

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2024)10-2992-16

论文引用格式: Li M H, Gao X, Xie Z X and Shen S H. 2024. Incremental scale estimation-based camera location recovery. Journal of Image and Graphics, 29(10):2992-3007(李梦晗, 高翔, 解则晓, 申抒含. 2024. 增量式尺度估计下的相机位置解算. 中国图象图形学报, 29(10):2992-3007)[DOI: 10. 11834/jig. 230745]

增量式尺度估计下的相机位置解算

李梦晗¹, 高翔^{2*}, 解则晓¹, 申抒含²

1. 中国海洋大学工程学院, 青岛 266100; 2. 中国科学院自动化研究所, 北京 100190

摘要: **目的** 全局式从运动恢复结构(structure from motion, SfM)通过运动平均一次性恢复所有相机的绝对位姿, 效率相对较高。运动平均中的平移平均主要负责解算相机在世界坐标系下的绝对位置, 其求解过程因尺度歧义性、估计敏感性和求解不确定性的影响而较为困难。本文提出了一种基于增量尺度估计的平移平均方法, 在消除尺度歧义性的同时提升了求解鲁棒性与准确性。**方法** 本文将平移平均问题解耦为3个子问题: 1) 局部绝对尺度的增量式估计; 2) 全局绝对尺度的增量式估计; 3) 基于 L_1 优化的尺度已知的绝对位置估计。**结果** 在1DSfM数据集上进行对比实验, 基线解算精度明显提升, 解算相机百分比的均值达到96%。当引入两种不同的绝对旋转进行计算时, 其绝对位置中值误差仅略差于BATA(bilinear angle-based translation averaging)与CReTA(correspondence reweighted translation averaging), 排名第3, 均值误差改善更为明显, 分别排名第1和第2。相较于原始方法, 本文方法在相机解算数量与位置解算精度上均有较大提升。**结论** 本文方法综合了尺度分离思想与增量式参数估计思想, 既消除了尺度歧义性, 又保证了鲁棒性与高效性, 求解所得的相机绝对位置稳定可靠。

关键词: 全局式从运动恢复结构; 平移平均; 尺度分离; 基线长度求解; 增量式参数估计

Incremental scale estimation-based camera location recovery

Li Menghan¹, Gao Xiang^{2*}, Xie Zexiao¹, Shen Shuhan²

1. College of Engineering, Ocean University of China, Qingdao 266100, China;

2. Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

Abstract: Objective The structure from motion (SfM) technique serves as the fundamental step in the sparse reconstruction process, finding extensive applications in remote sensing mapping, indoor modeling, augmented reality, and ancient architecture preservation. SfM technology retrieves camera poses from images, encompassing two main categories: incremental and global approaches. The global SfM, in contrast to the iterative nature of incremental SfM, simultaneously estimates the absolute poses of all cameras through motion averaging, resulting in relatively high efficiency. However, it still encounters challenges regarding robustness and accuracy. Rotation averaging and translation averaging constitute crucial components within the motion averaging. Compared with rotation averaging, translation averaging is more difficult due to the following three reasons: 1) Only relative translation directions could be recovered by essential matrix estimation and decomposition, i. e., the produced relative translations are scale ambiguous. 2) Only cameras in the same parallel rigid component could their absolute locations be uniquely determined by translation averaging, while for rotation averaging, the requirement simply degenerates to the connected component. 3) Compared with relative rotation, the estimation accuracy

收稿日期: 2023-10-17; 修回日期: 2024-01-04; 预印本日期: 2024-01-11

* 通信作者: 高翔 xiang.gao@ia.ac.cn

基金项目: 国家自然科学基金项目(62373349, U22B2055, 42076192)

Supported by: National Natural Science Foundation of China (62373349, U22B2055, 42076192)

of relative translation is more vulnerable to the feature point mismatches and more likely to be outlier contaminated. In traditional approaches, the translation averaging method based on scale separation (L_1 SE- L_1 TA) calculates the relative baseline length between cameras before estimating the absolute locations, eliminates the scale ambiguity, and the solving range is no longer constrained by the camera triplet, but its robustness and accuracy still need to be improved. Incremental translation averaging (ITA) introduces the idea of incremental parameter estimation into the translation averaging process for the first time, which has good robustness and high accuracy. However, its solving process depends on camera triplets and may suffer from degeneracy during collinear camera motion. To solve the above problems, this study proposes a translation averaging method based on incremental scale estimation (ISE- L_1 TA), which eliminates the scale ambiguity and enhances the method's robustness and result accuracy. **Method** Incremental SfM has been proven to be highly accurate and robust, making it a preferred choice for many applications. It has shown to be particularly effective in handling large datasets and overcoming the challenges posed by complex real-world scenarios. Recognizing its potential, researchers have sought to transfer the incremental parameter estimation ideology to other related tasks, such as incremental rotation averaging (IRA) and ITA. In particular, IRA is designed to estimate the camera absolute rotations incrementally and efficiently. Meanwhile, ITA is performed for the camera absolute locations, enabling it to handle outliers effectively and avoid error propagation. Overall, the adoption of incremental parameter estimation ideology for motion averaging tasks demonstrates the versatility and effectiveness of this approach. With its ability to handle complex datasets and overcome a range of challenges, the incremental parameter estimation ideology holds great promise for future research in the field of 3D reconstruction and beyond. In this study, ISE- L_1 TA is proposed by incorporating the scale separation strategy and incremental parameter estimation ideology. Specifically, the translation averaging problem is decomposed into three sub-ones and sequentially solved: 1) incremental estimation of local absolute scale, 2) incremental estimation of global absolute scale, 3) scale-aware absolute location estimation based on L_1 optimization. The input of our proposed method is the pairwise scale invariant feature transform point matches, and its output is the absolute camera locations. First, the relative motion between cameras is obtained by estimating and decomposing the essential matrix. Next, the two-view triangulation is performed to calculate the relative depths in the local coordinate system. On the basis of depth ratios, incremental estimations are conducted for the local and global absolute scales. Subsequently, the relative baseline length between cameras is computed, and rotation averaging is performed for absolute rotation estimation, enabling the final scale-aware absolute location estimation. **Result** We performed experimental tests to evaluate the selection of scale distance function and scale distance threshold. The experimental results confirmed that the normalized perfect square deviation function effectively eliminates the impact of scaling effects. Furthermore, the incremental scale estimation method shows good robustness and insensitivity to scale distance threshold and achieves remarkably higher baseline accuracy compared with L_1 SE. The experiments were conducted on the IDSfM dataset. In comparison with various state-of-the-art methods including bilinear angle-based translation averaging (BATA), correspondence reweighted translation averaging (CReTA), ITA, and L_1 SE- L_1 TA, our proposed method exhibited the following performance: 1) In terms of the number of cameras solved, the average percentage of successfully solved cameras using the proposed method is 96%. 2) The median error of absolute location estimation is slightly worse than that of BATA and CReTA and ranks third overall under different absolute rotations. 3) In terms of the mean error in absolute location estimation, the proposed method has remarkable advantages, ranking first and second respectively. Compared with the original L_1 SE- L_1 TA, the method in this study has a great improvement in the number of cameras solved and the accuracy of locations estimated. **Conclusion** The proposed method combines the concept of scale separation with incremental parameter estimation. By integrating these two ideas, our method effectively eliminates scale ambiguity while ensuring the effectiveness of outlier rejection and maintaining a concise solving process. As a result, the obtained absolute camera locations are stable and reliable.

Key words: global structure from motion; translation averaging; scale separation; baseline length computation; incremental parameter estimation

0 引言

三维重建技术作为计算机视觉领域的研究热点,在生物医学(姚砺等,2021;任非儿等,2021)、水下探测(王柯俨等,2022)和室内外场景建模(颜深等,2021;宁小娟等,2023)等方面应用广泛。传统的基于图像的大规模场景三维重建主要包括稀疏重建(Snavely等,2008)、稠密重建(Furukawa和Ponce,2010)、表面重建(Jancosek和Pajdla,2011)和纹理映射(Waechter等,2014)4个部分。从运动恢复结构(structure from motion, SfM)技术是稀疏重建中的核心,能够通过图像信息恢复相机位姿,主要包括增量式(Snavely等,2008)和全局式(Chatterjee和Govindu,2013)两大类。与增量式SfM迭代解算的方式不同,全局式SfM通过运动平均一次性恢复所有相机的绝对位姿,效率相对较高。

旋转平均(Gao等,2020;Purkait等,2020;Lee和Civera,2022;Li等,2022;Shi等,2022;Zhang等,2023)和平移平均(Wilson和Snavely,2014;Özyeşil和Singer,2015;Sweeney等,2015;Goldstein等,2016;Zhuang等,2018;Manam和Govindu,2022)是运动平均流程的两个关键步骤,旨在通过相机间的相对运动(Govindu,2006;Bourmaud,2016;Zhu等,2018)恢复各个相机的绝对位姿(Govindu,2001,2004;Govindu和Pooja,2014)。与旋转平均相比,平移平均更具挑战性,原因如下:1)估计并分解本质矩阵所得的相对平移(t_j)仅包含方向信息而不包含尺度信息,导致平移平均过程存在尺度歧义性(Nister,2004);2)在估计精度方面,相对平移比相对旋转更易受到特征误匹配的影响,对异常值更加敏感(Nister,2004);3)就平移平均而言,只有处于同一平行刚性分量中的相机才能被唯一确定绝对位置,而对于旋转平均,该项约束迅速退化为处于同一连通分量中(Özyeşi和Singer,2015)。值得注意的是,平移平均问题的求解约束本质上源于尺度歧义性的存在。如果事先已知相机间的基线长度,那么平移平均的可解性就等价于旋转平均的可解性。因此,在消除尺度歧义性影响的同时提升方法的鲁棒性对增强平移平均问题的可解性和准确性具有重要意义。现有的平移平均方法主要基于鲁棒优化和输入增强两个方面,在完成上述目标时仍面临许多挑战。

基于鲁棒优化的平移平均方法通过设计鲁棒的损失函数和相应的优化方案使其对相对平移异常值更加鲁棒。Govindu(2001)提出了一种目标函数最小化代数误差而非几何误差的平移平均方法,该方法最小化相对平移测量值与通过绝对位置反算的相对平移估计值之间的向量积,并通过迭代重加权的方式减小基线长度差异对估计结果的影响。Moulon等人(2013)构建了openMVG系统,将平移平均问题转化为 L_2 意义下的优化问题,该方法同步优化相机绝对位置、相对基线长度和所有约束的全局一致残差上界。Özyeşil和Singer(2015)的LUD(least unsquared deviations)算法和Goldstein等人(2016)的ShapeFit/Kick算法均设计了基于端点距离的损失函数,其不同之处在于,LUD算法通过迭代重加权最小二乘法(iteratively reweighted least squares, IRLS)最小化松弛端点距离,而ShapeFit/Kick算法使用交替向量乘法(alternating direction method of multipliers, ADMM)最小化相对平移估计值到相对平移测量值正交补空间的投影模长。同样借助IRLS进行求解的还有Zhuang等人(2018)提出的BATA(bilinear angle-based translation averaging)算法,该算法提倡基于角度目标进行平移平均,通过简单有效的双线性目标函数实现了精度较高的位置解算。上述方法在面对异常值时具有一定的鲁棒性,但其损失函数的设计较为复杂,优化方案相对繁琐。此外,在预处理(Moulon等,2013;Özyeşi和Singer,2015;Goldstein等,2016)或初始化操作(Zhuang等,2018)缺失的情况下,可能会导致估计结果精度较低甚至完全错误。

基于输入增强的平移平均方法通过滤除相对平移异常值,改善相对平移测量值或优化外极几何图(epipolar-geometry graph, EG)使其近似一致来减轻后续优化操作的负担。Wilson和Snavely(2014)提出的1DSfM算法将相对平移投影至若干一维方向上,根据异常值无法在所有方向上始终与其他投影保持一致的原理进行滤除,通过最小反馈弧集算法(minimum feedback arc set, MFAS)完成解算。Manam和Govindu(2022)提出的CReta(correspondence reweighted translation averaging)算法将特征对应权重更新、相对平移方向优化和相机绝对位置解算等步骤迭代进行,有效提升了平移平均的结果精度。Sweeney等人(2015)基于相机三元组进行图扩

增,根据循环一致性约束完成图优化,进而得到精确一致的EG子图,借助图优化策略解算相机的绝对位置。Kasten等人(2019)提出了一种基于矩阵秩约束的平移平均算法,通过迭代优化生成逐对一致的本质矩阵,同步解算相机的绝对旋转(R_i)和绝对位置(c_i),其后续改进算法能够适用于相机直线运动的情形。上述方法虽然已经取得了一定的研究进展,但其预先滤除和迭代优化等操作使得求解过程较为复杂。同时,上述方法的解算精度通常依赖于相机绝对旋转的解算精度,在鲁棒性、简洁性和高效性等方面仍有待提升。

为应对平移平均过程中的尺度歧义性问题,Cui和Tan(2015)以两视图重建时的特征点深度与实际基线长度成反比为出发点,设计了一种全新的基于尺度分离的平移平均流程 L_1 SE- L_1 TA。这一方法将复杂的平移平均问题解耦为3个简单的子问题,通过 L_1 优化(Candès等,2008)依次解算相机的局部绝对尺度、全局绝对尺度和全局绝对位置。根据局部绝对尺度与全局绝对尺度计算得到的基线长度消除了尺度歧义性,后续的绝对位置估计也因此得以简化。然而,尽管 L_1 优化对异常值的鲁棒性已得到证明,但若在求解过程中能够引入简洁有效的异常值滤除策略,其性能将获得进一步提升。受到增量式SfM(Snavely等,2008;Moulon等,2012;Wu,2013;Shah等,2014;Schönberger和Frahm,2016;Cui等,2017)流程的启发,Gao等人(2021,2022a,2022b,2023a,2024)将增量式参数估计思想引入到运动平均问题中,既保证了解算过程的鲁棒性,又提升了解算结果的准确性。其中,ITA(incremental translation averaging)算法(Gao等,2022b)首次通过增量式参数估计流程解决平移平均问题,在估计相机位置的同时滤除相对平移异常值。该算法无需对输入信息进行增强,优化过程简洁高效,但因尺度歧义性的存在,需要借助线性三焦张量求解器(Jiang等,2013)完成求解,因而受到相机三元组的限制,在相机共线运动时性能退化。以上述分析为基础,为了实现准确度高、鲁棒性好、简洁高效且不受相机运动形式约束的平移平均方法,本文借鉴上述两种方法的思想,将尺度分离与增量式参数估计同时引入到平移平均流程中,为相机局部绝对尺度与全局绝对尺度设计增量求解器,在估计相机尺度的同时滤除相对平移异常值,保证相机间基线长度的准确性,从而为后续

基于 L_1 优化的尺度已知的绝对位置估计提供相对精确的输入。

本文主要贡献如下:1)将增量式参数估计思想与尺度分离策略同时引入到平移平均过程中,通过增量式尺度估计方法解算相机的局部与全局绝对尺度,能够有效滤除尺度异常值,提升尺度估计方法的鲁棒性与平移平均结果的准确性;2)提出了一种用于计算尺度误差的尺度距离函数,即归一化完全平方差函数,可以有效消除尺度缩放效应对尺度误差计算所产生的不利影响;3)设计了一种基于中值标量对齐的基线精度评测方法,通过相对基线长度的解算精度反映各尺度解算方法的解算精度,从而实现对各尺度解算方法的有效性评估。

1 研究方法

本文方法以成对的SIFT(scale-invariant feature transform)特征匹配点(Lowe,2004;Muja和Lowe,2014;Schönberger和Frahm,2016)作为输入,输出相机在世界坐标系下的绝对位置,具体流程如图1所示。首先通过本质矩阵的估计与分解(Nister,2004;Zhao,2022)获取相机间的相对运动,随后进行两视图重建(Hartley和Sturm,1997;Lee和Civera,2019)并计算局部坐标系下的相对深度比,根据深度比依次对相机的局部绝对尺度与全局绝对尺度进行增量式估计,最后利用 L_1 优化解算相机的绝对位置。在进行位置解算前,本文通过IRA(incremental rotation averaging)算法(Gao等,2021)恢复相机的绝对旋转,并以此为前提将相对平移的方向对齐到全局坐标系下。在进行对比实验时,本文还根据Chatterje和Govindu(2013)提出的 L_1 GM算法恢复相机的绝对旋转,以确保实验结果的全面性与完整性。

1.1小节描述了尺度分离策略(Cui和Tan,2015)在平移平均问题上的应用与优势,对其可解范围进行了全面的分析,为本文方法的提出提供了理论依据;1.2小节详细介绍了基于增量式尺度估计的平移平均算法的具体流程,包括增量局部/全局尺度解算和尺度已知的绝对位置解算两部分。

1.1 尺度分离策略及相关分析

假设相机 j 和相机 k 与相机 i 之间存在共视特征

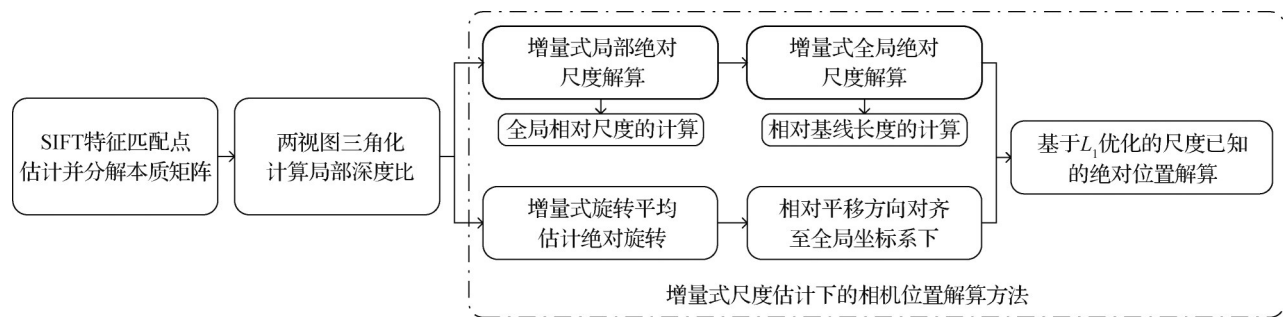


图1 增量式尺度估计下的相机位置解算方法流程

Fig. 1 Pipeline of camera location recovery based on incremental scale estimation

点,分别对相机对 (i, j) 和 (i, k) 对应的图像进行两视图重建,可以得到这些共视特征点在相机 i 坐标系下的局部深度 $\{d_m^j, d_m^k | m = 1, 2, \dots, n_{jk}^i\}$,其中, n_{jk}^i 表示共视匹配点的序号。通常情况下,由本质矩阵分解得到的相对平移存在尺度歧义问题,其模长满足 $\|\mathbf{t}_{ij}\|_2 = \|\mathbf{t}_{ik}\|_2 = 1$,使得 d_m^j 与 d_m^k 并不相等。因此,对于相机三元组 (i, j, k) 而言,可以通过 $d_{jk}^i = \text{med}(\{d_m^j/d_m^k\})$ 计算其局部相对深度比,其中, $\text{med}(\{x\})$ 表示序列 $\{x\}$ 的中值。根据两视图重建时局部尺度与局部深度成反比的关系,可得

$$\frac{s_{ik}^i}{s_{ij}^i} = d_{jk}^i \quad (1)$$

式中, s_{ij}^i 和 s_{ik}^i 分别表示相机 j 和相机 k 相对于相机 i 的局部绝对尺度。式(1)提出了基于尺度分离的平移平均流程的第1个子问题,即在局部相对深度给定的情况下解算局部绝对尺度。

在每个相机上分别进行局部尺度解算后,可以计算得到全局相对尺度 s_{ij} ,计算式为

$$s_{ij} = \frac{s_{ji}^j}{s_{ij}^i} \quad (2)$$

相机 i 与相机 j 的全局绝对尺度 s_i 和 s_j 与全局相对尺度 s_{ij} 满足如下关系,即

$$\frac{s_i}{s_j} = s_{ij} \quad (3)$$

由此可知,式(3)提出了基于尺度分离的平移平均流程的第2个子问题,即在全局相对尺度已知的情况下解算全局绝对尺度。

上述两个子问题完成后,利用局部绝对尺度与全局绝对尺度可以计算相机间的相对基线长度 b_{ij} ,计算式为

$$b_{ij} = \frac{1}{2}(s_i s_{ij}^i + s_j s_{ji}^j) \quad (4)$$

值得注意的是,在进行位置解算前,需要将所有相对平移的方向对齐到同一个全局坐标系下,即 $\mathbf{c}_{ij} = -\mathbf{R}_j^T \mathbf{t}_{ij}$ 。将基线长度 b_{ij} 引入到位置解算过程中,尺度已知的相对位置 $b_{ij} \mathbf{c}_{ij}$ 与绝对位置 \mathbf{c}_i 和 \mathbf{c}_j 满足如下约束,即

$$\mathbf{c}_j - \mathbf{c}_i = b_{ij} \mathbf{c}_{ij} = -b_{ij} \mathbf{R}_j^T \mathbf{t}_{ij} \quad (5)$$

至此,式(5)提出了基于尺度分离的平移平均流程的第3个子问题,即尺度已知的绝对位置解算。除上述内容外,更多关于尺度分离策略的具体细节可以参考Cui和Tan(2015)的研究。

根据式(1)和式(3)不难看出,这两个子问题在本质上是相同的,都是以两两标量间的比值关系为约束,计算一组新的标量值。Cui和Tan(2015)通过在等式两侧取对数的方式将原本的乘除关系转化为加减关系,进而将3个子问题统一构造为基于 L_1 优化的线性超定方程组求解问题。本文方法以此为基础,针对前两个子问题的共性设计增量式尺度求解器,进一步提升了算法的鲁棒性。为了更加深刻地理解尺度分离策略在可求解性上的表现,本文对其可求解范围进行了详细的分析。

传统的平移平均方法主要基于端到端的求解方式,缺乏对基线长度 b_{ij} 的先验知识,存在严重的尺度歧义问题,只有处于同一平行刚性分量中的相机才能被唯一确定其绝对位置(Özyeşil和Singer, 2015)。然而,对基于尺度分离的平移平均方法而言,由于位置解算时的基线长度已确定,可以通过相对简单的链式操作估计下一相机的绝对位置,因而其可求解范围被扩增至同一连通分量中,这证明了基于尺度分离的平移平均方法在可解性上优于传统方法。

需要进一步注意的是,式(5)中所包含的两个参

数 b_{ij} 和 c_{ij} 的存在条件具有一定的差异。当相机 i 与相机 j 对应图像之间的特征匹配点数目足够用于本质矩阵估计时, c_{ij} 通常存在。然而, b_{ij} 的存在条件却相对复杂。由式(2)一式(4)可以看出, 只有当局部绝对尺度 s_{ij}^i 和 s_{ij}^j 同时存在时, b_{ij} 才能被求解。进一步分析发现, 局部绝对尺度 s_{ij}^i 是否存在又受到式(1)的影响, 只有当相机 i 和相机 j 至少与一个其他相机 k 之间存在共视特征点且能求解局部深度比 d_{jk}^i 时, s_{ij}^i 才能被解算。鉴于上述分析, 本文在增量估计局部绝对尺度时, 为每个相机构建共视点连通图, 仅对处于最大连通分量(maximum connected component, MCC)中的相机进行解算。这一处理方式既能够为后续解算流程提供尽可能多的输入信息, 又能有效保证整个解算流程的简洁和高效。

1.2 基于增量式尺度估计的平移平均算法

1.2.1 增量式局部/全局绝对尺度估计

局部绝对尺度估计(local scale estimation, LSE)根据式(1)解算相机在局部坐标系下的绝对尺度, 而全局绝对尺度估计(global scale estimation, GSE)通过式(3)恢复相机在全局坐标系下的绝对尺度。尽管上述两个任务存在一定的差异, 但如1.1节所述, 其本质都是根据比值约束关系确定一组新的标量值。因此, 本文就LSE与GSE问题的共性展开分析, 提出了一种增量式尺度求解算法ISE, 并将其设计为通用的增量式尺度求解器。

式(1)和式(3)的比值关系均可以表示为 $c = ab$ (式中 a, b 和 c 均代表标量变量)的形式, 其不同之处主要体现在比值约束下的参数序数不一致。在式(1)中, 相对关系序数在后者作为绝对关系比例式中的分子, 而在式(3)中则恰好相反。为了更加直观地展现这两个子问题的共性, 方便后续对算法流程的表达, 本文用 x 代替 s 表示尺度标量, 并通过上标 P 对子问题1与子问题2进行区分, 当 $P = \text{LSA}$ (local scale averaging) 时表示局部尺度估计, 当 $P = \text{GSA}$ (global scale averaging) 时表示全局尺度估计, 式(1)和式(3)被重述为

$$x_{ij}^P = f_s(x_i^P, x_j^P) = \begin{cases} x_{ij}^{\text{LSA}} = \frac{x_j^{\text{LSA}}}{x_i^{\text{LSA}}} & P = \text{LSA} \\ x_{ij}^{\text{GSA}} = \frac{x_i^{\text{GSA}}}{x_j^{\text{GSA}}} & P = \text{GSA} \end{cases} \quad (6)$$

式中, 下标 s 代表尺度问题。由1.1节中的分析可知, 在进行尺度解算前, 需要通过MCC构建尺度解

算图 $G_s = (V_s, E_s)$, 式中 V_s 表示顶点集合, E_s 表示边集合。图 G_s 中包含 $|V_s|$ 个待求解的尺度标量 $\{x_i^P | v_i \in V_s\}$ 和 $|E_s|$ 个已确定的约束条件 $\{x_{ij}^P | e_{ij} \in E_s\}$ (LSE问题的约束条件为深度比 d_{jk}^i , GSE问题的约束条件为相对尺度 s_{ij}), v_i 表示顶点集中的顶点, e_{ij} 表示边集中的边。考虑到每条约束边所提供信息的准确程度不同, ISE算法采取了与IRA算法(Gao等, 2021)相似的处理方式, 将特征匹配点数目的多少加入到各边权重的计算中, 匹配点数目越多, 该边权重越大。具体而言, 在LSE问题中, 使用的是3个相机共视特征点的数目, 其中相机 i 和相机 j 用于局部尺度解算, 另一个相机用于尺度解算图的构建, 对应的特征点数目记做 n_{ij}^{LSA} ; 在GSE问题中, 使用的是两个相机 i 和 j 对应图像之间的特征匹配点数目, 记做 n_{ij}^{GSA} 。参考式(6), 本文将 n_{ij}^{LSA} 和 n_{ij}^{GSA} 统一记做 n_{ij}^P , 保证了问题求解的一致性。

虽然尺度分离策略(Cui和Tan, 2015)能够有效消除平移平均流程中的尺度歧义问题, 但最终计算所得的绝对尺度与相对基线长度并非真实值, 而是受到基准尺度值的影响, 与真实值之间相差一个尺度因子。在以相机 i 为中心构建的局部尺度解算图中, 假设 $\{s_{ij}^i\}$ 是一组有效解, 那么进行缩放并变化了 α_1 倍的 $\{\hat{s}_{ij}^i = \alpha_1 s_{ij}^i\}$ 也是一组有效解, 同样地, 当以相机 j 为中心构建局部尺度解算图时, $\{s_{ij}^j\}$ 和变化了 α_2 倍的 $\{\hat{s}_{ij}^j = \alpha_2 s_{ij}^j\}$ 均为有效解, 根据式(3), \hat{s}_{ij}^i 和 s_{ij}^j 之间的相互关系满足 $\hat{s}_{ij}^i = \frac{\alpha_2}{\alpha_1} s_{ij}^j$ 。在随后求解全局绝对尺度的过程中, 假设 $\{s_i\}$ 为一组有效解, 那么变化了 α_3 倍的 $\{\alpha_3 s_i\}$ 也同为有效解。综合上述关于局部和全局绝对尺度解的分析, 可以得到如下关系: $\hat{s}_{ij} = \frac{\hat{s}_i}{\hat{s}_j} = \frac{\alpha_2}{\alpha_1} s_{ij} = \frac{\alpha_2 \alpha_3 s_i}{\alpha_1 \alpha_3 s_j}$ 。将上述尺度关系代入式(4), 计算得到相对基线长度 \hat{b}_{ij} 与 b_{ij} 之间的相互关系: $\hat{b}_{ij} = \alpha_1 \alpha_2 \alpha_3 b_{ij}$ 。不难看出, 当基准尺度发生缩放变化时, 绝对尺度值与相对基线值也会随之改变, 这要求用于滤除尺度异常值的距离函数能够消除缩放变化所产生的影响, 本文选用的是归一化完全平方差函数, 具体为

$$d_R(x_i, x_j) = \frac{(x_i - x_j)^2}{x_i x_j} \quad (7)$$

在此基础上,本文的LSE与GSE问题定义为

$$\{x_i^{P*}\} = \arg \min \sum_{\substack{v_i, v_j \in V_s \\ e_{ij} \in E_s}} d_R(x_{ij}^P, f_s(x_i^P, x_j^P)) \quad (8)$$

式中, $\{x_i^{P*}\}$ 表示优化后的局部或全局绝对尺度。当且仅当 x_{ij}^P 满足 $d_R(x_{ij}^P, f_s(x_i^P, x_j^P)) < th_s^P$ 时,才将其作为相对尺度内值加入到解算过程中,其中, th_s^P 表示尺度误差阈值。式(8)所定义的优化问题及ISE算法中其余的优化问题均通过Ceres求解器进行解算。

与IRA算法(Gao等, 2021)和ITA算法(Gao等, 2022b)类似,本文的ISE算法也包含3个关键步骤。首先,基于局部优化的初始三元组选择旨在为增量求解器选取合适的初始种子;其次,基于加权支持集合的下一最优视图(next-best-view, NBV)选择优先考虑具有更高信息精度的相机,力求得到最为准确的最终解;最后,加权局部/全局优化和重新优化不仅能够提高结果的准确性,还可以增强方法的鲁棒性。

为了保证算法的高效性,在选择初始三元组时,ISE仅考虑尺度解算图中三边累计匹配点最多的前 n_{ini} (本文 $n_{ini} = 1000$) 个相机三元组,记做 $t_{ijk} \in T_s$, 其中 t_{ijk} 表示由相机 i, j 和 k 组成的相机三元组, T_s 表示所有被选三元组构成的集合。对于 T_s 中的每个三元组 t_{ijk} , 通过以下优化获取局部坐标系中3个相机的绝对尺度,具体为

$$\{x_i^{P*}, x_j^{P*}, x_k^{P*}\} = \arg \min \sum_{\substack{v_i, v_j \in V_s^{ijk} \\ e_{ij} \in E_s^{ijk}}} w_{ij}^0 d_R(x_{ij}^P, f_s(x_i^P, x_j^P)) \quad (9)$$

式中, V_s^{ijk} 和 E_s^{ijk} 表示三元组中的顶点集和边集, $w_{ij}^0 = n_{ij}^P e^{-d_R(x_{ij}^P, f_s(x_i^{P0}, x_j^{P0}))}$ 代表 e_{ij} 边上的初始权重,可以看出,权重的计算同时考虑了特征匹配点的数目 n_{ij}^P 和相对尺度测量值与计算值之间的误差 $d_R(\cdot, \cdot)$ 。 $\{x_i^{P0} = 1, x_j^{P0} = 1 * x_{ij}^P, x_k^{P0} = 1 * x_{ik}^P\}$ 表示绝对尺度 $\{x_i^P, x_j^P, x_k^P\}$ 的初始值,当 $P = \text{LSA}$ 时, $*$ 表示乘法,而当 $P = \text{GSA}$ 时, $*$ 表示除法。上述优化完成后,根据式(10)选取初始种子 $t_{i^*j^*k^*} = \{v_{i^*}, v_{j^*}, v_{k^*}\}$, 即

$$\{v_{i^*}, v_{j^*}, v_{k^*}\} = \arg \max_{\substack{t_{ijk} \in T_s \\ v_i, v_j \in V_s^{ijk} \\ e_{ij} \in E_s^{ijk}}} \sum n_{ij}^P e^{-d_R(x_{ij}^P, f_s(x_i^P, x_j^P))} \quad (10)$$

随后,增量式尺度估计流程的初始种子恢复被确定为 $\{x_{i^*}^{P*}, x_{j^*}^{P*}, x_{k^*}^{P*}\}$ 。

基于加权支撑集合的NBV选择是增量式参数

估计流程中的另一关键步骤。在这一步骤中,已估计绝对尺度的顶点集合记做 V_s^e , 未估计绝对尺度的顶点集合记做 V_s^u 。对 V_s^u 中的一个顶点 v_j 而言,它与 V_s^e 之间存在边集 E_s^j, E_s^j 中的每条边 e_{ij} 连接着顶点 v_j 与 V_s^e 中的顶点 v_i 。根据顶点 v_i 的绝对尺度对顶点 v_j 的绝对尺度进行预计算,计算式为

$$x_j^{Pi} = x_i^P * x_{ij}^P = \begin{cases} x_j^{\text{LSAi}} = x_i^{\text{LSA}} \times x_{ij}^{\text{LSA}} & P = \text{LSA} \\ x_j^{\text{GSAi}} = \frac{x_i^{\text{GSA}}}{x_{ij}^{\text{GSA}}} & P = \text{GSA} \end{cases} \quad (11)$$

所有预计算的尺度值构成顶点 v_j 的绝对尺度集 $\{x_j^{Pi}\}$ 。由于 $\{x_i^P\}$ 和 $\{x_{ij}^P\}$ 固有误差的存在,集合 $\{x_j^{Pi}\}$ 中的元素不可能完全相同。参考IRA算法与ITA算法的处理方式,对 $\{x_j^{Pi}\}$ 中的每个元素计算选择奖励 R_j^i , 计算式为

$$R_j^i = \sum_{\substack{v_k \in V_s^e \setminus \{v_i\} \\ e_{ij} \in E_s^j \setminus \{e_{ij}\}}} n_{ij}^P e^{-d_R(x_{ij}^P, f_s(x_i^P, x_j^P))} \quad (12)$$

具有最大选择奖励 R_j^{i*} 的元素 x_j^{mi*} 被选为顶点 v_j 的初始绝对尺度。随后,根据以下公式选取下一最优视图 v_{j^*} , 并将其从集合 V_s^u 移动到到集合 V_s^e 中, 即

$$v_{j^*} = \arg \max_{v_j \in V_s^u} (R_j^{i*}) \quad (13)$$

NBV的选取迭代进行,直到 V_s^u 中没有可以被继续解算的顶点。

每添加一次NBV均需进行一次局部优化。对于新选取的顶点 v_{j^*} , 其初始绝对尺度为 $x_{j^*}^P$, 在局部优化开始前,首先获取其内边集 $I(E_s^{j^*})$, 具体为

$$I(E_s^{j^*}) = \left\{ e_{ij} \mid v_i \in V_s^e, d_R(x_{ij}^P, f_s(x_i^P, x_j^P)) < th_s^P \right\} \quad (14)$$

随后,局部优化被执行,即

$$x_{j^*}^{P*} = \arg \min \sum_{\substack{v_i \in V_s^e \\ e_{ij^*} \in I(E_s^{j^*})}} w_{ij^*}^P d_R(x_{ij^*}^P, f_s(x_i^P, x_{j^*}^P)) \quad (15)$$

式中, $w_{ij^*}^P = n_{ij^*}^P e^{-d_R(x_{ij^*}^P, f_s(x_i^P, x_{j^*}^P))}$ 表示权重。与局部优化不同的是,只有当已估计绝对尺度的顶点的增长率超过 r_{inc} (本文 $r_{inc} = 30\%$) 时,全局优化才被执行。同样,需要获取内边集 $I(E_s^e \cup E_s^{j^*})$, 具体为

$$I(E_s^e \cup E_s^{j^*}) = \left\{ e_{ij} \mid v_i, v_j \in V_s^e \cup \{v_{j^*}\}, d_R(x_{ij}^P, f_s(x_i^P, x_j^P)) < th_s^P \right\} \quad (16)$$

随即进行全局优化,即

$$\{x_i^{p*}\} = \arg \min \sum_{\substack{v_i, v_j \in V^* \cup \{v_s\} \\ c_{ij} \in \{E_i^* \cup E_j^*\}}} w_{ij} d_{R^*}(x_i^p, f_s(x_i^p, x_j^p)) \quad (17)$$

式中, $w_{ij} = n_{ij}^p e^{-d_{R^*}(x_i^p, f_s(x_i^p, x_j^p))}$ 表示权重。全局优化结束后, 再次基于式(16)选取内边集合, 参考式(17)完成绝对尺度的重优化。

1.2.2 尺度已知的绝对位置解算

由于尺度解算时的增量式流程已经对异常值进行了有效滤除, 因此在位置解算时, 本文采用了与 Cui 和 Tan(2015) 相同的解算方法, 先构建线性超定方程组, 随后基于 L_1 范数优化求解。

当局部绝对尺度与全局绝对尺度的解算全部完成后, 根据式(4)计算相机间的基线长度 b_{ij} , 根据式(5)可得

$$R_j(c_i - c_j) = b_{ij} t_{ij} \quad (18)$$

MCC 中的每条边均对应一个上述方程, 将所有方程堆叠到一起构建如下线性超定方程组, 具体为

$$A_c x_c = b_c \quad (19)$$

式中, 下标 c 代表位置问题, A_c 是由非零元素 $\{R_j\}$ 和 $\{-R_j\}$ 构成的稀疏矩阵, x_c 对应待求解的绝对位置 $\{c_i\}$, b_c 表示已知尺度的相对平移 $\{b_{ij} t_{ij}\}$ 。令 $c_1 = \{0, 0, 0\}$, 通过以下优化完成对相机绝对位置的解算, 具体为

$$\arg \min_{x_c} \|A_c x_c - b_c\|_1 \quad (20)$$

需要注意的是, 本文的 L_1 优化问题均通过 OpenMVG 完成求解, 与 Cui 和 Tan(2015) 的研究略有不同。

2 实验结果与分析

2.1 实验环境

本文所使用的操作系统为 Ubuntu 20.04, 以 Visual Studio Code 为代码编辑器, 基于 C++ 语言进行开发。处理器型号为 Intel® Core™ i9-10900K, 主频 3.70 GHz。本文算法所依赖的第三方库主要包括 OpenCV、Eigen、Ceres、Boost 和 OpenMVG。

2.2 实验数据及评测方法

2.2.1 实验数据

本文实验数据分别来自 1DSfM 数据集 (Wilson 和 Snavely, 2014) 与 NotreDame (NTD) 数据集 (Sna-

vely 等, 2006), 详情可以参考 <https://www.cs.cornell.edu/projects/bigsfm/>。由 1.1 节内容可知, 无论是相对深度比的计算, 还是局部绝对尺度的估计, 均离不开相机三元组间特征点匹配信息的支撑, 但原始数据集中尚未给出相关内容。因此, 本文通过 COLMAP (collision-mapping) (Schönberger 和 Frahm, 2016) 工具箱对参与评测的各组数据进行了重新处理, 使用默认参数提取图像的特征点与匹配关系, 随后利用 OpenCV 库估计并分解本质矩阵 (Nister, 2004), 获取每个场景下相机之间的相对位姿, 生成新的 EG 图。

各组评测数据的元数据如表 1 所示, 表中 $|V|$ 和 $|V_{MCC}|$ 分别表示每个场景中原始的图像数量和参与计算的图像数量, $|E_{MCC}|$ 表示重新生成的 EG 图中的边数, \tilde{e}_R 和 \tilde{e}_T 则分别表示相对旋转和相对平移关于角度的中值误差。

表 1 1DSfM 数据集的元数据
Table 1 Metadata of the 1DSfM dataset

数据名称	$ V $	$ V_{MCC} $	$ E_{MCC} $	$\tilde{e}_R/(\circ)$	$\tilde{e}_T/(\circ)$
ALM	2 915	627	72 544	5.03	8.53
ELS	2 587	247	15 106	2.47	8.79
GDM	1 463	742	31 504	3.86	7.65
MDR	1 344	394	15 658	7.08	13.38
MND	2 298	474	37 063	2.34	5.32
NYC	2 550	376	15 104	3.72	7.13
PDP	2 251	354	21 434	2.10	4.17
PIC	7 351	2 508	267 246	4.68	10.23
ROF	2 364	1 134	61 713	4.41	7.89
TOL	1 576	508	22 716	3.63	4.48
TFG	15 685	5 433	1 056 735	4.19	9.52
USQ	5 961	930	33 942	4.23	23.52
VNC	6 288	918	72 884	3.58	8.35
YKM	3 368	458	18 417	3.74	7.20
NTD	553	553	68 146	2.82	5.06

2.2.2 评测方法

在传统平移平均算法中, 多以 Bundler (Snavely 等, 2008) 的相机标定结果作为相机位置的真值。然而, 近年来, COLMAP 因其频繁的更新与良好的性能已逐步成为应用最为广泛的开源 SfM 工具箱。因

此,为了能够对平移平均结果进行更为精确的评估,本文参照 Gao 等人(2023b)的工作,综合考虑上述两种标定结果,只有当 MCC 中的顶点满足如下条件时,才可以参与到最终的精度评测中,具体为

$$d_{s_0(3)}(\mathbf{R}_i^C, \mathbf{R}_i^B) < 5^\circ \cap d_{R^*}(\mathbf{c}_i^C, \mathbf{c}_i^B) < 0.5 \text{ m} \quad (21)$$

式中, $d_{s_0(3)}(\mathbf{R}_1, \mathbf{R}_2) = \arccos \frac{\text{tr}(\mathbf{R}_2 \mathbf{R}_1^T) - 1}{2}$ 表示相对旋转的角度误差, C^* 表示 COLMAP 的输出结果已对齐到 Bundler 带有真实物理尺度的输出结果(B)上, 5° 和 0.5 m 分别为旋转的误差阈值和位置的误差阈值。符合上述条件的所有顶点构成集合 $\mathbf{V}_c^{C \cap B}$, 将通过平移平均算法恢复的相机绝对位置与 $\{\mathbf{c}_j^C | v_j \in \mathbf{V}_c^{C \cap B}\}$ 统一到同一个坐标系中,并以中值误差(\bar{e}_c)与均值误差(\bar{e}_c)作为位置精度评测的主要参考标准。关于这一评测方法的更多细节可以参考 Gao 等人(2023b)的研究。

此外,由于本文基于尺度分离策略解决平移平均流程中的尺度歧义问题,并将增量式参数估计的思想引入到求解过程中,在解算相机绝对位置前预先计算相机间的相对基线长度,因此,这一参数的准确性也是本文一个极为重要的评测指标。在与另一种基于尺度分离策略的平移平均算法 $L_1\text{SA}-L_1\text{TA}$ 进行基线精度评测时,为了保证对比实验的全面性与合理性,本文设计了一种新的相对基线评测方案。

与位置评测方法相似,首先需要选取可以参与评测的基线真值,具体为

$$|b_{ij}^C - b_{ij}^B| < 0.5 \text{ m} \quad (22)$$

式中, $b_{ij}^C = \|\mathbf{c}_j^C - \mathbf{c}_i^C\|_2$ 和 $b_{ij}^B = \|\mathbf{c}_j^B - \mathbf{c}_i^B\|_2$ 表示由绝对位置真值计算得到的基线真值。考虑到尺度分离策略下的尺度值只能按比恢复,因而其估计所得的基线长度也需要与真值对齐。为了实现这一目的,本文进行了基于中值的标量对齐。对于待评估的算法 X , 根据式(4)计算得到所有相机间的基线长度后,将其与基线真值统一表示为 $\{b_{mn}^X, b_{mn}^C | e_{mn} \in \mathbf{E}_b^{C \cap B}\}$, 式中, $\mathbf{E}_b^{C \cap B}$ 是选取到的内边集合,下标 b 代表基线问题,进而得到一个比例中值 $\alpha_b = \text{med}(\{b_{mn}^C/b_{mn}^X\})$, 在此基础上,各算法相应的基线误差被确定为 $\{\alpha_b b_{mn}^X - b_{mn}^C\}$ 。最后,中值误差(\bar{e}_b)和均值误差(\bar{e}_b)成为基线精度评测的主要参考指标。

除上述评测指标外,考虑到可解范围的不同,本

文还记录了各平移平均算法下解算成功的相机数量 n , 并将其作为性能评测的指标之一。

2.3 实验结果与分析

2.3.1 距离函数选取实验及分析

本文 1.2 小节已经从理论上对选择归一化完全平方差函数作为尺度距离函数的优势进行了分析,为了进一步验证这种选择的合理性和有效性,以未考虑缩放效应的函数 $\dot{d}_R(x_i, x_j) = |x_i - x_j|$ 和考虑了缩放效应的函数 $\ddot{d}_R(x_i, x_j) = \left| \frac{x_i}{x_j} - 1 \right|$ 为参考,进行了尺度距离函数的选取实验。为了评测不同尺度距离函数对缩放效应抵抗能力的强弱,分别以 $\dot{d}_R(x_i, x_j)$, $\ddot{d}_R(x_i, x_j)$ 和 $d_R(x_i, x_j)$ 作为尺度距离函数,始终保持 LSE 问题中的基准尺度等于 GSE 问题中的基准尺度,并在每组实验中将该基准尺度依次设置为 0.1 , 1 和 10 , 记做 $s_0 = \{0.1, 1, 10\}$, 相对基线的中值误差与均值误差如表 2 和表 3 所示。

由表中数据可知,当使用考虑了缩放效应的距离函数 $\ddot{d}_R(x_i, x_j)$ 和 $d_R(x_i, x_j)$ 时,其基线计算结果优势明显。一方面,观察表中基线结果的中值误差,不论是距离函数的形式还是基准尺度的大小发生变化,其误差值均能保持在一个相对合理的范围内,这意味着增量式(局部/全局)尺度求解器能够得到合理有效的尺度解,并且在尺度距离函数和基准尺度发生变化时相对鲁棒。另一方面,观察表中的均值误差,当距离函数选用不考虑缩放效应的 $\dot{d}_R(x_i, x_j)$ 时,其均值误差显著升高,但当选用考虑了缩放效应的 $\ddot{d}_R(x_i, x_j)$ 和 $d_R(x_i, x_j)$ 时,其均值误差却基本保持稳定,仅在以 $\ddot{d}_R(x_i, x_j)$ 为距离函数且 $s_0 = 10$ 时发生较大波动。因此,基于上述现象与分析,本文选取考虑了缩放效应的 $d_R(x_i, x_j)$ 作为 ISE 算法中的尺度距离函数。

2.3.2 阈值参数设置实验及分析

除尺度距离函数的选择会对基线结果的精度产生重要影响外,尺度距离阈值(th_s^p)的设置也至关重要。当距离阈值过大时,解算过程无法对异常值进行有效的滤除,可能导致求解错误或精度下降,增量过程的鲁棒性受到影响;而当距离阈值过小时,虽然能够确保参与求解的信息准确可靠,但其严苛的选择条件将导致部分正确信息的丢失,进而影响到求

表2 不同尺度距离函数对相对基线长度中值误差 \tilde{e}_b 的影响Table 2 The impact of different scale distance functions on the median error (\tilde{e}_b) of relative baseline length

数据名称	/m								
	$d_R(x_i, x_j) = x_i - x_j $			$\ddot{d}_R(x_i, x_j) = \left \frac{x_i}{x_j} - 1 \right $			$d_R(x_i, x_j) = \frac{(x_i - x_j)^2}{x_i x_j}$		
	$s_0 = 0.1$	$s_0 = 1$	$s_0 = 10$	$s_0 = 0.1$	$s_0 = 1$	$s_0 = 10$	$s_0 = 0.1$	$s_0 = 1$	$s_0 = 10$
ALM	0.70	0.68	0.68	0.67	0.67	0.67	0.67	0.67	0.67
ELS	6.64	6.34	6.29	6.36	6.36	6.35	6.50	6.49	6.49
GDM	2.65	2.64	2.65	2.61	2.62	2.62	2.66	2.65	2.66
MDR	3.50	3.43	3.44	3.31	3.30	3.31	3.29	3.29	3.29
MND	0.79	0.78	0.78	0.78	0.77	0.78	0.78	0.78	0.77
NYC	1.15	1.10	1.10	1.07	1.07	1.06	1.07	1.07	1.07
PDP	1.41	1.40	1.40	1.39	1.39	1.38	1.39	1.39	1.39
PIC	1.59	1.57	1.56	1.54	1.53	1.54	1.53	1.53	1.53
ROF	3.94	3.77	3.63	3.21	3.22	3.21	3.02	3.02	3.01
TOL	3.46	3.33	3.33	3.35	3.33	3.35	3.31	3.31	3.31
TFG	5.33	5.28	5.32	5.46	5.48	5.48	5.23	5.23	5.23
USQ	4.05	3.98	3.97	4.02	4.01	4.00	4.05	4.05	4.05
VNC	2.78	2.73	2.73	2.71	2.71	2.72	2.70	2.70	2.70
YKM	2.24	2.20	2.18	2.09	2.10	2.09	2.02	2.03	2.03
NTD	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.38

注:加粗字体表示各行最优结果。

表3 不同尺度距离函数对相对基线长度均值误差 \bar{e}_b 的影响Table 3 The impact of different scale distance functions on the mean error (\bar{e}_b) of relative baseline length

数据名称	/m								
	$d_R(x_i, x_j) = x_i - x_j $			$\ddot{d}_R(x_i, x_j) = \left \frac{x_i}{x_j} - 1 \right $			$d_R(x_i, x_j) = \frac{(x_i - x_j)^2}{x_i x_j}$		
	$s_0 = 0.1$	$s_0 = 1$	$s_0 = 10$	$s_0 = 0.1$	$s_0 = 1$	$s_0 = 10$	$s_0 = 0.1$	$s_0 = 1$	$s_0 = 10$
ALM	2.62e+13	1.52e+05	2.93e+13	1.55	1.55	1.55	1.55	1.55	1.55
ELS	1.93e+09	1.68e+09	26.37	13.53	14.05	13.51	13.61	13.61	13.61
GDM	2.51e+06	4.50e+05	10.12	5.23	5.23	5.24	5.30	5.28	5.30
MDR	7.61e+07	97.01	11.67	10.57	10.57	10.57	10.61	10.61	10.61
MND	1.12e+11	2.41e+06	4.84e+04	1.22	1.22	1.22	1.20	1.20	1.20
NYC	4.09e+06	2 990.80	1.12e+06	1.81	1.81	1.81	1.80	1.80	1.80
PDP	4.21e+07	2 751.41	1.53e+06	2.62	2.62	2.62	2.64	2.64	2.64
PIC	2.41e+20	6.77e+10	1.56e+10	3.32	3.31	3.31	3.31	3.31	3.31
ROF	2.70e+18	4.75e+05	2.44e+06	7.62	7.63	2.77e+07	7.47	7.47	7.46
TOL	1.02e+12	8.78e+04	8.07e+08	6.63	6.62	6.63	6.42	6.42	6.42
TFG	1.55e+24	1.08e+15	2.21e+10	10.59	10.60	13.33	10.50	10.50	10.50
USQ	1.73e+06	9.43e+04	686.19	6.76	6.77	6.74	6.73	6.73	6.73
VNC	6.66e+21	1.45e+05	2.08e+11	4.40	4.40	4.41	4.38	4.38	4.38
YKM	4.77e+10	1.77e+11	1.02e+05	3.42	3.44	3.42	3.37	3.37	3.27
NTD	0.57	0.57	0.57	0.57	0.57	0.57	0.58	0.58	0.58

注:加粗字体表示各行最优结果。

解结果的完整性。在阈值参数设置实验中,通过大量实验发现,当 $10^{-6} \leq th_s^p \leq 10^0$ 时,增量式尺度估计所得的基线长度能够保持在合理的精度范围内,而当 $th_s^p < 10^{-6}$ 或 $th_s^p > 10^0$ 时,则可能出现求解错误或精度波动较大的情况。为了确定最终的尺度阈值,本文以基线求解精度为评测指标,分别选取 10^0 、 10^{-2} 、 10^{-4} 和 10^{-6} 为阈值参数,进行了相关的阈值参数设置实验,实验结果如表4所示。值得注意的是,在距离函数选取实验中,所有的尺度距离阈值均被设置为 10^{-2} ,而在本实验中,所有的尺度距离函数均被设置为 $d_R(x_i, x_j)$ 。由表4中的数据可以看出,当 th_s^p 分别设置为 10^0 、 10^{-2} 、 10^{-4} 和 10^{-6} 时,虽然基线求解精度的整体波动较小,但在 $th_s^p = 10^{-2}$ 时仍占有一定优势。在中值误差方面,除ELS(ellis island)与USQ(union square)两组数据外, $th_s^p = 10^{-2}$ 所对应的

表4 不同尺度距离阈值对相对基线长度中值误差 \tilde{e}_b 与均值误差 \bar{e}_b 的影响

Table 4 The impact of different scale distance thresholds on median and mean error (\tilde{e}_b & \bar{e}_b) of relative baseline length

数据名称	$th_s^p = 10^0$		$th_s^p = 10^{-2}$		$th_s^p = 10^{-4}$		$th_s^p = 10^{-6}$	
	\tilde{e}_b	\bar{e}_b	\tilde{e}_b	\bar{e}_b	\tilde{e}_b	\bar{e}_b	\tilde{e}_b	\bar{e}_b
ALM	0.68	1.59	0.67	1.55	0.67	1.55	0.67	1.55
ELS	6.56	13.62	6.49	13.61	6.33	13.55	6.35	13.53
GDM	2.75	5.19	2.65	5.28	2.90	5.42	3.16	5.80
MDR	3.33	10.64	3.29	10.61	3.32	10.61	3.33	10.59
MND	0.78	1.22	0.78	1.20	0.78	1.21	0.78	1.20
NYC	1.08	1.87	1.07	1.80	1.07	1.81	1.07	1.80
PDP	1.40	2.67	1.39	2.64	1.40	2.63	1.39	2.63
PIC	1.55	3.30	1.53	3.31	1.55	3.32	1.55	3.32
ROF	3.19	7.50	3.02	7.47	3.30	7.52	3.39	7.57
TOL	3.37	6.37	3.31	6.42	3.39	6.41	3.41	6.43
TFG	5.35	10.52	5.22	10.49	5.45	10.55	5.57	10.61
USQ	4.02	6.73	4.05	6.73	4.03	6.73	4.02	6.71
VNC	2.76	4.40	2.70	4.38	2.72	4.38	2.71	4.38
YKM	2.20	3.55	2.02	3.37	2.18	3.51	2.18	3.52
NTD	0.38	0.58	0.38	0.58	0.38	0.58	0.38	0.57

注:加粗字体分别表示各行中 \tilde{e}_b 和 \bar{e}_b 的最优结果。

结果均为最优值,而在均值误差方面,虽然该组结果仅在部分数据上达到最优,但其余各场景下的基线解算精度与最优精度之间仅有微小差异,求解所得的基线精度整体较高。因此,综合分析4组实验结果,本文将 $th_s^p = 10^{-2}$ 确定为最终的尺度距离阈值。

2.3.3 解算精度对比实验及分析

在精度对比实验中,由于本文方法源于Cui和Tan(2015)提出的 L_1 SE- L_1 TA算法,整体解算流程较为相似,主要的差异体现在尺度解算方式的不同上,因而首先针对这两种算法的尺度解算性能,也就是相对基线长度的精度进行了详细的对比实验,实验结果如表5所示。

表5 基线精度的对比

Table 5 Comparison of baseline accuracy

数据名称	L_1 SE- L_1 TA(Cui和Tan,2015)		本文方法	
	\tilde{e}_b	\bar{e}_b	\tilde{e}_b	\bar{e}_b
ALM	10.15	10.91	0.67	1.55
ELS	23.55	23.75	6.52	13.43
GDM	4.77	16.27	2.71	5.29
MDR	6.28	14.83	3.33	10.59
MND	7.38	9.73	0.82	1.22
NYC	3.16	8.60	1.07	1.79
PDP	11.05	13.12	1.48	2.66
PIC	9.98	10.60	1.59	3.28
ROF	10.29	20.77	3.22	7.55
TOL	9.31	27.67	3.47	6.51
TFG	30.86	34.21	5.97	10.53
USQ	6.72	9.68	4.18	6.56
VNC	14.04	16.18	2.78	4.34
YKM	4.07	10.84	2.07	3.43
NTD	5.54	7.11	0.40	0.60

注:加粗字体分别表示各行中 \tilde{e}_b 和 \bar{e}_b 的最优结果。

由表5数据可知,本文方法解算所得的相对基线长度比 L_1 SE- L_1 TA解算所得的相对基线长度更为准确,证明了ISE算法在尺度解算上的准确性与有效性。

随后,为了对基于增量尺度解算的平移平均流

程进行更加全面的性能评估,本文还将其与 L_1 SE- L_1 TA算法(Cui和Tan,2015)、BATA算法(Zhuang等,2018)、ITA算法(Gao等,2022b)和CReTA算法(Manam和Govindu,2022)的位置解算性能进行了对比,所使用的绝对旋转分别源于IRA算法(Gao等,2021)和 L_1 GM(L_1 Geman-McClure)算法(Chatterjee和Govindu,2013),参与对比的评测指标包括相机解算数量(n)和位置解算精度(\tilde{e}_c 和 \bar{e}_c),实验结果如表6和表7所示,表中最后一列AVG表示各算法在不同

评测指标下所有场景解算结果的均值。

由表6和表7数据可知,在相机解算数量上,本文方法略次于BATA与CReTA,优于ITA与 L_1 SE- L_1 TA。在绝对位置解算精度(\tilde{e}_c 和 \bar{e}_c)方面,由于本文方法在USQ场景下的表现略差,因而其中值误差整体排名在两组实验中低于BATA与CReTA,均值误差整体排名在第2组实验中略低于CReTA,但除该场景外,本文方法的解算精度基本优于BATA,并与CReTA效果相似。

表6 本文方法与其他几种主流平移平均方法的比较(绝对旋转:IRA(Gao等,2021))

Table 6 Comparison between the proposed method and several other state-of-the-art translation averaging methods (absolute rotations: IRA(Gao et al., 2021))

数据名称	L_1 SE- L_1 TA (Cui和Tan,2015)			BATA (Zhuang等,2018)			ITA (Gao等,2022b)			CReTA (Manam和Govindu,2022)			本文方法		
	n	\tilde{e}_c/m	\bar{e}_c/m	n	\tilde{e}_c/m	\bar{e}_c/m	n	\tilde{e}_c/m	\bar{e}_c/m	n	\tilde{e}_c/m	\bar{e}_c/m	n	\tilde{e}_c/m	\bar{e}_c/m
ALM	580	0.48	1.46	625	0.72	1.20	603	0.46	2.16	622	0.46	0.82	607	0.49	1.01
ELS	231	2.10	3.45	245	3.43	20.59	253	1.64	2.42	246	2.16	3.30	243	2.03	3.35
GDM	694	1.66	2.85	740	1.66	2.51	717	1.16	1.80	738	1.77	2.88	721	1.80	4.84
MDR	338	1.63	3.04	393	1.84	4.39	369	58.96	2.90e+05	391	0.84	1.74	352	1.08	3.00
MND	451	0.54	1.04	473	0.68	0.84	465	0.43	1.06e+04	472	0.52	0.78	463	0.44	0.58
NYC	330	0.59	1.31	374	0.66	0.97	355	0.57	1.00	373	0.72	0.92	349	0.60	1.07
PDP	346	0.90	2.24	353	0.86	3.47	347	0.83	33.80	353	0.87	3.23	352	0.91	2.25
PIC	2 327	1.41	2.58	2 502	1.09	1.82	2 443	0.80	44.89	2 485	1.01	1.83	2 441	1.23	1.91
ROF	1 091	5.14	11.13	1 133	7.24	9.86	1 113	8.36	17.91	1 132	5.19	12.13	1 119	4.49	9.95
TOL	475	2.52	3.90	505	2.75	3.56	482	1.94	3.46	505	2.33	3.17	486	2.37	3.30
TFG	5 173	5.76	8.63	5 432	3.93	5.80	5 349	2.57	1 420.22	5 404	3.75	7.19	5 358	5.07	8.61
USQ	864	12.23	13.57	929	2.53	3.18	872	2.33	1.91e+04	920	12.82	14.14	859	12.52	13.7
VNC	864	1.46	3.09	915	1.35	3.11	851	1.32	2.37	910	1.60	5.00	896	1.27	2.35
YKM	431	1.30	2.38	457	1.04	8.11	432	0.77	1.70	455	1.41	2.97	436	1.24	2.57
NTD	539	0.32	0.94	533	0.36	0.64	524	0.34	8.89	533	0.32	0.48	552	0.29	0.42
平均值	982	2.54	4.11	1 041	2.01	4.67	1 012	5.50	2.14e+04	1 036	2.38	4.04	1 016	2.39	3.93

注:加粗字体分别表示各行中 n , \tilde{e}_c 和 \bar{e}_c 的最优结果。

综合分析表6和表7的实验结果可知,相较于原始的基于尺度分离策略的平移平均算法(L_1 SE- L_1 TA),本文基于增量式尺度估计的平移平均算法(ISE- L_1 TA)能够有效提升相机的解算数量与位置的解算精度,进而证明了本文方法在尺度估计与位置解算上的鲁棒性与可靠性。根据本文方法解算所得的相机位置对所有场景进行重建,重建结果如图2所示,图中所有场景均被准确重建,进一步证明了本

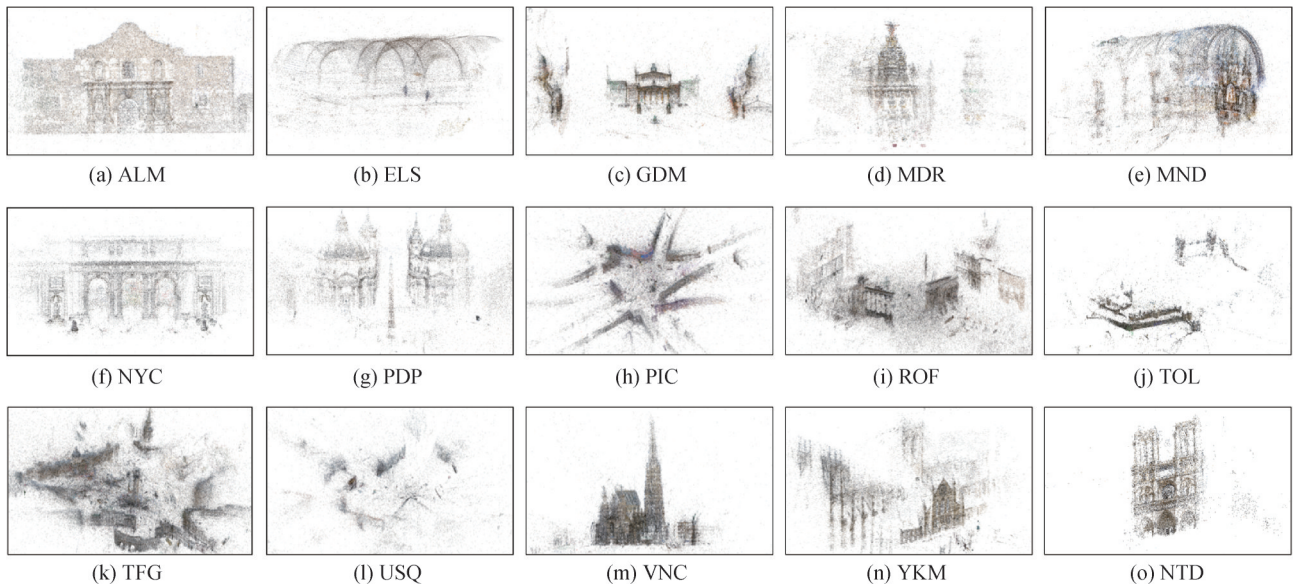
文方法的合理性与有效性。

3 结论

本文提出了一种基于增量式尺度估计的平移平均方法,将尺度分离和增量估计的思想同时引入到平移平均流程中。通过增量式尺度求解器解算相机的局部和全局绝对尺度,一方面可以消除相机三元

表7 本文方法与其他几种主流平移平均方法的比较(绝对旋转: L_1 GM(Chatterjee和Govindu, 2013))Table 7 Comparison between the proposed method and several other state-of-the-art translation averaging methods (absolute rotations: L_1 GM(Chatterjee and Govindu, 2013))

数据名称	L_1 SE- L_1 TA (Cui和Tan, 2015)			BATA (Zhuang等, 2018)			ITA (Gao等, 2022b)			CReTA(Manam和Govindu, 2022)			本文方法		
	n	$\bar{\epsilon}_c/m$	$\bar{\epsilon}_e/m$	n	$\bar{\epsilon}_c/m$	$\bar{\epsilon}_e/m$	n	$\bar{\epsilon}_c/m$	$\bar{\epsilon}_e/m$	n	$\bar{\epsilon}_c/m$	$\bar{\epsilon}_e/m$	n	$\bar{\epsilon}_c/m$	$\bar{\epsilon}_e/m$
ALM	580	0.49	1.55	625	0.69	1.16	605	0.46	2.36	618	0.41	0.75	607	0.48	0.98
ELS	231	2.06	3.46	245	3.44	20.72	235	1.62	2.44	244	2.02	2.78	243	2.08	3.35
GDM	694	2.83	13.21	740	3.11	17.25	716	2.23	48.22	731	2.15	18.15	721	2.31	14.92
MDR	338	1.07	2.90	393	1.46	4.16	369	611.44	3.36e+04	390	0.85	1.78	352	1.07	2.84
MND	451	0.54	1.04	473	0.68	0.85	459	0.47	1.50e+04	371	0.47	0.60	463	0.44	0.57
NYC	330	0.57	1.30	374	0.67	0.99	353	0.52	0.80	370	0.55	0.79	349	0.58	1.06
PDP	346	0.91	2.19	353	0.79	3.40	348	0.76	70.89	354	0.85	3.07	352	0.88	2.30
PIC	2 327	1.40	2.50	2 502	1.07	1.76	2 444	0.79	328.06	2 489	0.82	1.35	2 441	1.21	1.79
ROF	1 091	3.65	6.82	1 133	3.85	7.17	1 114	1.53	2.59	1 132	1.66	3.22	1 119	4.58	6.20
TOL	475	2.83	4.00	505	2.88	3.82	487	2.11	3.18	504	3.80	5.12	486	2.41	3.16
TFG	5 173	5.69	8.75	5 432	3.76	5.52	5 360	2.56	6.34	5 403	3.74	6.93	5 358	4.94	8.57
USQ	864	12.44	13.62	929	2.10	2.86	865	1.19	14.59	922	3.45	4.21	859	12.43	13.72
VNC	864	1.45	3.50	915	1.32	3.98	855	1.24	3.88	916	1.33	5.31	896	1.26	2.77
YKM	431	1.29	2.35	457	0.97	7.73	437	0.86	2.16	455	1.08	1.69	436	1.21	2.55
NTD	539	0.32	0.91	533	0.36	0.60	524	0.31	120.01	533	0.27	0.40	552	0.28	0.39
平均值	982	2.50	4.54	1 041	1.81	5.46	1 011	41.87	3 280.37	1 029	1.56	3.74	1 016	2.41	4.34

注:加粗字体分别表示各行中 n , $\bar{\epsilon}_c$ 和 $\bar{\epsilon}_e$ 的最优结果。图2 基于ISE- L_1 TA位置解算的稀疏重建点云图Fig. 2 Point cloud map of sparse reconstruction based on ISE- L_1 TA location recovery ((a) ALM; (b) ELS; (c) GDM; (d) MDR; (e) MND; (f) NYC; (g) PDP; (h) PIC; (i) ROF; (j) TOL; (k) TFG; (l) USQ; (m) VNC; (n) YKM; (o) NTD)

组对平移平均问题的可解性约束,使得位置解算在相机共线运动时仍能有序进行,缓解尺度歧义性所带来的不利影响;另一方面能够在尺度解算过程中有效滤除相对平移异常值,大大提升算法的鲁棒性。最终的实验结果验证了本文所提方法的有效性与准确性。

本文工作仍存在不足之处,虽然尺度分离策略的引入将位置解算范围从平行刚性分量扩展到了最大连通分量,但最终解算出的相机数量并非最优,估计所得的绝对位置准确性也有待提高。在未来的研究工作中,将考虑把增量式参数估计思想继续引入到尺度已知的绝对位置估计流程中,对引入后的算法能否提高位置的解算精度并增加相机的解算数量做出进一步的探究。

参考文献 (References)

- Bourmaud G. 2016. Online variational Bayesian motion averaging//Proceedings of the 14th European Conference on Computer Vision. Amsterdam, the Netherlands: Springer: 126-142 [DOI: 10.1007/978-3-319-46484-8_8]
- Candès E J, Wakin M B and Boyd S P. 2008. Enhancing sparsity by reweighted ℓ_1 minimization. *Journal of Fourier Analysis and Applications*, 14(5/6): 877-905 [DOI: 10.1007/s00041-008-9045-x]
- Chatterjee A and Govindu V M. 2013. Efficient and robust large-scale rotation averaging//Proceedings of 2013 IEEE International Conference on Computer Vision. Sydney, Australia: IEEE: 521-528 [DOI: 10.1109/ICCV.2013.70]
- Cui H N, Shen S H, Gao X and Hu Z Y. 2017. Batched incremental structure-from-motion//Proceedings of 2017 IEEE International Conference on 3D Vision. Qingdao, China: IEEE: 205-214 [DOI: 10.1109/3DV.2017.00032]
- Cui Z P and Tan P. 2015. Global structure-from-motion by similarity averaging//Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago, Chile: IEEE: 864-872 [DOI: 10.1109/ICCV.2015.105]
- Furukawa Y and Ponce J. 2010. Accurate, dense, and robust multiview stereopsis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(8): 1362-1376 [DOI: 10.1109/TPAMI.2009.161]
- Gao X, Luo J Z, Li K Q and Xie Z X. 2020. Hierarchical RANSAC-based rotation averaging. *IEEE Signal Processing Letters*, 27: 1874-1878 [DOI: 10.1109/LSP.2020.3031759]
- Gao X, Zhu L J, Cui H N, Xie Z X and Shen S H. 2022a. IRA++: distributed incremental rotation averaging. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(7): 4885-4892 [DOI: 10.1109/TCSVT.2021.3118883]
- Gao X, Zhu L J, Fan B, Liu H M and Shen S H. 2022b. Incremental translation averaging. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(11): 7783-7795 [DOI: 10.1109/TCSVT.2022.3183631]
- Gao X, Zhu L J, Xie Z X, Liu H M and Shen S H. 2021. Incremental rotation averaging. *International Journal of Computer Vision*, 129(4): 1202-1216 [DOI: 10.1007/s11263-020-01427-7]
- Gao X, Cui H N, Huang W T, Li M H and Shen S H. 2024. IRAv3+: hierarchical incremental rotation averaging via multiple connected dominating sets. *IEEE Transactions on Circuits and Systems for Video Technology*, 34(4): 3049-3055 [DOI: 10.1109/TCSVT.2023.3309661]
- Gao X, Cui H N, Li M H, Xie Z X and Shen S H. 2023a. IRAv3: hierarchical incremental rotation averaging on the fly. *IEEE Transactions on Circuits and Systems for Video Technology*, 33(4): 2001-2006 [DOI: 10.1109/TCSVT.2022.3217151]
- Gao X, Cui H N and Shen S H. 2023b. Incremental rotation averaging revisited and more: a new rotation averaging benchmark [EB/OL]. [2023-09-29]. <https://arxiv.org/pdf/2309.16924.pdf>
- Goldstein T, Hand P, Lee C, Voroninski V and Soatto S. 2016. ShapeFit and ShapeKick for robust, scalable structure from motion//Proceedings of the 14th European Conference on Computer Vision-ECCV 2016. Amsterdam, the Netherlands: Springer: 289-304 [DOI: 10.1007/978-3-319-46478-7_18]
- Govindu V M. 2001. Combining two-view constraints for motion estimation//Proceedings of 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Kauai, USA: IEEE: 218-225 [DOI: 10.1109/CVPR.2001.990963]
- Govindu V M. 2004. Lie-algebraic averaging for globally consistent motion estimation//Proceedings of 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, USA: IEEE: #1 [DOI: 10.1109/CVPR.2004.1315098]
- Govindu V M. 2006. Robustness in motion averaging//Proceedings of the 7th Asian Conference on Computer Vision-ACCV 2006. Hyderabad, India: Springer: 457-466 [DOI: 10.1007/11612704_46]
- Govindu V M and Pooja A. 2014. On averaging Multiview relations for 3D scan registration. *IEEE Transactions on Image Processing*, 23(3): 1289-1302 [DOI: 10.1109/TIP.2013.2246517]
- Kasten Y, Geifman A, Galun M and Basri R. 2019. Algebraic Characterization of Essential Matrices and Their Averaging in Multiview Settings//Proceedings of 2019 IEEE International Conference on Computer Vision. Seoul, Korea (South): IEEE: 5894-5902 [DOI: 10.1109/ICCV.2019.00599]
- Hartley R I and Sturm P. 1997. Triangulation. *Computer Vision and Image Understanding*, 68(2): 146-157 [DOI: 10.1006/cviu.1997.0547]
- Jancosek M and Pajdla T. 2011. Multi-view reconstruction preserving weakly-supported surfaces//Proceedings of 2011 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Colorado

- Springs, USA: IEEE: 1362-1376 [DOI: 10.1109/cvpr.2011.5995693]
- Jiang N J, Cui Z P and Tan P. 2013. A global linear method for camera pose registration//Proceedings of 2013 IEEE International Conference on Computer Vision. Sydney, Australia: IEEE: 481-488 [DOI: 10.1109/ICCV.2013.66]
- Lee S H and Civera J. 2019. Closed-form optimal two-view triangulation based on angular errors//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Seoul, Korea (South) : IEEE: 2681-2689 [DOI: 10.1109/ICCV.2019.00277]
- Lee S H and Civera J. 2022. HARA: a hierarchical approach for robust rotation averaging//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 15777-15786 [DOI: 10.1109/CVPR52688.2022.01532]
- Li H, Cui Z P, Liu S C and Tan P. 2022. RAGO: recurrent graph optimizer for multiple rotation averaging//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 15787-15796 [DOI: 10.1109/CVPR52688.2022.01533]
- Lowe D G. 2004. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2) : 91-110 [DOI: 10.1023/B:VISI.0000029664.99615.94]
- Manam L and Govindu V M. 2022. Correspondence reweighted translation averaging//Proceedings of the 17th European Conference on Computer Vision-ECCV 2022. Tel Aviv, Israel: Springer: 56-72 [DOI: 10.1007/978-3-031-19827-4_4]
- Moulon P, Monasse P and Marlet R. 2012. Adaptive structure from motion with a Contrario model estimation//Proceedings of the 11th Asian Conference on Computer Vision. Daejeon, Korea (South) : Springer: 257-270 [DOI: 10.1007/978-3-642-37447-0_20]
- Moulon P, Monasse P and Marlet R. 2013. Global fusion of relative motions for robust, accurate and scalable structure from motion//Proceedings of 2013 IEEE International Conference on Computer Vision. Sydney, Australia: IEEE: 3248-3255 [DOI: 10.1109/ICCV.2013.403]
- Muja M and Lowe D G. 2014. Scalable nearest neighbor algorithms for high dimensional data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(11) : 2227-2240 [DOI: 10.1109/TPAMI.2014.2321376]
- Ning X J, Gong L, Han Y, Ma T, Shi Z H, Jin H Y and Wang Y H. 2023. Semantic segmentation and model matching-integrated indoor scenario-relevant reconstruction method. *Journal of Image and Graphics*, 28(10) : 3149-3162 (宁小娟, 巩亮, 韩怡, 马婷, 石争浩, 金海燕, 王映辉. 2023. 结合语义分割与模型匹配的室内场景重建方法. *中国图象图形学报*, 28(10) : 3149-3162) [DOI: 10.11834/jig.220518]
- Nister D. 2004. An efficient solution to the five-point relative pose problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(6) : 756-770 [DOI: 10.1109/TPAMI.2004.17]
- Özyeşil O and Singer A. 2015. Robust camera location estimation by convex programming//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE: 2674-2683 [DOI: 10.1109/CVPR.2015.7298883]
- Purkait P, Chin T J and Reid I. 2020. NeuRoRA: neural robust rotation averaging//Proceedings of the 16th European Conference on Computer Vision-ECCV 2020. Glasgow, UK: Springer: 137-154 [DOI: 10.1007/978-3-030-58586-0_9]
- Ren F E, Liu T and Yang L. 2021. 3D reconstruction of a single plant leaf image. *Journal of Image and Graphics*, 26(11) : 2713-2722 (任非儿, 刘通, 杨龙. 2021. 单幅植物叶片图像的3维重建. *中国图象图形学报*, 26(11) : 2713-2722) [DOI: 10.11834/jig.200475]
- Schönberger J L and Frahm J M. 2016. Structure-from-motion revisited//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 4104-4113 [DOI: 10.1109/CVPR.2016.445]
- Shah R, Deshpande A and Narayanan P J. 2014. Multistage SFM: revisiting incremental structure from motion//Proceedings of the 2nd International Conference on 3D Vision. Tokyo, Japan: IEEE: 417-424 [DOI: 10.1109/3DV.2014.95]
- Shi Y P, Wyeth C and Lerman G. 2022. Robust group synchronization via quadratic programming//Proceedings of 2022 International Conference on Machine Learning. Beijing, China: ICML: 20095-20105
- Snaveley N, Seitz S M and Szeliski R. 2006. Photo tourism: exploring photo collections in 3D. *ACM Transactions on Graphics*, 25(3) : 835-846 [DOI: 10.1145/1141911.1141964]
- Snaveley N, Seitz S M and Szeliski R. 2008. Modeling the world from Internet photo collections. *International Journal of Computer Vision*, 80(2) : 189-210 [DOI: 10.1007/s11263-007-0107-3]
- Sweeney C, Sattler T, Höllerer T, Turk M and Pollefeys M. 2015. Optimizing the viewing graph for structure-from-motion//Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago, Chile: IEEE: 801-809 [DOI: 10.1109/ICCV.2015.98]
- Waechter M, Moehle N and Goesele M. 2014. Let there be color! large-scale texturing of 3D Reconstructions//Proceedings of the 13th European Conference on Computer Vision. Zürich, Switzerland: Springer: 836-850 [DOI: 10.1007/978-3-319-10602-1_54]
- Wang K Y, Huang S R and Li Y S. 2022. An optical reconstruction based underwater image analysis. *Journal of Image and Graphics*, 27(5) : 1337-1358 (王柯俨, 黄诗芮, 李云松. 2022. 水下光学图像重建方法研究进展. *中国图象图形学报*, 27(5) : 1337-1358) [DOI: 10.11834/jig.210819]
- Wilson K and Snaveley N. 2014. Robust global translations with 1DSfM//Proceedings of the 13th European Conference on Computer Vision. Zürich, Switzerland: Springer: 61-75 [DOI: 10.1007/978-3-319-10578-9_5]
- Wu C C. 2013. Towards linear-time incremental structure from motion//

- Proceedings of 2013 IEEE International Conference on 3D Vision. Seattle, USA: IEEE: 127-134 [DOI: 10.1109/3DV.2013.25]
- Yan S, Zhang M J, Fan Y C, Tan X H, Liu Y, Peng Y and Liu Y X. 2021. Progress in the large-scale outdoor image 3D reconstruction. *Journal of Image and Graphics*, 26(6): 1429-1449 (颜深, 张茂军, 樊亚春, 谭小慧, 刘煜, 彭杨, 刘宇翔. 2021. 大规模室外图像 3 维重建技术研究进展. *中国图象图形学报*, 26(6): 1429-1449) [DOI: 10.11834/jig.200842]
- Yao L, Zhang Y A, Zhang M X and Wan Y. 2021. The impact of joint axis angle prior on the results of 3D human body reconstruction. *Journal of Image and Graphics*, 26(12): 2918-2930 (姚砺, 张幼安, 张梦雪, 万燕. 2021. 关节轴角先验对 3 维人体重建结果的影响. *中国图象图形学报*, 26(12): 2918-2930) [DOI: 10.11834/jig.200348]
- Zhang G L, Larsson V and Barath D. 2023. Revisiting rotation averaging: uncertainties and robust losses//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE: 17215-17224 [DOI: 10.1109/CVPR52729.2023.01651]
- Zhao J. 2022. An efficient solution to non-minimal case essential matrix estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(4): 1777-1792 [DOI: 10.1109/TPAMI.2020.3030161]
- Zhu S Y, Zhang R Z, Zhou L, Shen T W, Fang T, Tan P and Quan L. 2018. Very large-scale global SfM by distributed motion averaging//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 4568-4577 [DOI: 10.1109/CVPR.2018.00480]
- Zhuang B B, Cheong L F and Lee G H. 2018. Baseline desensitizing in translation averaging//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 4539-4547 [DOI: 10.1109/CVPR.2018.00477]

作者简介

李梦晗,女,硕士研究生,主要研究方向为全局式从运动恢复结构。E-mail: limenghan@stu.ouc.edu.cn

高翔,通信作者,男,副研究员,主要研究方向为大规模从运动恢复结构和基于图像的场景三维建模。

E-mail: xiang.gao@ia.ac.cn

解则晓,男,教授,主要研究方向为机器视觉与视觉测量、精密激光扫描测量技术、水下三维视觉测量与三维重建技术。

E-mail: xiezexiao@ouc.edu.cn

申抒含,男,研究员,主要研究方向为三维计算机视觉、大规模场景三维重建和场景三维语义理解。

E-mail: shshen@nlpr.ia.ac.cn