

# 基于局部到全局优化策略的倾斜影像定向方法

叶培根, 杨泽, 孙岩标\*, 郝继贵

天津大学精密测试技术及仪器国家重点实验室, 天津 300072

**摘要** 作为影像位姿解算的传统方法, 光束平差法已被广泛应用于倾斜航空影像定向任务; 但其高度依赖良好初值的限制条件, 导致在导航系统缺失或者导航精度有限的飞行任务中常发生收敛缓慢或收敛失败的问题。因此, 为了在初值质量不佳或缺失的情况下进行有效迭代, 提出了一种基于局部到全局优化策略的倾斜图像定向方法来提高优化迭代解算的鲁棒性。所提方法首先根据最大重叠原则来构建“局部地图”(由 1 个正视图和 4 个斜视图构成), 然后通过优化这些局部地图来形成“全局地图”。在局部地图中, 图像同名点用作约束优化的观测值, 以此来获取每个局部地图的未知参数和对应的权值矩阵(信息矩阵)。在全局地图中, 局部地图的变量估计值及相应的权值矩阵一起构成全局优化问题的观测值和权值。通过一组大型仿真数据集和一组小型国际上公测的真实数据集, 对所提方法进行了测试。实验结果表明, 与传统方法相比, 所提方法在精度不丢失的前提下, 收敛速度和收敛性更优。

**关键词** 机器视觉; 倾斜摄影测量; 图像定向; 光束平差; 最小二乘估计; 全局优化

中图分类号 P231

文献标志码 A

doi: 10.3788/AOS202141.1415001

## Oblique Image Orientation Method Based on Local-to-Global Optimization Strategy

Ye Peigen, Yang Ze, Sun Yanbiao\*, Zhu Jigui

State Key Laboratory of Precision Measuring Technology and Instruments, Tianjin University, Tianjin 300072, China

**Abstract** The beam adjustment method has been widely used in oblique aerial image orientation tasks as a traditional image pose calculation method. However, it highly depends on a good initial value. This dependence often results in slow convergence or convergence failure in flight missions with missing navigation systems or limited navigation accuracy. Therefore, to effectively iterate when the initial value is missing or of poor quality, an oblique image orientation method based on a local-to-global optimization strategy is proposed in this paper. This strategy improves the robustness of the optimization iterative solution. In the proposed method, first, a “local map” (consisting of a nadir and four oblique images) according to the maximum overlap principle is constructed, and then, a “global map” is formed by optimizing these local maps. In the local map, the point of the image with the same name is used as the observation value of the constrained optimization to obtain the unknown parameters and corresponding weight matrix (information matrix) of each local map. In the global map, the estimates of variables and the corresponding weight matrix of the local map constitute the observations and weights of the global optimization problem. The proposed method was tested using a set of large simulated dataset and a set of small, real international public test dataset. The experimental results show that compared with the traditional methods, the proposed method can converge faster and more stably under the premise of no loss of accuracy.

**Key words** machine vision; oblique photogrammetry; image orientation; bundle adjustment; least-square estimation; global optimization

**OCIS codes** 150.1135; 110.6880; 100.6890; 100.2000

收稿日期: 2020-11-13; 修回日期: 2020-12-07; 录用日期: 2021-02-05

基金项目: 国家自然科学基金(52075382, 51721003, 41701522)

通信作者: \*yanbiao.sun@tju.edu.cn

# 1 引言

摄影测量因具有测量速度快、非接触、信息丰富等优点<sup>[1-2]</sup>,被广泛应用于三维模型重建<sup>[3]</sup>、目标位姿测量<sup>[4]</sup>等领域。倾斜影像定向是摄影测量在地理信息测量领域中的应用,其使用的倾斜多相机系统具有多个不同方向的相机,因此不仅捕获了更大区域的图像信息,还获取到了多个角度的场景图像,从而可以生成空间三维数据集。得益于此,倾斜摄影测量法<sup>[5]</sup>目前已被成功应用在城市三维建模<sup>[6]</sup>、点云分类等地理信息测绘场景中。

2015 年,来自丹麦和挪威的研究人员使用 1300 张倾斜图像成功建立了位于北格陵兰岛东部的基伦湖的三维地质模型,丰富了北极偏远地区的地理信息知识<sup>[7]</sup>。2009 年,位于阿拉斯加的里道特火山喷发,研究人员提出了一个分析熔岩穹顶生长的方法,该方法使用直升机观察和固定翼火山气勘测获得的倾斜数字图像,生成了一系列熔岩穹顶的数字高程模型(DEM),估算出了熔岩穹顶体积随时间的变化关系<sup>[8]</sup>。

尽管倾斜摄影测量法具有很多优点,并已成功应用于许多场景中,但它也面临一些挑战。例如,更大的视野使得图像匹配更加困难和不稳定,并且多相机的多角度结构使相机定位更加复杂且耗时。针对图像匹配这一挑战,许多研究人员提出了一系列新的算法<sup>[9-10]</sup>。另外,倾斜摄影中常用的光束平差(BA)法<sup>[11]</sup>也存在一定的局限性。

众所周知,当给定的初值误差较大时,传统的光束平差法会变得病态和发散<sup>[12]</sup>。例如,在没有高精度的 GPS 和 IMU 信息的情况下,仅依靠基于图像信息本身的方法来获得初值,当面对大规模和复杂的图像网络时,图像的测量噪声便会传播和累积,非

线性优化将变得异常困难。但幸运的是,少量图像的定向是较容易的。因为在少量图像中,干扰因素较少,从而保留了优化网络的简单性,因此可以相对容易地搜寻到相应目标函数的最小值。该假设为本文的研究提供了核心基础。本文从少量图像的局部优化开始,然后利用局部优化的结果对所有图像进行复杂的全局优化。

本文方法与传统的空中三角测量独立航带模型不同。独立航带模型通过各部分简单的旋转平移构建全局地图,这样每个局部的误差会累积和传播,导致最终解不优。所提方法通过消除刚性约束条件,将局部优化中得到的信息矩阵作为全局优化中的权值矩阵,允许误差在整个全局地图上进行权值调整,避免了局部误差的累积,最大程度地利用了观测信息。

## 2 原理与方法

### 2.1 局部地图的定义

五相机倾斜摄影系统是倾斜航空摄影测量中普遍使用的摄影系统,工作原理如图 1 所示。在侧面 4 个斜视相机倾斜角度较大的情况下,五相机系统单次观测所获得的 5 张图像之间没有重叠区域,即 5 张图像间将无法建立起相对约束关系,如图 1(b)所示。因此,如图 2 所示,按照重叠度最高的原则,选取每个曝光站处的正视图像和 4 个从其他曝光站处所获取的与其具有高重叠度的斜视图像构建“局部地图”。图 2(a)为 5 个曝光站捕获图像的示意图,其中图像 S1、S2、S3、S4 分别对应的 4 台斜视相机的观测方向都倾向图像 S0 所对应的正视相机,拍摄正视相机视角下方的对象,因此 S0~S5 这 5 张图像之间具有相当大的重叠区域,足以在估计 5 台摄像机之间相对姿态时提供强大的约束,这样的 5 张图像就构成了一个局部地图。

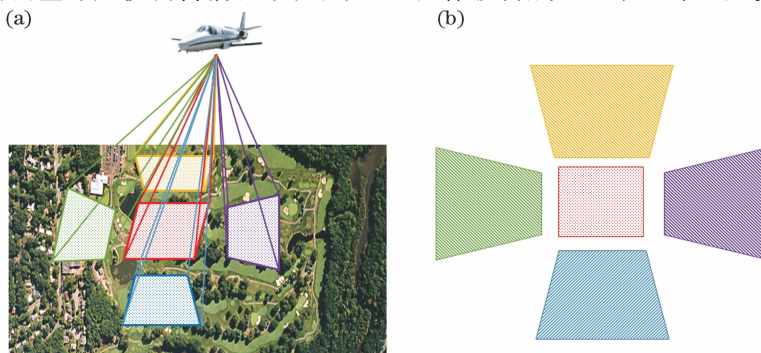


图 1 五相机倾斜摄影装置的工作原理。(a) 采自三维目标空间的 5 张图像;(b) 同一曝光站上捕获的 5 张图像的平面分布

Fig. 1 Working principle of five camera oblique photography device. (a) Five images captured in 3D object space;

(b) plane-directional distribution of five images captured on an exposure station

图 2(b) 给出了从 ISRPS EuroSDR 数据集<sup>[13]</sup> 生成的一个局部地图的示例。正视图像和斜视图像

之间的重叠度可以达到 70%，4 个斜视图像之间的重叠度可达到 40%。

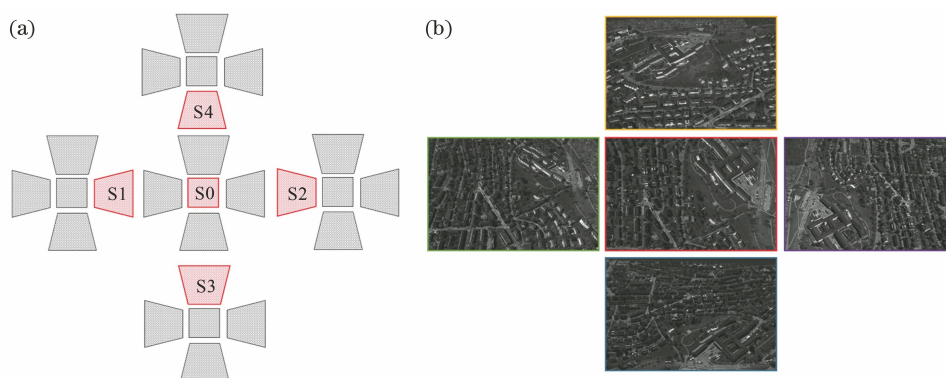


图 2 5 个不同曝光站的图片之间存在高重叠区域。(a) 5 个曝光站的平面分布；  
(b) 一个来自 ISRPS EuroSDR 数据集的局部地图示例

Fig. 2 There is a large overlap between the images on five different exposure stations. (a) Plane-directional distribution of the five exposure stations; (b) an example of local map from ISRPS EuroSDR dataset

### 2.2 优化策略的基本思想

在本研究中，相机位姿(为了方便描述，本文中某些地方也称为“图像位姿”)和空间特征点的三维坐标被联合定义为“地图特征”。在整个航空摄影过程中的所有曝光时刻的相机位姿(相对于全局坐标系的位姿)和空间特征点三维坐标(全局坐标系下的坐标)的集合称为“全局地图特征”；而上文所定义的仅包含 5 张图像的局部地图所对应的相机位姿(相对于其局部坐标系的位姿)和空间三维点坐标(其局部坐标系下的坐标)的集合则称为“局部地图特征”。

局部到全局优化(LGO)方法的原理结构如图 3 所示。第一阶段进行“局部优化”。按照局部地图的

定义，先将整个航空摄影过程中所获取到的图像生成为  $n$  个局部地图。利用光束平差法对每个局部地图内的相机位姿和空间三维点坐标进行估计，得到局部地图特征和信息矩阵，分别记为  $X_L^i$  和  $I_L^i$ ，其中  $i$  表示第  $i$  个局部地图， $i \in [1, n]$ 。

第二阶段进行“全局优化”。根据局部地图特征与由全局地图特征得到的局部特征之间的残差应该最小的原则，构建非线性最小二乘问题；然后根据对极约束，估计相邻正视图像之间的相对姿态、斜视相机相对于其对应正视图像的相对位姿，以获得全局优化最小二乘问题所需的初值；最后，利用高斯-牛顿(GN)法估计出定义在全局坐标系下的全局地图特征  $X_G$  和信息矩阵  $I_G$ 。

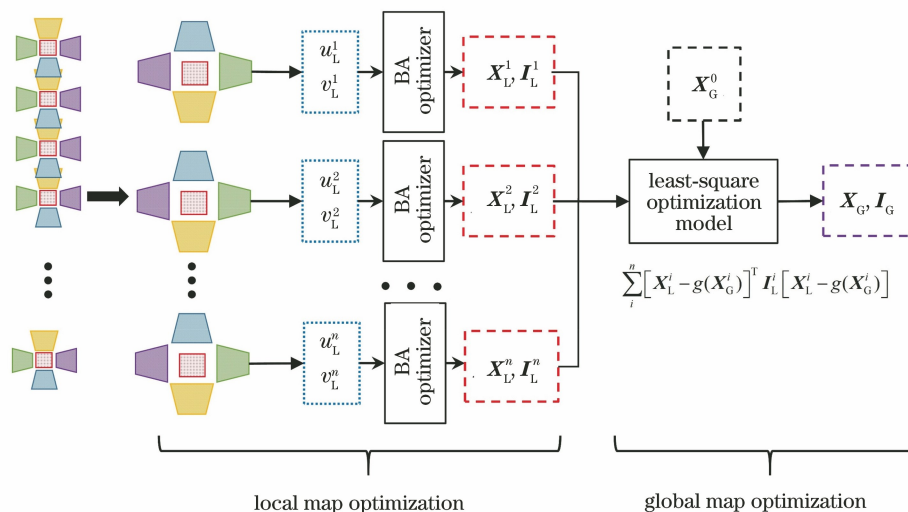


图 3 LGO 方法的原理结构

Fig. 3 Principle structure of the LGO method



### 2.3 局部优化

利用传统光束平差法估计局部地图特征  $\mathbf{X}_L^i$ 。每个局部地图都包含 1 个正视图和 4 个斜视图 (特别注意, 一个局部地图所包含的 5 个图像并不是来自于同一个曝光站的 5 个图像)。以正视图的中心为原点, 正视图中心和右斜视图中心在 XYZ 坐标系下的坐标作差, 令坐标差值最大的坐标轴方向上的距离始终固定为 1 (为了消除尺度信息的影响), 建立起局部地图坐标系。每个斜视图相对于正视图的位姿为  $(\mathbf{R}, \mathbf{t})$ , 其中  $\mathbf{R}$  和  $\mathbf{t}$  分别为旋转矩阵和平移向量。每个斜视图上所观测到的特征点的空间三维坐标为  $\mathbf{P}$ 。4 个斜视图位姿和 4 组特征点空间三维坐标则一起构成了局部地图特征  $\mathbf{X}_L^i$ 。对于每个斜视图, 都可以建立目标函数, 再利用传统光束平差法估计出未知参数  $\mathbf{X}$ ,  $\mathbf{X} = (\mathbf{R}, \mathbf{t}, \mathbf{P})$ 。目标函数表达式为

$$\|\boldsymbol{\varepsilon}\|_Q^2 = \sum [\mathbf{x} - f(\mathbf{X})]_Q^2, \quad (1)$$

式中:  $\mathbf{x}$  为从图像上获取到的特征点投影坐标, 即  $\mathbf{x} = [u, v]^T$ ; 协方差矩阵  $\mathbf{Q}$  用于描述每个观测值的相对不确定性。假设观测值受到的是独立均匀分布的高斯噪声的影响, 则协方差矩阵  $\mathbf{Q}$  可设置为单位矩阵  $\mathbf{I}$ 。

根据针孔相机模型, (1) 式中的观测函数可表示为

$$f(\mathbf{X}) = \begin{bmatrix} p_1/p_3 \\ p_2/p_3 \end{bmatrix}, \quad (2)$$

其中,

$$\begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = \mathbf{KR}(\mathbf{P} - \mathbf{t}), \quad (3)$$

式中:  $\mathbf{K}$  为相机内参矩阵。

此处使用高斯-牛顿法<sup>[14-15]</sup> 求解最优化问题。从初值  $\mathbf{X}_0$  开始, 通过迭代  $\mathbf{X}_k = \mathbf{X}_{k-1} + \Delta\mathbf{X}_k$  获得 (1) 式的解。其中  $\Delta\mathbf{X}_k$  表示每次迭代的状态向量的增量, 可通过增量方程得到。

$$\mathbf{J}^T \mathbf{Q}^{-1} \mathbf{J} \Delta\mathbf{X}_k = \mathbf{J}^T \mathbf{Q}^{-1} [\mathbf{x} - f(\mathbf{X})], \quad (4)$$

式中:  $\mathbf{J}$  为观测函数对于被估计参数  $\mathbf{X}$  的一阶导数;  $\mathbf{J}^T \mathbf{Q}^{-1} \mathbf{J}$  为法方程系数矩阵, 直接影响非线性优化的收敛性和收敛速度。由于最后一次迭代的法方程系数矩阵描述了被估计参数  $\mathbf{X}$  的不确定性, 因此计算它们的值, 然后将结果用作全局优化中所使用的加权矩阵。

### 2.4 全局优化

倾斜航空摄影的目标是构建全局地图, 而构建全局地图的核心就是得到全局地图的外部定向参数 (即全局地图特征中的位姿信息), 这也正是所提方法所要完成的“定向”任务。局部地图可以构建出全局地图, 自然地, 全局地图也可以提取出局部地图。因此, 如图 4 所示, 以由局部地图优化得到的局部地图特征  $\mathbf{X}_L^i$  与从全局地图中反投得到的局部地图特征  $g(\mathbf{X}_G^i)$  之间的残差应该最小为原则, 建立非线性最小二乘问题, 其中  $\mathbf{X}_G^i$  为全局地图特征  $\mathbf{X}_G$  中与第  $i$  个局部地图有关的地图特征元素。该最小二乘问题的目标函数为

$$\|\boldsymbol{\varepsilon}\|^2 = \sum_i^n [\mathbf{X}_L^i - g(\mathbf{X}_G^i)]^T \mathbf{I}_L^i [\mathbf{X}_L^i - g(\mathbf{X}_G^i)]. \quad (5)$$

对于任意第  $i$  个局部地图, (5) 式中的观测值  $\mathbf{X}_L^i$  由局部优化方法得到。为了利于优化迭代的数学计算过程, 全局优化过程中所用到的局部地图特

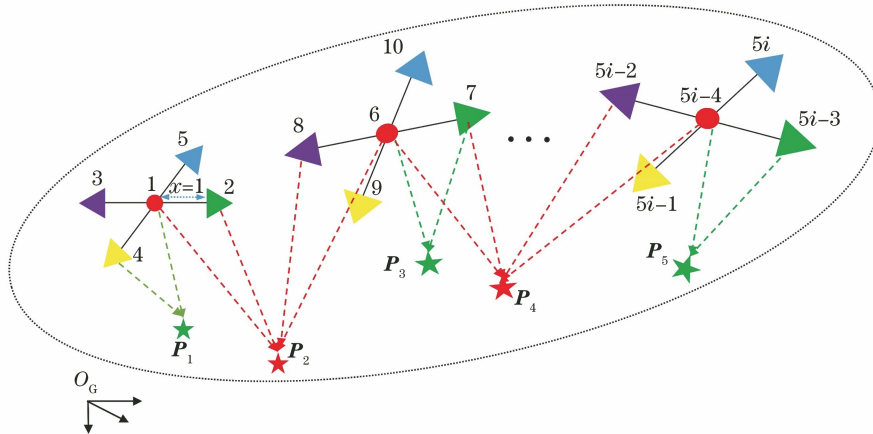


图 4 利用最小二乘优化将所有局部地图组合成全局地图

Fig. 4 Combining all local maps into a global map by a least-square optimization

征  $\mathbf{X}_L^i$  以欧拉角  $r$  的形式来表示姿态信息,而不以旋转矩阵  $\mathbf{R}$  来表示。因此,全局优化时需要注意两种形式之间的转换。相应观测函数  $g(\mathbf{X}_G^i)$  的表达式为

$$g(\mathbf{X}_G^i) = \begin{cases} \mathbf{r}_{5i-3}^L = r^{-1}(\mathbf{R}_{5i-3} \mathbf{R}_{5i-4}^T) \\ \mathbf{t}_{5i-3}^L = \mathbf{R}_{5i-4}(\mathbf{t}_{5i-3} - \mathbf{t}_{5i-4})/x_s \\ \mathbf{r}_{5i-2}^L = r^{-1}(\mathbf{R}_{5i-2} \mathbf{R}_{5i-4}^T) \\ \mathbf{t}_{5i-2}^L = \mathbf{R}_{5i-4}(\mathbf{t}_{5i-2} - \mathbf{t}_{5i-4})/x_s \\ \mathbf{r}_{5i-1}^L = r^{-1}(\mathbf{R}_{5i-1} \mathbf{R}_{5i-4}^T) \\ \mathbf{t}_{5i-1}^L = \mathbf{R}_{5i-4}(\mathbf{t}_{5i-1} - \mathbf{t}_{5i-4})/x_s \\ \mathbf{r}_{5i}^L = r^{-1}(\mathbf{R}_{5i} \mathbf{R}_{5i-4}^T) \\ \mathbf{t}_{5i}^L = \mathbf{R}_{5i-4}(\mathbf{t}_{5i} - \mathbf{t}_{5i-4})/x_s \\ \mathbf{P}_j^L = \mathbf{R}_{5i-4}(\mathbf{P}_j - \mathbf{t}_{5i-4})/x_s \end{cases}, \quad (6)$$

式中:  $\mathbf{R}_{5i-4}$  和  $\mathbf{t}_{5i-4}$  分别为第  $i$  个局部地图的正视相机相对于全局坐标系的旋转矩阵和平移矢量;  $(\mathbf{R}_{5i-3}, \mathbf{t}_{5i-3}), (\mathbf{R}_{5i-2}, \mathbf{t}_{5i-2}), (\mathbf{R}_{5i-1}, \mathbf{t}_{5i-1})$  和  $(\mathbf{R}_{5i}, \mathbf{t}_{5i})$  为在第  $i$  个局部图中的 4 个斜视图像相对于全局坐标系的旋转矩阵和平移矢量;  $(\mathbf{r}_{5i-3}^L, \mathbf{t}_{5i-3}^L), (\mathbf{r}_{5i-2}^L, \mathbf{t}_{5i-2}^L), (\mathbf{r}_{5i-1}^L, \mathbf{t}_{5i-1}^L)$  和  $(\mathbf{r}_{5i}^L, \mathbf{t}_{5i}^L)$  为从全局地图中提取的第  $i$  个局部图的 4 个斜视图像相对于局部坐标系的欧拉角和平移矢量;  $\mathbf{P}_j$  和  $\mathbf{P}_j^L$  分别为第  $j$  个空间特征点在全局地图坐标系和局部地图坐标系中的坐标;  $r^{-1}(\cdot)$  为从旋转矩阵得到欧拉角的函数。在局部地图坐标系的建立中,已将同一局部地图中的正视图像中心与右斜视图像中心两点坐标差值最大的方向上的距离固定为 1,因此(6)式中的平移矢量应除以  $x_s$ (此处假设  $X$  方向距离最长,故  $X$  方向需要固定为 1),使得其在  $X$  方向的长度为 1,以统一尺度。

为避免奇异,将初始局部地图的坐标系设置为全局地图坐标系,令第一个正视图像的位姿为  $\mathbf{R}_1$  ( $\mathbf{R}_1$  为单位矩阵),  $\mathbf{t}_1 = \mathbf{0}$ 。根据对极约束,计算相邻局部地图的正视图像之间的相对位姿。按照曝光点的顺序叠加相邻正视图像之间的相对姿态,可获得每幅正视图像相对于全局坐标系的位姿初值为

$$\mathbf{R}_{5i+1} = \mathbf{R}_{5i-4}^{5i+1} \mathbf{R}_{5i-4}, \mathbf{t}_{5i+1} = \mathbf{R}_{5i-4}^{5i+1} \mathbf{t}_{5i-4} + \mathbf{t}_{5i-4}, \quad (i = 1, 2, \dots, n-1), \quad (7)$$

式中:  $\mathbf{R}_{5i-4}^{5i+1}$  和  $\mathbf{t}_{5i-4}^{5i+1}$  分别为第  $i+1$  个正视图像相对于第  $i$  个正视图像的旋转矩阵和平移矢量。

全局优化中,最小二乘优化所需的位姿初值的计算表达式为

$$\begin{cases} \mathbf{R}_{5i-3} = \mathbf{r}_{5i-3}^i \mathbf{R}_{5i-4} \\ \mathbf{t}_{5i-3} = \mathbf{R}_{5i-4} \mathbf{t}_{5i-3}^i + \mathbf{t}_{5i-4} \\ \mathbf{R}_{5i-2} = \mathbf{r}_{5i-2}^i \mathbf{R}_{5i-4} \\ \mathbf{t}_{5i-2} = \mathbf{R}_{5i-4} \mathbf{t}_{5i-2}^i + \mathbf{t}_{5i-4} \\ \mathbf{R}_{5i-1} = \mathbf{r}_{5i-1}^i \mathbf{R}_{5i-4} \\ \mathbf{t}_{5i-1} = \mathbf{R}_{5i-4} \mathbf{t}_{5i-1}^i + \mathbf{t}_{5i-4} \\ \mathbf{R}_{5i} = \mathbf{r}_{5i}^i \mathbf{R}_{5i-4} \\ \mathbf{t}_{5i} = \mathbf{R}_{5i-4} \mathbf{t}_{5i}^i + \mathbf{t}_{5i-4} \\ \mathbf{P}_j = \mathbf{R}_{5i-4}^{-1} \mathbf{K}^{-1} \mathbf{p}_j^{5i-4} + \mathbf{t}_{5i-4} \end{cases}, \quad (8)$$

式中:  $\mathbf{p}_j^{5i-4}$  为第  $j$  个特征点在图像  $(5i-4)$  处的二维像素坐标;  $(\mathbf{r}_{5i-3}^i, \mathbf{t}_{5i-3}^i), (\mathbf{r}_{5i-2}^i, \mathbf{t}_{5i-2}^i), (\mathbf{r}_{5i-1}^i, \mathbf{t}_{5i-1}^i)$  和  $(\mathbf{r}_{5i}^i, \mathbf{t}_{5i}^i)$  分别为第  $i$  个局部地图的 4 个斜视图像相对于正视图像的位姿,它们已在局部地图参数初始化中由对极约束求得。

使用高斯-牛顿法来估计全局地图特征  $\mathbf{X}_G$ 。基于文献[16]的工作,本文使用以下指标来终止迭代:增量  $\Delta \mathbf{X}_k$  的大小小于  $10^{-10}$ ;目标函数值的变化量小于  $10^{-8}$ ;最大迭代次数为 200。

### 3 实验结果与分析

为了验证所提 LGO 方法的性能,针对一个仿真合成数据集和一组国际上公测的数据集进行了测试,并在精度和收敛性方面与传统光束平差法进行了比较。

#### 3.1 仿真数据集实验

模拟一个五相机倾斜系统,该系统位于海拔 1000 m 的高度,相机焦距为 53 mm,正方形像元大小为  $6 \mu\text{m}$ ,图像分辨率为  $9000 \text{ pixel} \times 6732 \text{ pixel}$ ,倾斜角度为  $45^\circ$ 。模拟出的地形具有 50 m 的高低起伏,如图 5 所示。

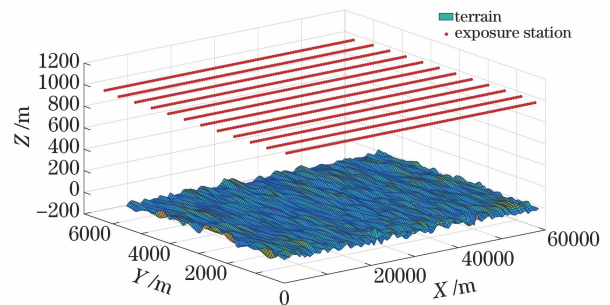


图 5 曝光站在所观测地形上空的空间分布  
Fig. 5 Spatial distribution of the exposure stations over the observed terrain

表 1 列出了仿真合成数据集的参数。60 km  $\times$  7 km 的摄影区域包含了 10 条航带,含有约  $5.4 \times$

$10^4$  个三维地面点, 约  $49 \times 10^4$  个二维像点。为了保证仿真数据与真实情况相近, 将 0.3 像素的高斯白噪声误差 ( $\sigma=0.3$ ) 添加到理论像点数据中。

表 1 大型仿真数据集的参数

Table 1 Parameters of the large scale simulated dataset

Parameter	Value
Area/(km×km)	$60 \times 7$
Number of track points	10
Number of images	5000
Number of local maps	1000
Number of 3D points	54337
Number of projection points	490086

### 3.1.1 精度比较

首先初始化全局地图的未知参数, 包括相机姿态和空间三维特征点。图 6 为用点表示的正视图像的轨迹。从图 6 可以看出, 随着大型数据集中参与计算的图像数目急剧增加, 定向轨迹的初始值的误

差累积现象也逐渐严重, 在轨迹末端的数据与真值存在较大偏离, 定向初值质量较差。

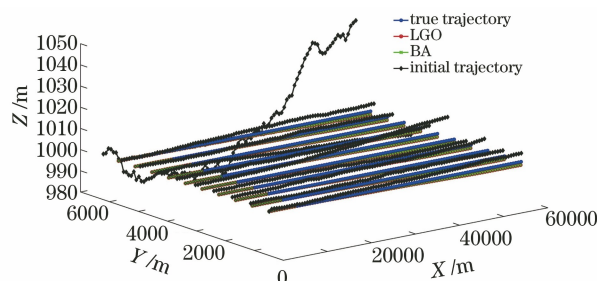


图 6 曝光站正视相机的轨迹

Fig. 6 Trajectories of the nadir cameras of the exposure stations

传统 BA 方法和 LGO 方法的收敛速度如图 7 所示, 在大型仿真数据集下, 标准 BA 方法无法收敛 (因此, 3.1 节中 BA 方法的估计结果都是将 LGO 方法估计值作为初值后所得), 而 LGO 方法经过 4 次迭代后完成收敛。

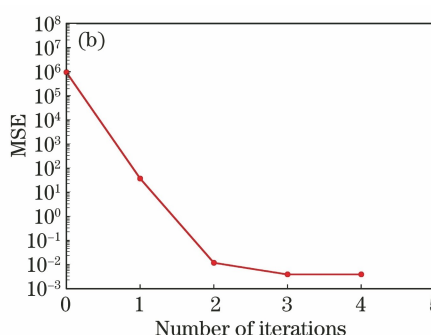
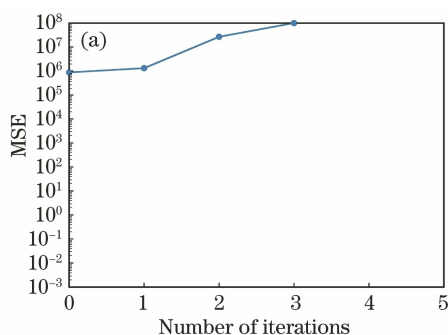


图 7 MSE 随迭代次数的变化。(a) BA 方法的 MSE; (b) LGO 方法的 MSE

Fig. 7 Change of MSE with the number of iterations. (a) MSE of the BA method; (b) MSE of the LGO method

从图 6 可以看到, 传统 BA 方法和 LGO 方法计算出的轨迹直观上重合。为了量化定向结果与真实轨迹间的误差, 绘制了估计轨迹相对于真实轨迹 (真

值) 的残差图, 如图 8 所示。在图 8 中, LGO 曲线位于 BA 曲线的下方, 因为两种方法的残差差异非常小, 所以 LGO 曲线几乎全被 BA 曲线遮挡, 显得不明显。

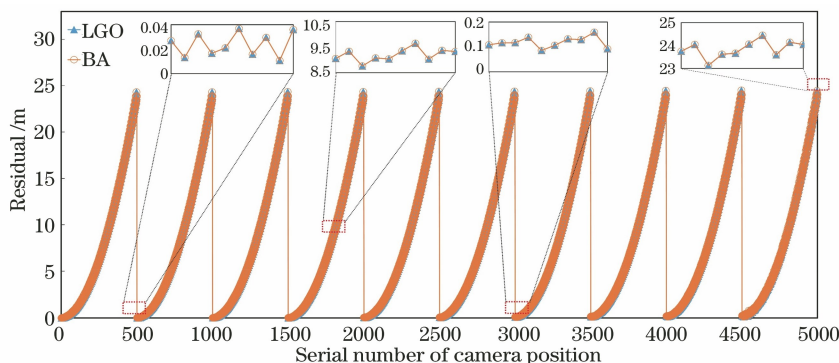


图 8 所有相机估计位置的残差

Fig. 8 Residual of all estimated camera positions

为了量化三轴定向参数的残差, 表 2 给出了由 BA 和 LGO 方法估计的所有相机位置参与计算得到

的均方根误差 (RMSE)。“3D”参数描述了 XYZ 三轴误差分量的总和, 两种方法之间的差异不大。

表 2 BA 和 LGO 估计的相机位姿的 RMSE

Table 2 RMSE of camera positions estimated by the BA and the LGO methods unit: m

Parameter	BA	LGO
X	9.6563	9.6516
Y	5.0619	5.0623
Z	0.8362	0.8568
3D	6.3131	6.3117

以上实验结果表明,尽管测试数据集有高达 5000 张影像、 $5 \times 10^4$  个三维点、 $49 \times 10^4$  个二维点,且测量误差在  $60 \text{ km} \times 7 \text{ km}$  的大型测量区域内累计,但所提方法在误差上相较于传统 BA 方法没有

表 3 大型仿真数据集存在初始化噪声的情况下 BA 和 LGO 方法的结果

Table 3 Results of BA and LGO methods on the large scale simulated dataset with initialization noise

Condition	BA method			LGO method		
	Initial MSE	Final MSE	Number of iterations	Initial MSE	Final MSE	Number of iterations
XYZ+5 m	$1.922 \times 10^3$	0.012	4	$2.352 \times 10^3$	0.004	3
XYZ+50 m	$1.998 \times 10^5$	0.012	5	$2.425 \times 10^5$	0.004	5
XYZ+100 m	$7.930 \times 10^5$	0.012	6	$1.288 \times 10^6$	0.004	6
XYZ+200 m	$3.594 \times 10^6$	Singular		$5.828 \times 10^6$	0.004	9
XYZ+300 m	$2.375 \times 10^{13}$	Singular		$3.188 \times 10^{13}$	Singular	
Ang+0.1 rad	$8.903 \times 10^5$	Singular		$9.805 \times 10^5$	0.004	5
Ang+0.2 rad	$3.823 \times 10^6$	Singular		$4.013 \times 10^6$	8.343	22
Ang+0.25 rad	$1.019 \times 10^7$	Singular		$6.496 \times 10^6$	Singular	
Ang+0.3 rad	$2.022 \times 10^7$	Singular		$9.538 \times 10^6$	Singular	
Ang+0.4 rad	$1.803 \times 10^{12}$	Singular		$1.677 \times 10^7$	Singular	

需要注意的是,对于 BA 方法来说,观测值为图像点,因此 BA 的 MSE 描述的是图像点的总重投影残差,由(1)式计算得出。而在 LGO 方法中,观测值为由局部优化得到的局部地图特征,因此 LGO 的 MSE 描述的是局部地图的观测值与局部地图的估计值之间的总残差,由(5)式计算得出。

如表 3 所示,在大型仿真合成数据集上,BA 方法在平移噪声  $\sigma_t$  为 5 m, 50 m 和 100 m 的初值下尚能完成收敛,但在存在任何旋转噪声的情况下都会收敛失败。与之相比,LGO 方法则出色地在平移噪声  $\sigma_t = 200 \text{ m}$ , 旋转噪声  $\sigma_r = 0.1 \text{ rad}$  时完成了收敛。不过,当初值质量进一步恶化时,LGO 也可能出现估计不准确甚至失败的风险。例如,当添加的旋转角度噪声达到  $0.2 \text{ rad}$  时,LGO 方法进行 22 次迭代后也只收敛到局部最小值 8.343,而不是全局最小值 0.004。

总的来说,在仿真数据下,所提 LGO 方法的收敛性能要强于传统 BA 方法的收敛性能。

损失,两种方法的残差值非常接近。

### 3.1.2 收敛性比较

比较两种方法对参数初值误差的容忍度,为此向正视相机的初始旋转角度和平移矢量添加不同级别的噪声,然后初始化所有斜视图像位姿和空间三维特征点坐标。

表 3 记录了引入噪声的大小以及两种方法的初始均方误差(MSE)、最终 MSE 和收敛完成所需的迭代次数。平移噪声由“XYZ+ $\sigma_t$ ”表示, $\sigma_t$  为引入的高斯噪声,“Ang+ $\sigma_r$ ”表示将大小为  $\sigma_r$  的高斯误差添加到旋转角度中。

### 3.2 真实数据集实验

在 ISPRS EuroSDR 项目发布的 Zurich 数据集<sup>[13]</sup>上测试所提方法与传统光束平差法。表 4 列出了 Zurich 数据集的详细参数。使用 Agisoft PhotoScan (<http://www.agisoft.com/>),从 135 张图像中提取了 225952 个像点(由 51672 个 3D 点投影所得)。

表 4 使用的 Zurich 数据集的全部参数

Table 4 Overall parameters of the used Zurich dataset

Parameter	Content
Camera type	Leica RCD30 Oblique Penta
Image size/(pixel×pixel)	9000×6732
Focal length/mm	53
Pixel size/mm	0.006
Platform height/m	1000
Title angles/(°)	35
Along-track overlap/%	70
Across-track overlap/%	50
Ground sample distance(GSD)/cm	6–12
Number of images	135
Number of 3D points	51672
Number of projection points	225952



图 9 给出了平面中 Zurich 数据的 135 张图像的分布,其中 27 个正视图像(圆形点标记)分布于三个轨道上。不同方向的箭头分别表示 4 个斜视图像。根据局部地图构建原则,选择来自 4 个不同曝光站上的斜视图像,它们和当前正视图像一起组成了一个局部地图。例如,第 1 个局部地图和第 26 个局部地图,它们都是由不同曝光站上的 5 个不同方

向的图像组成的,如图 9 所示,带箭头曲线指出局部地图的构成。

### 3.2.1 精度比较

图 10 为正视图像的初始轨迹和所有图像的初始轨迹。ISPRS EuroSDR 项目所发布的外部定向(EO)参数在本小节中视为真实姿态值,并用圆形实心点标记。

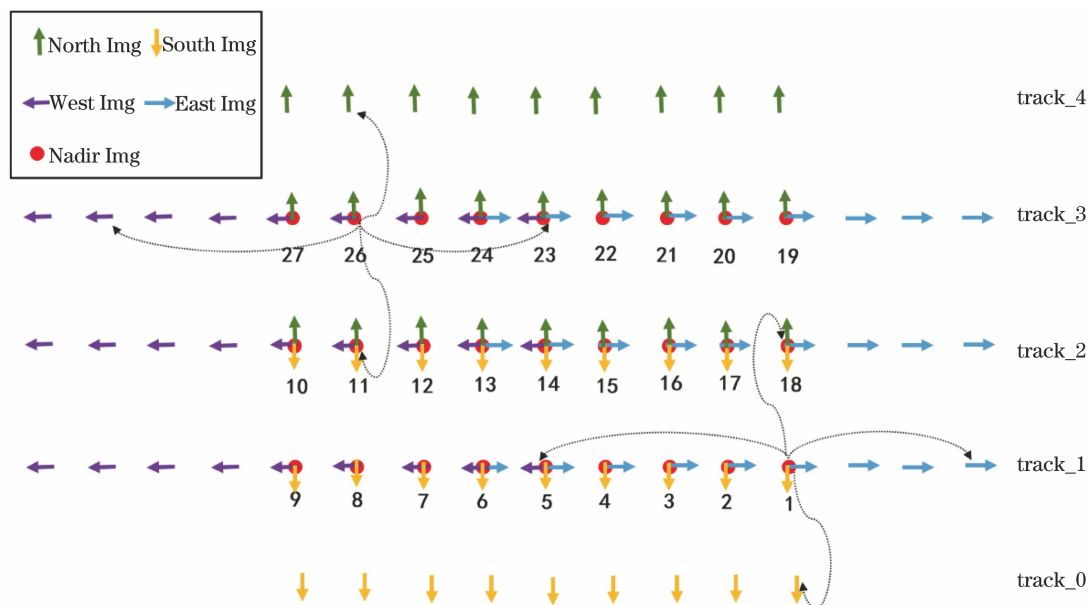


图 9 135 张 Zurich 图像的分布和局部地图的构成

Fig. 9 Distribution of the 135 Zurich images and the construction of local maps

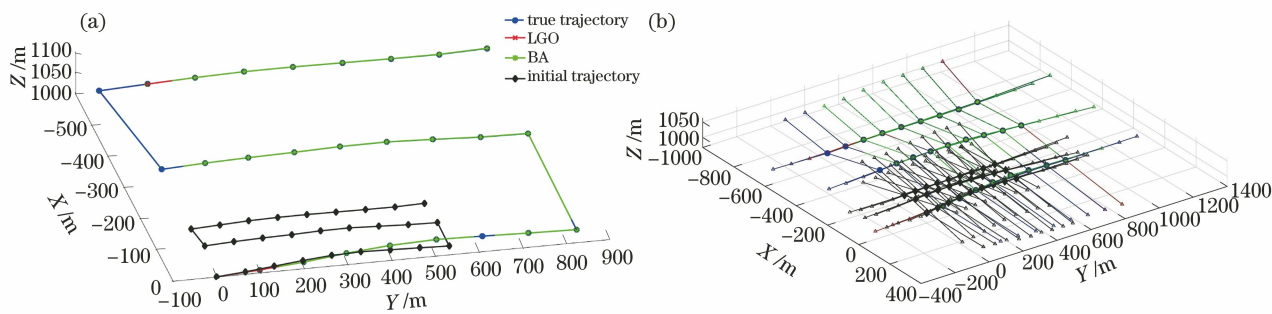


图 10 Zurich 数据的轨迹。(a)正视图像的轨迹;(b)正视和斜视图像的轨迹

Fig. 10 Trajectories for the Zurich data. (a) Trajectory of only nadir images; (b) Trajectory of both nadir and side images

表 5 中 BA 和 LGO 方法 MSE 的初始值可以通过(1)和(5)式的初始轨迹计算得出,分别为  $2.416 \times 10^6$  和  $1.838 \times 10^8$ 。BA 方法需要 8 次迭代才能收敛到最终 MSE(0.989),如图 11(a)所示。LGO 方法仅 4 次迭代便完成收敛,最终 MSE 为 0.282,如图 11(b)所示。表 5 中,BA 方法的最终 MSE 由(1)式计算出,描述了总重投影残差;LGO 方法的最终 MSE 由(5)式计算出,描述了局部地图的观测值与局部地图的估计值之间的总残差。

表 5 BA 方法和所提方法分别用于 Zurich 数据的参数

Table 5 Parameters of the BA method and the proposed method for Zurich data

Indicator	BA	LGO
Initial MSE	$2.416 \times 10^6$	$1.838 \times 10^8$
Final MSE	0.989	0.282
Number of iterations	8	4
RMSE along X-axis /m	0.1896	0.1817
RMSE along Y-axis /m	0.1704	0.1752
RMSE along Z-axis /m	0.1527	0.1535
RMSE in 3D space /m	0.1715	0.1705



从图 10 可以明显看出,平面图中,两种方法估计的轨迹都非常逼近真实值,误差在  $-0.5 \sim 0.5$  m 以内,且除了其中一个相机位置外,其他的相机位置的误差均小于 0.25 m。为了详细量化分析这两种方法的精度,给出分别在 X,Y,Z 方向上的估计位置与真实轨迹之间的残差,如图 12 所示。LGO 估计的第 42 个相机的位置精度比 BA 差,特别是 Z 方向残差。主要原因是:由于某些严

重的粗差像点存在,给第 8 张局部地图的局部优化引入了估计噪声;此外,该噪声也会被引入到全局优化中,从而当前相机位置的最终估计值产生轻微误差。但是,总体而言,这两种方法的图像定向精度是非常相似的。两种方法在 X 方向,Y 方向和 Z 方向上的 RMSE 差异也仅分别为 7.9 mm, 4.8 mm 和 0.8 mm;两种方法在三维空间上的位置 RMSE 也相似。

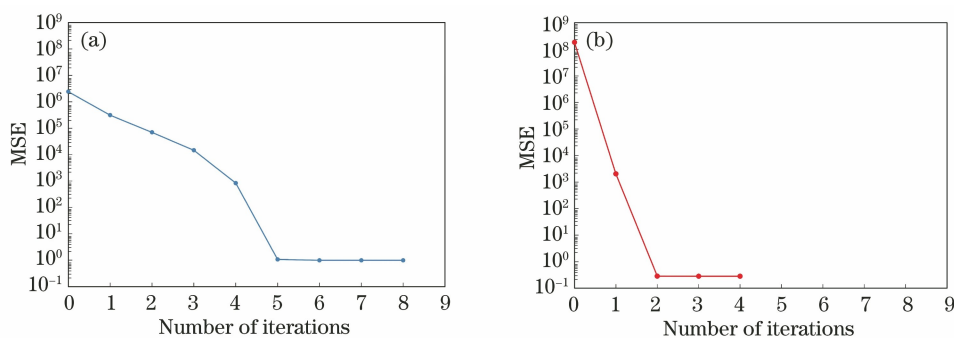


图 11 在 Zurich 数据集上,MSE 随迭代次数的变化。(a) BA 方法;(b) LGO 方法

Fig. 11 Change of MSE with the number of iterations on the Zurich dataset. (a) BA method; (b) LGO method

上面的分析结果表明,LGO 和 BA 方法在图像定向上具有相似的精度。

### 3.2.2 收敛性比较

为了测试 LGO 方法在初值存在噪声情况下的

收敛稳定性,将不同程度的噪声加在正视图像的旋转角度和平移向量上,得到含有不同程度噪声的初值。表 6 给出了两种方法的初始 MSE、最终 MSE 和迭代次数。

表 6 Zurich 数据集存在初始化噪声的情况下 BA 和 LGO 方法的结果

Table 6 Results of BA and LGO methods on the Zurich dataset with initialization noise

Condition	BA method			LGO method		
	Initial MSE	Final MSE	Number of iterations	Initial MSE	Final MSE	Number of iterations
XYZ+5 m	$4.337 \times 10^3$	0.989	5	$4.383 \times 10^3$	0.380	3
XYZ+50 m	$4.879 \times 10^5$	0.989	6	$4.818 \times 10^5$	0.380	3
XYZ+100 m	$1.352 \times 10^6$	0.989	8	$1.288 \times 10^6$	0.380	3
XYZ+200 m	$6.971 \times 10^6$	Singular		$5.498 \times 10^6$	0.380	3
XYZ+300 m	$3.687 \times 10^{12}$	Singular		$1.677 \times 10^7$	0.380	3
Ang+0.1 rad	$4.119 \times 10^5$	0.989	6	$4.003 \times 10^5$	0.380	5
Ang+0.2 rad	$1.897 \times 10^6$	0.989	7	$1.675 \times 10^6$	0.380	6
Ang+0.25 rad	$3.260 \times 10^6$	0.989	7	$2.593 \times 10^6$	0.380	7
Ang+0.3 rad	$4.419 \times 10^6$	0.989	8	$3.349 \times 10^6$	0.380	7
Ang+0.4 rad	$8.364 \times 10^8$	Singular		$6.680 \times 10^6$	0.380	8

如表 6 所示,当初始位置与真实轨迹之间的噪声大于  $\sigma=100$  m 和角度噪声  $\sigma=0.4$  rad 时,传统 BA 方法均收敛失败,而 LGO 方法在实验的所有情况下均稳定快速收敛成功。根据实验结果可以得出结论:在真实实验中,所提 LGO 方法在收敛速度和收敛性能上总体优于传统 BA 方法。

## 4 结 论

光束平差法是目前较先进的摄影测量图像定向方法,但在初始值较差的情况下,会收敛失败,尤其在倾斜图像应用和大型图像数据集应用中。通过引入辅助的传感器信息[例如来自惯性传感器(IMU)或全球导航卫星系统(GNSS)的信息],虽然可以在

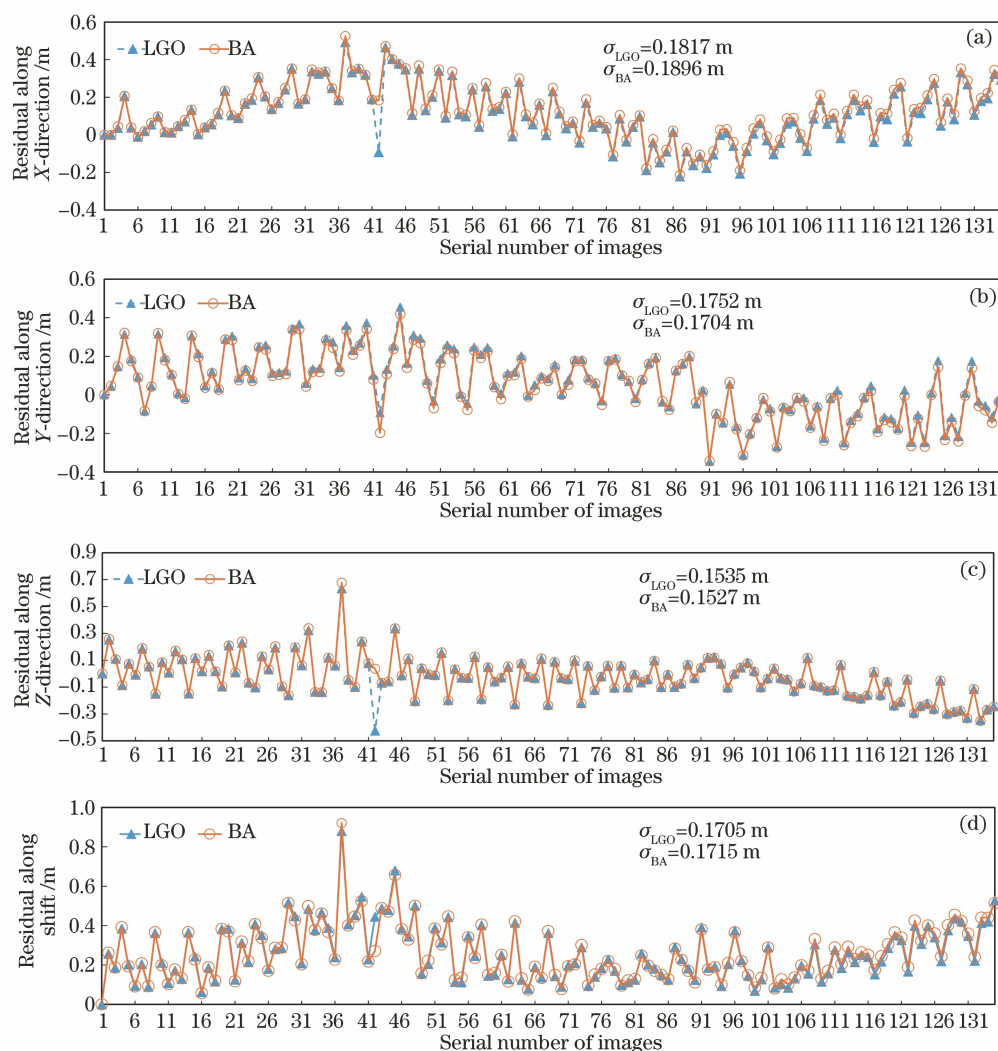


图 12 BA 和 LGO 方法分别估计的 135 张 Zurich 图像的残差。(a) X 方向;(b) Y 方向;(c) Z 方向;(d) 3D

Fig. 12 Residual of 135 Zurich images estimated by the BA and LGO methods. (a) X direction; (b) Y direction; (c) Z direction; (d) 3D

一定程度上克服初值问题,但总的来说仍然在许多情况下是不可行的,因此必须找到仅利用图像本身信息的解决方法。本文展示了一种提高非线性优化鲁棒性的方法,使得在 GNSS/IMU 信息缺失或 GNSS/IMU 提供导航定位信息不准确的情况下,也能对倾斜图像进行成功定向。

所提方法采用局部优化到全局优化的策略来实现图像定向,利用了局部小区域求解精度更高,鲁棒性更好的优点,将整个大的全局任务分成了多个局部任务来解决,并借助局部优化中所得信息矩阵在整个全局地图上对误差进行权值调整,避免了局部误差累积,这样既保证了精度,又使得小规模区域鲁棒性好的优点得以发挥。

将所提方法应用于一组仿真数据集和一组真实数据集进行实验与分析,可以看到,所提方法可以更

快、更稳定地收敛,说明所提方法对初值的质量要求更低,依赖程度更弱,并且完成优化任务所需要的迭代次数也相对更少。重要的是,所提方法在估计结果的精度上也并没有任何损失。所提方法相对于传统 BA 方法收敛鲁棒性强的主要原因是,由局部优化产生的局部地图的几何形状非常接近真值,因此大大减小了全局优化过程中的解搜索空间。所提方法之所以能够保证与传统单次 BA 方法同样的精度,是因为尽管前后进行了两次优化,但第一级局部优化得到的信息矩阵也按相应顺序引入到了第二级全局优化中,从而避免了其他多级优化方法中相关信息丢失的问题。

#### 参 考 文 献

- [1] Li Z Y, Lin J R, Sun Y B, et al. A method for

- measuring circular pose based on line structured light sensor [J]. *Acta Optica Sinica*, 2020, 40 (15): 1512004.
- 李志宇, 林嘉睿, 孙岩标, 等. 一种基于线结构光传感器的圆位姿测量方法 [J]. *光学学报*, 2020, 40 (15): 1512004.
- [2] Zhu J G, Li Y J, Ye S H, et al. Study on single camera simulating stereo vision measurement technology [J]. *Acta Optica Sinica*, 2005, 25 (7): 943-948.
- 郝继贵, 李艳军, 叶声华, 等. 单摄像机虚拟立体视觉测量技术研究 [J]. *光学学报*, 2005, 25(7): 943-948.
- [3] Xiao X W. Oblique photogrammetry based scene 3D reconstruction with structure sensing functions [J]. *Acta Geodaetica et Cartographica Sinica*, 2019, 48 (6): 802.
- 肖雄武. 具备结构感知功能的倾斜摄影测量场景三维重建 [J]. *测绘学报*, 2019, 48(6): 802.
- [4] Chen Y, Lin J R, Gao Y, et al. A hybrid of vision and inclination sensor method for estimating relative pose [J]. *Acta Optica Sinica*, 2015, 35 (12): 1212003.
- 陈洋, 林嘉睿, 高扬, 等. 视觉与倾角传感器组合相对位姿测量方法 [J]. *光学学报*, 2015, 35(12): 1212003.
- [5] Yuan X X, Chen S Y, Zhong C. Oblique aerial image relative orientation based on fundamental matrix [J]. *Geomatics and Information Science of Wuhan University*, 2016, 41(8): 995-1000.
- 袁修孝, 陈时雨, 钟灿. 基于基础矩阵的倾斜航摄影像相对定向方法 [J]. *武汉大学学报·信息科学版*, 2016, 41(8): 995-1000.
- [6] Gao Y L. 3D reconstruction of oblique photogrammetry enhanced by mobile lidar mapping systems [D]. Wuhan: Wuhan University, 2017: 1-17.
- 高云龙. 基于车载点云增强的倾斜摄影建筑物三维模型重建 [D]. 武汉: 武汉大学, 2017: 1-17.
- [7] Svennevig K, Guarnieri P, Stemmerik L. From oblique photogrammetry to a 3D model: structural modeling of Kilen, eastern North Greenland [J]. *Computers & Geosciences*, 2015, 83: 120-126.
- [8] Diefenbach A K, Bull K F, Wessels R L, et al. Photogrammetric monitoring of lava dome growth during the 2009 eruption of Redoubt Volcano [J]. *Journal of Volcanology and Geothermal Research*, 2013, 259: 308-316.
- [9] Remondino F, Spera M G, Nocerino E, et al. State of the art in high density image matching [J]. *The Photogrammetric Record*, 2014, 29(146): 144-166.
- [10] Hu H, Ding Y L, Zhu Q, et al. Stable least-squares matching for oblique images using bound constrained optimization and a robust loss function [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2016, 118: 53-67.
- [11] Long C Y, Zhu J G, Guo Y, et al. Study on close-range photogrammetry based on nonparameteric measurement model [J]. *Acta Optica Sinica*, 2014, 34(12): 1215004.
- 隆昌宇, 郝继贵, 郭寅, 等. 基于非参数测量模型的摄影测量方法研究 [J]. *光学学报*, 2014, 34(12): 1215004.
- [12] Triggs B, McLauchlan P F, Hartley R I, et al. Bundle adjustment: a modern synthesis [M] // Triggs B, Zisserman A, Szeliski R. *Vision algorithms: theory and practice*. Heidelberg: Springer, 2000, 1883: 298-372.
- [13] Cavegn S, Haala N, Nebiker S, et al. Benchmarking high density image matching for oblique airborne imagery [J]. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2014, XL-3: 45-52.
- [14] Dan Foresee F, Hagan M T. Gauss-Newton approximation to Bayesian learning [C] // *Proceedings of International Conference on Neural Networks (ICNN'97)*, June 12, 1997, Houston, TX, USA. New York: IEEE Press, 1997: 1930-1935.
- [15] Zhao S H, Li Y Y, Cao J, et al. A BFGS-corrected Gauss-Newton solver for bundle adjustment [J]. *Acta Scientiarum Naturalium Universitatis Pekinensis*, 2020, 56(6): 1013-1019.
- 赵帅华, 李言言, 曹健, 等. 基于 BFGS 修正的高斯牛顿光束法平差解算方法 [J]. *北京大学学报(自然科学版)*, 2020, 56(6): 1013-1019.
- [16] Zhao L, Huang S D, Sun Y B, et al. ParallaxBA: bundle adjustment using parallax angle feature parametrization [J]. *The International Journal of Robotics Research*, 2015, 34(4/5): 493-516.