结果如图11所示(左侧为航拍图像,右侧为卫星图像;绿色线段表示

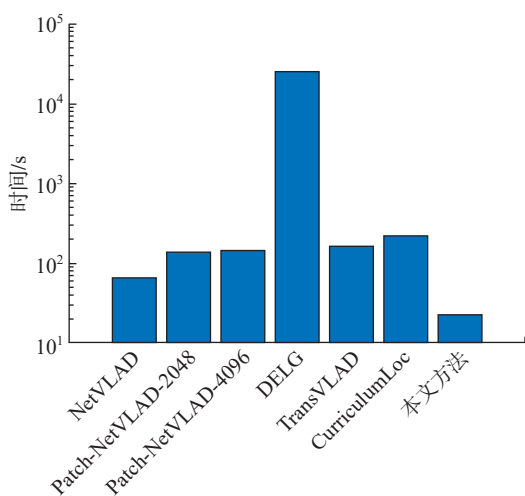


图10 运算速度对比结果

Fig.10 Comparison results of computational speed

表2 消融试验结果 (DIST=20 m)

Table 2 Results of ablation experiments (DIST=20 m)

方法	ALTO				时间/s
	R@1/%	R@5/%	R@10/%	R@20/%	
SPLG	37.0	82.0	91.5	95.0	95.0
SPLG+SMP	37.0	82.0	91.5	95.0	64.9
SPLG+LPIPS (Alex)	77.5	85.5	87.5	88.0	24.1
SPLG+LPIPS (Alex)+SMP	77.5	85.5	87.5	88.0	22.5

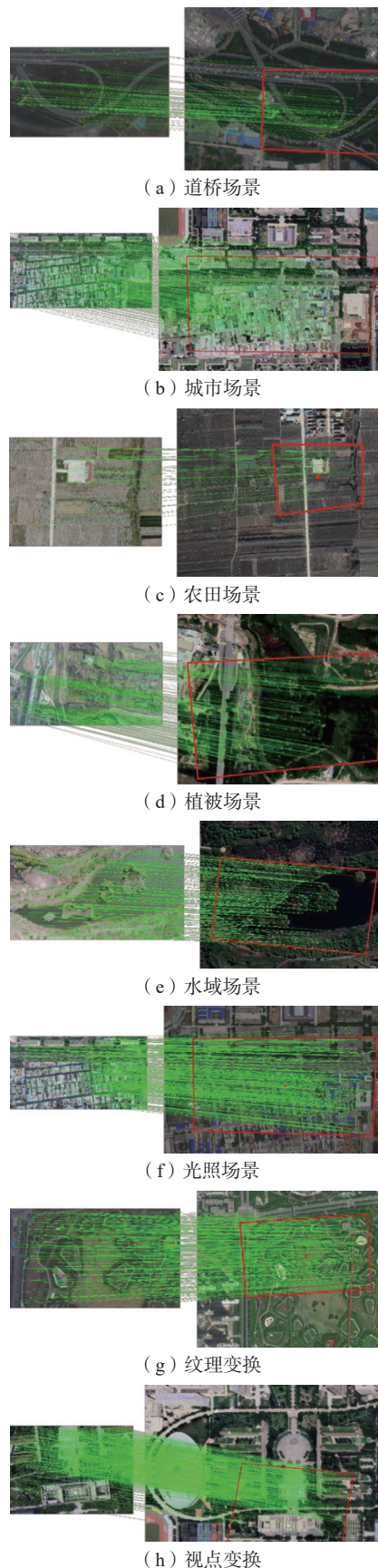


图11 实际航拍影像测试可视化结果

Fig.11 Visualization results of real-world aerial image tests

稀疏特征匹配模块得到的特征点对应关系,红色方框为两张图像共视区域,红色实心圆点为航拍图像中心视点在卫星图像上的映射)。从图 11 可以看出,本文方法在各种场景及变换条件下均能很好地实现图像配准,证明了本文方法在无人机视觉定位任务上的有效性与鲁棒性。

3 结论

本文针对 GNSS 拒止环境下的无人机视觉地理定位任务,提出了一种基于稀疏特征匹配算法的解决方案,在无人机视觉定位任务的重要指标 $R@1$ 下,该方案的定位精度显著优于其他现有方法。

(1) 本文提出了一种卫星地图预处理方法,可有效降低无人机飞行期间的计算负载,提升系统推理速度。

(2) 采用 LPIPS 指标快速过滤无用卫星图像,显著降低后续特征匹配模块计算量,降低误匹配率。

(3) 采用稀疏特征匹配方法,相较于密集匹配方法计算量更小,且基于卷积神经网络的特征提取模块更易于移动端计算平台的部署。

参考文献

[1] 张涛, 芦维宁, 李一鹏. 智能无人机综述[J]. 航空制造技术, 2013, 56(12): 32–35.

ZHANG Tao, LU Weining, LI Yipeng. Intelligent UAV[J]. Aeronautical Manufacturing Technology, 2013, 56(12): 32–35.

[2] 陶于金, 李沛峰. 无人机系统发展与关键技术综述[J]. 航空制造技术, 2014, 57(20): 34–39.

TAO Yujin, LI Peifeng. Development and key technology of UAV[J]. Aeronautical Manufacturing Technology, 2014, 57(20): 34–39.

[3] PSIAKI M L, HUMPHREYS T E. GNSS spoofing and detection[J]. Proceedings of the IEEE, 2016, 104(6): 1258–1270.

[4] XU Y F, SHAMSOLMOALI P, GRANGER E, et al. TransVLAD: Multi-scale attention-based global descriptors for visual geolocalization[C]//2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). Waikoloa, Piscataway, NJ: IEEE, 2023: 2839–2848.

[5] ARANDJELOVIC R, GRONAT P, TORII A, et al. NetVLAD: CNN architecture for weakly supervised place recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, Piscataway, NJ: IEEE, 2016: 5297–5307.

[6] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 2002, 86(11): 2278–2324.

[7] CAO B Y, ARAUJO A, SIM J. Unifying deep local and global features for image search[C]//Computer Vision—ECCV 2020. Cham: Springer International Publishing, 2020: 726–743.

[8] HAUSLER S, GARG S, XU M, et al. Patch-NetVLAD: Multi-scale fusion of locally-global descriptors for place recognition[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, IEEE, 2021: 14136–14147.

[9] VASWANI A, SHAZEER N M, PARMAR N, et al. Attention is

all you need[C]//Neural Information Processing Systems. New York, 2025.

[10] HU B N, CHEN L, CHEN R J, et al. CurriculumLoc: Enhancing cross-domain geolocalization through multistage refinement[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 5615914.

[11] CISNEROS I, YIN P, ZHANG J, et al. ALTO: A large-scale dataset for UAV visual place recognition and localization[EB/OL]. 2022: 2207.12317. <https://arxiv.org/abs/2207.12317v1>.

[12] LIU Z, LIN Y T, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, QC, Piscataway, NJ: IEEE, 2021: 9992–10002.

[13] KRIZHEVSKY A. One weird trick for parallelizing convolutional neural networks[EB/OL]. 2014: 1404.5997. <https://arxiv.org/abs/1404.5997v2>.

[14] ZHANG R, ISOLA P, EFROS A A, et al. The unreasonable effectiveness of deep features as a perceptual metric[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, Piscataway, NJ: IEEE, 2018: 586–595.

[15] 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, Piscataway, NJ: IEEE, 2018: 337–33712.

[16] LINDENBERGER P, SARLIN P E, POLLEFEYS M. LightGlue: Local feature matching at light speed[C]//2023 IEEE/CVF International Conference on Computer Vision (ICCV). Paris, Piscataway, NJ: IEEE, 2023: 17581–17592.

[17] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: From error visibility to structural similarity[J]. IEEE Transactions on Image Processing, 2004, 13(4): 600–612.

[18] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[EB/OL]. 2014: 1409.1556. <https://arxiv.org/abs/1409.1556v6>.

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

[20] RUBLEE E, RABAU D V, KONOLIGE K, et al. ORB: An efficient alternative to SIFT or SURF[C]//2011 International Conference on Computer Vision. Barcelona, Piscataway, NJ: IEEE, 2011: 2564–2571.

[21] FISCHLER M A, BOLLES R C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography[J]. Readings in Computer Vision, 1987, 30(6): 726–740.

[22] ZHENG L, SHEN L Y, TIAN L, et al. Scalable person re-identification: A benchmark[C]//2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Piscataway, NJ: IEEE, 2015: 1116–1124.

[23] 秦剑琪, 蓝朝桢, 崔志祥, 等. 一种面向无人机绝对定位的卫星基准影像检索方法[J]. 武汉大学学报(信息科学版), 2023, 48(3): 368–376.

QIN Jianqi, LAN Chaozhen, CUI Zhixiang, et al. A satellite reference image retrieval method for unmanned aerial vehicle absolute positioning[J]. Geomatics and Information Science of Wuhan University, 2023, 48(3): 368–376.

通信作者: 罗远哉, 教授, 博士, 研究方向为多媒体检索与人工智能。

(责编 七七)