

基于并行化处理的无人机影像三维重建算法

陈怀圆^{1,2}, 党建武^{1,2*}, 岳彪^{2,3}, 杨景玉³

¹兰州交通大学光电技术与智能控制教育部重点实验室, 甘肃 兰州 730070;

²轨道交通信息与控制国家级虚拟仿真实验教学中心, 甘肃 兰州 730070;

³兰州交通大学电子与信息工程学院, 甘肃 兰州 730070

摘要 针对增量运动恢复结构(SFM)算法在重建大规模无人机影像数据集时效率低、易产生场景漂移的问题,提出一种可并行化处理的增量SFM重建算法。首先,利用词汇树图像检索结果约束图像特征匹配的空间搜索范围,提高图像特征匹配的效率和精度。其次,综合考虑特征匹配数量和无人机平台获取的全球定位系统(GPS)信息构建无向加权场景图,并选用归一化割算法将场景图划分为多个相互重叠的子集。然后,将每个子集分布在多核CPU上并行执行增量SFM重建算法。最后,基于子集间公共重建点和强相关子集优先被合并的策略实现子集合并。此外,结合GPS信息为光束法平差(BA)代价函数添加位置约束项,有效消除每次执行BA优化引入的误差。为了验证所提算法的有效性,在3个无人机数据集上进行实验,实验结果表明,所提算法相比原始增量SFM重建算法不仅显著提高了位姿估计和场景重建的效率,而且合理优化了重建结果的精度。

关键词 增量运动恢复结构; 光束法平差; 词汇树图像检索; 归一化割; 场景合并

中图分类号 TP751

文献标志码 A

DOI: 10.3788/LOP231339

Three Dimensional Reconstruction Algorithm of Unmanned Aerial Vehicle Images Based on Parallel Processing

Chen Huaiyuan^{1,2}, Dang Jianwu^{1,2*}, Yue Biao^{2,3}, Yang Jingyu³

¹Key Laboratory of Optoelectronic Technology and Intelligent Control, Ministry of Education, Lanzhou Jiaotong University, Lanzhou 730070, Gansu, China;

²National Virtual Simulation Experimental Teaching Center of Rail Transit Information and Control, Lanzhou 730070, Gansu, China;

³School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, Gansu, China

Abstract A parallelizable incremental structure from motion (SFM) recovery reconstruction algorithm is employed to address low efficiency and susceptibility to scene drift when reconstructing large-scale unmanned aerial vehicle image datasets. First, the vocabulary tree image retrieval results are used to constrain the spatial search range and improve the efficiency of image feature matching. Second, by considering the feature matching number and the global positioning system (GPS) information obtained by the drone platform, an undirected weighted scene map is constructed, and a normalized cut algorithm is selected to divide the scene map into multiple overlapping subsets. Further, each subset is distributed on multicore central processing units (CPUs), and the incremental SFM reconstruction algorithm is executed in parallel. Finally, based on the strategy of common reconstruction points between subsets and priority merging of strongly correlated subsets, subset merging is achieved. In addition, combining GPS information to add positional constraints to the bundle adjustment (BA) cost function eliminates the errors introduced by each BA optimization execution. To verify the effectiveness of the algorithm, experiments are conducted on three unmanned aerial vehicle datasets. The experimental results show that the proposed algorithm not only significantly improves the efficiency of pose estimation and scene reconstruction compared with the original incremental SFM reconstruction algorithm but also reasonably optimizes the accuracy of the reconstruction results.

Key words increasing structure from motion; bundle adjustment; vocabulary tree image retrieval; normalized cut; scene merging

收稿日期: 2023-05-19; 修回日期: 2023-06-29; 录用日期: 2023-07-24; 网络首发日期: 2023-08-15

基金项目: 2022年度中央引导地方科技发展资金项目(22ZY1QA002)、甘肃省教育科技自主创新项目(2021jyjbgs-05)、甘肃省军民融合专项(2020JG01)、甘肃省重点研发计划(21YF5GA158)、甘肃省知识产权计划项目(21ZSCQ013)

通信作者: *dangjw@mail.lzjtu.cn

1 引言

近年来,运动恢复结构(SFM)已成为无人机影像三维重建的核心技术^[1]。与传统的 positioning and orientation system (POS) 辅助空中三角测量(AT)不同,SFM 技术不需要任何辅助信息就可以同时从重叠无序的无人机图像中恢复相机位姿和三维空间点。根据参数初始化策略,SFM 可分为 3 类:增量 SFM、全局 SFM 和混合型 SFM^[2]。与其他方法相比,增量 SFM 通过添加新注册图像进行参数初始化和迭代执行光束法平差(BA)优化参数,对异常值具有较强的鲁棒性,但这种方法牺牲了重建效率^[3]。对此,一些学者对增量 SFM 算法进行了改进。Sun 等^[4]提出一种简化的光束法平差模型(RBA),该模型减少未知参数的数量和法向方程的维数,大大降低计算复杂度和存储开销。Zheng 等^[5]提出一种快速准确的 BA 优化方法,在 LM 算法框架下,同时采用预处理共轭梯度(PCG)迭代求解方程,通过点重采样和法矩阵压缩来降低计算复杂度。Chen 等^[6]将大规模 SFM 问题看作一个图问题,在场景图划分和子集合并两个步骤分别构造图,引导子集扩展和合并。尹双双等^[7]利用无人机平台获取的全球定位系统(GPS)信息,优化影像匹配、BA 等步骤,有效提高重建模型精度。郭复胜等^[8]提出一种基于辅助信息的无人机图像批处理三维重建算法,在效率方面大幅提高。Ye 等^[9]提出一种新颖高效的 BA 求解器 CoLi-BA,通过紧凑的线性化和重排序显著提高 BA 优化的速度。

尽管这些方法对增量 SFM 重建的改进是有效的,但在处理大规模无人机数据集时,增量 SFM 重建算法仍存在一些问题。一方面,对于特征匹配,为了

构建稳定的几何约束,无人机需获取较多数量的多视图图像并且在航向和旁向轨迹配置高重叠度,导致图像搜索匹配对的组合复杂度变高,特征匹配效率降低。另一方面,由于 BA 优化随新注册图像不断迭代执行,BA 优化的效率随图像数量的增加而显著降低。此外,由于前期特征匹配工作误差的影响,每次迭代执行 BA 优化都会引入新误差,可能因为误差累积过大使得全局 BA 优化无法进行,最终出现场景漂移问题。

针对上述问题,本文对原始增量 SFM 重建算法进行改进,提出一种可并行化处理的增量 SFM 重建算法。通过词汇树检索在为每张图像选取相关图像集后进行特征匹配计算、场景图构造,并将场景图划分为多个重叠子集并行处理,根据子集间公共重建点对各子集的重建模型进行合并。结合无人机平台获取的 GPS 信息为光束法平差代价函数引入绝对 GPS 误差项。实验结果表明,所提算法不仅显著提升了重建程序的效率,而且合理地提高了重建精度。

2 增量 SFM 算法

增量 SFM 重建的总体流程如图 1 所示,主要分为两个阶段,同名点获取和增量重建。第 1 阶段的目的是为每张图像搜索可靠、分布良好的特征匹配对,第 2 阶段的目的是估计相机位姿和三维空间点坐标。首先,通过特征提取技术计算出图像特征点的位置以及对应描述子;然后,对每张图像进行特征匹配。尺度不变特征变换(SIFT)由于其对图像旋转、尺度差异以及光照变化的不变性成为被广泛应用的特征提取与匹配技术。所提算法采用 SIFTGPU^[10]进行特征提取与匹配。

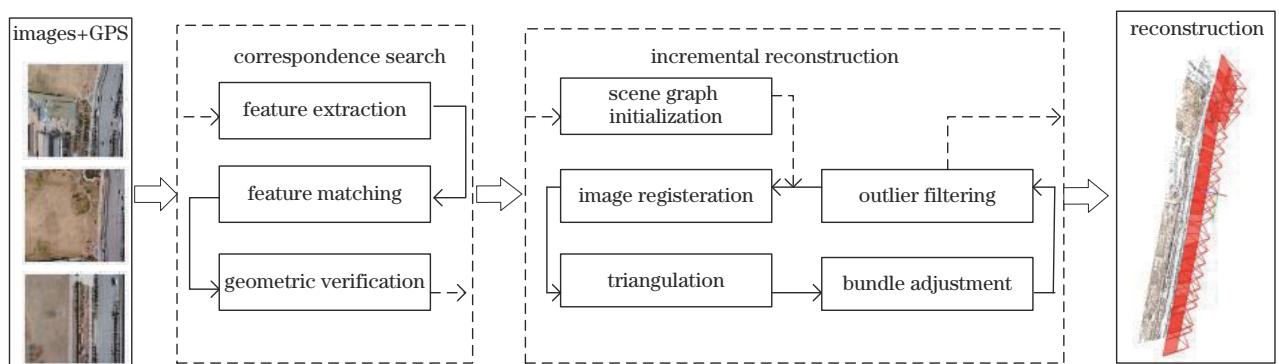


图 1 增量 SFM 重建流程

Fig. 1 The incremental SFM reconstruction process

获取同名点后,执行增量 SFM 恢复相机位姿信息与场景三维空间点。首先,将第 1 阶段经过几何验证后的匹配对表示为场景图,场景图的顶点表示一张待重建图像,边连接几何验证后的图像对。然后,选择两个具有足够多均匀匹配的图像对进行增量 SFM 重建实现初始化,得到初始重建模型。其次,不断从剩余图

像中选择新的注册图像,并通过三角化求解出新加入图像的位姿信息和三维空间点信息,同时交替执行局部和全局光束法平差对恢复出的相机位姿和三维空间点进行优化。

针对局部和全局光束法平差,将优化恢复出的相机位姿和三维空间点的问题公式化为最小化重投影误

差,即最小化恢复出的三维空间点投影到二维像面的投影点与其对应特征点之间的距离。图 2 为重投影误差示意图。

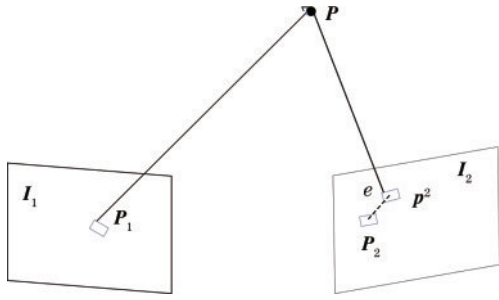


图 2 重投影误差示意图

Fig. 2 Diagram of the reprojection error

图 2 中,特征点 P_1 和 P_2 是图像 I_1 和 I_2 上的一对匹配点,由匹配点 P_1 、 P_2 恢复出的三维空间点为 P ,利用图像 I_1 对应的相机内参矩阵可将点 P 投影到图像 I_2 上得到二维投影点 p^2 。投影点 p^2 与特征点 P_2 的欧氏距离即

为重投影误差。因而光束法平差的主要任务是通过迭代更新相机位姿和三维场景点坐标来达到最小化重投影误差的目的。BA 优化的代价函数为

$$E_{\text{vis}} = \min_{C_j, X_i} \sum_{i=1}^n \sum_{j=1}^m \rho_{ij} \| p(C_j, X_i) - x_{ij} \|^2, \quad (1)$$

式中: X_i 、 C_j 分别表示恢复出的三维空间点 i 的坐标和相机 j 的参数; $p(C_j, X_i)$ 表示三维空间点 X_i 在相机 C_j 视角下的投影点; x_{ij} 是对应的特征点坐标; $\|\cdot\|$ 表示 L^2 范数; ρ_{ij} 是用于减小离群值影响的指示函数,如果点 X_i 在相机 C_j 中可见,则 $\rho_{ij}=1$, 否则 $\rho_{ij}=0$ 。

3 所提方法

考虑到相关性低甚至不相关的两张图像之间几何约束性较小,在原始增量 SFM 算法的基础上,引入词汇树图像检索技术并结合无人机平台获取的 GPS 信息提出可并行化处理的增量 SFM 算法,具体流程如图 3 所示:

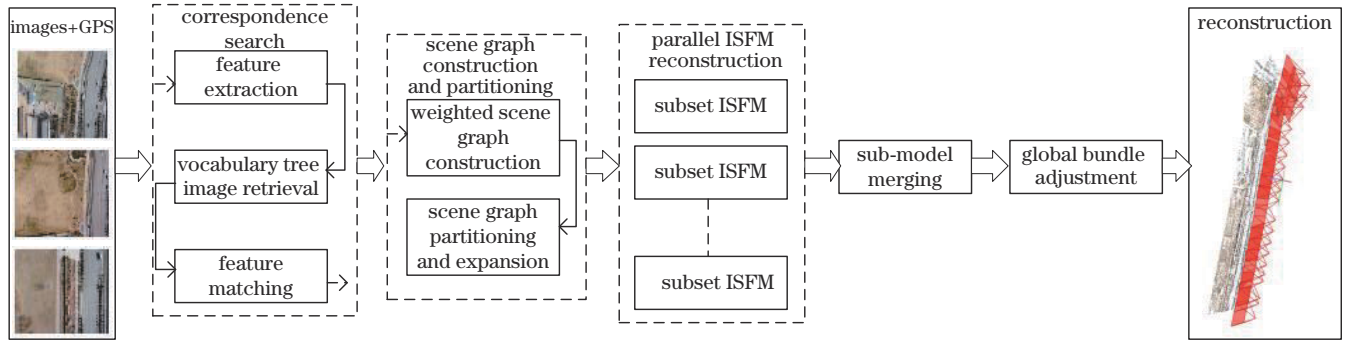


图 3 并行增量 SFM 重建算法流程

Fig. 3 Flow of the parallel incremental SFM reconstruction algorithm

首先,对带有 GPS 信息的多视无人机图像进行同名点获取。与原算法不同的是,在进行特征匹配前,采用词汇树图像检索技术为每张待匹配图像选取相关图像集,使每张待匹配图像只与所选择的相关图像进行特征匹配计算,相比对任意两张图像均进行匹配计算,利用词汇树检索结果约束空间搜索范围的匹配对选择策略可避免与一些相关性低甚至不相关图像进行匹配计算,降低匹配对的组合复杂度,提高匹配效率。然后,在词汇树检索基础上综合考虑特征匹配数量和无人机平台获取的 GPS 信息构建无向加权场景图,并对场景图进行划分与扩展,之后每个子集分布在多核 CPU 上并行执行增量 SFM。这样不仅可以充分利用计算资源,而且光束法平差步骤的优化参数随图像规模的减小显著降低,从而加速重建过程。最后,基于子集间的公共重建点对各子集重建结果进行合并。此外,由于重建出的各个场景存在不同的误差,所以对合并后的整体模型执行全局光束法平差进行优化。

3.1 基于词汇树图像检索场景图构造

场景图作为并行增量 SFM 重建划分与合并的基本结构,在构建场景图前,采用基于词汇树的图像检索方法进行相关图像对的选取,为场景图构造提供可靠的图像匹配对。词汇树图像检索技术的核心思想是为每张查询图像选择与其相似性得分较高的图像,并利用倒排索引技术加速相似性计算,每张图像通过词汇树索引表示为一个 bag-of-words (BoW) 向量,将寻找匹配对的问题转化为检索与查询图像之间具有最接近 BoW 向量的图像^[11-12]。使用一个预训练与发布开放的词汇树,它有 256000 个视觉词汇,这些视觉词使用来自互联网社区的图像进行训练,如图 4 所示,黄色边框的图像为查询图像,其余绿色边框图像为黄色边框图像检索到的相关图像。

在进行相关图像对选择后,对每个查询图像与其所检索数据库中的相关图像使用特征匹配算法进行匹配。然后,开始场景图构建,场景图的节点表示图像,边表示所连图像之间的关系。对场景图进行扩展,即综合考虑图像间特征匹配数量和无人机平台提供的



图 4 查询图像基于词汇树的检索结果

Fig. 4 The retrieval results of the query image based on the vocabulary tree

GPS 信息为场景图的边计算权重,增加对应边 e_{ij} 的连接强度,具体表示如式(2)所示:

$$W(e_{ij}) = R_{ew} W_{inlier} + (1 - R_{ew}) W_{gps}, \quad (2)$$

式中: e_{ij} 表示图像 i 和 j 相连的边; W_{inlier} 表示根据特征匹配的数量为场景图中边计算的权重; W_{gps} 表示根据 GPS 坐标信息为场景图的边计算的权重,两张图像的 GPS 坐标距离越近,权重越高; R_{ew} 为两项的权重比,大小位于 0~1 之间,本研究中设定为 0.5。

3.2 场景图划分与并行重建

场景图构造完成后,选用归一化割(Ncut)算法^[13]对场景图进行划分,并通过 Zhu 等^[14]提出的两个约束条件对划分过程进行约束:大小约束给出每个子集中

图像的上界;完整性约束定义为 $\eta_{(i)} = \frac{\sum_{j \neq i} |S_i \cap S_j|}{S_i}$, 其

中, S_i 和 S_j 分别代表子集 i 和 j , 保证子集间有足够的公共图像进行可靠合并。归一化割算法易于切割权值较小的边,生成内部连接紧凑的子集。定义场景图 $G = (V, E)$, 其中, V 代表顶点, E 代表边。假设需要将场景图划分为互不相交的两个子集 A 和 B , 那么归一化割可表示为

$$Ncut(A, B) = \frac{cut(A, B)}{assoc(A, V)} + \frac{cut(A, B)}{assoc(B, V)}, \quad (3)$$

$$assoc(A, V) = \sum_{u \in A, t \in V} \omega(u, t), \quad (4)$$

$$assoc(B, V) = \sum_{v \in B, t \in V} \omega(v, t), \quad (5)$$

式中: $cut(A, B)$ 为子集 A 和子集 B 所有相连边的权重和; $assoc(A, V)$ 和 $assoc(B, V)$ 分别为子集 A 和子集

B 中的点与场景图中的所有其他点的相连边的权重和。

场景图划分后,每张图像仅被分配到一个子集中。归一化割的结果如图 5 所示。

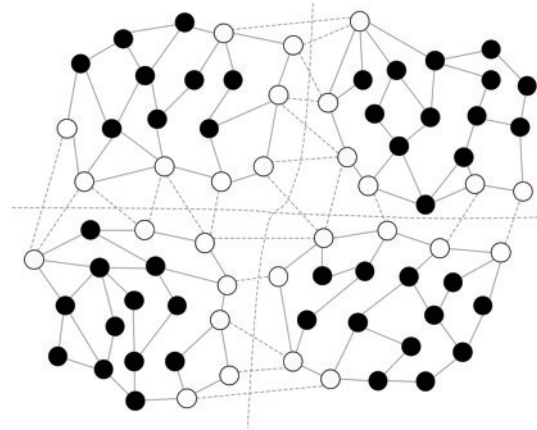


图 5 归一化割结果

Fig. 5 Normalized cut result

场景图划分步骤中,设定每个子集最大节点数为 K , 迭代执行归一化割算法,将超过设定最大节点数 K 的任何子集 S_i 划分为两个互不重叠的子集 S_{i1}, S_{i2} , 直到所有子集包含节点数都不超过 K 。此外,为了方便子集合并,对划分后的每个子集进行扩展,扩展结果如图 6 所示,扩展步骤中,将相邻两个子集之间在归一化割步骤中丢失的边 E_{ij} 按边权重 W_{ij} 降序排序,然后将排名前 N 的边 E_{ij} 所连接的两个顶点 V_i, V_j 同时添加到相邻的两个子集中。通过上述步骤划分场景图,将相连边权重更高的图像划分到同一子集中,不同子集

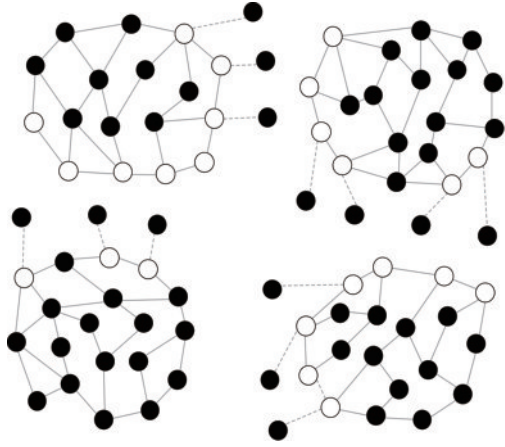


图 6 子集扩展

Fig. 6 Subset expansion

之间具有更小的相关性,不仅保证各子集被分配到多核 CPU 上并行执行增量 SFM 而且也后续每个子集执行增量 SFM 提供足够的重叠度。

3.3 子集合并

并行重建后,每个子集重建的三维模型都在独立的坐标系下,因此需要将每个子集的重建模型配准到同一坐标系下生成完整的重建模型。假设模型 Q_s 和 Q_r 分别为两个重建子集,并且这两个子集之间有 n 个公共重建点,分别用 $\{s_i\}$ 和 $\{r_i\}$ 表示,定义 Q_r 为参考合并子集,则从子集 Q_s 到子集 Q_r 的变换由相似性变换

$T = [\lambda R | t]$ 表示。 T 可以通过最小化目标函数来计算,具体如式(6)所示:

$$e^2(\mathbf{R}, \mathbf{t}, \lambda) = \frac{1}{n} \sum_{i=1}^n \|r_i - (\lambda R s_i + \mathbf{t})\|^2, \quad (6)$$

式中: \mathbf{R} 是旋转矩阵; \mathbf{t} 是平移向量; λ 是缩放参数; e^2 是在变换 T 下将结构 $\{s_i\}$ 变换为 $\{r_i\}$ 的均方误差 (MSE), 可以使用 SVD 算法^[15] 进行鲁棒估计。

根据子集间几何相关性确定子集合并对和合并顺序,选择与其他子集最相关的子集作为参考合并子集,并按照强相关子集优先被合并的策略进行合并,这里,通过子集间公共重建点的数量表示各子集间的几何相关性。例如,如果子集及其邻近子集之间有效重建点数量为 $Q_1, Q_2, Q_3, \dots, Q_r$, 并根据几何相关性指数降序排列。将 Q_1 作为初始参考合并子集,首先将 Q_1 与其最相关子集 Q_2 合并并标记为已处理,合并后的子集命名为 Q_{12} , 然后将 Q_{12} 作为新的参考合并子集,继续寻找下一个最佳目标合并子集 Q_i , 如果一个子集与新参考合并子集 Q_{12} 之间有最高的几何相关性,则该子集被认为是下一个最佳目标合并子集,之后将其与新参考合并子集 Q_{12} 进行合并并标记为已处理,重复上述过程,直到所有子集都被标记为已处理。如图 7 所示,以子集 Q_1 为初始参考合并子集,给出了 5 个子集 $\{Q_1, Q_2, \dots, Q_5\}$ 的合并过程。

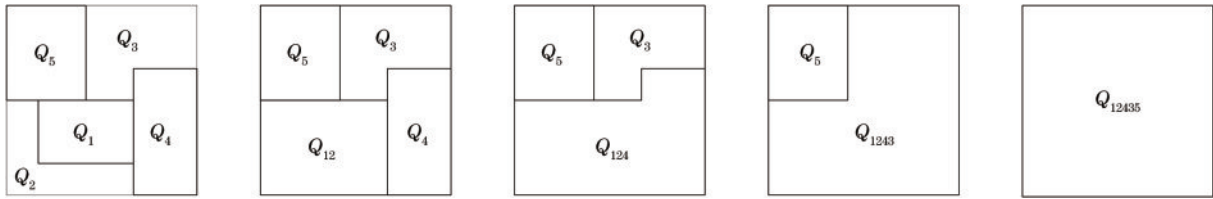


图 7 子集合并过程

Fig. 7 Subset merging process

3.4 GPS 信息辅助光束法平差

光束法平差通过最小化重投影误差达到优化恢复相机位姿和三维空间点坐标的目的。但由于无人机拍摄大规模图像的特点以及前期特征匹配步骤带来的误差,每次迭代执行光束法平差都会引入新的误差,这种情况下,如果仅仅依赖重投影误差最小化,最终会因为误差累积过大使得光束法平差的代价函数无法收敛到全局最小值,最终出现场景漂移。为了避免出现场景漂移,将无人机平台获取的 GPS 信息与光束法平差结合,为光束法平差的代价函数添加新约束,即引入绝对 GPS 误差项,同时减少图像重投影误差和相机中心距离误差,获得同时满足视觉约束和位置约束的最优三维空间点坐标和相机位姿,改进后光束法平差的代价函数是 GPS 绝对误差与最小化重投影误差的线性组合,通过最小化光束法平差代价函数实现。GPS 绝对误差项的表达式为

$$E_{\text{gps}} = \sum_{i=1}^n \vartheta \times \text{Huber}(\mathbf{C}_i - \mathbf{P}_i^{\text{gps}}), \quad (7)$$

加入式(7)后,代价函数为

$$E(\mathbf{C}, \mathbf{X}) = \sum_{i=1}^n \sum_{j=1}^m \rho_{ij} \|p(\mathbf{C}_j, \mathbf{X}_i) - x_{ij}\|^2 + \sum_{i=1}^n \vartheta \times \text{Huber}(\mathbf{C}_i - \mathbf{P}_i^{\text{gps}}), \quad (8)$$

式中:式(8)第 1 项为重投影误差,表示视觉上的约束关系;第 2 项为绝对 GPS 误差,表示位置上的约束关系; ϑ 为补偿系数; \mathbf{C}_i 表示相机 i 的预测位置; $\mathbf{P}_i^{\text{gps}}$ 表示该相机 i 实际对应的 GPS 位置; $\text{Huber}(\mathbf{C}_i - \mathbf{P}_i^{\text{gps}})$ 是两者之差的 Huber 损失函数, Huber 损失函数是一种平滑的损失函数,能够在存在离群值的情况下仍然保持良好的性能。引入 GPS 绝对误差限制了预测相机中心点的位置,即约束了相机运动轨迹,在增量重建过程中,降低了预测的相机运动轨迹与真实 GPS 轨迹之间

的距离,有效消除了累积误差,避免了场景漂移现象。

4 实验结果与分析

4.1 实验数据与环境

在实验中,使用 3 个不同数量级的无人机数据集对所提并行化增量 SFM 重建算法的有效性进行验证,分别为 1 个自采数据集 Dataset1 和 2 个从网络下载的无人机数据集 NPU_phantom3 和 NPU_Central^[16],如图 8 所示。其中:Dataset1 数据集为使用大疆无人机对某公路

区域实地拍摄的图像集,包含 225 张图片,存在航向重叠率 75%,不存在旁向重叠率;NPU_phantom3 与 NPU_Central 为 2 个开源无人机数据集,分别包含 538 和 1790 张图片,存在航向重叠率 80%,旁向重叠率 60%。并且 3 个数据集都包含无人机平台获取的 GPS 信息与 EXIF 文件,且 3 个数据集均不包含任何地面控制点。

此外,所有实验都在相同实验环境中进行: Intel Core i7-8700, 32 GB 内存, 3.19 GHz CPU, 6 GB Nvidia GeForce GTX 1060 显卡, Ubuntu 18.04 系统。



图 8 实验数据集。(a)Dataset1;(b)NPU_phantom3;(c)NPU_Central
Fig. 8 Experimental datasets. (a) Dataset1; (b) NPU_phantom3; (c) NPU_Central

4.2 性能评估

为了对所提并行化增量 SFM 算法的性能进行有效评估,选择 3 个先进的开源增量 SFM 框架在上述 3 个数据集上进行对比实验。3 个增量 SFM 框架分别为 ColMap^[17]、VisualSFM^[18] 及 Metashape^[19],其中,ColMap 和 VisualSFM 是使用 C++ 编程实现的两个开源增量 SFM 框架, Metashape 是一个商业软件包。3 个开源框架在运行过程中均未使用 GPS 辅助信息,此外,在保证实验环境一致、相机模型一致以及优化的相机内部参数一致的情况下进行所有实验。性能评价指标为平均重投影误差(MMSE)、平均轨迹长度(L)以及重建点个数(N_{point})。平均重投影误差以像素为单位,是重建出三维空间点投影到像平面上的投影点和其对应二维特征点之间的距离,距离越小重建精度越高。平均轨迹长度是三维空间点平均在二维像面中观测值个数,通过特征点数量除以重建点个数来计算,值越大,表示重建出三维空间点可靠性越高。表 1 为 4 种算法在 3 个数据集上的结果。

从表 1 可以看出,相较于对比方法,所提方法精度最高,获得了最小的重投影误差,在 3 个数据集上的平均重投影误差分别为 0.337、0.391、0.410 pixel。Metashape 利用 POS 数据和多尺度策略进行匹配对选择,恢复出最多的三维空间点,但该方法在 3 个数据集上的表现不稳定,对于 Dataset1 和 NPU_phantom3 数据集,平均重投影误差仅次于所提方法,但对于图像规模较大的数据集 NPU_Central,平均重投影误差较大。VisualSFM 在特征匹配过程中采用两两匹配对的选择方式,使得特征匹配过程中引入不匹配对,导致特征点误匹配,恢复出的三维空间点数量最少并且精度最低,

表 1 4 种算法性能对比

Table 1 Performance comparison of the four algorithms

Dataset	Method	N_{point}	L	MMSE / pixel
Dataset1	ColMap	104810	4.209	0.583
	VisualSFM	67152	4.217	0.619
	Metashape	183697	4.799	0.461
	Proposed method	138759	4.481	0.337
NPU_phantom3	ColMap	427735	9.441	0.602
	VisualSFM	336042	8.316	0.741
	Metashape	598442	9.542	0.539
	Proposed method	461943	8.138	0.391
NPU_Central	ColMap	762733	24.161	0.539
	VisualSFM	601922	21.014	0.822
	Metashape	1031644	24.050	0.692
	Proposed method	870532	23.150	0.410

这也可以通过图 9 的重建点云来验证。此外,由于所提方法利用 GPS 信息对光束法平差步骤进行约束,所以平均重投影误差相比其他 3 个方法小。3 个对比方法和所提方法(使用 CloudCompare 点云可视化工具)重建出的点云结果如图 9 所示。

4.3 效率评估

从第 3.1 节可知,增量式 SFM 算法的时间开销主要表现在特征匹配和光束法平差迭代优化两个步骤,因此,在效率评估中,以特征匹配和光束法平差的时间开销作为效率评价指标。光束法平差的时间开销如表 2 所示,在实验过程中记录了 4 种方法在 4 个数据集上 BA 优化步骤耗费的时间(T_{ba}),并且为了控制实验变量,除场景图划分和合并外,所提方法的其

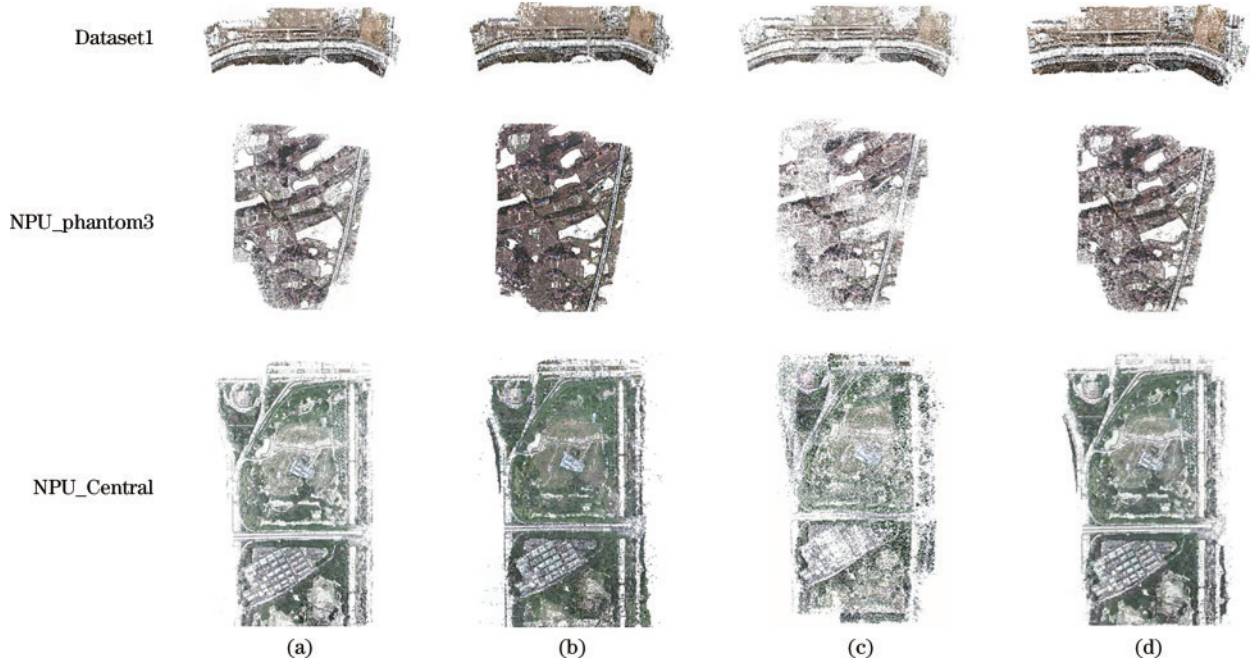


图 9 重建可视化结果。(a)ColMap;(b)Metashape;(c)VisualSFM;(d)所提方法

Fig. 9 Reconstructed visualization results. (a) ColMap; (b) Metashape; (c) VisualSFM; (d) proposed method

表 2 BA 优化运行时间

Table 2 BA optimization runtime unit: min

Method	Dataset1	NPU_phantom3	NPU_Central
ColMap	23.062	89.401	493.849
VisualSFM	6.334	21.662	137.961
Metashape	34.183	63.854	211.402
Proposed method	8.327	18.392	100.631

他实现部分配置了相同的默认参数。特征匹配步骤的时间开销如表 3 所示,实验过程中主要记录了所提方法使用基于词汇树匹配策略的时间(T_{voc})开销和使用传统匹配策略的时间(T_w)开销。为了有效对比,除匹配对选择策略外,其余部分方法保持一致。

表 3 特征匹配运行时间

Table 3 Feature matching run time unit: min

Method	T_w	T_{voc}
Dataset1	10.052	6.547
NPU_phantom3	49.819	17.799
NPU_Central	263.951	56.872

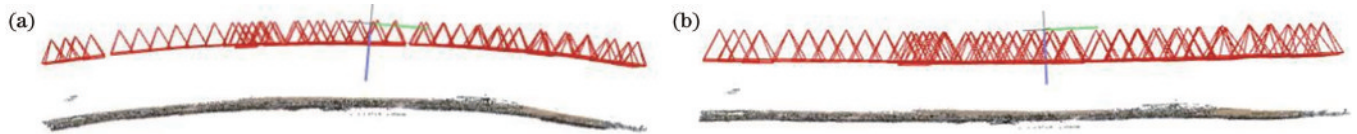


图 10 有无 GPS 辅助重建结果。(a)无 GPS;(b)有 GPS

Fig.10 Reconstruction results with or without GPS assistance. (a) Without GPS; (b) with GPS

对于 Dataset1 这种缺乏旁向运动约束的无人机影像数据集,在没有 GPS 信息辅助的情况下,光束法平差过程仅通过视觉约束对重建结果进行优化,难以消除重建过程中不断积累的误差,最终造成影像重建误

从表 2 可以看出,在图像数量不太大的情况下,3 个对比算法的时间开销都在可接受范围内,例如,在 Dataset1 数据集上 Colmap、VisualSFM 和 Metashape 的时间开销分别为 23.062 min、6.334 min 和 34.183 min。然而,随着图像数量的增加,效率急剧下降,尤其当数据集 NPU_Central 包含 1790 张图片时,ColMap 算法 BA 优化时间开销达到 493.849 min,约为所提方法的 4.9 倍。而所提方法在处理大规模无人机数据集时,由于并行化计算资源被充分利用以及 BA 优化计算复杂度随图像规模减小而显著降低,取得了最高的效率。

从表 3 可以看出,由于所提方法在通过词汇树检索获得每个待匹配图像的相关图像集后完成特征匹配,相比传统的穷举暴力匹配,显著降低了匹配对的组合复杂度,提高了特征匹配步骤的效率。

4.4 GPS 信息辅助光束法平差实验

为了验证在有无 GPS 信息辅助下光束法平差对重建结果精度的影响,对 Dataset1 数据集在有无 GPS 信息辅助下的重建结果进行对比,并将其可视化(侧视图)结果进行展示,可视化结果如图 10 所示。

差不断积累,出现场景漂移现象,如图 10(a)所示,重建结果弯曲。而在有 GPS 信息辅助的情况下,在 BA 优化过程中,通过视觉约束和位置约束共同对重建结果进行优化,可以使得重建结果在满足视觉约束的同

时尽可能靠近真实重建结果, 累积误差被有效消除, 避免了场景漂移。

为了进一步分析 GPS 信息辅助光束法平差的效果, 将数据集 Dataset1 在有无 GPS 信息辅助下恢复出的相机运动轨迹与相机真实对应的 GPS 轨迹通过轨迹可视化(EVO)工具进行对齐, 实验中指定真实 GPS 轨迹为参考轨迹, 可视化结果如图 11 所示。

其中, 灰色虚线轨迹代表真实的 GPS 轨迹, 彩色

线轨迹代表所提方法预测的轨迹, 右边颜色柱为对应帧之间的误差, 误差越大颜色越红, 误差越小颜色越蓝。从结果可以看出: 在无 GPS 信息辅助的情况下, 如图 11(a) 所示, 两条轨迹之间间隔较大, 最大误差达到 5.9 m; 而在有 GPS 信息辅助的情况下, 如图 11(b) 所示, 由于所提方法为 BA 优化代价函数添加了位置约束项, 约束了相机运动轨迹, 使得两条轨迹基本重合, 有效消除了误差累积问题, 避免了场景漂移。

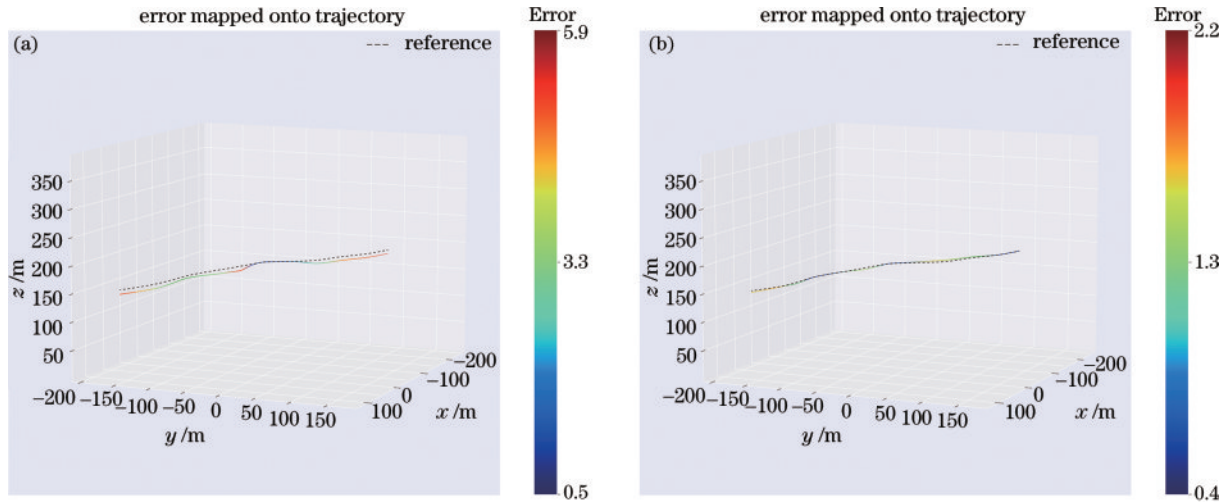


图 11 轨迹对齐结果。(a)无 GPS;(b)有 GPS

Fig. 11 Trajectory alignment results. (a) Without GPS; (b) with GPS

5 结 论

针对增量 SFM 算法在处理大规模无人机影像时效率低、易产生场景漂移的问题, 提出一种可并行化处理的增量 SFM 算法, 其核心思想是在通过词汇树检索为每张图像选择相关图像集后进行特征匹配和场景图构造, 并设计实现了一种划分-合并策略的增量 SFM 流程。此外, 为光束法平差的代价函数添加位置约束项, 同时考虑相机中心距离误差。在 3 个数据集上的实验结果表明, 与其他增量 SFM 相比, 所提并行化增量 SFM 重建算法在图像特征匹配和光束法平差这两个最耗时步骤的效率显著提升, 并且取得了较高的精度。同时, 使用 GPS 信息辅助光束法平差有效消除了误差累积问题, 避免了场景漂移, 实现了大规模无人机影像高效精确的三维重建。

参 考 文 献

- [1] Jiang S, Jiang W S. Efficient SfM for oblique UAV images: from match pair selection to geometrical verification[J]. *Remote Sensing*, 2018, 10(8): 1246.
- [2] Cui H N, Gao X, Shen S H, et al. HSfM: hybrid structure-from-motion[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 2393-2402.
- [3] Jiang S, Jiang C, Jiang W S. Efficient structure from motion for large-scale UAV images: a review and a comparison of SfM tools[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, 167: 230-251.
- [4] Sun Y B, Sun H B, Yan L, et al. RBA: reduced Bundle Adjustment for oblique aerial photogrammetry[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2016, 121: 128-142.
- [5] Zheng M T, Zhang F Y, Zhu J F, et al. A fast and accurate bundle adjustment method for very large-scale data[J]. *Computers & Geosciences*, 2020, 142: 104539.
- [6] Chen Y, Shen S H, Chen Y S, et al. Graph-based parallel large scale structure from motion[J]. *Pattern Recognition*, 2020, 107: 107537.
- [7] 尹双双, 邓非. GPS 改进的移动球形全景影像三维重建方法[J]. *测绘科学*, 2020, 45(5): 13-22.
- [8] 郭复胜, 高伟. 基于辅助信息的无人机图像批处理三维重建方法[J]. *自动化学报*, 2013, 39(6): 834-845.
- [9] Ye Z C, Li G L, Liu H M, et al. CoLi-BA: compact linearization based solver for bundle adjustment[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2022, 28(11): 3727-3736.
- [10] Wu C. SIFTGPU: a GPU implementation of david lowe's scale invariant feature transform (SIFT)[EB/OL]. (2017-06-

- 19)[2023-03-02]. <https://github.com/pitzer/SIFTGPU>.
- [11] Jiang S, Jiang W S, Guo B X. Leveraging vocabulary tree for simultaneous match pair selection and guided feature matching of UAV images[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2022, 187: 273-293.
- [12] Nister D, Stewenius H. Scalable recognition with a vocabulary tree[C]//2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), June 17-22, 2006, New York, NY, USA. New York: IEEE Press, 2006: 2161-2168.
- [13] Shi J B, Malik J. Normalized cuts and image segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(8): 888-905.
- [14] Zhu S Y, Shen T W, Zhou L, et al. Parallel structure from motion from local increment to global averaging [EB/OL]. (2017-02-28)[2023-02-03]. <https://arxiv.org/abs/1702.08601>.
- [15] Umeyama S. Least-squares estimation of transformation parameters between two point patterns[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1991, 13(4): 376-380.
- [16] Bu S H, Zhao Y, Wan G, et al. Map2DFusion: real-time incremental UAV image mosaicing based on monocular SLAM[C]//2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), October 9-14, 2016, Daejeon, Korea (South). New York: IEEE Press, 2016: 4564-4571.
- [17] Schönberger J L, Frahm J M. Structure-from-motion revisited[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 4104-4113.
- [18] VisualSFM: a visual structure from motion system[EB/OL]. [2022-01-30]. <http://www.cs.washington.edu/homes/ccwu/vsfm>.
- [19] Agisoft. Agisoft metashape homepage[EB/OL]. (2021-12-18)[2023-01-30]. <http://www.agisoft.com>.