引文格式:姜三,马一尘,江万寿,等.不依赖固定锚点合并的无人机影像并行化SfM方法[J].武汉大学学报(信息科学版),2026,51(6):1194-1203.DOI:10.13203/j.whugis20240347

Citation:JIANG San,MA Yichen,JIANG Wanshou,et al.Parallel SfM for UAV Images via Anchor-Free Merging[J].Geomatics and Information Science of Wuhan University,2026,51(6):1194-1203.DOI:10.13203/j.whugis20240347

不依赖固定锚点合并的无人机影像并行化SfM方法

姜 三[1,2,3] 马一尘[3] 江万寿[4] 李清泉[1,2]

1 深圳大学广东省城市空间信息工程重点实验室,广东 深圳,518060

2 深圳大学自然资源部大湾区地理环境监测重点实验室,广东 深圳,518060

3 自然资源信息管理与数字孪生工程软件教育部工程研究中心,湖北 武汉,430074

4 武汉大学测绘遥感信息工程全国重点实验室,湖北 武汉,430079

摘 要:增量式运动恢复结构(structure from motion,SfM)并行化方法中,存在场景图构建忽略影像局部连接强度、子场景(空间关联的影像集合)合并路径长和效率低等问题,使其难以满足大场景影像处理需求。提出了一种不依赖固定锚点(参考模型)合并的无人机影像并行化SfM方法。设计了顾及影像局部连接强度约束的场景图构建方法,有效提升了场景图分块的紧凑性,保证了分块内部影像SfM重建的高稳健性;提出了不依赖固定锚点的子场景并行化合并方法,实现合并路径自适应调整与多子场景并行化合并,有效提高了子场景合并的可靠性和效率。利用真实无人机影像进行实验,结果表明,所提方法能够实现9倍加速比的子场景合并,且相对定向和绝对定向性能与现有方法相当。所提方法已成功应用于覆盖范围约50 km^{2}近9万张无人机影像的并行化SfM重建。

关键词:数字摄影测量;无人机遥感;运动恢复结构;三维重建

中图分类号:P237 **文献标识码**:A **收稿日期**:2025-01-15

DOI:10.13203/j.whugis20240347 **文章编号**:1671-8860(2026)06-1194-10

Parallel SfM for UAV Images via Anchor-Free Merging

JIANG San[1,2,3] *MA Yichen*[3] *JIANG Wanshou*[4] *LI Qingquan*[1,2]

1 Guangdong Key Laboratory of Urban Informatics, Shenzhen University, Shenzhen 518060, China

2 Key Laboratory for Geo-Environmental Monitoring of Great Bay Area, Ministry of Natural Resources, Shenzhen University, Shenzhen 518060, China

3 Engineering Research Center of Natural Resource Information Management and Digital Twin Engineering Software, Ministry of Education, Wuhan 430074, China

4 State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, 430079, China

Abstract: **Objectives**: Incremental structure from motion (SfM) has become the widely used workflow for aerial triangulation of unmanned aerial vehicle (UAV) images. However, it suffers from significant issues of error accumulation and efficiency degradation. The existing parallelization methods for SfM suffer from issues such as the scene graph construction ignoring the local connection strength of images, long merging paths for sub-scenes (spatially related image collections), and low efficiency, making it difficult to meet the requirements of large-scale scene image processing. **Methods**: This paper proposes a parallelized SfM method for UAV images that does not rely on fixed anchor points (reference models). The proposed solution develops a scene graph construction algorithm that considers local connection strength constraints, effectively improving the compactness of scene graph partitioning and ensuring the high robustness of SfM reconstruction within partitions. Then, it proposes a parallelized sub-scene merging method that does not rely on

基金项目:国家自然科学基金(42371442);深圳市科技计划资助(JCYJ20250604181614019)。

第一作者:姜三,博士,副教授,研究方向为航空摄影测量和三维重建。jiangsan@szu.edu.cn

通信作者:李清泉,教授,中国工程院院士。liqq@szu.edu.cn

fixed anchor points, enabling adaptive adjustment of merging paths and parallel merging of multiple sub-scenes, significantly enhancing the reliability and efficiency of sub-scene merging. To avoid the influence of outliers in corresponding geometric elements, this paper employs the random sample consensus hypothesis-verification framework for robust parameter estimation and designs a transformation model estimation based on a bidirectional mean square reprojection error metric. **Results:** Experiments using real UAV images demonstrate that the proposed method achieves a 9 times speedup in sub-scene merging, with relative and absolute orientation performance comparable to the existing methods. And the proposed method has also been successfully applied to the parallel SfM reconstruction of nearly 90 000 UAV images covering an area of about 50 km^2. **Conclusions:** Considering the metric of efficiency and precision, the proposed method can be a validate solution for the efficient SfM reconstruction for large-scale UAV datasets.

Key words: digital photogrammetry; UAV remote sensing; structure from motion(SfM); 3D reconstruction

由于具备硬件成本低、数据获取的空间和时间分辨率高等优势，无人机(unmanned aerial vehicle，UAV)已成为最重要的摄影测量平台[1]，并在城市实景三维[2]、文物数字化保护[3]、电力走廊巡检[4]等领域得到广泛应用。尽管高精度实时动态测量(real-time kinematic，RTK)可实现厘米级定位，UAV平台往往没有搭载高精度定向传感器，导致影像定向精度低，无法实现直接定位定向。近年来，运动恢复结构(structure from motion，SfM)中增量式SfM(incermental SfM，ISfM)不依赖未知数良好初值且很好适应非固定航带的数据采集，已成功应用于有序或无序UAV影像的定位定向[5]，包括商业软件Agisoft Metashape和开源软件COLMAP。然而，为了实现未知数初值的稳健解算，ISfM迭代进行影像定向和光束法平差(bundle adjustment，BA)，存在误差累积和效率下降显著的问题，无法满足大规模UAV影像定位定向需求[6-7]。

为了提升ISfM处理大规模场景的能力，现有工作已从BA效率提升开展了系列研究，包括基于CPU/GPU混合加速的BA效率提升[8]，以及顾及多镜头倾斜摄影相机内在约束的BA模型简化方法[9]。但是，这类方法并没能从本质上解决大场景ISfM的顺序性约束。另一方面，基于分块-合并策略的并行化SfM技术近年来也得到了越来越多的研究。其核心思想是将大场景分割为几何连接强、BA优化效率高的小场景，通过并行化ISfM重建和子场景稳健合并实现整体加速，有效降低BA迭代优化的误差累积和效率下降问题[10-13]。并行化SfM主要包括3个关键步骤：(1)基于影像连接关系构建场景图$G=\{V,E\}$。(2)场景图分块$\{G_1,G_2,\cdots,G_n\}$与并行化ISfM重建$\{M_1,M_2,\cdots,M_n\}$。(3)利用子场景连接约束实现模型合并，得到最终模型$M=\{M_1\cap M_2\cap M_n\}$。基于上述分块-合并策略，近年来研究者已提出了面向大规模场景的有效SfM方法。文献[14]提出分区优化混合式SfM方法，主要解决了不稳定分区检测和分区可靠合并的难题；文献[15]提出顾及场景连通性的混合式SfM方法，解决分区影像内连接松散、分区合并稳健性差的问题；文献[16]提出基于加权连通支配集的并行化SfM方法，利用全局约束模型提升子场景合并精度。

尽管分块-合并策略显著提升了ISfM处理大场景影像的适用性，但是现有方法仍然存在两个方面的重要缺陷。一方面，场景图G的构建往往采用全局策略，对边$e_{i,j}$计算边权值$w_{i,j}$，即任一连接边$e_{i,j}$需要与场景图G中所有边比较来量化其重要性。常用方法包括利用影像对的匹配点数量及其在影像平面的分布计算边权值[17-18]。然而，ISfM迭代查找下一张最佳影像进行定位定向，意味着待处理的下一张影像与已有模型的局部连接强度具有更重要的价值。另一方面，现有方法往往采用基于固定锚点或者无固定锚点策略进行子模型$\{M_1,M_2,\cdots,M_n\}$合并。其中，基于固定锚点方法以一个子模型为参考模型，并计算其他子模型到参考模型的变换参数。例如，文献[11]将连接边最多的子场景作为参考模型，文献[16]提取场景的加权连通支配集作为参考模型。这类方法的最大问题在于合并路径没有考虑各个子场景的连接强度，容易导致后续子场景的合并失败。同时各个子场景的合并互相依赖，难以实现并行化。相反，无固定锚点方法则是计算任意两个子模型的变换参数，迭代进行场景合并[17,19]。现有方法均采用顺序合并策略，存在合并效率低的问题，无法处理大场景UAV影像。

针对上述问题，本文提出了一种不依赖固定锚点合并的UAV影像并行化SfM方法。针对现

有全局边加权策略无法有效适应ISfM重建特性的问题,设计了顾及影像局部连接强度约束的场景图构建方法,有效提升了场景图分块的紧凑性。针对现有子场景合并方法存在的误差累积和效率低等问题,提出了不依赖固定锚点的子场景并行化合并方法,实现合并路径自适应调整、多子场景并行化合并,有效提升了ISfM的精度和效率。利用多组大规模UAV影像进行了实验,以验证本文方法的可靠性。

1 方法原理

图1展示了本文提出的不依赖固定锚点合并的UAV影像并行化SfM方法技术流程。整个流程的输入数据仅为UAV影像$A=\{a_i\}$。关键步骤包括:(1)联合局部聚合描述符向量(vector of locally aggregated descriptors,VLAD)全局描述子与分层可导航小世界网络(hierarchical navigable small world,HNSW)图索引技术,高效检索得到影像匹配对$P=\{p_{ij}\}$,并采用顾及局部连接强度的边加权策略构建场景图$G=\{V,E\}$。(2)利用图切割算法将场景图G聚类分割为尺寸相当、连接紧凑的子场景$C=\{C_1,C_2,\cdots,C_n\}$,并通过并行化处理得到各个子模型$\{M_1,M_2,\cdots,M_n\}$。(3)挖掘子场景的独立特性,基于双向重投影误差最小化的几何变换测度,实现无固定锚点约束的子场景并行化合并,得到最终的完整模型$M=\{M_1\cup M_2\cup M_n\}$。值得注意的是,本文所提方法假设相机参数已经提前检校,因而在场景合并过程中没有进行相机参数优化。

1.1 顾及局部连接强度的场景图构建

分块并行化SfM首先需要构建表示影像连接关系的场景图$G=\{V,E\}$。其中,$V=\{v_i\}$和$E=\{e_{i,j}\}$分别表示顶点集合和边集合,对应影像$A=\{a_i\}$及其具有重叠的匹配对$P=\{p_{i,j}\}$。尽管现有UAV系统具备高精度RTK定位能力,但是其定向精度仍然较低。同时,贴近或者优视摄影等非常规航线的UAV摄影测量也日渐普及,导致仅依赖影像位置信息约束难以获取精确、合理的影像匹配对。因此,本文采用影像检索方法进行高效影像匹配对检索。该方法利用VLAD全局描述子加速影像索引构建,利用HNSW图索引结构实现高维描述子最近邻检索。具体步骤包括:(1)编码本构建。利用K-means++算法将训练特征描述子划分为k个子集,得到包含k个单词的编码本$D=\{d_1,d_2,\cdots,d_k\}$。(2)VLAD向量聚合。使用编码本C对影像的特征描述子向量进行聚合,得到每张影像的VLAD向量。(3)HNSW索引与检索。利用影像的VLAD向量构建图索引结构,并利用索引实现高效描述子检索,得到具有空间重叠的影像匹配对$P=\{p_{ij}\}$,引导后续的特征匹配。本文采用SIFTGPU进行特征匹配,并使用随机抽样一致性(random sample consensus,RANSAC)基础矩阵估计实现粗差剔除。

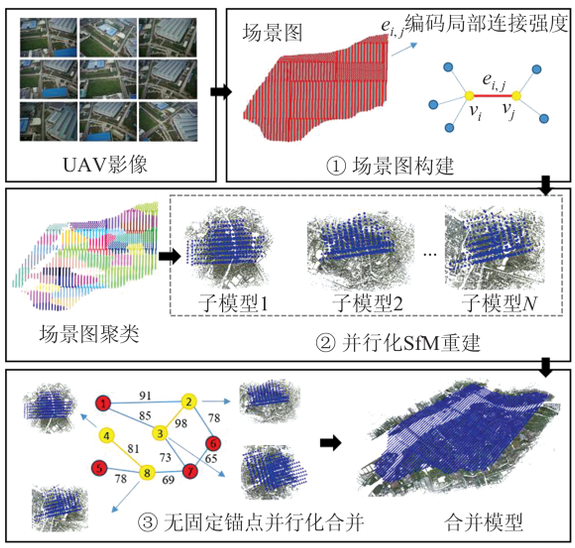


图1 本文提出的UAV影像并行化SfM技术流程

Fig. 1 Workflow of the Proposed Parallel SfM for UAV Images

利用检索得到的匹配对P和特征匹配结果可以构建场景图G。假设场景包含n张影像$A=\{a_i\}$,m个匹配对$P=\{p_{i,j}\}$,则场景图$G=(V,E)$可由以下方法构建:为影像集合I中每一张影像i_i添加一个顶点v_i,则可构成顶点集合$V=\{v_i\}$;为匹配对集合P中每一对匹配对$p_{i,j}$添加一条边$e_{i,j}$以连接顶点v_i和v_j,则可构成边集合$E=\{e_{i,j}\}$。同时,为边$e_{i,j}$赋权值$w_{i,j}$,用来量化影像对的连接强度,保证后续场景分块重建和子场景合并的可靠性。考虑到在ISfM重建中,局部连接强度决定了下一张最优影像的选择及其连接路径。本文设计了顾及局部连接强度的边权值$w_{i,j}$计算方法,原理参见图2。具体计算如下:

$$w_{i,j}=\frac{1}{2}\left(\frac{N_{i,j}}{\sum_{k=1}^{K}N_{i,k}}+\frac{N_{i,j}}{\sum_{l=1}^{L}N_{j,l}}\right) \tag{1}$$

式中,$N_{i,j}$是影像a_i与影像a_j的匹配点数量;$\sum_{k=1}^{K}N_{i,k}$是影像i_i与其在G中所有相邻影像的匹配点数量之和;$\sum_{l=1}^{L}N_{j,l}$同理。因此,式(1)所计算的

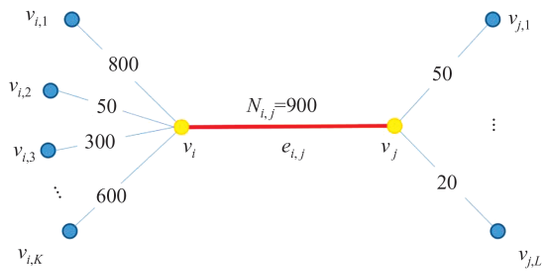


图 2　边权值计算示意图

Fig. 2　Illustration of Edge Weight Calculation

边权值 $w_{i,j}$ 编码了连接边 $e_{i,j}$ 相较于影像 i_i 与影像 i_j 的所有相邻边的重要性。图 2 中 $K=4$,$L=2$,边权值 $e_{i,j}=0.5\times0.34+0.5\times0.93$,那么,$0.5\times0.34$ 和 0.5×0.93 则分别表示连接边 $e_{i,j}$ 对影像 i_i 和影像 i_j 的重要性,即连接边 $e_{i,j}$ 对影像 i_j 更重要,因为 $e_{i,j}$ 是其所有连接边中最稳定边。采用本文的加权策略能够获得更高的边权值,保证了后续图割分块中 v_i 和 v_j 能够分到同一个子图,即提升了 v_j 重建的成功率。

1.2　场景图分块与并行化 ISfM 重建

利用构建的场景图 $G=\{V,E\}$ 可实现场景分块与并行化 ISfM 重建。与现有大部分工作类似,本文采用归一化切割(normalized cut,NC)算法进行场景分块[11]。NC 算法通过最小化切割边权值实现场景分块,保证分块子场景内部连接边的可靠性。如图 3 所示,场景分割为两个子场景 C_i 和 C_j,同时,连接边 E 划分为场景内边 E_{intra} 和场景间边 E_{inter},其中,E_{intra} 用于各个子场景的 ISfM 重建;E_{inter} 用于子场景的合并处理。

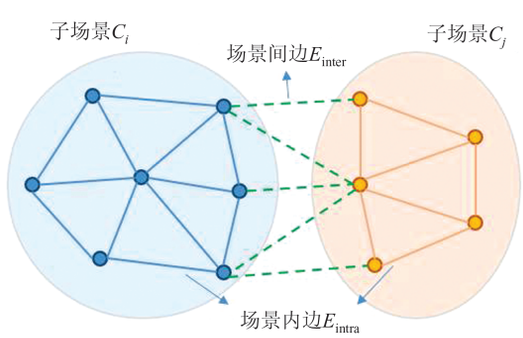


图 3　场景图分块示意图

Fig. 3　Illustration of Match Graph Clustering

已有研究通常采用尺寸约束 C_{size} 和完整性约束来划分子场景大小及其重叠度[11,18]。前者保证子场景影像数量合适且大小相当;后者保证子场景之间有足够多的重叠影像以利于子场景合并。考虑到子场景之间具有数量足够的公共三维点及过多重叠影像导致的额外计算,本文场景分块算法仅使用尺寸约束,且 $C_{\text{size}}=500$。基于上述方法可将场景图 G 分割为尺寸相当的子场景 $\{C_1,C_2,\cdots,C_n\}$。对于分割的所有子场景,本文进一步采用多线程技术,实现并行化重建,得到所有子场景的 ISfM 重建模型 $\{M_1,M_2,\cdots,M_n\}$。

1.3　无固定锚点约束的子场景并行合并

对于并行化 ISfM 重建得到的子模型 $\{M_1,M_2,\cdots,M_n\}$,本文提出了无固定锚点约束的子场景并行合并方法。与现有方法相比,本文所提方法既避免了依赖固定锚点策略导致的合并路径长、误差累积问题,还有效解决了现有无固定锚点策略导致的顺序合并效率低的问题。所提方法的关键技术包括子模型的稳健参数估计和有效子场景并行合并。

1.3.1　子模型的稳健参数估计

场景合并的思路是利用公共影像位姿和三维点坐标,将重建的子模型转换到统一的全局坐标系。子模型的转换参数可表示为相似变换 $T=[\lambda R|t]$。其中,R 和 t 分别代表旋转矩阵和平移向量;λ 是缩放因子。假设参考模型 M_r 和源模型 M_s 之间存在 n 个对应几何元素,分别表示为 $\{r_i\}$ 和 $\{s_i\}$,那么,M_s 到 M_r 的转换参数可通过最小化均方误差计算:

$$e^2(\boldsymbol{R},\boldsymbol{t},\lambda)=\frac{1}{n}\sum_{i=1}^{n}\left\| r_i-(\lambda\boldsymbol{R}s_i+\boldsymbol{t})\right\|^2 \tag{2}$$

式中,e^2 是在变换矩阵 T 下 s_i 到 r_i 的均方根误差。式(2)所表示的最小化问题可通过奇异值分解(singular value decomposition,SVD)算法求解。考虑到本文重建的子场景不存在公共影像位姿,因此将子模型的公共三维点作为对应几何元素进行参数估计。

为了避免对应几何元素中粗差影响,本文采用 RANSAC 假设-验证框架进行稳健参数估计。首先,从公共三维点中随机选择 3 对,并利用式(2)计算相似变换矩阵 T。然后,采用下式所示的双向均方重投影误差测度计算 RANSAC 的几何估计残差 e_i^2:

$$e_i^{\,2}=\frac{1}{m+l}\left(\sum_{j=1}^{m}\left\| P_c(C_j,T_{s_i})-x_{i,j}\right\|^2+\sum_{j=1}^{l}\left\| P_c(C_jT^{-1}r_i)-x_{i,j}\right\|^2\right) \tag{3}$$

式中,m 和 l 分别表示参考模型和源模型中可观测到三维点 r_i 和 s_i 的影像数量;$P_c(C_j,T_{s_i})$ 表示相似变换后三维点 T_{s_i} 在相机 C_j 上的投影点;$x_{i,j}$ 表示三维点 r_i 在相机 C_j 上对应的影像点坐标。本文设定最大残差 e_i 为 2.0 像素,用于区分相似变换

T下的内点和外点数量,并保留内点数最多的变换参数为最佳参数T^*。迭代执行上述步骤,直到达到最大迭代次数。

1.3.2 *有效子场景并行合并*

随着UAV影像规模增加,子场景数量会越来越多,场景间的连接变得越来越复杂。找到合适的参考模型并确定合理的子模型合并顺序变得十分重要。因此,本文提出无锚点约束的子场景合并方法,利用场景间边E_{inter}实现并行化的高效场景合并。

无固定锚点并行合并方法的核心思想是迭代并行化合并相互独立的子场景对。该方法首先构建子场景连接图$G'=(V',E')$表示子场景间的连接关系。假设子场景集合为$C=\{C_i\}$,则以每个子场景C_i为图添加一个顶点v_i',组成顶点集合$V'=\{v_i'\}$;为每对相邻子场景C_i和C_j添加一条边$e_{i,j}'$,组成边集合$E'=\{e_{i,j}'\}$。与此同时,为每条边赋予权值$w_{i,j}'$,表示相应子场景间的连接强度。假设子场景C_i和C_j的场景间边集记为$E_{\text{inter}}^{i,j}=\{e_k^{i,j}\}$,$k=1,2,\cdots,K$,那么,子场景间的连接强度可表示为:

$$w_{i,j}'=\sum_{k=1}^{K}w_k^{i,j} \tag{4}$$

式中,$w_k^{i,j}$为对应边$e_k^{i,j}$的边权值。因此,上述子场景间的连接强度定义为场景间边集合的边权值总和,如图4所示,其中黄色为当前执行合并的子场景,红色为未合并的子场景。无固定锚点约束的子场景并行合并步骤为:(1)根据边权值对边集E'降序排序,并依次检查每一条边是否为有效场景对。为了避免场景合并冲突,有效场景对需满足条件:任何有效场景对的顶点都不是其他有效场景对的相邻顶点。对于有效场景对,将其添加到$P'=\{p_{ij}'\}$。(2)根据上文子模型的稳健参数估计,并行化合并P'包含的有效场景对。为了提高合并效率,将较大的子模型设置为参考模型M_r,另一个子模型设置为源模型M_s。(3)更新场景图G'的边权值,包括合并子场景的连接边剔除(图4(b)中子场景2和3)和相邻子场景的连接边合并(图4(b)中子场景1与子场景2和3)。迭代执行上述有效场景对查找、子场景合并以及场景图边权值更新,最终所有子场景合并为一个场景,得到最终模型$M=\{M_1\cap M_2\cap M_n\}$。

1.3.3 *参数估计的约束条件*

为了进一步提升RANSAC参数估计效率,上述并行化合并算法需要考虑两个关键因素:(1)随着子场景合并,参考模型$\{r_i\}$和源模型$\{s_i\}$的公共三维点数量n显著增加。(2)UAV影像的高重叠度导致三维点的可观测影像数量m和l较大。上述因素导致RANSAC稳健估计效率低。因此,本文对参数进行了限制:(1)考虑到重投影误差小的三维点属于可靠对应点的概率更高,本文对初始公共三维点按照重投影误差升序排列,并根据实验结果选择前10 000个三维点作为参数估计的样本点。(2)考虑到ISfM重建已进行严格粗差剔除以及上述重投影误差排序,本文在计算几何估计残差e_i^2时,将三维点投影到第一张观测影像,即限定式(3)中$m=1$和$l=1$。

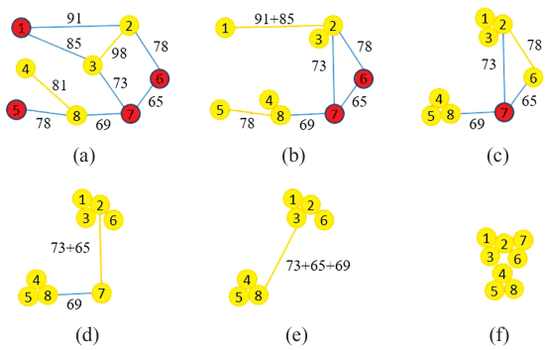


图4 子场景并行化合并示意图

Fig. 4 Illustration of Parallel Merging of Sub-Scenes

2 实验与分析

2.1 实验数据

本文选取了3个实验区域的UAV倾斜影像数据集,实验数据集的详细信息如表1所示。实验区域1拍摄于含有密集低层建筑的大学校园,采用单相机系统采集了3 743张影像,地面分辨率(ground sample distance, GSD)为2.6 cm。实验区域2采用单相机的优视摄影测量技术[20],采集了一栋结构复杂的校园建筑,共4 030张影像,GSD为1.2 cm。同时,为了验证绝对定向精度,在该实验区域采集了26个地面控制点(ground control points, GCPs)。实验区域3采用经典的五镜头倾斜摄影测量系统,采集了包含低矮植被和建筑的郊区场景。本文将该区域划分为两个不同规模的数据集,分别包含21 654和87 197张影像,用于大场景三维重建性能验证。3个实验区域的正射影像如图5所示。

2.2 结果与分析

2.2.1 *边加权策略对场景图构建的影响分析*

首先分析顾及局部连接强度的边加权策略对场景图构建的影响。本文对比分析了另外两个常用的边加权策略:(1)匹配点数量的边加权

表 1　实验数据集详细信息

Table 1　Details of Experimental Datasets

数据集	飞行高度/m	相机型号	相机数量/个	焦距长度/mm	倾斜角度/(°)	影像数量/张	影像尺寸/像素	地面分辨率/cm
1	80.0	DJI FC6310R	1	24.0	0	3 743	5 472×3 648	2.6
2	—	DJI Zenmuse P1	1	35.0	—	4 030	8 192×5 460	1.2
3和4	87.0	Sony ILCE 7R	5	35.0	0/45(下视/倾斜)	21 654/87 197	6 000×4 000	1.2



(a) 数据集1　(b) 数据集2　(c) 数据集3和4

图 5　4组实验数据集正射图

Fig. 5　Orthophotos of the Four Datasets

$e_{i,j}^{i_n}=N_{i,j}/N_{\max}$。其中，$N_{i,j}$和$N_{\max}$分别是匹配对$p_{i,j}$的匹配点数量和所有匹配对$\{p_{i,j}\}$的最大内点数量。(2)匹配点数量及分布的边加权$e_{i,j}^{i_n}=R^{*}e_{i,j}^{i_n}+(1-R)^{*}e_{i,j}^{d_i}$。其中，$e_{i,j}^{d_i}=(C\times H_i+C\times H_j)/(A_i+A_j)$是匹配点影像分布加权项，根据匹配点在两张影像所占据的最小凸多边形面积$C\times H_i$和$C\times H_j$，以及影像平面的面积A_i和A_j计算得到；R是$e_{i,j}^{i_n}$和$e_{i,j}^{d_i}$的加权因子，在本文中设置为0.5。

本文使用两个指标进行评估，即重建影像数量和重建三维点数量。3种方法构建场景图均采用本文所提出的并行化ISfM方法进行重建。实验结果如表2所示，可以看出：(1)本文方法在3个数据集上的表现最佳，重建影像数量分别达到3 734、3 987和21 642张；重建三维点数量分别达到1 041 872、1 639 651和10 807 368个。(2)匹配点数量方法在数据集1上重建影像数量仅有3 592张，主要原因是边加权导致场景分块不紧凑，影响分块重建。图6展示了不同方法的场景分块可视化结果。可以看出，与本文方法的紧凑分块相比，其他两种方法均存在空间较远的影像划分到同一分块(图6(a))，且存在狭长分块(图6(b))的问题。主要原因是本文提出的边加权策略很好地考虑了影像的局部连接强度，保证了关键连接边包含的顶点能够划分到同一个子场景，提升后续影像定位定向可靠性。相反，匹配点数量方法、匹配点数量及分布方法主要考虑连接边的全局贡献，即利用最大匹配点数进行边加权，容易削弱部分关键连接边的重要性。综上所述，本文方法能够实现紧凑的场景分块，提升整体重建完整度。

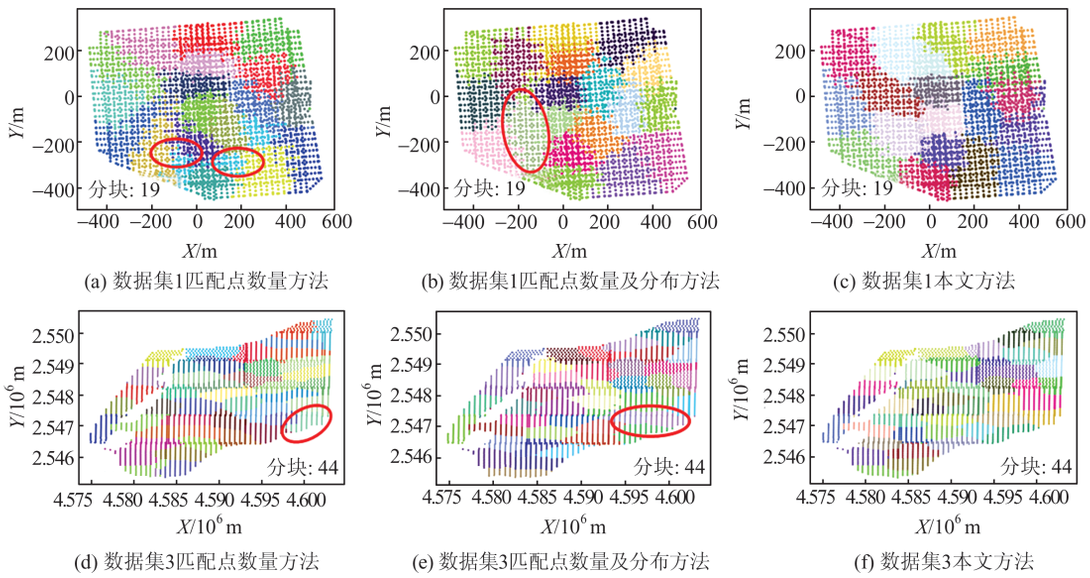
表 2　不同边加权方法的SfM重建统计结果

Table 2　Statistical Results of SfM Reconstruction for Different Edge Weighting Methods

方法	数据集1		数据集2		数据集3	
	重建影像数量/张	重建三维点数量/个	重建影像数量/张	重建三维点数量/个	重建影像数量/张	重建三维点数量/个
匹配点数量方法	3 592	999 602	3 711	1 547 028	21 576	10 709 637
匹配点数量及分布方法	3 712	1 027 807	3 793	1 541 896	21 538	10 617 461
本文方法	3 734	1 041 872	3 987	1 639 651	21 642	10 807 368

2.2.2　并行化合并方法的性能对比分析

基于本文的边加权方法构建场景图后，对比分析了不同的场景合并方法，包括基于树结构分层(Hier)的合并方法[21]、基于最小高度树(minimum height tree，MHT)的合并方法[11]，以及基于固定锚点(weakly connected dominating set，WCDS)的合并方法[16]。同时，使用重建影像数量、合并耗时以及最大内存占用3个指标进行结果评估。所有方法采用NC算法分块。实验结果如表3所示，可知：(1)对于小场景数据集1和2，所有方法均能实现完整重建，但是MHT在数据集2上丢失影像较多，主要原因是数据集2的结



(a) 数据集1匹配点数量方法 (b) 数据集1匹配点数量及分布方法 (c) 数据集1本文方法

(d) 数据集3匹配点数量方法 (e) 数据集3匹配点数量及分布方法 (f) 数据集3本文方法

图6 数据集1和3场景分块示意图

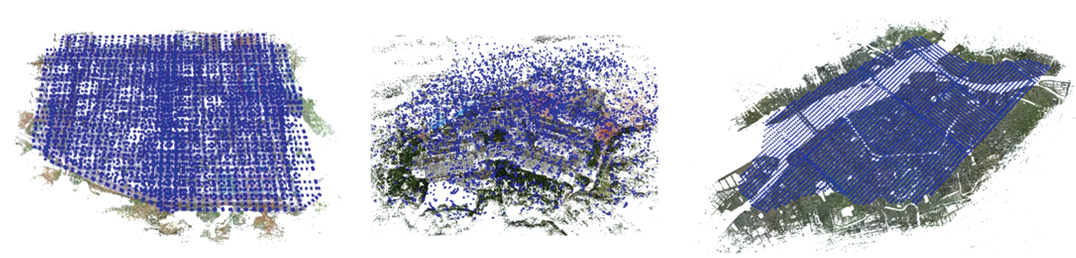
Fig. 6 Scene Clustering for Dataset 1 and Dataset 3

表3 不同子场景合并方法的性能统计

Table 3 Performance Statistics of Different Sub-Scene Merging Methods

方法	数据集1			数据集2			数据集3		
	重建影像数量/张	合并耗时/min	最大内存占用/GB	重建影像数量/张	合并耗时/min	最大内存占用/GB	重建影像数量/张	合并耗时/min	最大内存占用/GB
Hier	3 734	2.03	3.97	3 991	4.10	5.99	19 761	148.60	23.90
MHT	3 736	2.78	7.48	3 635	3.67	7.30	—	—	—
WCDS	3 734	25.20	2.14	4 028	80.10	2.51	—	—	—
本文方法	3 734	0.66	1.91	3 987	2.19	2.24	21 642	16.20	11.67

构较复杂,导致MHT合并的误差累积大。(2)对于大场景数据集3,Hier和本文方法能够实现较好重建。但是MHT和WCDS合并失败,主要原因是影像数量较大、子场景较多,一方面导致基于固定锚点的MHT合并路径较长,累积误差大;另一方面导致WCDS依赖的全局模型误差大,子场景合并难度大。(3)在合并耗时方面,本文方法用时最少,分别为0.66 min、2.19 min和16.20 min;(4)在最大内存占用方面,本文方法的最大内存占用最少,分别为1.91 GB、2.24 GB和11.67 GB。与Hier相比,本文方法内存节约率达到51.2%~62.6%。图7展示了本文方法的重建结果,可以看出,本文方法能够在保证内存占用和时间消耗性能良好的情况下完成可靠的场景合并。



(a) 数据集1 (b) 数据集2 (c) 数据集3

图7 本文方法在3个数据集上的重建结果

Fig. 7 Reconstruction Results of the Proposed Method for Three Datasets

2.2.3 相对定向和绝对定向对比分析

为评估本文方法的整体性能,首先对比分析4个SfM方法相对定向的重建结果,包括开源软件包COLMAP和DAGSfM、商业软件Agisoft Metashape和Pix4Dmapper。本文使用效率、精度和完整度3个指标进行性能评估。其中,效率指

整体时间消耗;精度指 BA 优化的重投影误差;完整度包括重建影像数量和三维点数量。实验结果如表 4 所示,可以看出:(1)效率方面,本文方法耗时最少,在 3 个数据集中分别为 18.1 min、39.7 min 和 990.4 min;与 COLMAP 相比,本文方法加速比最高可达 3.5。(2)精度方面,Pix4Dmapper 的精度最高,在数据集 1 和 2 上重投影精度分别为 0.318 像素和 0.327 像素;本文方法的精度与 COLMAP 相当。(3)完整度方面,利用数据集 2 的 GCPs 进行重建模型的绝对定向精度评价。本文方法和 COLMAP 的性能最佳,在数据集 1 和 2 上,重建影像数量占比超过 99%。但是在数据集 3 上,COLMAP 重建影像数量占比为 91.3%,本文方法重建影像数量占比高达 99.9%。综上所述,对于相对定向,本文方法可实现最高的重建效率和完整度。

表 4　不同方法相对定向统计结果

Table 4　Statistical Results of Relative Orientation for Different Methods

方法	数据集 1				数据集 2				数据集 3			
	效率/min	精度/像素	三维点数量/个	重建影像数量/张	效率/min	精度/像素	三维点数量/个	重建影像数量/张	效率/min	精度/像素	三维点数量/个	重建影像数量/张
COLMAP	36.4	0.422	1 036 605	3 734	68.1	0.556	1 538 331	3 991	3 481.6	1.350	7 621 218	19 761
DAGSfM	62.5	0.653	1 282 296	3 736	101.1	0.638	1 415 671	3 634	—	—	—	—
Agisoft Metashape	50.0	0.957	1 764 717	3 741	186.0	1.140	1 536 021	4 030	—	—	—	—
Pix4Dmapper	298.2	0.318	468 254	3 620	636.4	0.327	726 366	3 909	—	—	—	—
本文方法	18.1	0.508	1 041 872	3 734	39.7	0.641	1 639 651	3 987	990.4	0.799	10 807 638	21 642

本文选取 3 个分布均匀点作为地面控制点,其余点作为精度检查点。使用地面控制点与 SfM 重建模型点在 X、Y、Z 3 个方向残差的最大值、平均值和标准差作为评价指标。利用本文方法与 COLMAP、DAGSfM 和 Pix4Dmapper 进行对比,还设计了在本文方法基础上增加了完整性约束进行对比,即场景分块时设置子场景重叠影像数为 20。实验结果统计如表 5 所示,检查点残差如图 8 所示。从表 5 可以看出:(1)Pix4Dmapper 的精度最高,残差标准差在 X、Y、Z 3 个方向上分别为 0.013 m、0.016 m 和 0.019 m。(2)本文方法的精度与 DAGSfM 相当。本文方法的残差平均值分别为 0.030 m、0.043 m 和 0.041 m。分析图 8 可知,本文方法在某些检查点残差较大,主要原因是分块间无重叠影像。增加重叠影像后,增加完整性约束的本文方法中残差较大的点消失,且残差最大值、平均值、标准差均有所减小。综上所述,本文方法能够生成精确的定位定向结果。

表 5　不同方法绝对定向统计结果/m

Table 5　Statistical Results of Absolute Orientation for Different Methods/m

方法	残差最大值			残差平均值			残差标准差		
	X	Y	Z	X	Y	Z	X	Y	Z
COLMAP	0.101	0.188	0.276	0.044	0.054	0.049	0.053	0.060	0.081
DAGSfM	0.112	0.074	0.113	0.043	0.035	0.039	0.053	0.041	0.035
Pix4Dmapper	0.028	0.036	0.048	0.010	0.012	0.015	0.013	0.016	0.019
本文方法	0.100	0.252	0.140	0.030	0.043	0.041	0.042	0.065	0.054
增加完整性约束的本文方法	0.134	0.094	0.086	0.033	0.030	0.046	0.042	0.041	0.040

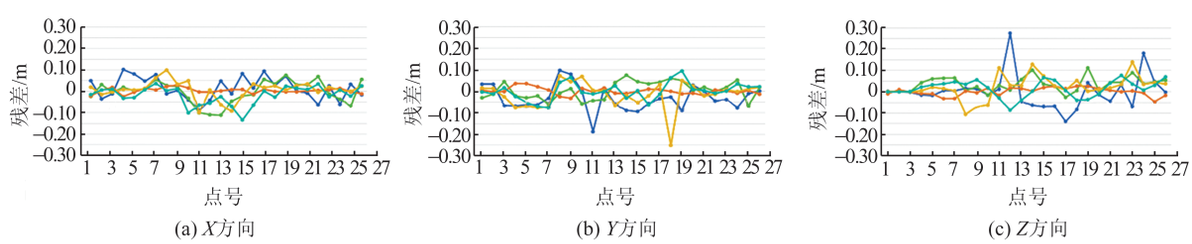
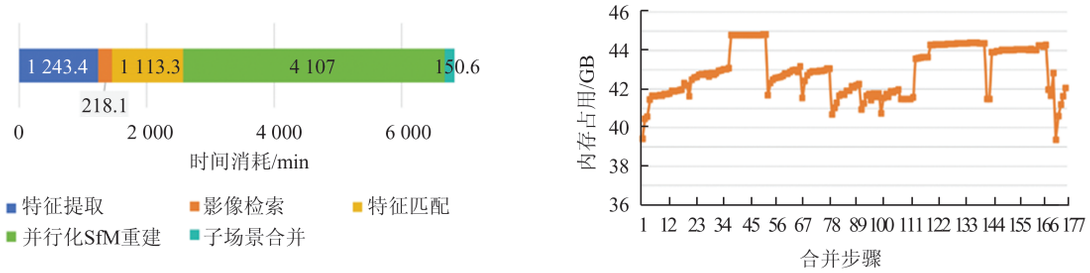


图 8　检查点残差分布图

Fig. 8　Individual Residual Plots of Check Points

2.2.4 大场景三维重建性能分析

利用数据集4验证本文并行化SfM方法处理大场景的性能。数据集4包含近9万张影像,覆盖范围约50 km²。实验结果表明,上述商业和开源软件均处理该数据集失败。为了详细分析不同步骤的性能,将本文方法划分为特征提取、影像检索、特征匹配、并行化SfM重建和子场景合并5个步骤。图9统计了本文方法的时间和内存消耗。本文方法总体时间消耗为113.87 h。特征提取、特征匹配和并行化SfM重建是最耗时的3个环节,其时间消耗分别为1 243.4、1 113.3和4 107 min,子场景合并的时间消耗为150.6 min,且峰值内存消耗小于45.0 GB。图10展示了数据集4的重建结果,包含86 748张重建影像和39 869 508个三维点,验证了本文方法对大场景的适应能力。



(a) 三维重建整体耗时 (b) 内存消耗

图9 时间和内存消耗统计

Fig. 9 Statistics of Time and Memory Consumption



图10 数据集4并行化重建结果

Fig. 10 Parallel SfM Result of Dataset 4

3 结 语

针对现有方法存在的场景图构建和子场景合并存在的问题,本文提出了一种不依赖固定锚点合并的UAV影像并行化SfM方法。一方面,针对SfM重建中下一张最优影像选择的特点,设计了顾及影像局部连接强度的场景图边加权方法,实现场景图的紧凑分块,保证了分块内部影像SfM重建的高稳健性。进一步建立了不依赖固定锚点的子场景并行化合并方法,实现合并路径自适应调整、多子场景并行化合并,有效提升了ISfM的精度和效率。利用多组大场景UAV影像进行实验。结果表明,本文所提的并行化SfM方法可显著提升子场景合并的稳健性和效率,降低内存消耗,成功应用于大规模UAV影像重建。后续研究将聚焦于特征匹配效率提升和分布式计算实现并行化子场景重建。

参 考 文 献

[1] JIANG S, SfMJIANG W S, WANG L Z. Unmanned Aerial Vehicle-Based Photogrammetric 3D Mapping: A Survey of Techniques, Applications, and Challenges[J]. *IEEE Geoscience and Remote Sensing Magazine*, 2022, 10(2): 135-171.

[2] WANG X, XIANG H Y, NIU W Y, et al. Oblique Photogrammetry Supporting Procedural Tree Modeling in Urban Areas[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2023, 200: 120-137.

[3] MURTIYOSO A, GRUSSENMEYER P. Documentation of Heritage Buildings Using Close-range UAV Images: Dense Matching Issues, Comparison and Case Studies[J]. *The Photogrammetric Record*, 2017, 32(159): 206-229.

[4] JIANG S, JIANG W S, HUANG W, et al. UAV-Based Oblique Photogrammetry for Outdoor Data Acquisition and Offsite Visual Inspection of Transmission Line[J]. *Remote Sensing*, 2017, 9(3): 278.

[5] ZHANG Q, TIAN Y Q, DING Y F, et al. CoopHance: Cooperative Enhancement for Robustness of Deep Learning Systems[C]//ACM SIGSOFT International Symposium on Software Testing and

Analysis, Seattle, USA, 2023.

[6] 许志华, 吴立新, 刘军, 等. 顾及影像拓扑的 SfM 算法改进及其在灾场三维重建中的应用[J]. 武汉大学学报(信息科学版), 2015, 40(5): 599-606.
XU Zhihua, WU Lixin, LIU Jun, et al. Modification of SfM Algorithm Referring to Image Topology and Its Application in 3-Dimension Reconstruction of Disaster Area[J]. *Geomatics and Information Science of Wuhan University*, 2015, 40(5): 599-606.

[7] 陈武, 姜三, 李清泉, 等. 无人机影像增量式运动恢复结构研究进展[J]. 武汉大学学报(信息科学版), 2022, 47(10): 1662-1674.
CHEN Wu, JIANG San, LI Qingquan, et al. Recent Research of Incremental Structure from Motion for Unmanned Aerial Vehicle Images[J]. *Geomatics and Information Science of Wuhan University*, 2022, 47(10): 1662-1674.

[8] WU C C, AGARWAL S, CURLESS B, et al. Multicore Bundle Adjustment[C]//EEE/CVF Conference on Computer Vision and Pattern Recognition, Colorado Springs, USA, 2011.

[9] 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.

[10] VERHOEVEN G. Taking Computer Vision Aloft-Archaeological Three-Dimensional Reconstructions from Aerial Photographs with Photoscan[J]. *Archaeological Prospection*, 2011, 18(1): 67-73.

[11] CHEN Y, SHEN S H, CHEN Y S, et al. Graph-Based Parallel Large Scale Structure from Motion [J]. *Pattern Recognition*, 2020, 107: 107537.

[12] JU Y, HUANG Y H, ZHENG J T, et al. Multithread Parallel Algorithm for Reconstructing 3D Large-Scale Porous Structures[J]. *Computers & Geosciences*, 2017, 101: 10-20.

[13] 姜三, 马一尘, 李清泉, 等. 无序无人机影像的并行化 SfM 三维重建方法[J]. 测绘学报, 2024, 53(5): 946-958.
JIANG San, MA Yichen, LI Qingquan, et al. Parallel SfM-Based 3D Reconstruction for Unordered UAV Images[J]. *Acta Geodaetica et Cartographica Sinica*, 2024, 53(5): 946-958.

[14] LIU Z D, QV W, CAI H L, et al. An Efficient and Robust Hybrid SfM Method for Large-scale Scenes [J]. *Remote Sensing*, 2023, 15(3): 769.

[15] CUI H N, GAO X, SHEN S H, et al. HSfM: Hybrid Structure-from-Motion[C]//IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017.

[16] JIANG S, LI Q Q, JIANG W S, et al. Parallel Structure from Motion for UAV Images via Weighted Connected Dominating Set[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5413013.

[17] XU B, ZHANG L, LIU Y X, et al. Robust Hierarchical Structure from Motion for Large-Scale Unstructured Image Sets[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2021, 181: 367-384.

[18] ZHU S Y, SHEN T W, ZHOU L, et al. Parallel Structure from Motion from Local Increment to Global Averaging[PP/OL]. V3. arVix (2017-06-05)[2024-04-18]. https://arxiv.org/abs/1702.08601.

[19] KLODT M, VEDALDI A. Supervising the New with the Old: Learning SfM from SfM[C]//European Conference on Computer Vision, Munich, Germany, 2018.

[20] LI Q Q, HUANG H, YU W S, et al. Optimized Views Photogrammetry: Precision Analysis and a Large-Scale Case Study in Qingdao[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2023, 16: 1144-1159.

[21] SCHÖNBERGER J L, FRAHM J M. Structure-from-motion Revisited[C]//IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016.