

一种基于场景图分割的混合式多视图三维重建方法

薛俊诗¹ 易辉¹ 吴止媛¹ 陈向宁¹

摘要 针对大范围三维重建, 重建效率较低和重建稳定性、精度差等问题, 提出了一种基于场景图分割的大范围混合式多视图三维重建方法. 该方法首先使用多层次加权核 K 均值算法进行场景图分割; 然后, 分别对每个子场景图进行混合式重建, 生成对应的子模型, 通过场景图分割、混合式重建和局部优化等方法提高重建效率、降低计算资源消耗, 并综合采用强化的最佳影像选择标准、稳健的三角测量方法和迭代优化等策略, 提高重建精度和稳健性; 最后, 对所有子模型进行合并, 完成大范围三维重建. 分别使用互联网收集数据和无人机航拍数据进行了验证, 并与 1DSFM、HSFM 算法在计算精度和计算效率等方面进行了比较. 实验结果表明, 本文算法大大提高了计算效率、计算精度, 能充分保证重建模型的完整性, 并具备单机大范围场景三维重建能力.

关键词 机器视觉, 多视图三维重建, 场景图分割, 核 K 均值算法, 迭代优化, 混合式重建

引用格式 薛俊诗, 易辉, 吴止媛, 陈向宁. 一种基于场景图分割的混合式多视图三维重建方法. 自动化学报, 2020, 46(4): 782–795

DOI 10.16383/j.aas.c180155

A Hybrid Multi-View 3D Reconstruction Method Based on Scene Graph Partition

XUE Jun-Shi¹ YI Hui¹ WU Zhi-Huan¹ CHEN Xiang-Ning¹

Abstract To solve the problem of low computational efficiency and poor stability of large scale 3D reconstruction, a novel hybrid scheme of large scale reconstruction was proposed. Scene graph was partitioned by multi-level weighted kernel K-means algorithm at first; then sub-scenes were reconstructed by hybrid reconstruction producing sub-models, in which improved optimal image selection criteria, robust triangulation methods and iterative optimization strategies were adopted, and the computational efficiency was improved by using strategies of scene graph partition, hybrid reconstruction and partial bundle adjustment (BA); Finally, All sub-models were merged into the final reconstruction result. Experiments were performed using images collected from the internet and UAV aerial images respectively, and comparison was made with 1DSFM and HSFM in terms of computation accuracy and computation efficiency. Experimental results demonstrate the proposed algorithm greatly improves computational efficiency and computational accuracy, fully ensures the integrity of the reconstructed scene and is able to reconstruct large scale scene in single computer.

Key words Machine vision, 3D reconstruction, scene graph partition, kernel K-means, iterative optimization, hybrid reconstruction

Citation Xue Jun-Shi, Yi Hui, Wu Zhi-Huan, Chen Xiang-Ning. A hybrid multi-view 3D reconstruction method based on scene graph partition. *Acta Automatica Sinica*, 2020, 46(4): 782–795

多视图三维重建是指利用相机从不同视角拍摄的影像, 经过特征点提取与匹配、相机姿态估计、三角测量和光束法平差(Bundle adjustment, BA)等步骤, 恢复相机的位置和姿态, 重建场景的三维结构的过程. 近年来, 大范围三维重建技术受到广泛的关注, 而从运动中恢复结构(Structure from motion,

SFM)是实现三维重建的一项关键技术, 常应用于城市级的三维重建中. 大范围三维重建技术在文物保护、自动驾驶、即时地图构建、军事侦察和情报获取^[1]等方面具有重大的应用价值.

目前, 常用的三维重建算法主要包含增量式重建、全局式重建和混合式重建三类^[2], 其中增量式三维重建是最早得到广泛研究和应用的重建方式, 它通过选择初始像对开始重建, 并以迭代的方式添加影像进行配准, 同时计算新增空间点的三维坐标, 最后通过光束法平差实现相机参数和场景三维结构的优化. Bundler 和 VSFM^[3]是典型增量式重建系统, 具有较高的重建精度和重建稳定性. 然而, 初始像对和最佳视图的选择对重建的精度和稳健性具有重要的影响. 当重建的范围较大时, 重建精度会随着重建

收稿日期 2018-03-20 录用日期 2018-09-27
Manuscript received March 20, 2018; accepted September 27, 2018

国家高技术研究发展计划(863)计划(2014AA7031072E), 军队探索项目(7131145)资助

Supported by National High Technology Research and Development Program of China (863 Program) (2014AA7031072E), Exploration Project of the Army (7131145)

本文责任编辑 吴毅红
Recommended by Associate Editor WU Yi-Hong

1. 航天工程大学航天信息学院 北京 101416
1. School of Space Information, Space Engineering University, Beijing 101416

影像的增加而降低,重建的场景结构还可能出现偏移. VSFM 和 ISFM^[4] 提出使用重三角测量来消除累积误差,但随着影像数目的进一步增加,其效果受到限制.同时,由于增量式重建需要反复进行 BA 优化,系统的重建效率较低、对计算资源消耗较大.文献 [5-7] 提出利用凝聚聚类算法采用自上而下的分层重建法,提高重建效率,避免误差累积,实现大范围三维建模,但会造成场景完整度的缺失. Agarwal 等^[8] 设计了著名的“一天重建罗马”三维重建系统,首次实现了增量式三维重建的分布式处理,但该方法需要消耗大量的计算资源. Heinly 等^[9] 综合使用标志影像聚类^[10-11] 和增强词汇树表示^[12] 等方法,提出了基于流的三维重建架构,可进行世界规模的三维重建,但必须反复进行影像间重叠关系的检测,场景重建的完整性有待进一步改善.文献 [13-14] 分别提出了利用 POS 信息的重建算法,能够高效地进行大范围场景的高精度重建.文献 [15] 提出分块-重建-合并的增量式三维重建策略,但只适用于具有坐标信息的影像.

全局式三维重建在进行大范围场景重建中具有重建效率高、不存在误差积累等优势,近年来,得到了较大的发展.全局式重建从对极几何场景中同时估计所有相机的初始姿态和中心位置,在完成三角测量后,只进行一次全局光束平差,最终得到相机的位姿和场景的三维结构. 1DSFM^[16] 和 OpenMVG 是典型的全局重建系统,主要由旋转平均和平移平均两个过程组成. Govindu^[17] 首次提出利用旋转平均进行相机姿态估计,以四元数对旋转矩阵进行参数化,采用奇异值分解法得到旋转矩阵的最小二乘解. Martinec 等提出使用 Frobenius 范数求解旋转平均问题^[18],而文献 [19-20] 提出了基于旋转流形李代数结构的最小化 L_1 范数求解方法,这两种方法能够一定程度提高旋转平均求解算法的抗噪声性能,但求解稳定性都有待提高. Jiang 等提出了一种针对全局相机姿态配准的线性求解方法^[21-22],而 Arie 等提出使用本质矩阵分解的方法求解全局相机平移矩阵^[23],这两种方法进一步提高了全局法的求解效率,但依赖高精度的对极几何关系,对外点极为敏感. Wilson 等构建的 1DSFM 重建系统^[16] 通过将平移平均求解映射为一维排序问题,缓解了噪声影响,却带来了重建稳定性的降低.文献 [24] 将旋转矩阵参数化为单位球体上的一组离散标签,然后利用马尔科夫随机场离散置信传播求解旋转平均问题,该方法能够在一定程度克服噪声影响,且适用于大范围重建,但其实现困难、计算复杂度高.文献 [25] 采用鲁棒性更好的 L_1 范数,每次只更新一个旋转参数,迭代进行,直到达到全局最优.该方法能够充分利用李代数特性进行鲁棒估计,但在进行大范

围三维重建时计算精度和效率都较低.可见,全局式重建法能够同时从场景图中估计所有相机的位置和姿态,可以极大地提高计算效率,但是重建精度受噪声影响较大,重建的稳定性存在一定的不足.

混合式重建法在近几年才开始得到一定的发展.它主要是通过结合增量式与全局式重建的各自优势,实现场景的快速重建.目前,大部分混合式三维重建算法只适用于小场景或者序列影像.为解决大范围重建问题, Havlena 等^[26] 和 Bhowmick^[27-28] 等利用场景图的二分结构,提出了分而治之的混合式三维重建方法,但在场景图分割过程中造成子场景图与匹配链之间大量的连接影像的丢失,影响了重建精度和重建的完整性.文献 [2] 提出了一种基于自适应聚类分层的混合重建方法,首先以全局的方式估计摄像机的旋转,然后以增量式重建方法求解相机中心,在保证重建精度的前提下大大提高了重建效率,但重建完整性也有待提高.

综上所述,设计适用于大范围场景的三维重建系统,需要综合考虑系统的重建效率、重建精度和重建完整性等各方面内容.本文的主要贡献在于:1) 为提高大范围场景三维重建效率,提出了一种基于多层次加权核 K-means 场景分割的混合式三维重建算法,通过对场景图进行简化-分割-优化,实现大范围重建区域的场景图分割;2) 对分割后的子场景图进行混合式重建,提出一种强化的最佳影像选择标准,增强重建的鲁棒性,并使用局部优化策略,进一步提高了系统重建效率.另外,在混合式重建阶段,采用文献 [29] 中基于 L_1 优化算子的全局旋转平均估计方法,提高系统的可扩展性及对外点的鲁棒性,并使用最小化最大重投影误差法进行三角测量,以较低的计算成本实现完整的场景重建;在完成相机中心估计后,主要根据文献 [4] 中的异常点过滤、重三角测量和迭代优化方法,进一步降低系统误差积累风险,提高重建的精度和稳健性.

1 系统架构

本文提出的基于场景图分割的三维重建系统流程如图 1 所示,主要包含特征点提取与匹配、几何校验与场景图建立、场景图分割、子场景图混合式重建和子模型合并等模块.

1) 特征点提取与匹配

目前常用的特征点提取算法有 SIFT、SURF、KAZE 和 BRIEF 等^[30],为了提高重建效率,本文采用 SIFTGPU 进行特征点提取.特征点提取完成后,大规模无序影像的特征点匹配往往要耗费大量的计算时间.本文选用文献 [31] 提出的匹配方法,使用词汇树进行特征点匹配,以满足大规模重建需求.

2) 几何校验与场景图建立

SFM 通过估计特征点在不同影像之间的投影关系进行几何校验. 像对之间的几何关系一般由如下变换关系描述: 单应矩阵描述了像对之间匹配对应点的空间变换关系; 本质矩阵或基础矩阵描述了像对之间的对极几何关系^[32]. 如果一个变换关系能够将一幅影像上足够多的点映射到另一幅影像上, 一般可认为实现了匹配几何验证. 通常在进行变换关系估计时必须进行鲁棒性估计, LO-RANSAC 是最常用的算法.

在完成匹配校验后, 对于重建影像序列, $I = \{I_1, I_2, \dots, I_N\}$, 任意两张影像 I_i, I_j 间的匹配点数目为 N_{ij} , 无匹配点时 $N_{ij} = 0$, 影像 I_i, I_j 对应相机之间的相对旋转矩阵为 R_{ij} , 平移向量为 t_{ij} . 在构建场景图时, 首先挑选一张影像作为初始影像, 将其旋转矩阵设为 3×3 单位矩阵, 位置设为 $[0, 0, 0]$, 这时, 任一影像 i 可根据相对位置、姿态关系确定一个未经优化的初始相机位置 m_i . 在构建场景图 $\Gamma = \{v, \xi\}$ 时, 每一个顶点 $m_i \in v$ 表示相机 $C_i \in C$ 的位置, 每一条边 $e_{ij} \in \xi$ 表示两相机 C_i, C_j 之间的连接关系. 对于加权场景图 $\Gamma = (v, \xi, W)$, W 为 $|v| \times |v|$ 的邻接矩阵, 其中, 任意元素 w_{ij} 为两顶点连线的权值, 本文中取 $w_{ij} = N_{ij}$, 当没有连接时, $w_{ij} = 0$.

3) 场景图分割

多层次加权核 K-means 场景图分割算法是一种自上而下的分割方法, 所有的数据初始化为类, 然后进行递归地分割. 如图 2 所示, 该算法主要由简化分层、层内分割与分割优化三个过程组成. 该内容将在第 3 部分进行详细论述.

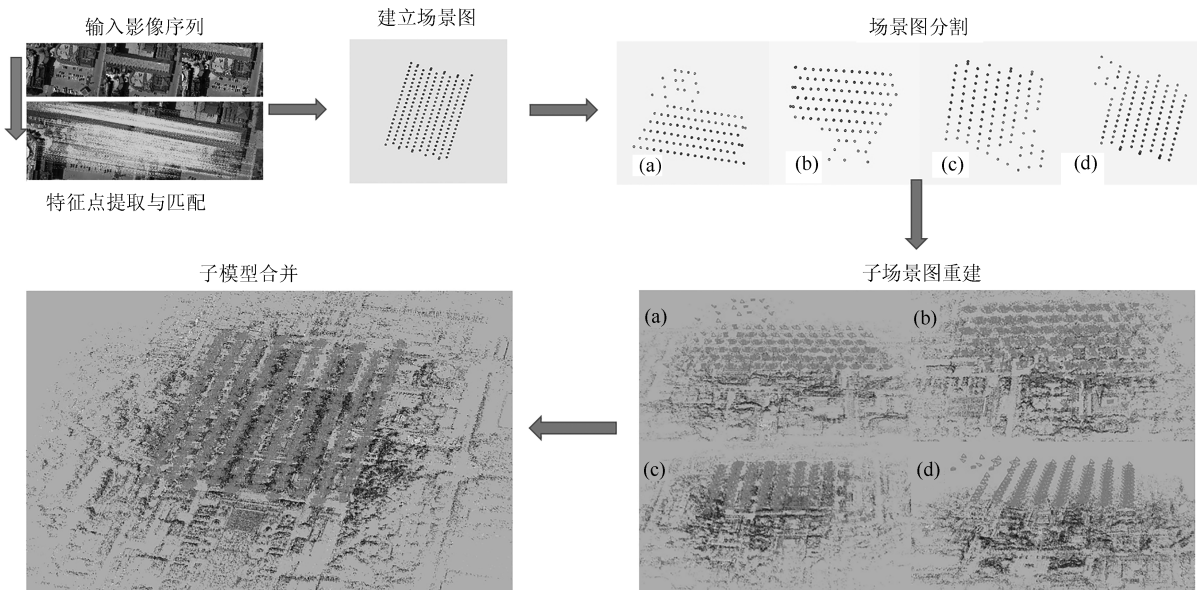


图 1 基于场景分割的三维重建流程图

Fig. 1 3D reconstruction based on scene partition pipeline

4) 子模型重建

对分割后的子场景图, 采用混合式方法进行重建. 首先是对子场景图进行全局式旋转平均鲁棒估计, 得到相机间的旋转矩阵. 然后选择初始像对, 完成初始重建. 迭代进行下述操作: 选择最佳添加影像, 然后进行相机中心估计, 并通过三角测量得到新的场景点. 当子模型增加到一定程度 (大于 50 时), 选择采用部分邻域的形式, 进行 BA 优化.

5) 子模型合并

子场景图重建后, 得到的子模型在不同的坐标系内. 为了得到完整的三维重建, 必须对所有的子模型进行配准, 使它们在一个统一的坐标内. 本文在进行场景图分割过程中, 在子场景图之间设置一定的重叠影像, 在子模型合并时, 可利用这些重叠影像, 计算不同子模型之间的相似变换与平移变换 (最少需要三张影像). 在进行配准时, 选择重建影像数目最多的子模型作为基准, 直接通过使用改进的 ICP (Iterative closest point) 算法^[33] 完成不同子模型之间的配准与合并.

6) BA 优化

Levenberg-Marquardt 算法是目前进行 BA 全局优化问题最常用的算法, 它能够充分利用多视图几何矩阵的稀疏性, 进行舒尔补问题求解. 目前, Ceres solver 库^[34] 中集成了该算法, 能够适用于各种非线性优化问题的求解. 本文也采用该方法进行 BA 优化. 除在子场景图混合式场景中使用 BA 优化外, 在完成子模型合并后, 还需要对整个重建场景进行 BA 优化.

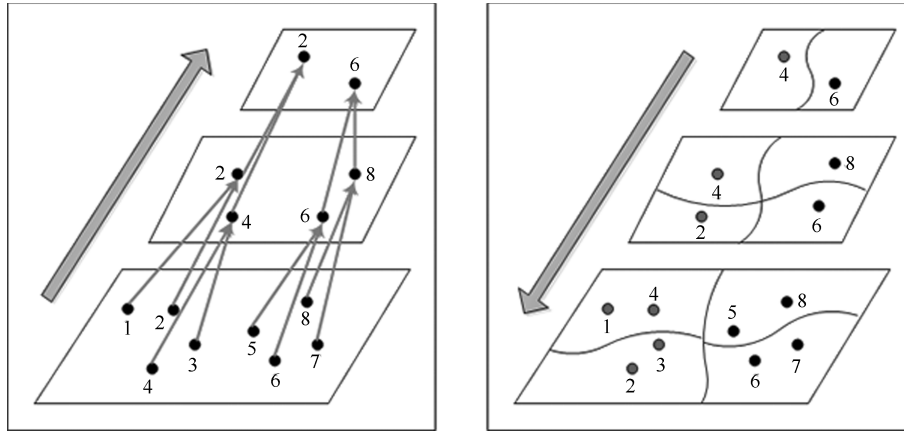


图 2 多层次场景分割示意图

Fig. 2 Diagram of multi-level scene segmentation partition

2 场景图分割

2.1 分割原则

1) 大小适宜: 分割应尽量满足每个子场景图中的相机数量适中, 各个子场景图中相机数目接近, 但不必相等. 首先, 大小适中的影像集可以有效地避免反复 BA 优化和可能的偏移现象, 同时, 使重建系统不受计算资源限制; 其次, 各子场景图中影像数目接近能够提高系统的可扩展性, 并能够在一定程度上提高效率. 综上所述, 场景图分割应满足以下条件:

$$\begin{cases} \forall \Gamma_k \in \Gamma, & Num(C_k) \leq H_0 \\ \forall \Gamma_i, \Gamma_j \in \Gamma, & |Num(C_i) - Num(C_j)| \leq L_0 \end{cases} \quad (1)$$

其中, H_0 与 L_0 分别为子场景中影像数目上限与影像数目差值下限. 随着集群影像数量的增加, 由局部增量式 SFM 计算得到的重投影误差首先显著下降, 然后趋于稳定. 因此, 子场景中影像数目可在较大范围内变化, 本文选择 $H_0 = 500, L_0 = 150$ 来进行精度和效率之间的平衡.

2) 场景完整性: 引入场景的完整性主要是保证相机与相机之间的有效连接, 从而确保子场景重建后, 子模型与子模型能够完整拼接. 为保证重建的完整性, 本文在子场景图之间设置一定的重叠相机, 如图 3 所示. 但重叠相机的数目不宜过多, 否则会影响计算效率, 且难以满足子场景图的大小约束. 本文在后续重建中选择重叠影像的数目 $N_O = 40$.

2.2 分层简化

分层简化阶段主要是通过不断合并初始场景图 Γ_0 的顶点, 实现多层次场景图的构建, 层与层之间顶点数递减, 其主要过程如下: 首先设置简化层级 L , 则本文选择 $L = |\nu_0|/300$, 通过简化可得到每一层的场景图 $\Gamma_0^*, \Gamma_1^*, \dots, \Gamma_L^*$, 其中

$|\nu_0| > |\nu_1^*| > \dots > |\nu_{L-1}^*|$. 如图 2 所示, 在逐层简化过程中, 由 Γ_i^* 中的一个或几个顶点组合生成 Γ_{i+1}^* 中的一个顶点, 新生成的顶点的权值为生成该点的权值之和. 对于初始场景, 所有的顶点都是未标记的. 随机访问任一顶点 x , 如果 x 是未标记的, 那么将 x 与其连接的所有顶点中挑选未标记的权值最大的点 y 合并, 并标记 x 与 y ; 如果与 x 连接的所有顶点都已被标记, 则 x 单独保留, 并进行标记. 当所有顶点都标记后, 意味着完成了本层简化, 并得到下一层的场景图. 依次进行, 直到所有的层级构建完毕. 同时, 为了防止过度简化, 当最低一级的顶点数目低于 200 时, 则不再进行简化. 最终, 得到最简化的场景图 Γ_{L-1}^* 中只包含少量的顶点. 在该层直接采用核 K-means 算法进行分割, 可以加速 K-means 算法的执行速度, 克服 K-means 收敛于局部最小的缺点.

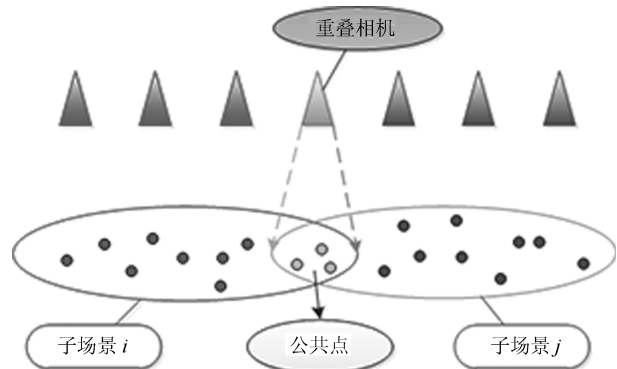


图 3 子场景分割示意图

Fig. 3 Diagram of Sub-scene Relationship

2.3 场景图分割

场景图分割的主要任务是将输入的场景图 Γ 划分为 k 个相互之间有一定重叠的子场景 $\Gamma_1, \Gamma_2, \dots, \Gamma_k$. 在每一层, 本文使用加权核 K-means 算法进行分割. 核函数可采用任意

常用的核函数, 对结果影响不大. 本文在后续实验过程中采用高斯函数, 即: $K(m_i, m_j) = \exp(-\|m_i - m_j\|/2\alpha^2) = K_{ij}$, 其中表示 2-范数, α 为系数. 根据 Mercer 定理, 存在映射 ϕ : $K(m_i, m_j) = \phi(m_i)\phi(m_j)$ ^[28]. 对于待分割场景图 $\Gamma(v, \xi, W)$, 从最粗化层 Γ_L^* 中随机选择 K 个顶点作为初始聚类中心 $D = \{m_1^*, \dots, m_k^*\}$; 然后, 遍历所有顶点, 根据式 (2) 判断每一个顶点的所属聚类.

这时, 每一个聚类中都存在一定数目的顶点, 根据 $\phi(m_c) = \frac{\sum_{m_i \in \Gamma_c} w_i \phi(m_i)}{\sum_{m_i \in \Gamma_c} w_i}$ 重新计算每一个聚类的中心, 并对聚类中心 D 进行更新, 然后, 再次重复进行上述步骤, 直到所有聚类中心都不再进行更新. 以最粗化 Γ_L^* 层的 K 个聚类中心为 Γ_{L-1}^* 的初始化中心, 然后重新进行上述步骤, 直到完成 Γ_{L-1}^* 的分割. 以此类推, 最终实现原始场景图的分割.

$$D(\{\Gamma_c\} |_{c=1}^k) = \min_{\nu_1, \dots, \nu_k} \sum_{c=1}^k \sum_{m_i \in v_c} w_i \|\phi(m_i) - \phi(m_c^*)\|^2 = \min_{\nu_1, \dots, \nu_k} \sum_{c=1}^k \sum_{m_i \in v_c} w_i \left(K_{ii} - \frac{2 \sum_{m_j \in v_c} w_j K_{ij}}{\sum_{m_j \in v_c} w_j} + \frac{2 \sum_{m_l, m_j \in v_c} w_j w_l K_{jl}}{\left(\sum_{m_j \in v_c} w_j \right)^2} \right) \quad (2)$$

算法 1. 场景图多层次核聚类分割算法伪码实现

输入: 影像集 $I = \{I_1, I_2, \dots, I_N\}$, 影像间相对旋转矩阵 R_{ij} , 平移向量 t_{ij} , 匹配点数目 N_{ij} .

过程:

- 1 计算影像 i 初始位置 $\{m_1, m_2, \dots, m_N\}$
- 2 构建加权场景图 $\Gamma = \{v, \xi, W\}$, 其中 $v = \{m_1, m_2, \dots, m_N\}$, $e_{ij} \in \xi$, $w_{ij} = N_{ij} \in W$
- 3 构建多层次场景图: $\Gamma_0^*, \Gamma_1^*, \dots, \Gamma_{L-1}^*$, 令 $l = L$
- 4 初始化聚类中心 $D = \emptyset$
- 5 While($l - -$)
- 6 If 初始聚类中心 $D = \emptyset$
- 7 从场景图 Γ_l^* 中随机选取 k 个顶点加入 D 中, $D = \{m_1^*, \dots, m_k^*\}$
- 8 Else continue;
- 9 Repeat
- 10 For $j = 1 : N$
- 11 根据式 (2) 计算 m_j 与各聚类中心的距离 d_{ji}
- 12 通过加权距离最小化原则, 判断所属聚类: $\lambda = \arg \min_{0 \leq i \leq k} w_j d_{ji}$

- 13 将 m_j 划入对应的聚类: Γ_λ^*
- 14 End for
- 15 For $i = 1 : k$
- 16 更新聚类中心: $\phi'(m_c) = \frac{\sum_{m_i \in \Gamma_\lambda^*} w_i \phi(m_i)}{\sum_{m_i \in \Gamma_\lambda^*} w_i}$
- 17 If $\phi(m_c^*) = \phi'(m_c)$ do nothing
- 18 If $\phi(m_c^*) \neq \phi'(m_c)$ 更新 $\phi(m_c^*) = \phi'(m_c)$, 记录边缘点变化, 加入 U_j
- 19 End For
- 20 Until 所有 $\phi(m_j)$ 没有更新
- 21 End if
- 22 对 U_j 中元素进行排序, 取前 50 作为最终元素, 删除其他元素
- 23 End While

输出: 分割后的场景图 $\{\Gamma_1, \Gamma_2, \dots, \Gamma_k\}$, 子场景图之间的重叠影像索引 $\{U_1, U_2, \dots, U_N\}$.

2.4 分割优化

对于给定的简化场景图 Γ_{i+1}^* , 通过如下方式得到场景图 Γ_i^* : 1) 如果场景图 Γ_{i+1}^* 中的某一顶点在子场景图 c 中, 那么场景图 Γ_i^* 形成该顶点的顶点组合都在子场景图 c 中. 2) 如果场景图中的某一场景图中的顶点超过了规定的数目, 则进行核 K-means 分割, 如图 2 所示. 3) 以场景图 Γ_{i+1}^* 分割得到的质心为初始点, 重新进行核均值分割, 得到优化后的结果. 重复以上步骤, 直到完成对场景图 Γ_0 的分割. 由于在每一层进行分割时, 都有很好的初始值, 所以整个优化过程会很快的收敛.

为了对完成重建后的各个子模型进行拼接, 并保证拼接后的模型的完整性, 需要在场景分割过程中设置场景图之间的重叠影像. 尤其是在实际的分割过程中, 进行优化时, 往往只存在边界的变化, 即原属于某一个子场景图的边界点经过优化后, 变为另一个子场景图的边界点. 在这个变化过程中, 我们记录下这类点, 并对其按照边界点权值的积 $m \cdot n$ 最大进行排序, $m = \sum_{j \in \Gamma_i} w_{ij}$, $n = \sum_{j \in \Gamma_{i+1}} w_{ij}$, Γ_i, Γ_{i+1} 为相邻子场景, 从排序中选择前 50 作为子场景重叠相机. 场景图分割效果如图 1 所示, 输入数据为大雁塔航拍数据, 在建立场景图后, 在建立场景图后, 通过场景分割, 可得到 4 个子场景图, 然后分别进行混合式重建生成子模型, 在完成子模型合并后, 实现整个重建的三维重建.

综上所述, 场景图多层次加权 K-means 分割算法的伪码流程如算法 1 所示.

3 子场景混合式重建

在完成场景图分割后, 需进行场景图的混合式重建, 重建流程主要包含全局旋转矩阵估计、两视图

重建、最佳视图选择、相机中心估计、三角测量和迭代优化等模块,如图4所示。

3.1 旋转矩阵估计

对于某一子场景图 Γ_k , 相机 i 与相机 j 之间的相对旋转矩阵为 R_{ij} , 相机 i 的绝对旋转矩阵为 R_i , 于是有: $R_{ij} = R_j R_i^{-1}, \forall \{i, j\} \in \xi_k$. 所有的旋转矩阵构成特殊正交群 $SO(3)$, 具有黎曼流形的可微性质. $SO(3)$ 作为一个李群, 可使用该条件提高旋转平均的求解效率. 设 $\omega = \theta n$, 其中, θ 为绕单位坐标轴 n 的旋转角度, $so(3)$ 为李代数, 于是有 $R = e^{[\omega]_{\times}} \in SO(3)$, 其中 $[\omega]_{\times}$ 为 ω 的反对称矩阵. 可定义全局旋转表示 $R_{\text{global}} = \{R_1, \dots, R_N\}$, 其对应的角表示形式 $\omega_{\text{global}} = [\omega_1, \dots, \omega_N]^T$. 相对旋转矩阵的一阶李代数近似可写为:

$$\omega = \omega_j - \omega_i = [0, \dots, -I \dots I, \dots, 0] \omega_{\text{global}} = A_{ij} \omega_{\text{global}} \quad (3)$$

于是有 $A \omega_{\text{global}} = \omega_{\text{rel}}$, 其中 A 为所有系数矩阵 A_{ij} 的堆叠, ω_{rel} 为所有相对旋转角观测值的堆叠. 本文选用对噪声鲁棒性更好的 L_1 范数进行求解, 于是有:

$$\arg \min_{\omega_{\text{global}}} \|A \omega_{\text{global}} - \omega_{\text{rel}}\|_{L_1} \quad (4)$$

在得到旋转矩阵的初值后, 采用迭代加权最小二乘法进一步得到全局旋转平均的优化求解^[35]. 使用该方法进行全局旋转矩阵估计, 具有较高的计算效率, 且对噪声有很强的鲁棒性, 能够为后续重建提供高精度的相机位姿值.

3.2 最佳新增影像

新增加影像的选择标准主要考虑三方面的内容: 1) 已重建模型在新增影像中有足够多的可视点; 2) 可视点在新增影像中的分布情况; 3) 基线距离长, 在得到相机的旋转矩阵之后, 可以通过计算两个相机匹配特征点之间的角度来衡量两个相机之间的基线距离, 具体来说, 如果相机 i 与相机 j 存在匹配对

应点 (p_i, p_j) , 且旋转矩阵分别为 R_i, R_j , 那么它们之间的夹角为 $\arccos(R_i^T p_i, R_j^T p_j)$, 可依此来衡量两个相机之间的基线距离: 夹角越大, 基线距离越长. 设, S_i, Q_i 分别为影像 I_i 中已重建场景的可见点的数目和提取的特征点数目, A_i 是影像 I_i 的整体面积, 是可见点的凸包面积. 定义最佳影像选择指数如下:

$$a_i = \frac{1}{3} \frac{S_i}{Q_i} + \frac{1}{3} \frac{CH(S_i)}{A_i} + \frac{1}{3} \frac{\arccos(R_i^T p_i, R_j^T p_j)}{180^\circ} \quad (5)$$

显然, a_i 在范围 $[0, 1]$ 内, 值越大, 表明候选影像更适合作为最佳影像. 为了进一步提高计算效率, 首先, 对所有未添加影像进行筛选, 确定候选影像, 只有满足一定可视点阈值条件 (大于等于 100) 的影像才可作为候选影像; 其次, 对候选影像按照场景可视点的数目多到少进行排序, 选择可视点最多的前 5 个作为候选影像, 其他的删除; 最后, 计算 5 张影像的选择指数 a_i , 选择值最大的候选影像作为最佳视图. 通过使用强化的最佳影像选择策略, 可以提高重建算法的稳健性.

3.3 相机中心估计

如图5所示, 新增影像的特征点可分为两部分: 一部分是重叠点, 这些点是已重建场景在新增影像中的可视点, 可用来恢复相机的中心位置; 另一部分是新增点, 这些点在完成新增影像配准后通过三角测量计算得到. 经过全局旋转平均后, 可得到任意新增相机 i 的旋转矩阵 R_i , 利用重叠点 $X_p = [x_p, y_p, z_p]^T, p = 1, \dots, N$, 通过最小化 3D 场景点的重投影误差和求解新增相机 i 的平移向量 t , 目标函数为:

$$\arg \min_{t_i} \sum_{p=1}^M \left\| u_p - \frac{r_{i1}^T X_p + t_{i1}}{r_{i3}^T X_p + t_{i3}}, v_p - \frac{r_{i2}^T X_p + t_{i2}}{r_{i3}^T X_p + t_{i3}} \right\| \quad (6)$$

其中, $R_i = [r_{i1}, r_{i2}, r_{i3}]^T, t = [t_{i1}, t_{i2}, t_{i3}]^T$. 2D 点 (u_p, v_p) 为空间点 X_p 对应的特征点. 设已注册影像 j 的相机中心的平移向量为 t_j , 新增影像与已注册影像 j 间的相对平移向量为 t_{jk} , 新增影像的旋转矩阵

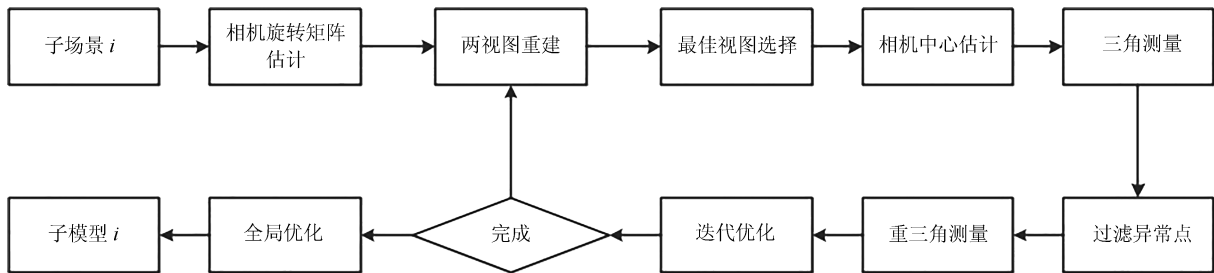


图4 子场景图混合式重建流程图

Fig. 4 Hybrid reconstruction of sub-scene graph pipeline

R_k 已在第 3.1 节通过全局旋转平均得到, 可通过获得新增影像的相机中心的平移向量的初值 \hat{t}_k . 在实际重建过程中, 与新增影像有重叠关系的影像数目可能较多, 可选择与新增影像之间匹配点数目最多的 8 张影像按照上述方法求得相机中心位置的初值后, 取平均作为新增影像的相机中心的初始位置. 这样, 可加快式 (6) 求解时的收敛速度. 使用该方法进行相机中心估计, 能够提高重建精度和系统的稳健性.

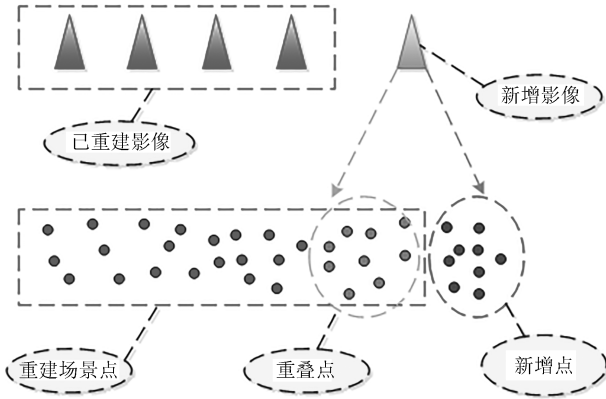


图 5 新增影像重建点分布情况示意图

Fig. 5 Distribution of the points in selected image

3.4 三角测量

三角测量是指在给定相机投影矩阵的情况下, 从影像 2D 特征点中估计场景 3D 结构的过程^[36]. 经过相机位姿估计, 可以得到相机 i 的投影矩阵 $P_i = [R_i | -t_i]$, $i = 1, 2, \dots, N$, 对于任意 2D-3D 匹配对应点 $[u_p, v_p]$ 与 $X_p \in \mathbf{R}^3$, 提出使用最小化最大重投影误差计算新增 3D 点 X 的坐标值, 目标函数为:

$$\arg \min_X \max_{p=1}^N \left\| u_p - \frac{r_{i1}^T X + t_{i1}}{r_{i3}^T X + t_{i3}}, v_p - \frac{r_{i2}^T X + t_{i2}}{r_{i3}^T X + t_{i3}} \right\|$$

$$\text{s.t. } r_{i3}^T X + t_{i3} > 0, \forall i = 1, \dots, N$$

(7)

使用本文方法进行三角测量, 能够均匀分布重投影误差, 进一步避免误差积累, 提高重建的稳健性, 同时降低外点率.

3.5 局部优化

目前, 三维重建在进行 BA 优化时, 一般采用全局优化的形式进行, 即所有已重建的影像全部参与优化. 每添加一张影像就需要进行一次全局优化, 大大降低了系统的重建效率. 实际上, 新添加影像与已重建影像之间的连接关系图 6 所示, 一级连接影像是与新添加影像由直接匹配对应点的影像, 次级连接影像是除新添加影像外与一级连接影像有直接匹

配对点的影像, 依次类推. 依据文献 [37] 中相关结论, 提出使用一级连接影像及部分二级连接影像参与 BA 优化. 具体方法如下:

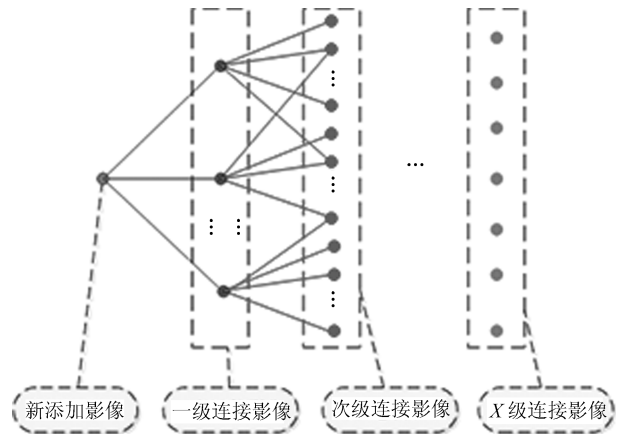


图 6 重建结果中的影像连接关系

Fig. 6 The image connecting relationship of 3D reconstruction result

在完成新增影像配准和三角点测量后, 根据 2D-2D 匹配链可得到已重建影像中与新增影像直接相连的一级影像, 通过一级连接影像可获得次级连接影像. 本文选择参与局部优化的影像数目为 50, 每次优化前, 确定参与优化的影像集合 Q_i : 首先从一级连接影像中选择, 如果一级连接影像不足, 则从次级连接影像中按照可视点的由多到少进行补足. 于是, BA 优化函数可写为:

$$g(R, t, X) = \sum_{i \in Q_i} \sum_{j=1}^M \|u_{ij} - P_i X_{ij}\|^2 \quad (8)$$

其中, u_{ij} 为特征点, X_{ij} 为 u_{ij} 对应的 3D 重建点. 求解上式非线性方程, 使用最为广泛的是 Levenberg-Marquardt 算法, 它是介于牛顿法与梯度下降法之间的一种非线性优化方法, 对于过参数化问题不敏感, 能有效处理冗余参数问题, 使代价函数陷入局部极小值的机会大大减小. Ceres solver 库包含了 Levenberg-Marquardt 算法的求解方案, 本文使用该库对式 (8) 进行求解.

为进一步提高重建的精度与稳健性, 本文采用文献 [4] 中的外点滤除、重三角测量和迭代优化策略. 首先在经过全局 BA 后, 还是会存在一些投影误差较大的观测点需要滤除. 同时, 对于重建的三维点, 通过设置所有对应影像的最小三角测量角度进行几何配置的优化^[3]. 其次, 为降低场景偏移风险, 提高场景重建的完整性, 通过 2D-2D 匹配链, 检索前期三角测量失败的 2D 点, 重新进行三角测量, 并保留低于重投影滤波阈值的点. 最后, 进行迭代优

化, 即反复进行外点滤除、重三角测量和 BA 优化直到重建场景中不再有异常点.

4 实验验证

为验证本文重建系统的重建效率、重建精度和重建的完整性, 共进行 4 组实验, 并分别与 1DSFM、HSFM 进行了对比. 实验 1, 采用互联网数据 Rome Forum 与航测数据 (西安大雁塔), 验证本文场景图分割算法的有效性. 实验 2 采用互联网下载数据的 4 组数据进行三维重建, 分别为龙泉寺^[38]、Yorkminster、Piccadilly 和 Trafalgar^[22], 其中龙泉寺数据集规模较小, 且相机排列规范, 相片尺度一致性好, 重建难度 Yorkminster、Piccadilly 为中等规模数据集, Yorkminster 数据集影像间明暗变化大, 很多影像存在较严重遮挡, 且影像间连接较弱, 重建难度大; Trafalgar 为较大规模数据集, 对单机重建具备一定的挑战性. 实验 3 采用航测数据集, 分别为西安大雁塔、大连市与某城市市区, 不使用辅助数据直接进行三维重建, 并完成后续重建步

骤, 从而更直观的观测重建质量. 实验 4 采用采用我国嵩山地区带控制点航测数据集和带地理参考信息的 Quad 数据集进行三维重建, 进一步验证本文提出算法的精度. 实验环境为 8 核 2.8 GHz Intel Core i7-4900MQ CPU 的计算机, 算法采用 C++ 实现.

4.1 分块效果评估

大雁塔数据共包含 1045 张影像, Rome Forum 共包含 2364 张影像, 重建时, 设置场景图分割数目为 4, 子场景图之间的重叠影像数目为 50. 可见, 不同子场景图重建后对应的区域位置不同, 但相互间保持一定的重叠. 其次, 从表 1 中可见, 分割后子场景图中的影像数目较为均衡, 能够保证重建效率, 分散重建误差; 合并前后重投影误差没有出现较大的变化, 证明了算法的有效性和场景的稳定性. 从图 8 中可见, 对于互联网数据, 子场景图中的相机位置与重建后的位置相差较大, 只使用相对位置信息进行场景图分割势必会造成较大的误差, 重建后的场景的完整性也就无法得到保证. 使用本文算法得到的

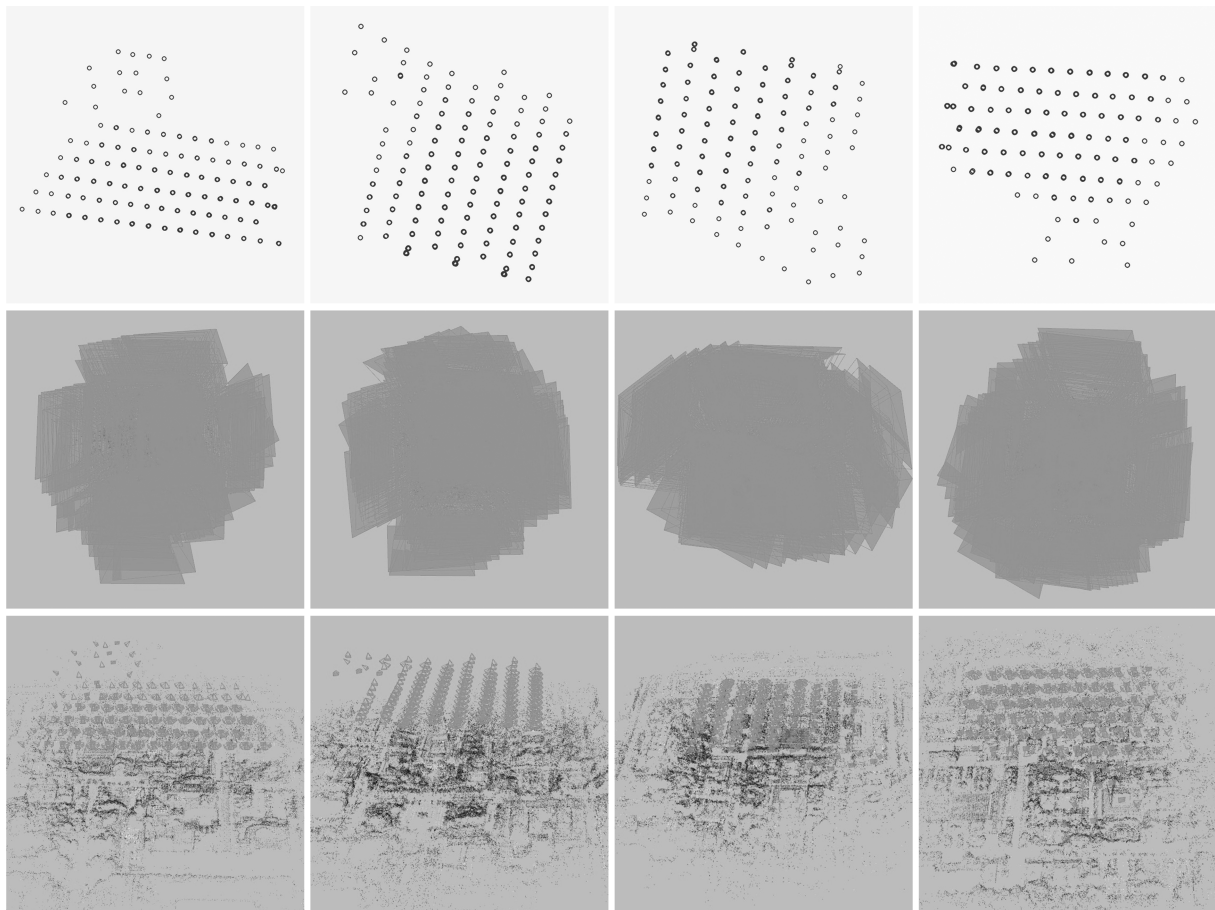


图 7 大雁塔场景图分割重建结果 (上: 分割后子场景图; 中: 子场景图相机地面投影; 下: 子场景图重建结果)

Fig. 7 Reconstruction results of DAYANTA based on scene graph partition (up: sub-scene graph; middle: camera ground projection in sub-scene graph; below: reconstruction result of sub-scene graph)

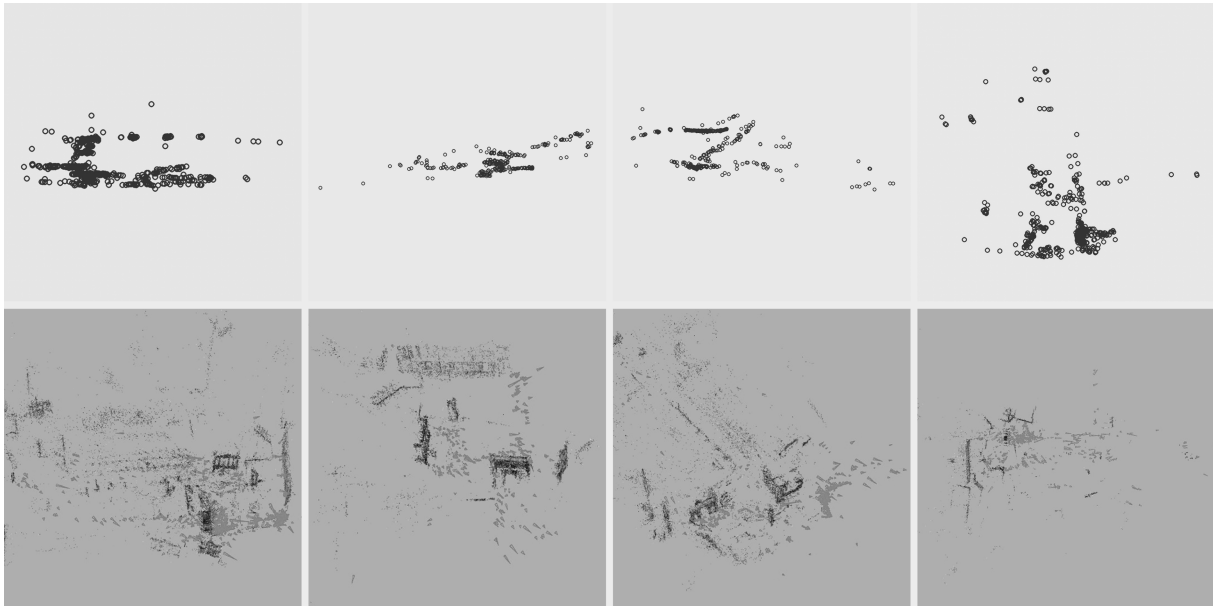


图 8 场景图分割重建结果 (上: 分割后子场景图; 下: 子场景图重建结果)

Fig. 8 Reconstruction results of Rome Forum based on scene graph partition (up: sub-scene graph; below: reconstruction result of sub-scene graph)

表 1 合并前后重建结果对比

Table 1 Comparison of reconstruction results before and after merging

数据集	子场景序号	合并前		合并后	
		N_R	A_{RE}	N_R	A_{RE}
大雁塔	1	296	0.416	1 045	0.455
	2	323	0.463		
	3	268	0.451		
	4	253	0.437		
Rome Forum	1	470	0.577	1 475	0.572
	2	386	0.556		
	3	355	0.583		
	4	431	0.572		

表中, N_R 为重建影像数目; A_{RE} 表示重投影误差, 单位 (pixel).

重建结果, 子场景图覆盖区域分布均匀, 最终模型重建精度较高.

4.2 重建效率和完整性

图 9(a)~(d) 分别为龙泉寺、Yorkminster、Piccadilly 和 Trafalgar 数据集的重建结果, 其中锥体表示重建得到的相机位置和姿态, 浅色点云代表场景的稀疏结构. 可见, 使用本文算法得到的重建结果, 相机位姿合理, 场景立面完整、结构清晰, 外点数目极少.

通过表 2 可见, 除龙泉寺数据外, 使用本文算法重建后, 注册影像数目多于 1DSFM 与 HSFM 重建结果, 表明本文算法对场景的重建完整性更好, 尤

其是对于存在一定挑战的 Yorkminster 数据集, 本文算法在保证重建精度的前提下, 共重建影像 1712 张, 重建结果更完整. 在重建效率方面, 本文算法与 HSFM 场景效率远高于 1DSFM, 同时, 以 Trafalgar 数据集为例, 本文算法重建时间为 438.443 min, 少于 HSFM 重建的时间 (481.857 min), 且本文重建注册的影像更多. 由于本文采用分割重建策略, 受计算资源限制较小, 可使用单机进行大范围场景的三维重建.

实验 3 采用三组航拍影像进行三维重建, 三组影像集分别为西安大雁塔、大连市与某城市市区, 全部采用五镜头相机获取. 使用本文算法完成稀疏重建后, 分别使用半全局匹配算法进行稠密重建、使用泊松算法进行表面重建, 完成纹理映射后结果如图 10 所示. 可见, 使用本文算法重建后的场景真实性良好, 场景内建筑、车辆等物体清晰、结构完整.

4.3 重建精度

根据表 2, 对各数据集重建的重投影误差进行对比可见, 本文算法重建精度大大优于 1DSFM 算法, 对比 HSFM 算法, 也有一定的优势. 为进一步验证本文算法的重建精度, 使用本文算法对带地面控制点的高山航测数据集与带地理参考信息的 Quad 数据集进行三维重建, 重建结果如图 11 所示.

高山数据集共有 3910 张影像, 影像尺寸为 10328×7760 , 采用航测飞机搭载 AMC580 五镜头相机进行影像获取, 飞行高度约为 900 m, 像元分辨率为 $0.0052 \text{ mm} \times 0.0052 \text{ mm}$, 其中相机 1 为下

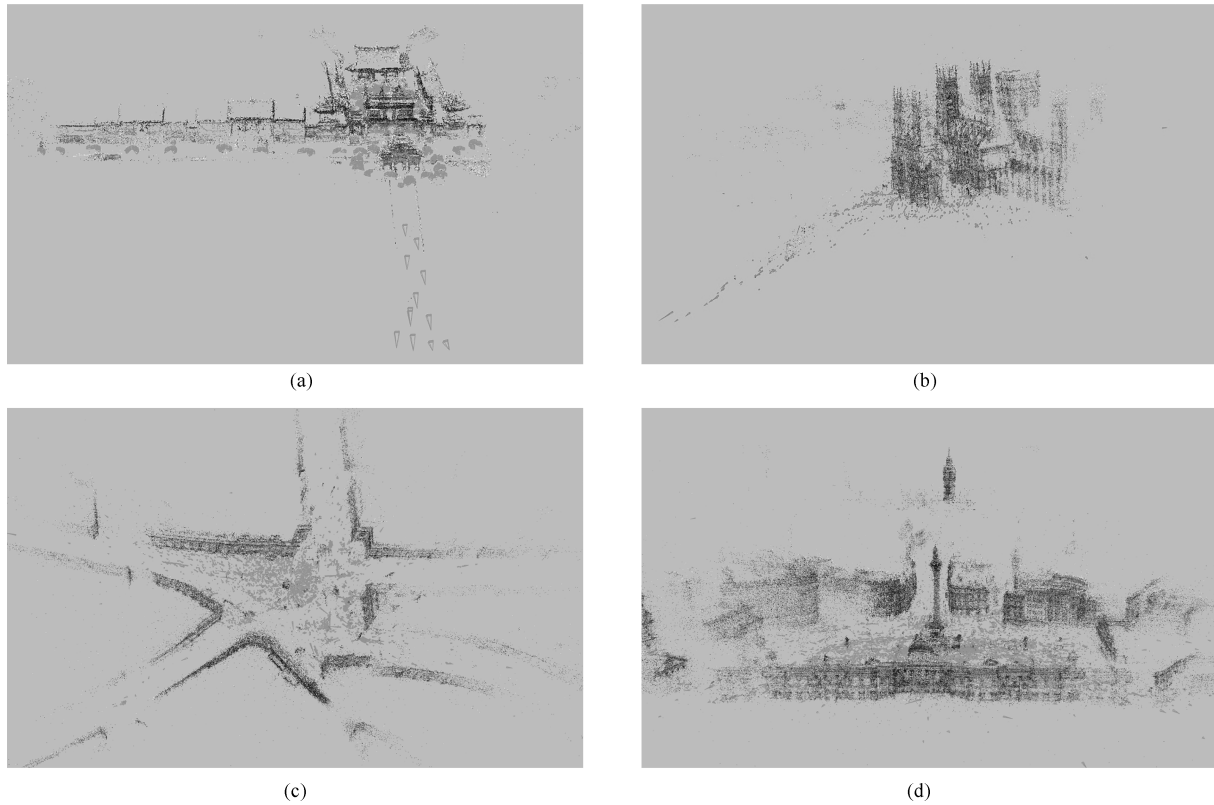


图 9 互联网数据集三维重建结果

Fig. 9 3D reconstruction result of datasets downloaded from internet

表 2 1DSFM、HSFM 与本文算法的不同数据集三维重建结果对比

Table 2 Comparison of 3D reconstruction result of different dataset using 1DSFM, HSFM and Ours

数据集 名称	1DSFM				HSFM				Ours	
	N_D	N_R	T_A	A_{RE}	N_R	T_A	A_{RE}	N_R	T_A	A_{RE}
龙泉寺	443	406	25.386	0.841	417	19.661	0.717	413	16.815	0.711
Yorkminster	3 368	1 176	93.910	0.736	1 472	64.726	0.628	1 712	65.803	0.607
Piccadilly	7 351	6 445	476.781	1.194	6 791	269.561	0.865	6 979	231.794	0.822
Trafalgar	15 683	9 384	765.685	1.173	11 943	481.857	0.837	12 741	438.443	0.816
西安大雁塔	1 045	1 043	58.357	0.617	1 045	31.637	0.510	1 045	26.946	0.455
大连市市区	4 900	4 900	327.625	1.397	4 900	231.394	1.478	4 900	197.872	1.353
某城市市区	15 750	NA	NA	NA	15 745	845.632	1.359	15 750	785.451	1.211

表中, N_D 为数据集中的影像数目; N_R 为重建影像数目; T_A 表示重建时间, 单位 (min); A_{RE} 表示重投影误差, 单位 (pixel).

视相机, 焦距为 55 mm, 侧视相机焦距为 80 mm.

由图 11 (a) 可见, 重建场景中航线稳定, 相机位姿合理, 物体结构清晰. 选择部分控制点作为选取部分控制点作为检查点, 得到的部分误差结果如表 3 所示, 三维重建场景的三维平均中误差为 0.4238 m, 平面平均中误差为 0.1873 m, 通过换算可得, 三维中误差优于 5 个像素, 平面中误差优于 2.5 个像素, 部分重投影误差分布如图 12 所示, 特征点与重投影点基本完全重合, 重建精度较高.

表 3 嵩山地区部分误差结果

Table 3 Partial error result of the Songshan area

序号	ΔX	ΔY	ΔZ	RMS
1	0.2727	0.2337	0.2512	0.4383
2	0.1222	0.1053	0.1597	0.3370
3	0.1725	0.3045	0.3623	0.5037
4	0.2045	0.3643	0.5239	0.6703
5	0.1380	0.1419	0.4578	0.4987

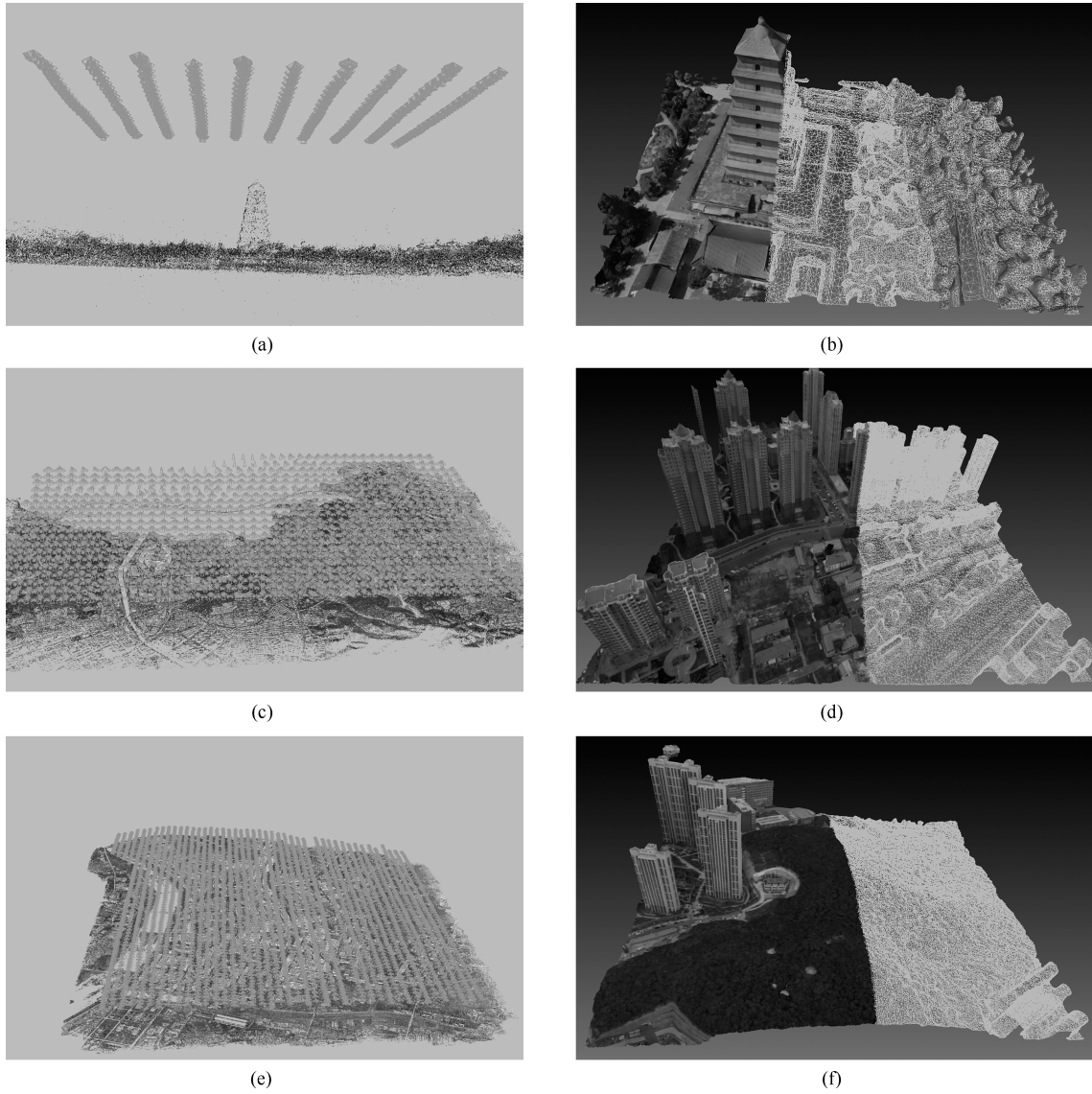
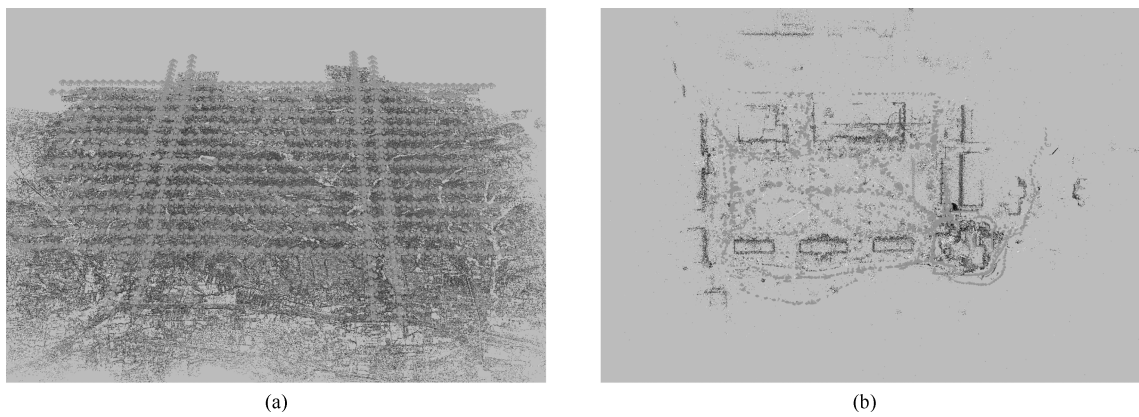


图 10 航测数据集 (西安大雁塔、大连市、某城市市区) 三维重建结果

Fig. 10 3D reconstruction result of aerial images (Da-Yan Tower in Xi'an, Dalian City, and A City Center)



(a) (b)

图 11 稀疏重建结果 (左: 嵩山, 右: Quad)

Fig. 11 Sparse reconstruction results (left: Songshan; right: Quad)



图 12 嵩山地区部分区域特征点与重投影点分布图 (× 表示特征点位置, + 表示重投影点位置)

Fig. 12 Distribution of feature points and re-projection points in the Songshan areas (×: feature points, +: re-projection points)

Quad 数据集提供了 348 张影像的相机位置参考数据, 采用部分相机位置信息进行坐标转换, 然后从未参与坐标转换的影像中选择部分相机位置作为检查点, 得到的部分误差结果如表 4 所示, 经过统计分析, 使用本文方法进行三维重建得到的最终结果的中误差为 1.0155 m, 优于 HSFM 的 1.03 m.

表 4 Quad 数据集部分误差结果

Table 4 Partial error result of Quad dataset

序号	ΔX	ΔY	ΔZ	RMS
1	0.2727	0.2337	0.2512	0.4383
2	0.1222	0.1053	0.1597	0.3370
3	0.1725	0.3045	0.3623	0.5037
4	0.2045	0.3643	0.5239	0.6703
5	0.1380	0.1419	0.4578	0.4987

综上所述, 本文提出的三维重建算法能够实现大范围场景三维重建, 并在重建精度、重建完整性和重建效率方面较 1DSFM 和 HSFM 算法有较大提高.

5 结论

本文提出了一种高效、高精度的三维重建方法, 该方法通过使用多层次加权核 K-means 算法进行场景图分割, 可实现大范围场景的高效重建, 同时, 提出使用混合式重建和局部优化等方法进一步提高了算法的重建效率. 通过综合使用强化的最佳影像

选择标准、稳健的三角测量方法、迭代优化等策略大大提高了算法的重建精度和稳健性. 经实验验证, 该算法能够进行航测数据和互联网收集数据的三维重建, 重建效率、重建精度和重建的完整性较 1DSFM 和 HSFM 算法有较大的提高.

References

- 1 Wang Xue, Shi Jian-Bo, Park Hyun-Soo, Wang Qing. Synchronization of Video Sequences Through 3D Trajectory Reconstruction. *Acta Automatica Sinica*, 2017, **43**(10): 1759–1772
(王雪, Shi Jian-Bo, Park Hyun-Soo, 王庆. 基于运动目标三维轨迹重建的视频序列同步算法. *自动化学报*, 2017, **43**(10): 1759–1772)
- 2 Cui H, Gao X, Shen S. HSFM: hybrid structure-from-motion. In: *Proceedings of the 2017 Conference on Computer Vision and Pattern Recognition*. IEEE, 2017. 2393–2402
- 3 Wu C. Towards linear-time incremental structure from motion. In: *Proceedings of In3D Vision-3DV 2013, 2013 International Conference*. IEEE, 2013. 127–134
- 4 Schonberger J L, Frahm J M. Structure-from-motion revisited. In: *Proceedings of the IEEE Computer Vision and Pattern Recognition*. IEEE, 2016. 4104–4113
- 5 Toldo R, Gherardi R, Farenzena M, Fusiello A. Hierarchical structure-and-motion recovery from uncalibrated images. *Computer Vision and Image Understanding*, 2015, **140**: 127–143
- 6 Song Zheng-Xi, Zhang Ming-Huan. 3D Reconstruction on Unmanned Aerial Video by Using Patch Clustering Matching Method. *Journal of Northwestern Polytechnical University*, 2016, **34**(4): 731–737

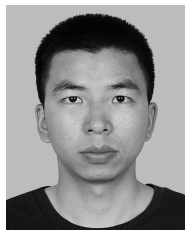
- (宋征玺, 张明环. 基于分块聚类特征匹配的无人机航拍三维场景重建. 西北工业大学学报, 2016, **34**(4): 731–737)
- 7 Cefalu A, Haala N, Fritsch D. Hierarchical structure from motion combining global image orientation and structureless bundle adjustment. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 2017, **42**: 535
- 8 Agarwal S, Furukawa Y, Snavely N. Building Rome in a day. *Communications of the ACM*, 2011, **54**(10): 105–112
- 9 Heinly J, Schonberger J L, Dunn E. Reconstructing the world in six days. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition, 2015: 3287–3295
- 10 Wang Wei, Gao Wei, Zhu Hai, Hu Zhan-Yi. Rapid and robust piecewise planar reconstruction of urban scenes. *Acta Automatica Sinica*, 2017, **43**(4): 674–684
(王伟, 高伟, 朱海, 胡占义. 快速鲁棒的城市场景分段平面重建. 自动化学报, 2017, **43**(4): 674–684)
- 11 Özyeşil O, Voroninski V, Basri R, Singer A. A survey of structure from motion. *Acta Numerica*. 2017, **26**: 305–64
- 12 Simone B, Ciocca G, Marelli D. Evaluating the performance of structure from motion pipelines. *Journal of Imaging*, 2018, **4**(8): 98
- 13 Cui H, Shen S, Gao W, Hu Z. Efficient large-scale structure from motion by fusing auxiliary imaging information. *IEEE Transactions on Image Processing*. 2015, **24**(11): 3561–3573
- 14 Zhu S, Shen T, Zhou L, Zhang R, Wang J, Fang T, Quan L. Accurate, scalable and parallel structure from motion [Ph.D. dissertation], Hong Kong University of Science and Technology, China, 2017
- 15 Yang Y, Chang MC, Wen L, Tu P, Qi H, Lyu S. Efficient large-scale photometric reconstruction using Divide-Recon-Fuse 3D Structure from Motion. In: Proceedings of the 13th IEEE International Conference on Advanced Video and Signal Based Surveillance. IEEE, 2016. 180–186
- 16 Wilson K, Snavely N. Robust global translations with 1dsfm. In: Proceedings of the 2014 European Conference on Computer Vision, Cham: Springer, 2014. 61–75
- 17 Govindu V M. Lie-algebraic averaging for globally consistent motion estimation. In: Proceedings of the 2004 Computer Vision and Pattern Recognition, IEEE, 2004. 684–691
- 18 Martinec D, Pajdla T. Robust rotation and translation estimation in multiview reconstruction. In: Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2007. 1–8
- 19 Jensen S H, Del Bue A, Doest M E, Aanæs H. *A Benchmark and Evaluation of Non-Rigid Structure from Motion*. arXiv preprint arXiv: 1801.08388, 2018.
- 20 Chatterjee A, Madhav Govindu V. Efficient and robust large-scale rotation averaging. In: Proceedings of the 2013 IEEE International Conference on Computer Vision, IEEE, 2013. 521–528
- 21 Jiang N, Cui Z, Tan P. A global linear method for camera pose registration. In: Proceedings of the 2013 IEEE International Conference on Computer Vision, IEEE, 2013. 481–488
- 22 Guo Fu-Sheng, Gao Wei. Batch reconstruction from UAV images with prior information. *Acta Automatica Sinica*, 2013, **39**(6): 834–845
(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. 自动化学报, 2013, **39**(6): 834–845)
- 23 Arie-Nachimson M, Kovalsky S Z, Kemelmacher-Shlizerman I, Singer A, Basri R. Global motion estimation from point matches. In: Proceedings of the 2012 Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission, IEEE, 2012. 81–88
- 24 Crandall D, Owens A, Snavely N, Huttenlocher D. Discrete-continuous optimization for large-scale structure from motion. In: Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2011. 3001–3008
- 25 Cui Z. *Global Structure-from-Motion and Its Application* [Ph.D. dissertation], Applied Sciences: School of Computing Science, 2017
- 26 Havlena M, Torii A, Pajdla T. Efficient structure from motion by graph optimization. In: Proceedings of the European Conference on Computer Vision, Berlin: Springer, 2010. 100–113
- 27 Bhowmick B, Patra S, Chatterjee A, Govindu VM, Banerjee S. Divide and conquer: efficient large-scale structure from motion using graph partitioning. In: Proceedings of the Asian Conference on Computer Vision, Cham: Springer, 2014. 273–287
- 28 Agudo A, Moreno-Noguer F. A scalable, efficient, and accurate solution to non-rigid structure from motion. *Computer Vision and Image Understanding*, 2018, 167: 121–133
- 29 Dhillon I S, Guan Y, Kulis B. Weighted graph cuts without eigenvectors a multilevel approach. *IEEE transactions on pattern analysis and machine intelligence*, 2007, **29**(11): 1–14
- 30 Karami E, Prasad S, Shehata M. Image matching using SIFT, SURF, BRIEF and ORB: performance comparison for distorted images. arXiv preprint arXiv: 1710.02726, 2017.
- 31 Li X, Larson M, Hanjalic A. Pairwise geometric matching for large-scale object retrieval. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2015. 5153–5161
- 32 Sweeney C, Sattler T, Hollerer T, Turk M, Pollefeys M. Optimizing the viewing graph for structure-from-motion. In: Proceedings of the 2015 IEEE International Conference on Computer Vision, IEEE, 2015. 801–809
- 33 Wei Sheng-Bin, Wang Shao-Qing, Zhou Chang-He, Liu Kun, Fan Xin. An iterative closest point algorithm based on biunique correspondence of point clouds for 3D reconstruction. *Acta Optica Sinica*, 2015, **35**(5): 244–250
(韦盛斌, 王少卿, 周常河, 刘昆, 范鑫. 用于三维重建的点云单应性迭代最近点配准算法. 光学学报, 2015, **35**(5): 244–250)

- 34 Agarwal S, Mierle K. Ceres Solver: Tutorial & Reference. Google Inc, 2012, 2:72.
- 35 Hartley R, Aftab K, Trunpf J. L_1 rotation averaging using the Weiszfeld algorithm. In: Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2011. 3041–3048
- 36 Zhang Q, Chin T J. Coresets for Triangulation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017.
- 37 Shah R, Chari V, Narayanan PJ. *A Unified View-Graph Selection Framework for Structure from Motion*. arXiv preprint arXiv: 1708.01125, 2017.
- 38 National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences. datasets for 3D reconstruction [Online], available: <http://vision.ia.ac.cn/data>, January 3, 2018



薛俊诗 航天工程大学航天信息学院博士研究生. 主要研究方向为计算机视觉, 摄影测量与遥感. 本文通信作者.
E-mail: xueao2015@sina.com
(**XUE Jun-Shi** Ph.D. candidate at the School of Space Information, Space Engineering University. His research interest covers computer vision, photogrammetry, and remote sensing. Corresponding author of this paper.)

photogrammetry, and remote sensing. Corresponding author of this paper.)



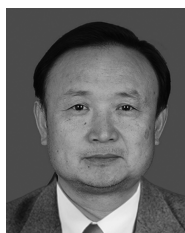
易辉 航天工程大学航天信息学院博士研究生. 主要研究方向为计算机视觉, 摄影测量与遥感.
E-mail: 18810962910@163.com

(**YI Hui** Ph.D. candidate at the School of Space Information, Space Engineering University. His research interest covers computer vision, photogrammetry, and remote sensing.)



吴止镭 航天工程大学航天信息学院博士研究生. 主要研究方向为遥感图像处理. E-mail: wuzhihuan@hotmail.com

(**WU Zhi-Huan** Ph.D. candidate at the School of Space Information, Space Engineering University. His main research interest is remote sensing image processing.)



陈向宁 航天工程大学航天信息学院教授. 主要研究方向为计算机视觉, 图像处理, 机器学习.

E-mail: laser115@163.com
(**CHEN Xiang-Ning** Professor at the School of Space Information, Space Engineering University. His research interest covers computer vision, image processing, and machine learning.)