

分类号\_\_\_\_\_

密级\_\_\_\_\_

U D C\_\_\_\_\_

编号\_\_\_\_\_

# 中国海洋大学

## 博士后研究工作报告

基于图像的大规模场景三维重建

---

中的运动平均化方法研究

---

高 翔

工作完成日期 2019年8月—2022年2月

---

报告提交日期 2022年4月

---

中国海洋大学（山东）

2022年4月



基于图像的大规模场景三维重建  
中的运动平均化方法研究

MOTION AVERAGING IN IMAGE-BASED  
LARGE-SCALE SCENE 3D RECONSTRUCTION

博 士 后 姓 名 高 翔

流动站（一级学科）名称 中国海洋大学水利工程

专 业（二级学科）名称 模式识别与智能系统

研究工作起始时间 2019 年 8 月 30 日

研究工作期满时间 2022 年 2 月 28 日

中国海洋大学人事处（山东）

2022 年 4 月



## 内 容 摘 要

运动平均化是全局式从运动恢复结构技术中的核心步骤，其将匹配图像对之间的相对旋转与平移作为输入，输出相机的绝对位姿，具有误差均摊、求解高效等特点。然而，图像特征误匹配导致的相对位姿测量外值以及在不同问题规模下的外极几何图整体化处理，对现有运动平均化方法在高效性与鲁棒性方面提出了挑战。对此，本文在大规模运动平均化的高效性与鲁棒性问题研究方面开展了一些初步探索，将增量式参数估计，随机抽样一致性，分而治之策略等思想融入到运动平均化过程，在滤除相对位姿测量外值的同时估计相机的绝对位姿，以期实现简单、高效、精确、鲁棒的大规模运动平均化。本文的工作主要包括如下三方面：

1. 为实现更加准确、鲁棒的绝对旋转估计，受已广泛研究的增量式从运动恢复结构方法启发，提出了一种新型的旋转平均化方法，增量式旋转平均化。与传统的旋转平均化方法相比，该方法不再同时估计所有相机的绝对旋转，而是采用增量式参数估计流程，依次估计各个相机的绝对旋转。另外，为进一步提升其精度与鲁棒性，在旋转平均化方法中还引入了若干同样受增量式从运动恢复结构方法启发的关键技术，包括初始三元组与下一最优视图的选取策略，以及迭代过程中采用的加权局部、全局优化，重旋转平均化技术等。在实验验证环节，通过参数设置实验，对关键技术消融实验，以及对比实验，验证了该方法的有效性。
2. 针对现有旋转平均化方法大都是在整个外极几何图上进行绝对旋转估计，效率较低，且在大规模与高噪声情形下的鲁棒性较差的问题，基于分治策略提出了一种新型旋转平均化流程：层级式旋转平均化。基于该流程，可将原始的单个规模更大、难度更高的旋转平均化问题转化为多个规模更小、难度更低的子问题。在此基础上，该流程可以分别求解各子问题并将求解结果融合以获取原始问题的解。另外，将两种现有旋转平均化方法融入到上述流程之中形成了两种新型旋转平均化算法。实验验证环节对上述层级式旋转平均化流程以及基于此流程构建的两种旋转平均化方法进行了测评，验证了其有效性。

3. 由于尺度歧义性、估计敏感性、以及求解不确定性，平移平均化问题相对于旋转平均化问题更加困难。为求解该问题，与增量式平均化类似，基于增量式参数估计思想，提出了一种新型的增量式平移平均化方法。该方法结合了增量式参数估计流程的精度高与鲁棒性强以及全局式运动平均化方法的简洁与高效的优点。与传统平移平均化同时估计所有相机绝对位置的方式不同，增量式平移平均化采用增量式计算的方式对其进行估计。该方法既能对测量外值鲁棒，实现参数的精确估计，又因其较少依赖其它鲁棒性操作，更加简洁与高效。在实验验证环节，通过参数设置实验，对关键技术的消融实验，以及对比实验，验证了该方法的有效性。

关键词：全局式从运动恢复结构，旋转平均化，平移平均化，增量式参数估计，分治策略

## Abstract

Motion averaging is a key step in global Structure from Motion (SfM). It takes relative rotation and translation of matched image pairs as input, and outputs the absolute orientations and locations of cameras, which features of uniform error distribution and high solution efficiency. However, the relative pose measurement outlier caused by image feature mis-matching and the Epipolar-geometry Graph (EG) integrative processing of different problem scales challenge the existing motion averaging methods in terms of efficiency and robustness. To deal with this issue, this thesis conducts some initial explorations in the efficiency and robustness problems of large-scale motion averaging, which integrates several thoughts, such as incremental parameter estimation, RANdom SAmples Consensus (RANSAC), and divide-and-conquer strategy, into the motion averaging process, to simultaneously filter relative pose measurement outliers and estimate absolute camera poses, in order to achieve simple, efficient, accurate, and robust large-scale motion averaging. The three main works of this thesis are summarized as follows:

- In order to achieve more accurate and robust estimation of absolute rotations, inspired by the extensively studied incremental SfM methods, a novel rotation averaging method, incremental rotation averaging, is proposed. Compared with the traditional rotation averaging methods, it no longer estimates all the cameras' absolute rotations simultaneously, but adopts the incremental parameter estimation pipeline, to estimate the absolute rotation of each camera individually. In addition, in order to improve its accuracy and robustness further, several key techniques which are also inspired by incremental SfM methods, including initial triplet and next-best view selection strategies and weighted local or

global optimization and re-rotation averaging during the iteration process, are involved into the incremental rotation averaging. In the experimental verification section, the effectiveness of the proposed method is verified by parameter setting experiments, ablation studies of key techniques, and comparative experiments.

- Aiming at the efficiency problem caused by performing absolute rotation estimation in the entire EG and the robustness problem during large scale and high noise situations of the existing rotation averaging methods, a novel rotation averaging pipeline, hierarchical rotation averaging, is proposed based on divide-and-conquer strategy. This novel pipeline could convert the original single larger and more difficult rotation averaging problem into several smaller and easier sub-problems. On this basis, this pipeline could solve each sub-problem respectively and integrate their results to obtain the solution of the original problem. In addition, two novel rotation averaging methods are brought up by integrating two existing rotation averaging methods into the proposed novel pipeline. The experimental verification section evaluates the above rotation averaging pipeline and the two rotation averaging methods constructed based on the pipeline, to demonstrate their effectiveness.
- Translation averaging is more difficult than rotation averaging due to scale ambiguity, estimation sensitivity, and solution uncertainty. To tackle this tough problem, similar to the method of incremental rotation averaging, a novel translation averaging method, termed as incremental translation averaging, is proposed based on the thought of incremental parameter estimation. It combines the advantages of high accuracy and strong robustness in incremental parameter estimation pipeline and those of high simplicity and efficiency in global motion averaging approach. Unlike the traditional translation averaging methods which estimate all the absolute camera locations simultaneously, the incremental translation averaging method computes them in an incremental way. The proposed

method is robust to measurement outliers and accurate in parameter estimation, and is simple and efficient because of its less dependency on additional robust operations. In the experimental verification section, the effectiveness of the proposed method is verified by parameter setting experiments, ablation studies of key techniques, and comparative experiments.

Keywords: Global structure from motion, rotation averaging, translation averaging, incremental parameter estimation, divide-and-conquer strategy



# 目 次

1	绪论	1
1.1	研究背景	1
1.2	研究现状	4
1.2.1	旋转平均化	4
1.2.2	平移平均化	7
1.3	预备知识	9
1.3.1	旋转平均化	9
1.3.2	平移平均化	9
1.4	评测基准	11
1.4.1	评测环境	11
1.4.2	评测数据	11
1.4.3	评测方法	14
1.5	内容安排	15
1.5.1	研究内容	16
1.5.2	结构安排	16
2	增量式旋转平均化	17
2.1	引言	17
2.2	方法介绍	17
2.2.1	基于局部优化的初始三元组选取	18
2.2.2	基于加权支撑集合的下一最优视图选取	19
2.2.3	加权局部/全局优化与重旋转平均化	20
2.3	实验验证	21
2.3.1	参数设置测评	22
2.3.2	消融实验	23
2.3.3	种子敏感度验证	25
2.3.4	对比实验	26
2.4	本章小结	30
3	层级式旋转平均化	31
3.1	引言	31

3.2	方法介绍	32
3.2.1	基于受限社区发现的外极几何图聚类	33
3.2.2	类内子图旋转平均	33
3.2.3	基于投票单旋转平均的类间相对旋转估计	34
3.2.4	类间子图旋转平均	35
3.2.5	旋转全局对齐优化	35
3.3	实验验证	36
3.3.1	图结构对比实验	36
3.3.2	HRRA 验证实验	37
3.3.3	IRA++验证实验	39
3.4	本章小结	42
4	增量式平移平均化	45
4.1	引言	45
4.2	方法介绍	46
4.2.1	基于局部优化的初始四元组选取	46
4.2.2	基于加权支撑集合的下一最优视图选取	47
4.2.3	加权局部/全局优化与重平移平均化	49
4.3	实验验证	50
4.3.1	参数设置测评	50
4.3.2	消融实验	51
4.3.3	种子敏感度验证	53
4.3.4	对比实验	54
4.4	本章小结	57
5	总结与展望	59
5.1	总结	59
5.2	展望	59
	参考文献	61
	致谢	65
	博士生期间发表的学术论文, 专著	67
	博士后期间发表的学术论文, 专著	69
	个人简历	71
	永久通信地址	73

# 1. 绪论

## 1.1 研究背景

基于图像的大规模场景三维重建流程主要包括稀疏重建<sup>[1]</sup>、稠密重建<sup>[2]</sup>、表面重建<sup>[3]</sup>、纹理映射<sup>[4]</sup>四个部分（图 1-1）。作为稀疏重建背后的关键技术，从运动恢复结构（**Structure from Motion, SfM**）<sup>[1,5~9]</sup>是整个重建流程的核心，其在遥感测绘<sup>[10]</sup>、室内建模<sup>[11]</sup>、增强现实<sup>[12]</sup>、古建保护<sup>[13]</sup>等领域有着重要的应用（图 1-2）。SfM 将采集的图像集合作为输入，输出相机的内参数（焦距、主点等），绝对位姿（包括位置与朝向）以及由稀疏空间点表示的场景三维结构（图 1-3<sup>[14]</sup>）。

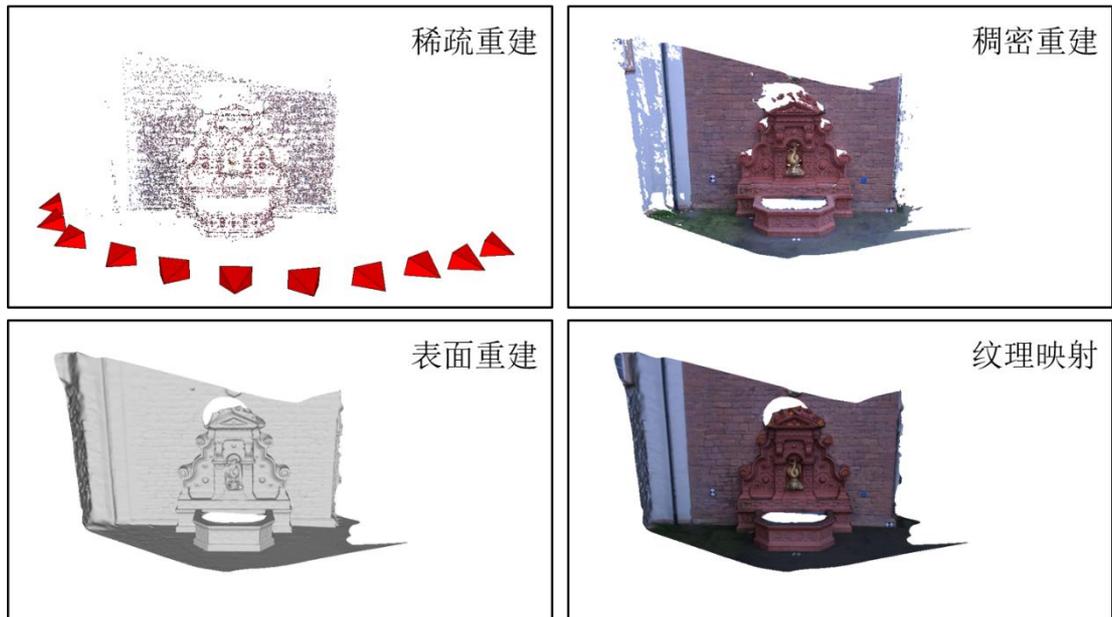


图 1-1 基于图像的大规模场景三维重建流程

SfM 技术主要包括对应关系获取以及相机位姿初始化两部分。其中，对应关系获取通常包括基于局部纹理的特征提取<sup>[15~16]</sup>，基于词袋（**Bog of Word, BoW**）<sup>[17]</sup>模型与近似最近邻（**Approximate Nearest Neighbor, ANN**）<sup>[18]</sup>的特征匹配，以及基于基本矩阵（**fundamental matrix**）<sup>[19]</sup>与本质矩阵（**essential matrix**）<sup>[20]</sup>估计的几何验证。相关的理论与技术研究较为成熟，流程较为固定。

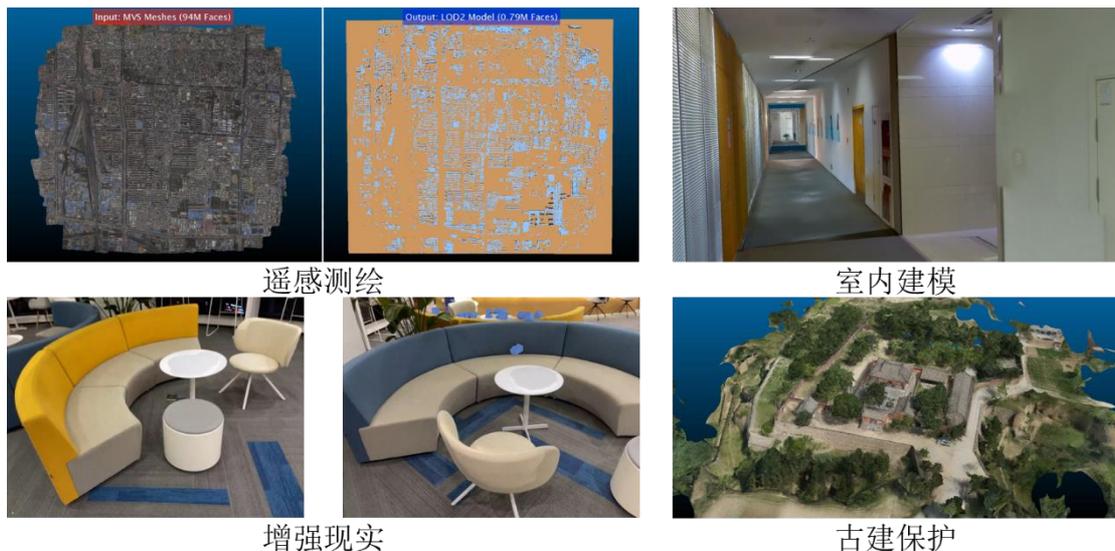


图 1-2 SfM 的主要应用领域举例

对于 SfM 中的相机位姿初始化方式，其大致可分为增量式<sup>[1, 5-6]</sup>与全局式<sup>[7-9]</sup>两类。增量式 SfM 通过迭代式地进行相机位姿估计与场景结构扩展实现相机绝对位姿的初始化。在每次迭代过程中，增量式 SfM 会分别引入随机抽样一致性（RANdomSAmple Consensus, RANSAC）<sup>[21]</sup>以及捆绑调整（BundlE AdjustmEnt, BA）<sup>[22]</sup>技术以应对不可避免的特征误匹配以及不够精确的相机位姿与场景结构估计值，其常见流程如图 1-4 所示。对于全局式 SfM，相机的绝对位姿通常采用运动平均化（motion averaging）技术进行初始化，在此基础上仅通过一次性多视图三角测量（triangulation）及全局捆绑调整对相机参数与场景结构进行优化，其常见流程如图 1-5 所示。相较于全局式 SfM，尽管效率较低，由于在其中频繁使用基于 RANSAC 的模型估计以及基于 BA 的参数优化，通常认为增量式 SfM 的精度与鲁棒性更好。

运动平均化包括旋转平均化（rotation averaging）与平移平均化（translation averaging）两步，其任务分别是通过给定的两两相机之间的相对旋转与相对平移测量值，估计各相机的绝对旋转与位置。其中相对旋转与平移的测量值通常通过本质矩阵的估计与分解<sup>[20]</sup>获取。绝大部分的全局式 SfM 方法<sup>[23~24]</sup>在进行相机位姿初始化时采用依次进行旋转平均化与平移平均化的方式，并且为简化问题，在进行平移平均化时通常将由旋转平均化得到的相机绝对旋转固定并引入其中。



图 1-3 SfM 的输入与输出举例

然而，由于外极几何图（Epipolar-geometry Graph, EG）中的相对旋转存在因特征误匹配导致的测量外值，旋转平均化任务本身并不容易。这种情况在从因特网上下载得到的图像集合中更为显著<sup>[1]</sup>。为应对该情况，现有旋转平均化方法或力图设计鲁棒的损失函数，使得估计绝对旋转时的优化过程更加鲁棒<sup>[7, 25~28]</sup>，或尝试开发有效的外值滤除策略以净化被外值污染的外极几何图<sup>[29~31]</sup>。然而，上述方法在精度与鲁棒性等方面仍存在不足。

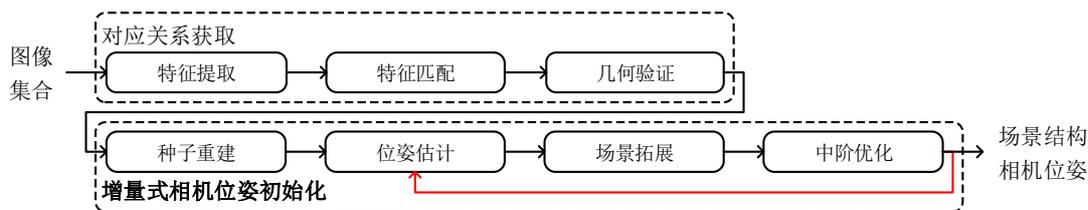


图 1-4 增量式 SfM 常见流程图

对于平移平均化，通常认为相对于旋转平均化而言其难度更大，原因有如下三点：1）不同于相对旋转，对于相对平移，本质矩阵中仅包含其方向信息，即获取的相对平移具有尺度不确定性<sup>[20]</sup>；2）相对于相对旋转，相对平移的估计精度更容易受特征误匹配的影响<sup>[20]</sup>；3）对于平移平均化方法，仅可唯一确定处于同一平行刚性分支中的相机绝对位置<sup>[32]</sup>，而旋转平均化方法可唯一确定处于同一连通分支中的相机绝对旋转。针对上述问题，现有平移平均化方法主要聚焦于：

1) 设计恰当的损失函数形式及其优化方式<sup>[32~35]</sup>, 2) 开发外极几何图的滤除或优化策略<sup>[9, 24]</sup>, 3) 尝试借助一些其他辅助信息, 例如特征点轨迹<sup>[23, 36]</sup>、相机三元组<sup>[8, 37]</sup>、矩阵秩约束<sup>[38~39]</sup>等。然而, 对于复杂损失形式及其优化方法, 预处理操作, 或附加信息等的过多依赖使得上述方法较为复杂且效率较低, 并且也在精度与鲁棒性等方面存在不足。

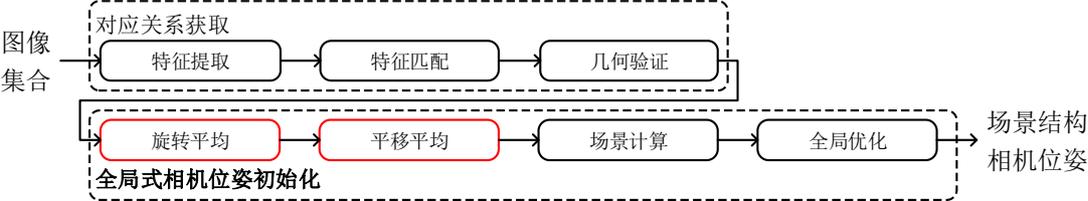


图 1-5 全局式 SfM 常见流程图

另外, 现有的旋转平均化与平移平均化方法大多将外极几何图看作一个整体进行处理。然而随着待重建场景规模的增大, 其外极几何图中的顶点与边的数量也相应变大, 这会导致现有方法效率明显下降。与此同时, 在规模更大的外极几何图上更容易出现顶点的度 (degree) 差异较大, 即不同顶点连接紧密程度差异较大的情况, 这会使得外极几何图的运动平均化可解性变差<sup>[40]</sup>。

针对上述在现有旋转与平移平均化方法中存在的问题, 本文在大规模运动平均化的高效性与鲁棒性问题研究方面开展了一些初步探索, 将增量式参数估计, 随机抽样一致性, 分而治之策略等思想融入到运动平均化过程, 在滤除相机的相对旋转、平移测量外值的同时估计相机的绝对朝向与位置, 以期实现简单、高效、精确、鲁棒的大规模运动 (旋转、平移) 平均化。

## 1.2 研究现状

本节对现有的具有代表性的旋转平均化与平移平均化方法分别进行了概述。

### 1.2.1 旋转平均化

现有的旋转平均化方法主要包括基于鲁棒损失的, 基于外值滤除的, 基于增量估计的, 以及基于深度学习的方法四类, 分别概述如下。

#### (1) 基于鲁棒损失的旋转平均化

基于鲁棒损失的旋转平均化方法<sup>[7, 25~28]</sup>尝试设计鲁棒损失函数并将其引入到参数优化过程,以实现在相对旋转测量外值存在的情况下的绝对旋转的鲁棒估计。Hartley 等人<sup>[25]</sup>在 $\ell_1$ 范数意义下对旋转平均化问题进行了研究,他们提出了一种基于 Weiszfeld 算法的旋转平均化方法。在参数估计的过程中,他们采取了一种分布式平均化的方法。Crandall 等人<sup>[26]</sup>提出了一种两步式旋转平均化方法。在第一步中,他们采用基于 Markov 随机场的离散置信传播方法对绝对旋转进行初始化。在第二步中,他们采用基于 Levenberg-Marquardt 算法的非线性优化方法对绝对旋转进行调优。Chatterjee 与 Govindu<sup>[7]</sup>也提出了一种两步式旋转平均化方法,他们首先将 $\ell_1$ 范数下的最优解作为绝对旋转初值,然后利用基于 Geman-McClure 损失的迭代重加权最小二乘 (Iteratively Reweighted Least Squares, IRLS) 算法对该初值进行优化以实现高效、鲁棒的绝对旋转求解。之后,上述作者<sup>[27]</sup>又对其前述工作<sup>[7]</sup>进行了扩展。在其扩展工作中,第一步的绝对旋转初始化与前述工作一致,对于第二步基于 IRLS 的绝对旋转优化,作者对不同的鲁棒损失函数进行了较为详尽的测评,最终得出的结论为 $\ell_{\frac{1}{2}}$ 损失函数性能最优。Shi 等人<sup>[28]</sup>提出了一种新型鲁棒优化框架,消息传播最小二乘 (Message Passing Least Squares, MPLS)。作者将其用作对于旋转平均化任务中的 IRLS 的替代品,并对其相对于 IRLS 的优越性进行了理论分析。尽管上述方法在一定程度上对相对旋转测量外值较为鲁棒,这些方法中的优化方式对初始化效果依赖较为严重,并且由于其相对复杂的问题定义形式,效率通常较低。

## (2) 基于外值滤除的旋转平均化

相对于上述基于鲁棒损失的方法,基于外值滤除的旋转平均化方法<sup>[29~31]</sup>在原理上更加直观。此类方法通常遵循先清理后优化的策略,即先探测并滤除原始外极几何图上的相对旋转测量外值,后利用现有的旋转平均化方法处理净化后的外极几何图,实现绝对旋转的估计。Govindu<sup>[29]</sup>提出了一个基于 RANSAC 的旋转平均化方法,该方法采用随机深度优先搜索的方式在外极几何图上产生随机生成树,并将随机生成树用于绝对旋转的初始化以及相对旋转的外值探测。Zach 等人<sup>[30]</sup>提出了一种用于相对旋转测量外值滤除的回环一致性约束。该工作对外极几何图中多个回路的统计信息进行了收集,并将收集到的

统计信息送入置信传播公式中实现测量外值的识别。Cui 等人<sup>[31]</sup>提出了一个基于多棵正交最大生成树（Orthogonal Maximum Spanning Trees, OMSTs）的旋转平均化方法。基于从统计学角度的对相对旋转测量值的精度与特征点匹配对数呈正相关的观察，作者通过特征点匹配对数对原始的外极几何图进行加权，并从中选取了多棵 OMSTs，以获取连接足够紧密并且精度相对更高的相对旋转集合用于后续的旋转平均化。然而，相对旋转外值滤除与绝对旋转估计是一个“鸡生蛋还是蛋生鸡”的问题，这是因为当且仅当已经对绝对旋转进行了精确估计后才能实现对相对旋转测量外值的有效滤除。

### (3) 基于增量估计的旋转平均化

在旋转平均化方法中，还存在一类基于增量估计的方法<sup>[41~42]</sup>，该类方法通常集成于同时定位与构图（Simultaneous Localization And Mapping, SLAM）系统之中。Bustos 等人<sup>[41]</sup>对其 SLAM 系统中的各关键帧迭代进行局部旋转平均化，并且当检测到回环时进行全局旋转平均化。该方法中的局部与全局平均化均采用方法[7]实现。Chng 等人<sup>[42]</sup>提出的方法在原理上与方法[41]类似，两者的主要区别在于在进行局部旋转平均化时，方法[42]估计的绝对旋转不再锚定于当前局部外极几何图的第一个顶点，而是锚定于多个相邻顶点上以缓解链式效应。然而，上述两方法均未借助中间结果进行外值探测与滤除，并且它们频繁调用[7]中的较为复杂的优化方法致使其效率较低。另外，上述两方法的基本假设使得它们仅适用有序图像集合，而无法处理无序图像集合，如 1DSfM 数据集<sup>[9]</sup>。

### (4) 基于深度学习的旋转平均化

近年来，借助于深度神经网络强大的推断能力，出现了若干基于深度学习的旋转平均化方法<sup>[43~44]</sup>。Purkait 等人<sup>[43]</sup>提出了 NeuRoRA，一种基于图神经网络的两步式旋转平均化方法。其第一步是通过一个外极几何图清理网络探测相对旋转测量外值，第二步是采用一个微调网络来优化相机绝对旋转。然而，上述网络并非可以端到端训练的，并且这种分步式方法存在步骤二对步骤一的初始化结果非常敏感的情况。Yang 等人<sup>[44]</sup>提出了一种可以端到端训练的旋转平均化神经网络。该方法设计了一种多源传播器（Multi-Source Propagator, MSP），可以通过将图卷积迭代式应用于输入的外极几何图来计算相机绝对

旋转的初值。另外，该方法还设计了一个可以利用图像内容与特征点对应关系的外形几何融合网络，用于评测外极几何图中各条边的质量，以更好地实现基于 MSP 的绝对旋转初始化。然而，基于深度学习的方法对训练数据依赖较为严重，使得此类方法在面对全新测试环境时泛化性与鲁棒性相对较差。

## 1.2.2 平移平均化

基于依赖的不同鲁棒性操作，现有的平移平均化主要包括基于鲁棒优化的，基于预先处理的，以及基于信息增强的方法三类。值得注意的是，本节讨论的绝大多数方法均包含不止一种鲁棒性操作，而此处仅讨论其最有代表性的。

### (1) 基于鲁棒优化的平移平均化

基于鲁棒优化的平移平均化方法<sup>[32~35]</sup>尝试设计合适的损失函数形式以及对应的优化方式以使得此类方法对相对平移测量值更加鲁棒。Govindu<sup>[33]</sup>提出了一种基于最小化相对平移测量值与根据绝对平移估计值反算的相对平移估计值之间向量积的平移平均化方法，该方法的目标函数最小化的是代数误差而非几何误差。并且，该方法采用了一种专门设计的迭代重加权方式以期减小因基线长度的差别导致的对相机绝对位置估计结果的影响。LUD<sup>[32]</sup>与 ShapeFit/Kick<sup>[34]</sup>算法均设计了基于端点距离的损失函数，其中 LUD<sup>[32]</sup>采用的是最小化松弛端点距离而 ShapeFit/Kick<sup>[34]</sup>采用的是最小化相对平移估计值到相对平移测量值生成空间的正交补（the orthogonal complement of the span）的投影模长。上述两种算法分别采用 IRLS 与交替方向乘子法（Alternating Direction Method of Multipliers, ADMM）进行求解。通过观察到基于角度的损失函数对相对平移测量外值更加鲁棒，Zhuang 等人<sup>[35]</sup>提出了 BATA 算法，该算法也通过 IRLS 进行求解并且可以取得更高精度的估计结果。尽管上述方法对测量外值相对较为鲁棒，其中定义的损失函数形式通常较为复杂，这会导致这类方法的优化方式也相应较为复杂。另外，这类方法通常要求进行预处理<sup>[32, 34]</sup>或者初始化<sup>[35]</sup>，否则产生的估计结果精度会较差。

### (2) 基于预先处理的平移平均化

基于预先处理的平移平均化方法<sup>[9, 24]</sup>企图探测并滤除相对平移

测量外值<sup>[9]</sup>或者对外极几何图进行优化使其近似一致<sup>[24]</sup>。通过引入上述预处理步骤，后续的相机绝对位置估计步骤变得更加简单、直接。Wilson 等人<sup>[9]</sup>提出了 1DSfM 算法，用于相对平移测量值外值滤除。该方法的基本思想是通过将相对平移投影至若干所谓的一维方向上，而在此若干次投影过程中，相对平移测量值外值并不会一直与其他相对平移测量值在端点投影顺序上保持一致。然后，上述外值滤除问题可归结为一个组合排序问题并通过最小反馈弧集（minimum feedback arc set）算法进行求解。Sweeney 等人<sup>[24]</sup>提出了一种用于全局式 SfM 的外极几何图优化方法。该方法采用基于三元组的图扩增方法构建了原始外极几何图的一个子图，并引入回环一致性约束对该子图进行了优化。然而，上述预处理操作以及它们用到的额外信息（方法[9, 24]中的特征点轨迹以及方法[24]中的相机三元组）会使得上述方法的简洁性与效率变差。另外，根据[9]与[24]中的报导，若不进行最终的全局捆绑调整，由于其中相对简单的优化方式<sup>[9, 24]</sup>以及不完全一致的外极几何图<sup>[24]</sup>，这两种方法精度相对较差。

### (3) 基于信息增强的平移平均化

基于信息增强的平移平均化尝试借助一些其他辅助信息，诸如特征点轨迹<sup>[23, 36]</sup>、相机三元组<sup>[8, 37]</sup>、矩阵秩约束<sup>[38~39]</sup>等，以获取更加精确的平移平均化结果。在[23]与[36]中，特征点轨迹被分别用于构建稀疏深度图与相机位置约束。而在[37]与[8]中，相机三元组被分别用于消除相机位置的尺度不确定性与优化相对平移。尽管上述操作提高了方法的鲁棒性，但对于基于特征点轨迹与相机三元组的方法分别引入了更多的参数与约束，从而使得上述方法的简洁性与效率变差。除了特征点轨迹与相机三元组，矩阵秩约束也是一种针对平移平均化问题的常用辅助信息。Kasten 等人<sup>[38]</sup>提出了一种本质矩阵平均化方法，该方法可以基于矩阵秩约束同时估计相机的绝对旋转与绝对位置。随后，该方法在工作[39]中进行了拓展，使其能应用于相机未标定以及直线运动的情形。尽管此类新型流程可以对旋转平均化与平移平均化问题进行同时求解，其推导与优化过程相对来讲均较为复杂，并且这类方法的估计精度还不足以彻底取代目前常见的先旋转后平移的运动平均化流程。

### 1.3 预备知识

本节对旋转平均化与平移平均化所用到的预备知识分别进行简述。对于旋转平均化与平移平均化问题，其主要输入均为外极几何图，记为 $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ ，该图中共有 $|\mathcal{V}|$ 个相机顶点以及 $|\mathcal{E}|$ 条外极几何边。

#### 1.3.1 旋转平均化

对于旋转平均化问题， $\mathcal{G}$ 中的每个顶点 $v_i \in \mathcal{V}$ 分别对应各相机的绝对旋转 $\mathbf{R}_i$ ，每条边 $e_{ij} \in \mathcal{E}$ 分别对应各匹配图像对之间的相对旋转测量值 $\mathbf{R}_{ij}$ 。在此基础上，可对旋转平均化问题进行定义：

$$\{\mathbf{R}_i^*\} = \arg \min \sum_{\substack{v_i, v_j \in \mathcal{V} \\ e_{ij} \in \mathcal{E}}} \rho \left( d^R(\mathbf{R}_{ij}, \mathbf{R}_j \mathbf{R}_i^T) \right)$$

其中， $\{\mathbf{R}_i^*\}$ 为最终估计的相机绝对旋转集合， $\rho(\cdot)$ 为用于鲁棒优化的损失函数， $d^R(\cdot, \cdot)$ 为用于计算两个旋转矩阵之间距离的距离函数。由于本文涉及到的旋转平均化算法对测量外值较为鲁棒，在此选用简单、常用的 $\ell_2$ 范数作为损失函数；另外，对于距离函数，此处选择绝大多数相关工作<sup>[7, 27, 31, 45]</sup>同样使用的角度距离函数 $d_\theta^R(\cdot, \cdot)$ ，其定义如下：

$$d_\theta^R(\mathbf{R}_1, \mathbf{R}_2) = \frac{\|\log(\mathbf{R}_2 \mathbf{R}_1^T)\|_F}{\sqrt{2}} = \cos^{-1} \frac{\text{tr}(\mathbf{R}_2 \mathbf{R}_1^T) - 1}{2}$$

其中， $\log(\mathbf{R}) = [\boldsymbol{\omega}]_\times$ ， $[\boldsymbol{\omega}]_\times$ 为向量 $\boldsymbol{\omega}$ 的反对称矩阵形式， $\boldsymbol{\omega} = \theta \mathbf{n}$ 为旋转矩阵的轴角表示<sup>[27]</sup>。基于上述距离函数，可对相对旋转 $\mathbf{R}_{ij}$ 进行内外值判：当条件 $d_\theta^R(\mathbf{R}_{ij}, \mathbf{R}_j^* \mathbf{R}_i^{*T}) < \theta_{th}^R$ 满足时， $\mathbf{R}_{ij}$ 为内值，其中 $\theta_{th}^R$ 为旋转角度距离阈值。根据上述描述，本文中的旋转平均化问题可进一步定义如下：

$$\{\mathbf{R}_i^*\} = \arg \min \sum_{\substack{v_i, v_j \in \mathcal{V} \\ e_{ij} \in \mathcal{E}}} d_\theta^R(\mathbf{R}_{ij}, \mathbf{R}_j \mathbf{R}_i^T)^2$$

#### 1.3.2 平移平均化

对于平移平均化问题， $\mathcal{G}$ 中的每个顶点 $v_i \in \mathcal{V}$ 分别对应各相机的绝对位置 $\mathbf{c}_i$ ，每条边 $e_{ij} \in \mathcal{E}$ 分别对应各匹配图像对之间的相对平移测

量值 $\mathbf{t}_{ij}$ 。另外需要注意的是，在进行平移平均化之前，需要根据由旋转平均化得到的绝对旋转估计值 $\{\mathbf{R}_i^*\}$ ，将定义于两视图局部坐标系下的相对平移 $\mathbf{t}_{ij}$ 变换到全局坐标系下，即 $\mathbf{c}_{ij}$ 。其计算方法定义如下：

$\mathbf{c}_{ij} = -\mathbf{R}_j^{*\top} \mathbf{t}_{ij}$ 。在此基础上，可定义本文中的旋转平均化问题：

$$\{\mathbf{c}_i^*\} = \arg \min \sum_{\substack{v_i, v_j \in \mathcal{V} \\ e_{ij} \in \mathcal{E}}} d_{ch}^t \left( \mathbf{c}_{ij}, \frac{\mathbf{c}_j - \mathbf{c}_i}{\|\mathbf{c}_j - \mathbf{c}_i\|_2} \right)^2$$

其中， $\{\mathbf{c}_i^*\}$ 为最终估计的相机绝对位置集合， $d_{ch}^t(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|_2$ 为 $S^2$ 空间上的弧距，其同样用于 1DSfM 算法<sup>[9]</sup>。另外，本文采用如下方法对 $\mathbf{c}_{ij}$ 进行内外值判别：当条件 $d_{\theta}^t \left( \mathbf{c}_{ij}, \frac{\mathbf{c}_j - \mathbf{c}_i}{\|\mathbf{c}_j - \mathbf{c}_i\|_2} \right) < \theta_{th}^t$ 满足时， $\mathbf{c}_{ij}$ 为内值，其中 $d_{\theta}^t(\mathbf{u}, \mathbf{v}) = \cos^{-1}(\mathbf{u} \cdot \mathbf{v})$ 为 $S^2$ 空间上的角度距离， $\theta_{th}^t$ 为平移角度距离阈值。

另外，不同于旋转平均化问题可通过简单链式计算( $\mathbf{R}_j = \mathbf{R}_{ij} \mathbf{R}_i$ )获取新的绝对旋转估计值，为对相机绝对位置进行序列式估计，需要引入相机三元组以消除相对平移中的尺度不确定性。本文采用的是已知绝对旋转的线性三焦张量算法<sup>[37]</sup>。具体来说，对于一个相机三元组 $\mathbf{t}_{ijk}$ 来说，给出其全部三个相对平移测量值 $\{\mathbf{c}_{ij}, \mathbf{c}_{ik}, \mathbf{c}_{jk}\}$ 以及三个绝对相机位置 $\{\mathbf{c}_i, \mathbf{c}_j, \mathbf{c}_k\}$ 中的任意两个，另外一个可以通过一种高效且几何意义上最优的方式进行计算。以 $\mathbf{c}_k$ 为例，其可通过下式对其进行计算：

$$\mathbf{c}_k = \frac{1}{2} (\mathbf{R}_i(\theta'_i) s_{ij}^{ik} (\mathbf{c}_j - \mathbf{c}_i) + \mathbf{R}_j(-\theta'_j) s_{ij}^{jk} (\mathbf{c}_i - \mathbf{c}_j) + \mathbf{c}_i + \mathbf{c}_j)$$

其中， $\mathbf{R}_i(\theta'_i)$ 为一个绕轴 $\mathbf{c}_{ij} \times \mathbf{c}_{ik}$ 逆时针旋转 $\theta'_i$ 角度的旋转矩阵， $\theta'_i$ 为 $\mathbf{c}_{ij}$ 与 $\mathbf{c}_{ik}$ 之间的夹角， $s_{ij}^{ik}$ 为基线长度 $\|\mathbf{c}_k - \mathbf{c}_i\|_2$ 与 $\|\mathbf{c}_j - \mathbf{c}_i\|_2$ 之间的比值，即 $s_{ij}^{ik} = \frac{\|\mathbf{c}_k - \mathbf{c}_i\|_2}{\|\mathbf{c}_j - \mathbf{c}_i\|_2} = \frac{\sin \theta'_j}{\sin \theta'_k}$ 。与之类似， $\mathbf{c}_i$ 与 $\mathbf{c}_j$ 也可分别通过下式进行计算：

$$\begin{aligned} \mathbf{c}_i &= \frac{1}{2} (\mathbf{R}_j(\theta'_j) s_{jk}^{ij} (\mathbf{c}_k - \mathbf{c}_j) + \mathbf{R}_k(-\theta'_k) s_{jk}^{ik} (\mathbf{c}_j - \mathbf{c}_k) + \mathbf{c}_j + \mathbf{c}_k) \\ \mathbf{c}_j &= \frac{1}{2} (\mathbf{R}_i(-\theta'_i) s_{ik}^{ij} (\mathbf{c}_k - \mathbf{c}_i) + \mathbf{R}_k(\theta'_k) s_{ij}^{jk} (\mathbf{c}_i - \mathbf{c}_k) + \mathbf{c}_i + \mathbf{c}_k) \end{aligned}$$

## 1.4 评测基准

本节对本文中所用到的评测环境，评测数据，以及基于此的评测方法进行简要介绍。

### 1.4.1 评测环境

本文提出的旋转平均化与平移平均化方法均在 Windows 10 平台下基于 Visual Studio 2019 开发，使用的 CPU 型号为 Intel I7-9750H、主频为 2.60GHz，设备内存为 32GB。

### 1.4.2 评测数据

本文使用的测评数据为 1DSfM 数据集<sup>[9]</sup>，该数据集包含 14 组测评数据，所有图像均从因特网上下载获得，每组数据对应一个建筑场景。各组数据匹配图像对之间的特征匹配对数以及相对旋转、平移测量值均有提供。另外，增量式 SfM 中的代表性方法 Bundler<sup>[1]</sup>的相机标定结果在 1DSfM 数据集中也有提供，在此将其相机绝对旋转与绝对位置的估计结果分别作为旋转平均化与平移平均化的评测真值。

表 1-1 1DSfM 数据集元数据

	$ V $	$ V_{GT} $	$ \mathcal{E} $	$\tilde{n}_{ij}$	$\bar{n}_{ij}$	$\tilde{r}_{ij}^R$	$\bar{r}_{ij}^R$	$\tilde{r}_{ij}^t$	$\bar{r}_{ij}^t$
ALM	627	577	97206	105	192	2.78°	9.09°	4.65°	18.80°
ELS	247	227	20297	106	160	2.89°	12.50°	8.75°	36.66°
GDM	742	677	48144	73	144	12.30°	33.33°	26.87°	52.65°
MDR	394	341	23784	61	128	9.34°	29.30°	13.46°	36.30°
MND	474	450	52424	180	310	1.67°	7.51°	3.33°	18.61°
NYC	376	332	20680	80	167	4.22°	14.14°	7.18°	28.86°
PDP	354	338	24710	87	128	1.81°	8.38°	3.07°	21.75°
PIC	2508	2152	319257	56	97	4.93°	19.09°	2.92°	7.71°
ROF	1134	1084	70187	65	188	2.97°	13.83°	4.01°	30.19°
TOL	508	472	23863	81	220	2.60°	11.58°	2.63°	19.92°
TFG	5433	5058	680012	71	109	3.01°	8.62°	6.56°	23.67°
USQ	930	789	25561	87	150	3.61°	9.02°	20.32°	43.22°
VNC	918	836	103550	229	408	2.59°	11.26°	4.24°	24.47°
YKM	458	437	27729	112	245	2.68°	11.16°	3.40°	21.20°

1DSfM 数据集中各测试数据的元数据如表 1-1 所示，各测试数据的名称缩写参照文献[27]，表格中的 $|V|$ ， $|\mathcal{E}|$ ，以及 $|V_{GT}|$ 分别表示外极几何图中的顶点与边的数量，以及带有真值的顶点数量。 $\tilde{n}_{ij}$ 与

$\bar{n}_{ij}$  分别表示各测试数据匹配图像对之间的特征匹配对数的中值与均值,  $\bar{r}_{ij}^R$  与  $\bar{r}_{ij}^r$  分别表示相对旋转测量值的误差中值与均值,  $\bar{r}_{ij}^t$  与  $\bar{r}_{ij}^T$  分别表示相对平移测量值的误差中值与均值。由表 1-1 可知: 1) 在大部分情形下, 相对旋转与相对平移测量值的精度均与特征匹配对数正相关; 2) 整体上讲, 相对旋转测量值的精度高于相对平移测量值。另外, 图 1-6 以及图 1-7 展示了各测试数据采用开源三维重建软件 COLMAP<sup>[6]</sup> 进行稀疏重建的定性结果。其中, 红色四棱锥代表估计的相机绝对位姿。

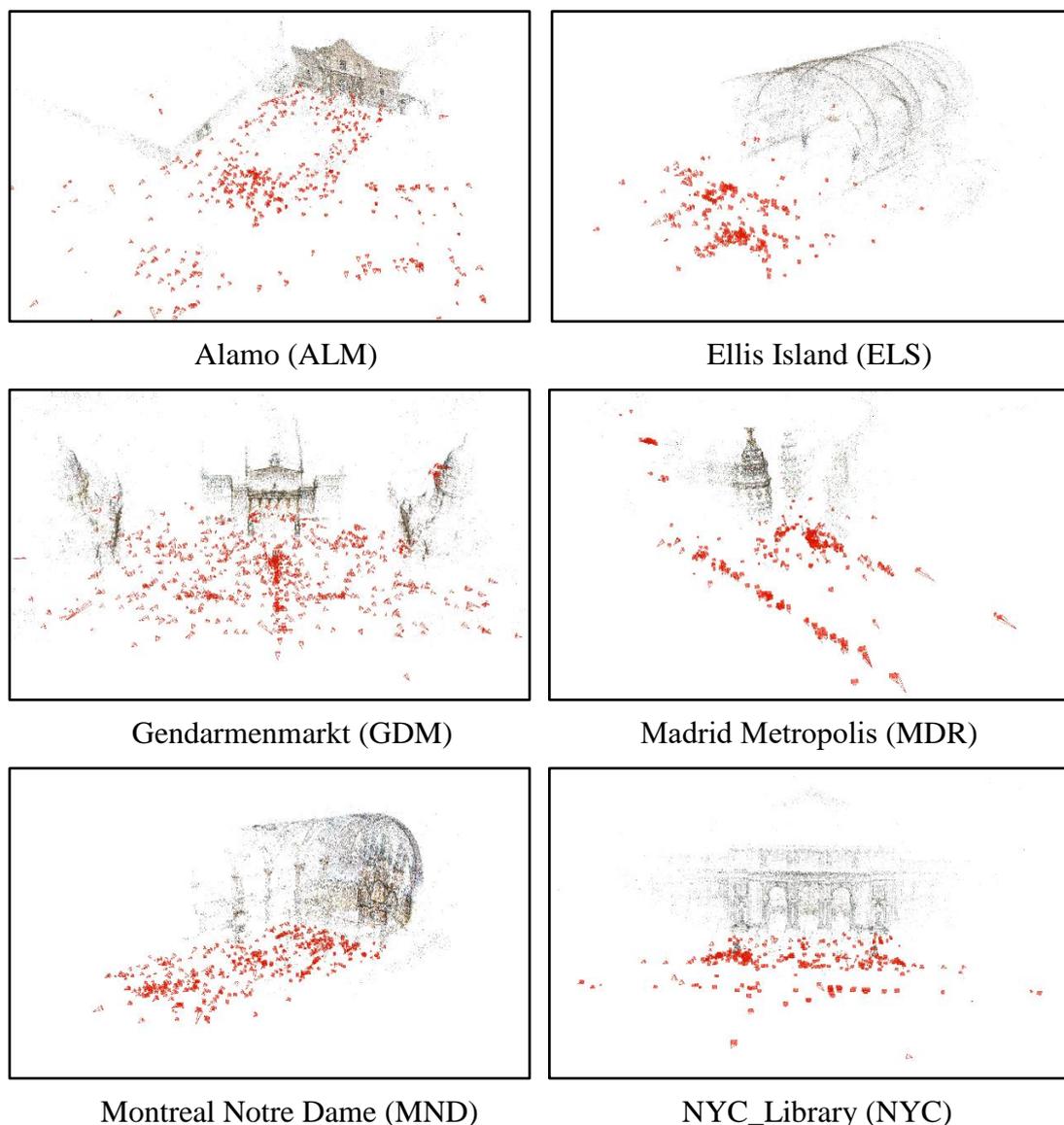
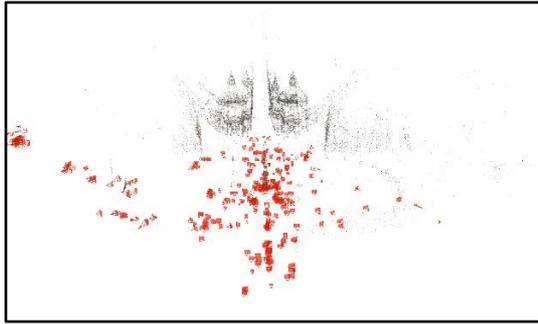
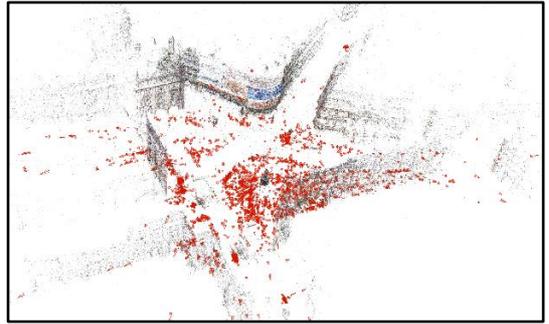


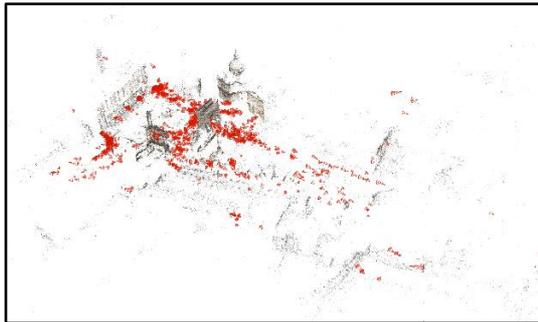
图 1-6 1DSfM 数据集 COLMAP 稀疏重建结果



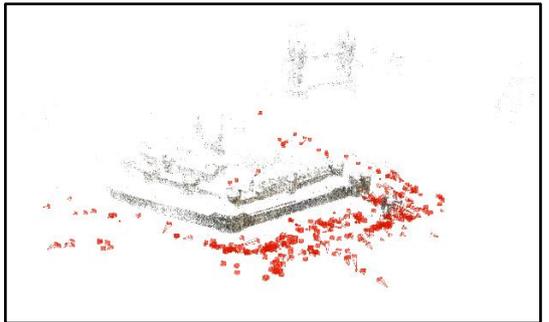
Piazza del Popolo (PDP)



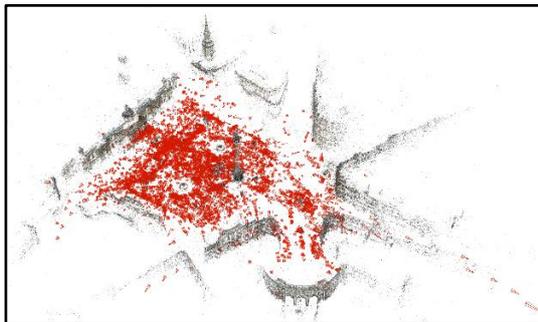
Piccadilly (PIC)



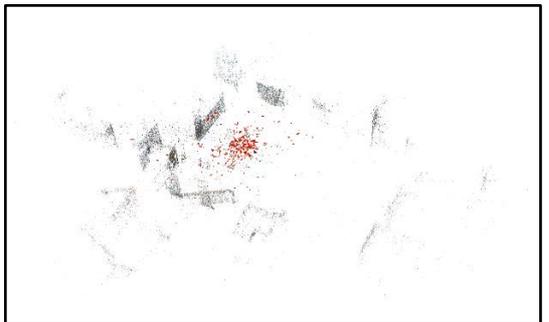
Roman Forum (ROF)



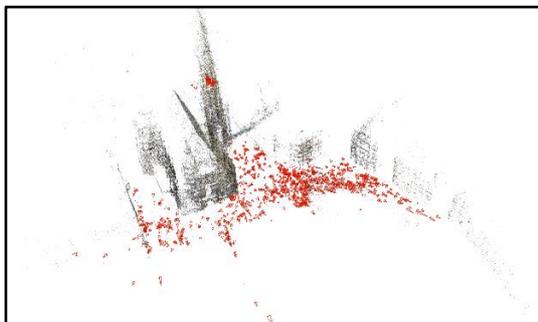
Tower of London (TOL)



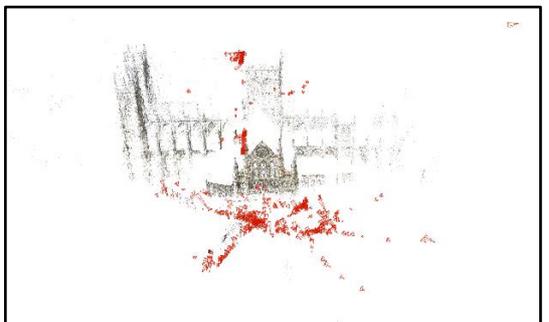
Trafalgar (TFG)



Union Square (USQ)



Vienna Cathedral (VNC)



Yorkminster (YKM)

图 1-7 1DSfM 数据集 COLMAP 稀疏重建结果 (续)

### 1.4.3 评测方法

在对本文方法进行测评时，主要包括两方面：外值滤除效果与参数估计精度，分别介绍如下。

#### (1) 外值滤除效果评测方法

由于本文提出的一系列旋转平均化与平移平均化方法均符合在滤除测量外值的同时估计参数这一方式，在此先对本文中用到的外值滤除效果评测方法进行介绍。首先可以根据相机绝对旋转与位置真值以及相对旋转、平移测量值获取相对旋转、平移测量内值真值集合， $\mathcal{E}_{GT}^R$ 以及 $\mathcal{E}_{GT}^t$ ，计算过程分别如下：

$$\begin{cases} \mathcal{E}_{GT}^R = \left\{ d_{\theta}^R \left( \mathbf{R}_{ij}, \mathbf{R}_j^{GT} \mathbf{R}_i^{GT T} \right) < \theta_{th}^R \right\} \\ \mathcal{E}_{GT}^t = \left\{ d_{\theta}^t \left( \mathbf{c}_{ij}, \frac{\mathbf{c}_j^{GT} - \mathbf{c}_i^{GT}}{\|\mathbf{c}_j^{GT} - \mathbf{c}_i^{GT}\|_2} \right) < \theta_{th}^t \right\} \end{cases}$$

for  $v_i, v_j \in \mathcal{V}_{GT}, e_{ij} \in \mathcal{E}$

其中， $\{\mathbf{R}_i^{GT}\}$ 以及 $\{\mathbf{c}_i^{GT}\}$ 分别为相机绝对旋转以及绝对位置的真值集合。另外，对于某种旋转平均化或者平移平均化方法X，可获取其估计结果对应的相对旋转或相对平移测量内值集合 $\mathcal{E}_X^R$ 以及 $\mathcal{E}_X^t$ ，计算过程分别如下：

$$\begin{cases} \mathcal{E}_X^R = \left\{ d_{\theta}^R \left( \mathbf{R}_{ij}, \mathbf{R}_j^X \mathbf{R}_i^{X T} \right) < \theta_{th}^R \right\} \\ \mathcal{E}_X^t = \left\{ d_{\theta}^t \left( \mathbf{c}_{ij}, \frac{\mathbf{c}_j^X - \mathbf{c}_i^X}{\|\mathbf{c}_j^X - \mathbf{c}_i^X\|_2} \right) < \theta_{th}^t \right\} \end{cases}$$

for  $v_i, v_j \in \mathcal{V}, e_{ij} \in \mathcal{E}$

其中， $\{\mathbf{R}_i^X\}$ 以及 $\{\mathbf{c}_i^X\}$ 分别为方法X获取的相机绝对旋转以及绝对位置的估计值集合。在此基础上，可通过精度、召回率、以及F-Score三个评测指标对不同方法的外值滤除效果进行评测。上述评测指标分别计算如下：

$$P = \frac{|\mathcal{E}_X^{(R|t)} \cap \mathcal{E}_{GT}^{(R|t)}|}{|\mathcal{E}_X^{(R|t)}|}, R = \frac{|\mathcal{E}_X^{(R|t)} \cap \mathcal{E}_{GT}^{(R|t)}|}{|\mathcal{E}_{GT}^{(R|t)}|}, F = \frac{2PR}{P + R}$$

#### (2) 参数估计精度评测方法

为对旋转平均化结果进行精度评测，本文参照文献[7]与[27]，首

先在绝对旋转真值与估计值集合之间估计一个最优的相对旋转；然后利用该相对旋转将绝对旋转估计值逐一对齐至对应绝对旋转真值，并估计它们之间的角度误差；最后，基于某种评测指标，即可对旋转平均化结果进行评测。与文献[7]，[25]，[27]一致，本文选择采用绝对旋转的中值估计误差作为评测指标。尽管在某些文献中也会采用均值估计误差<sup>[31, 40, 46]</sup>，作者相信由于对于估计的绝对旋转外值更为鲁棒，此处采用的基于中值误差的旋转平均化精度评测方法相对来讲更为合理。

对于平移平均化，首先需要注意的是，在进行精度计算与比较时，采用的平移平均化结果均为最终的全局捆绑调整之前的结果，并且以 **Bundler** 估计的相机绝对位置作为真值。原因有如下三点：1) 由于此处的研究问题为平移平均化，而非 SfM，因此对不同平移平均化方法的直接输出（即捆绑调整前的相机定位结果）进行报导与比较相对来讲更为合理；2) 绝大多数旋转平均化方法<sup>[23~24, 32, 34~36, 39, 47~48]</sup>均未提供开源代码，并且均以 **Bundler** 结果作为真值，因此对于本文来说，采用同样的真值来源（**Bundler** 结果）是一种水到渠成并且似乎是唯一可行的方式；3) **Bundler** 的相机定位精度高于进行最终全局式捆绑调整之前的任何平移平均化方法是确定无疑的，但是该精度优势对于捆绑调整之后的情况不一定还能继续维持，这种情况之下将 **Bundler** 结果继续作为真值是不恰当的。

为对平移平均化方法进行精度评测，与现有的绝大多数平移平均化方法<sup>[9, 23~24, 32, 34~36, 38]</sup>类似，本文首先估计一个在绝对位置真值集合与估计值集合之间的全局相似变换；然后利用该相似变化将绝对位置估计值逐一对齐至对应绝对位置真值，并估计它们之间的位置误差；最后，基于某种特定评测指标，即可对平移平均化结果进行评测。与文献[24]，[36]，[48]一致，本文选择采用绝对位置的中值估计误差作为评测指标。尽管在某些文献中也会采用均值估计误差，作者相信由于对于估计的绝对位置外值更为鲁棒，此处采用的基于中值误差的平移平均化精度评测方法相对来讲更为合理。

## 1.5 内容安排

本节对本文的研究内容与结构安排进行简要介绍。

## 1.5.1 研究内容

本文的研究内容主要包括如下三个方面：

1) 针对外极几何图中相对旋转测量外值过多，现有方法对其鲁棒性较差的问题，提出了基于增量式参数估计的增量式旋转平均化方法，在滤除相对旋转测量外值的同时对绝对旋转进行估计；

2) 针对在大规模场景下增量式旋转平均化方法与基于 RANSAC 旋转平均化方法分别存在的误差累积、拓展性差以及最小集大、一致集获取困难等问题，将分而治之的思想引入上述两方法，提出了分布增量式旋转平均化与层级 RANSAC 旋转平均化方法，提升了算法的精度、效率、鲁棒性以及拓展性；

3) 针对平移平均化的特点与难点，将增量式旋转平均化拓展至平移平均化问题，实现了对该问题的精确、高效、简洁、鲁棒的求解。

## 1.5.2 结构安排

本文的结构安排如下：

第一章为绪论，介绍了本文的研究背景，研究现状，预备知识，对提出方法的评测基准，以及对本文研究内容与结构安排进行简介；

第二章为增量式旋转平均化方法，介绍了如何将增量式 SfM 中采用的增量式参数估计流程引入旋转平均化问题，以对其进行精确、鲁棒、高效求解；

第三章为层级式旋转平均化方法，介绍了基于分治策略的思想的层级式旋转平均化流程，以及基于该流程的两种新型层级式旋转平均化方法及其优势；

第四章为增量式平移平均化方法，介绍了如何将增量式旋转平均化方法拓展至平移平均化问题，以及该方法相对于现有方法的优势；

第五章为总结与展望，对本文工作进行了总结，并对未来可能的研究方向进行了展望。

## 2. 增量式旋转平均化

### 2.1 引言

为了实现更加准确、鲁棒的绝对旋转估计，受已广泛研究的增量式 SfM 方法启发，本章提出了一种新型的旋转平均化方法。与传统的旋转平均化方法相比，该方法不再同时估计所有的绝对旋转，而是采用增量式的参数估计流程，依次估计各个相机的绝对旋转。该流程与增量式 SfM 类似，因此将其命名为增量式旋转平均化。另外，为进一步提升其精度与鲁棒性，在旋转平均化方法中还引入了若干同样受增量式 SfM 方法启发的关键技术，包括初始三元组以及下一最优视图（Next Best View, NBV）的选取策略，以及迭代过程中采用的加权局部、全局优化，重旋转平均化技术等。

本章工作的主要贡献包括以下两点：1) 受增量式 SfM 流程启发提出了一种增量式旋转平均化方法，该方法具有与增量式 SfM 类似的参数估计准确以及对测量外值鲁棒的特点；2) 参考增量式 SfM 中的关键技术提出了几种应用于旋转平均化过程的关键技术，进一步推进了旋转平均化结果。据作者所知，该方法为首个将增量式参数估计流程应用于有序及无序图像集合的绝对旋转估计方法，并继承了该流程准确、鲁棒的特点。

### 2.2 方法介绍

本章提出的增量式旋转平均化方法的输入为匹配图像对之间的相对旋转测量值，记作 $\{\mathbf{R}_{ij} | e_{ij} \in \mathcal{E}\}$ 。匹配图像对通常利用局部特征匹配算法获取，如 SIFT 算法<sup>[15]</sup>。此外，考虑到 Cui 等人<sup>[31]</sup>相对旋转的精度与特征匹配的数量在统计意义上呈正相关的观点，各匹配图像对之间的特征匹配对数也作为本方法的输入，记作 $\{n_{ij} | e_{ij} \in \mathcal{E}\}$ 。与其它旋转平均化方法相同，本方法的输出为每个相机优化后的绝对旋转： $\{\mathbf{R}_i^* | v_i \in \mathcal{V}\}$ 。图 2-1 展示了本章提出的增量式旋转平均化方法的基本流程：1) 首先，利用基于局部优化的初始三元组选择策略进行初始种子视图的选取和构建；2) 其次，根据基于加权支撑集合的下一最优视图选择策略决定顶点的添加顺序；3) 在完成下一最优视图的选

取和初始化后，对绝对旋转进行加权局部或者全局优化；4) 在每次执行全局优化后再进行一次重旋转平均化，进一步提升估计结果的准确性和鲁棒性。

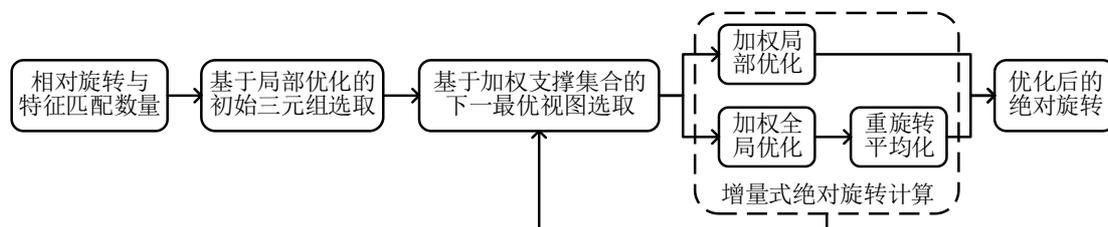


图 2-1 增量式旋转平均化流程图

## 2.2.1 基于局部优化的初始三元组选取

初始种子的选择既是增量式从运动恢复结构技术中至关重要的一步<sup>[1, 31, 49]</sup>，也是增量式旋转平均化方法中至关重要的一步。选择拥有最大特征匹配对数的图像对是实现这一步骤最为直接的方法之一。然而，当场景存在重复纹理或者对称结构的情况下，基于局部特征的图像匹配结果有时并不可靠。另外，与图像对相比，图像三元组在外点滤除和模型估计方面具有更强的鲁棒性<sup>[8, 24, 37]</sup>。考虑到以上因素的影响，本章使用图像三元组代替图像对作为增量式旋转平均化方法的初始种子，其选取策略描述如下。

在介绍初始三元组选择策略之前，应当注意到，一个外极几何图中通常包含许多相机三元组，当顶点数量很大时上述现象更为显著。为使初始三元组的选取过程更加高效，本章方法中只考虑特征匹配数量最大的前 $n_1$ 条边所构成的三元组，被选中的三元组集合记作 $t_{ijk} \in \mathcal{T}^*$ 。

对于 $\mathcal{T}^*$ 中的一个三元组 $t_{ijk}$ ，为获得相机在局部坐标系 $\{v_i, v_j, v_k\}$ 中的绝对旋转，需要在这一局部坐标系下进行如下基于三元组的优化：

$$R_i^*, R_j^*, R_k^* = \arg \min \sum_{\substack{v_i, v_j \in \mathcal{V}_{t_{ijk}} \\ e_{ij} \in \mathcal{E}_{t_{ijk}}}} \left( \omega_{ij}^0 d_{\theta}^R(R_{ij}, R_j R_i^T) \right)^2$$

其中， $\mathcal{V}_{t_{ijk}}$ 与 $\mathcal{E}_{t_{ijk}}$ 分别表示为三元组 $t_{ijk}$ 构成的顶点与边集合，另外， $\omega_{ij}^0 = n_{ij} \cos d_{\theta}^R(R_{ij}, R_j^0 R_i^{0T})$ 表示三元组各条边的初始权重，该权重

既考虑到了特征匹配数目也考虑到了相对旋转误差。可以看出，当特征匹配数越多，相对旋转误差越小时，该边对应的权重越大。 $\omega_{ij}^0$ 中的 $\mathbf{R}_i^0, \mathbf{R}_j^0, \mathbf{R}_k^0$ 代表三元组 $t_{ijk}$ 中绝对旋转的初始估计，可以通过该三元组的最大生成树来获取。上述非线性加权最小二乘问题都是由 Ceres Solver<sup>1</sup>进行求解。

完成上述优化后， $\mathcal{T}^*$ 中每个三元组的选择代价 $C_{ijk}$ 均可以通过 $C_{ijk} = \omega_{ij}^* + \omega_{ik}^* + \omega_{jk}^*$ 计算，其中， $\omega_{ij}^* = n_{ij} \cos d_{\theta}^R(\mathbf{R}_{ij}, \mathbf{R}_j^* \mathbf{R}_i^{*\top})$ 代表优化后各条边的权重。在此之后，可通过如下方式确定选取的初始三元组：

$$i^*, j^*, k^* = \arg \max \{C_{ijk} | t_{ijk} \in \mathcal{T}^*\}$$

在此基础上，对应于初始三元组的三个优化后的绝对旋转 $\mathbf{R}_{i^*}^*, \mathbf{R}_{j^*}^*, \mathbf{R}_{k^*}^*$ 构成了增量式旋转平均化方法的初始种子。

## 2.2.2 基于加权支撑集合的下一最优视图选取

下一最优视图的选取是增量式 SfM 中的另一关键技术<sup>[1, 6, 50]</sup>，在增量式旋转平均化的实现流程中也需要被认真考虑。下一最优视图可以简单地选取为与当前已估计绝对旋转的相机中连接边数最多的相机。然而，在外极几何图中，边集 $\mathcal{E}$ 中不同的边 $e_{ij}$ 对应不同的相对旋转测量误差，在选择下一最优视图时不同的边不应将其同等对待。而上述误差可以利用边上特征匹配数量和反算的旋转误差来衡量。为了增强鲁棒性，本章提出了一种基于加权支持集合的下一最优视图选择策略，具体内容描述如下：

在外极几何图中，将已进行绝对旋转估计的顶点和未进行绝对旋转估计的顶点分别记为 $v_i \in \mathcal{V}_1$ 和 $\mathcal{V}_2$ ，可知 $\mathcal{V}_1 \cup \mathcal{V}_2 = \mathcal{V}$ 。下一最优视图的选取是指从 $\mathcal{V}_2$ 选择一个合适的顶点，将其移至 $\mathcal{V}_1$ 中以使得增量式参数估计的流程可以更加鲁棒。为了提高选择下一最优视图的效率，本章方法只考虑 $\mathcal{V}_2$ 中与 $\mathcal{V}_1$ 中顶点共享最多边的前 $n_2$ 个顶点。被选中的顶点记作 $v_m \in \mathcal{V}_2^* \subseteq \mathcal{V}_2$ 。

对于 $\mathcal{V}_2^*$ 中的每个顶点 $v_m$ ，获取其与 $\mathcal{V}_1$ 之间的边集并将其记作 $e_{im} \in \mathcal{E}_{1m}$ 。对于边集 $\mathcal{E}_{1m}$ 中的每条边 $e_{im}$ ，顶点 $v_m$ 的绝对旋转可以通

<sup>1</sup> <http://www.ceres-solver.org/>

过 $\mathbf{R}_m^i = \mathbf{R}_{im} \mathbf{R}_i$ 计算。理想情况下，旋转集合 $\{\mathbf{R}_m^i | e_{im} \in \mathcal{E}_{1m}\}$ 中的每个旋转值 $\mathbf{R}_m^i$ 应该相等。但在实际情况下，受到相对旋转集合 $\{\mathbf{R}_{im}\}$ 中测量误差和绝对旋转集合 $\{\mathbf{R}_i\}$ 中估计误差的影响，这一现象不会发生。为了进行下一最优视图的选取，此处对 $\{\mathbf{R}_m^i\}$ 中每个绝对旋转的选取代价计算如下：

$$C_m^i = \sum_{\substack{v_j \in \mathcal{V}_1 \\ e_{jm} \in \mathcal{E}_{1m}}} n_{jm} \cos d_\theta^R(\mathbf{R}_m^i, \mathbf{R}_m^j)$$

然后，集合 $\{\mathbf{R}_m^i\}$ 中的代表性绝对旋转 $\mathbf{R}_m^{i^*}$ 通过下式选取：

$$i^* = \arg \max \{C_m^i | e_{jm} \in \mathcal{E}_{1m}\}$$

最终，下一最优视图被选为：

$$m^* = \arg \max \{C_m^{i^*} | v_m \in \mathcal{V}_2^*\}$$

即本策略选中的下一最优视图对应的顶点为 $v_{m^*}$ ，其绝对旋转可初始化为 $\mathbf{R}_{m^*}^{i^*}$ 。

本章方法提出得基于加权支撑集合的下一最优视图选取策略通过特征匹配数目和旋转误差对外极几何边进行加权，因而对相对旋转测量外值的鲁棒性更强。

### 2.2.3 加权局部/全局优化与重旋转平均化

完成下一最优视图的选择后，顶点 $v_{m^*}$ 的绝对旋转已初始化为 $\mathbf{R}_{m^*}^{i^*}$ 。对当前估计的绝对旋转进行局部或全局优化，可以使估计结果更加精确和鲁棒。其中，局部优化是指仅优化最新估计的绝对旋转 $\mathbf{R}_{m^*}$ 而固定其它相机的绝对旋转，而全局优化表示对当前所有已估计的绝对旋转 $\{\mathbf{R}_i | v_i \in \mathcal{V}_1 \cup \{v_{m^*}\}\}$ 同时进行优化。考虑到效率问题，本章方法借鉴 Wu 等人<sup>[5]</sup>的思想，在常规情况下仅进行局部优化，而只有在当前已估计的绝对旋转的相机数量增长到一定的比率 $r$ 时，才进行全局优化。与下一最优视图的选择相似，此处局部优化和全局优化的执行也采用加权的方式进行。此外，为了应对增量式参数估计流程中的漂移问题，在每次全局优化完成后，还需对外极几何图中已估计绝对旋转的子图进行重旋转平均化，这一思想与 VSfM<sup>[5]</sup>中的重三角测量类似。本章提出的加权局部/全局优化与重旋转平均化技术详述如下。

### (1) 加权局部优化

在完成下一最优视图的选择和初始化的基础上,能够通过如下方式获取 $\mathcal{E}_{1m^*}$ 中的内值边集合 $\mathcal{E}_{1m^*}^{\parallel}$ :

$$\mathcal{E}_{1m^*}^{\parallel} = \{d_{\theta}^R(\mathbf{R}_{im^*}, \mathbf{R}_m^{i^*} \mathbf{R}_i^T) < \theta_{th}^R\}$$

for  $v_i \in \mathcal{V}_1, e_{im^*} \in \mathcal{E}_{1m^*}$

其中,  $\parallel$ 代表内值。之后,采用如下方式对绝对旋转 $\mathbf{R}_{m^*}$ 进行优化:

$$\mathbf{R}_{m^*}^* = \arg \min \sum_{\substack{v_i \in \mathcal{V}_1 \\ e_{im^*} \in \mathcal{E}_{1m^*}^{\parallel}}} \left( \omega_{im^*}^{\parallel} d_{\theta}^R(\mathbf{R}_{im^*}^{\parallel}, \mathbf{R}_m^{i^*} \mathbf{R}_i^T) \right)^2$$

其中,各内值边对应的权重 $\omega_{im^*}^{\parallel} = n_{im^*}^{\parallel} \cos d_{\theta}^R(\mathbf{R}_{im^*}^{\parallel}, \mathbf{R}_m^{i^*} \mathbf{R}_i^T)$ 。

### (2) 加权全局优化

加权全局优化与加权局部优化类似,首先通过下式从所有当前估计的绝对旋转的边集合 $\mathcal{E}_1 \cup \mathcal{E}_{1m^*}$ 中获取内值边集合 $(\mathcal{E}_1 \cup \mathcal{E}_{1m^*})^{\parallel}$ :

$$(\mathcal{E}_1 \cup \mathcal{E}_{1m^*})^{\parallel} = \{d_{\theta}^R(\mathbf{R}_{ij}, \mathbf{R}_j \mathbf{R}_i^T) < \theta_{th}^R\}$$

for  $v_i, v_j \in \mathcal{V}_1 \cup \{v_{m^*}\}, e_{ij} \in \mathcal{E}_1 \cup \mathcal{E}_{1m^*}$

其中,  $\mathcal{E}_1$ 代表 $\mathcal{V}_1$ 中所有顶点之间的边构成的集合。之后,绝对旋转集合 $\{\mathbf{R}_i | v_i \in \mathcal{V}_1 \cup \{v_{m^*}\}\}$ 通过如下方式进行优化:

$$\{\mathbf{R}_i^*\} = \arg \min \sum_{\substack{v_i, v_j \in \mathcal{V}_1 \cup \{v_{m^*}\} \\ e_{ij} \in (\mathcal{E}_1 \cup \mathcal{E}_{1m^*})^{\parallel}}} \left( \omega_{ij}^{\parallel} d_{\theta}^R(\mathbf{R}_{ij}^{\parallel}, \mathbf{R}_j \mathbf{R}_i^T) \right)^2$$

其中,各内值边对应的权重 $\omega_{ij}^{\parallel} = n_{ij}^{\parallel} \cos d_{\theta}^R(\mathbf{R}_{ij}^{\parallel}, \mathbf{R}_j \mathbf{R}_i^T)$ 。

### (3) 重旋转平均化

得到经过全局优化的绝对旋转集合 $\{\mathbf{R}_i^*\}$ 后,再利用加权全局优化中的方法分别重新计算内边集合以及对当前已估计的绝对旋转集合进行再一次的全局优化。

## 2.3 实验验证

本节对本章中提出的增量式旋转平均化方法进行了较为详尽的验证,包括参数设置评测,关键技术的消融实验,种子敏感度验证,以及关于相对旋转测量外值滤除与绝对旋转估计精度、效率的对比实验等。具体过程详述如下。

### 2.3.1 参数设置评测

在本章提出的增量式旋转平均化方法中主要有 4 个预先定义参数，包括：1) 在进行初始三元组选取时的预选边数 $n_1$ ，2) 在进行下一最优视图选取时的预选顶点数 $n_2$ ，3) 旋转角度距离阈值 $\theta_{th}^R$ ，以及 4) 用于控制局部与全局优化操作切换的增长比率 $r$ 。其中，本章方法对 $n_1$ 与 $n_2$ 值的变化并不敏感，在实验中分别将其设为 100 与 10。然而，参数 $\theta_{th}^R$ 与 $r$ 的值会对最终的旋转平均化结果有着显著影响。因此，此处先对上述两参数的设置进行评测。

表 2-1 旋转角度距离阈值 $\theta_{th}^R$ 与增长比率 $r$ 的参数设置评测结果

Data	Estimation error in degrees with different $\theta_{th}^R$					Elapsed time in seconds with different $\theta_{th}^R$				
	1°	2°	3°	4°	5°	1°	2°	3°	4°	5°
ALM	0.80	0.81	0.83	0.84	0.86	10	14	16	16	18
ELS	0.42	0.45	0.51	0.53	0.53	2	3	3	4	4
MDR	3.45	3.12	0.85	0.85	0.86	2	2	3	3	4
MND	0.42	0.47	0.51	0.54	0.56	7	7	8	8	9
NYC	0.86	0.98	1.00	1.11	1.18	2	2	3	3	3
PDP	0.72	0.82	0.90	0.87	0.92	3	3	4	4	4
PIC	1.43	1.51	1.67	1.78	1.90	97	139	165	187	201
ROF	1.47	1.30	1.51	2.33	2.39	8	11	12	13	14
TOL	1.95	2.29	2.45	2.61	2.73	2	3	3	3	3
TFG	3.57	6.14	3.30	7.97	8.19	444	532	620	729	782
USQ	5.91	4.18	4.40	4.75	4.87	4	5	4	7	5
VNC	0.76	0.90	1.02	1.08	1.12	20	24	22	28	37
YKM	1.44	1.50	1.57	1.60	1.68	4	4	5	6	6

Data	Estimation error in degrees with different $r$					Elapsed time in seconds with different $r$				
	110%	120%	130%	140%	150%	110%	120%	130%	140%	150%
ALM	0.82	0.83	0.83	0.83	0.83	27	21	19	16	16
ELS	0.50	0.50	0.50	0.51	0.51	5	4	4	3	3
MDR	0.81	0.82	0.81	0.85	1.04	4	4	3	3	3
MND	0.51	0.50	0.51	0.51	0.51	14	10	8	8	8
NYC	1.03	1.04	1.02	1.00	1.04	5	4	4	3	3
PDP	0.89	0.90	0.88	0.90	0.90	7	5	5	4	4
PIC	1.63	1.65	1.66	1.67	1.67	335	217	180	165	155
ROF	2.29	2.22	1.51	1.51	1.51	21	15	13	12	11
TOL	2.44	2.44	2.43	2.45	2.45	5	4	3	3	3
TFG	7.91	3.43	7.90	3.30	7.86	1059	828	671	620	555
USQ	4.30	4.40	4.47	4.40	4.41	6	5	4	4	3
VNC	1.01	1.02	1.00	1.02	1.03	39	29	24	22	20
YKM	1.58	1.58	1.58	1.57	1.58	9	6	5	5	5

在参数设置评测的过程中，采取的测评方式是固定其中一个参数的值而改变另一个，并在各测试数据上进行旋转平均化操作，评测结果如表 2-1 所示。其中，参数 $\theta_{th}^R$ 的变化范围与步长分别为 $[1^\circ, 5^\circ]$ 以及 $1^\circ$ ；参数 $r$ 的变化范围与步长分别为 $[110\%, 150\%]$ 以及 $10\%$ 。表 2-1 的上半部分与下半部分分别为 $\theta_{th}^R$ 与 $r$ 的参数设置测评结果；其左半部分与右半部分分别为以角度为单位的估计误差中值以及以秒为单位的运行时间。另外需要注意，在对参数 $\theta_{th}^R$ 进行设置评测时， $r$ 的值固定为 $140\%$ ；在对参数 $r$ 进行设置评测时， $\theta_{th}^R$ 的值固定为 $3^\circ$ 。

由表 2-1 可知，本章方法的估计精度对 $\theta_{th}^R$ 的取值较为敏感。对于大部分测试数据，如 ALM, ELS, MND, NYC, PIC, TOL, VNC, 以及 YKM, 随着 $\theta_{th}^R$ 值的增加，本方法估计精度下降。因此，对于上述测试数据，更加严格的外值滤除策略可以获得更加干净的外极几何图以及更加精确的绝对旋转估计结果。然后，对于其它测试数据，尤其是 MDR, TFG, 以及 USQ, 本章方法的绝对旋转估计精度与 $\theta_{th}^R$ 的取值变化并无明显关系。这是由于对于这几组测试数据，严格的角距离阈值产生了过多的假阴性 (false-negative) 相对旋转测量值误判，这会导致在增量式计算的过程中绝对旋转的优化陷入局部极小值。另外，对于所有的测试数据，运行时间均会随着 $\theta_{th}^R$ 值的增大而变长。这是由于更大的 $\theta_{th}^R$ 会在整个流程的参数估计与优化时引入更多的约束条件，进而使得整个流程效率降低。因此，为权衡估计精度与效率，本章实验中将 $\theta_{th}^R$ 的值设置为 $3^\circ$ 。

对于参数 $r$ 来讲，由表 2-1 可知，在大部分测试数据上本章方法的估计精度对于 $r$ 的取值不太敏感。然而，TFG 数据是一个例外。随着 $r$ 的取值变化，在 TFG 上的估计精度呈现出了一种不规则的变化趋势。由表 1-1 可知，TFG 是 1DSfM 数据集中规模最大的测试数据，因此，对于同样的增长比率 $r$ ，在 TFG 上两次相邻的全局优化之间通过增量式计算估计的绝对旋转数量更多。因此，规模越大的测试数据对 $r$ 的取值越敏感。另外，显而易见的是，随着 $r$ 取值的增大，所有测试数据上的运行时间均会减小。基于表 2-1 中的评测结果，为权衡估计精度与效率，本章实验中将 $r$ 的值设置为140%。

另外，由表 2-1 可知，不管是对于 $\theta_{th}^R$ 还是 $r$ ，均不存在能够使全部测试数据取得最优旋转平均化结果的全局统一的参数设置。然而，作者认为对于某一参数在所有测试数据上的全局统一的参数设置可以提升方法的鲁棒性。

### 2.3.2 消融实验

基于上述参数设置测评，在此通过若干消融实验验证本章方法提出的各关键技术的有效性，包括基于局部优化的初始三元组选取 (Local Optimization-based Initial Triplet Selection, LOITS)，基于加权支撑集的下一最优视图选取 (Weighted Supporting Set-based NBV

Selection, WSSNS), 以及加权局部/全局优化 (Weighted Local/Global Optimization, WL/GO) 与重旋转平均化 (Re-Rotation Averaging, RRA) 等。各消融实验结果如表 2-2 所示。在进行结果分析前, 先对表 2-2 中不同的消融实验的情形简介如下。

表 2-2 不含 LOITS, WSSNS, 加权, WGO, WLO, 或 RRA 情形下的消融实验结果, 其中 w/o 表示不包含某关键技术

Data	Estimation error in degrees (elapsed time in seconds) on different ablation study situations						
	w/o LOITS	w/o WSSNS	w/o weighting	w/o RRA	w/o WGO	w/o WLO	Our IRA
ALM	0.83(13)	0.83(11)	0.94(15)	0.87(11)	0.92(5)	1.15(5)	0.83(16)
ELS	0.51(2)	0.50(2)	0.51(3)	0.52(2)	0.57(1)	0.78(1)	0.51(3)
MDR	0.82(2)	1.09(2)	0.81(3)	1.91(2)	2.38(1)	2.83(1)	0.85(3)
MND	0.51(6)	0.51(5)	0.53(8)	0.54(5)	0.58(3)	0.89(2)	0.51(8)
NYC	1.00(2)	1.01(2)	1.06(4)	1.10(2)	1.09(1)	1.46(1)	1.00(3)
PDP	0.90(3)	0.90(3)	0.85(4)	0.92(3)	0.93(1)	1.29(1)	0.90(4)
PIC	1.66(154)	1.71(135)	1.84(163)	1.77(125)	1.87(61)	1.99(58)	1.67(165)
ROF	1.65(12)	2.02(9)	1.55(13)	1.75(9)	2.01(5)	2.46(5)	1.51(12)
TOL	2.45(3)	2.48(3)	2.19(4)	2.60(2)	2.70(1)	2.93(1)	2.45(3)
TFG	3.30(671)	3.13(599)	5.84(648)	2.62(475)	4.53(308)	5.76(304)	3.30(620)
USQ	4.40(3)	4.54(3)	3.29(4)	4.60(3)	4.80(1)	4.20(1)	4.40(4)
VNC	1.01(20)	1.03(15)	1.18(21)	1.04(15)	1.01(8)	1.41(7)	1.02(22)
YKM	1.57(5)	1.57(4)	1.47(5)	1.66(4)	1.77(1)	1.89(1)	1.57(5)

**不含 LOITS:** 在这种情形下, 不进行 LOITS 操作, 而将其利用特征匹配对数最多的图像对进行初始化。

**不含 WSSNS:** 在这种情形下, 不进行 WSSNS 操作, 下一最优视图选取为与当前已估计的绝对旋转集合中加权共享边数最多的相机。

**不含加权:** 在这种情形下, 特征匹配对数与相对旋转误差在整个流程中均不再考虑, 即对外极几何图上的所有边同等对待。具体来说, 在此情形下将 2.2.3 节中的  $\omega_{im}^{\parallel}$  与  $\omega_{ij}^{\parallel}$  的值均置为 1。

**不含 RRA:** 在这种情形下, 在每次加权全局优化后不再进行重旋转平均化操作。

**不含 WGO:** 在这种情形下, 在增量式绝对旋转计算的过程中, 加权全局优化与重旋转平均化均不再进行, 即在整个过程中仅对最新估计的绝对旋转进行加权局部优化。

**不含 WLO:** 在这种情形下, 在增量式绝对旋转计算的过程中, 不再有优化操作, 即各新计算的绝对旋转仅将其定为通过基于加权支撑集合的下一最优视图选取过程获取的初值。

由表 2-2 可知, 一方面来讲, 在绝大多数的测试数据上, 绝对旋转估计精度在全部消融实验情形下均有所下降, 这证实了本章方法提出的各关键技术的有效性; 另一方面来讲, 显而易见的是, 与原始方法比较, 对于绝大多数的消融实验情形, 运行时间也均有下降。需要

注意的是，本章方法的估计精度仅稍高于不含 LOITS 的情形（在所有测试数据上的平均精度比较： $1.59^\circ$  vs.  $1.60^\circ$ ），这说明本章方法并不过于依赖于初始种子的选取与构建。对于不含 RRA, WGO, 或 WLO 的消融实验情形，其估计精度依次下降且运行时间也依次缩短。这是由于在上述三种情形中优化操作与计算量依次减少。另外还有一点值得注意的是，对于测试数据 TFG 来讲，其在不含 LOTIS 的消融实验情形下的运行时间长于原始本章方法。上述现象的一个可能原因是本章提出的初始三元组选取以及加权策略会加速优化过程的收敛。

### 2.3.3 种子敏感度验证

众所周知，增量式 SfM 对于初始种子选取较为敏感，为验证本章方法对初始种子选取的敏感程度，在此进行了初始图像对敏感度实验验证，实验结果如表 2-3 所示。其中， $\mathcal{E}^I$ 与 $\mathcal{E}^O$ 分别表示外极几何图内值与外值边集， $n_{ij}$ 与 $r_{ij}$ 分别表示选取的外极几何边上的特征匹配对数与相对旋转误差， $\tilde{r}$ 表示旋转平均化估计的绝对旋转中值误差。

表 2-3 初始图像对选取敏感度实验结果， $r_{ij}$ 及 $\tilde{r}$ 的单位：角度

Data	$\arg \max\{n_{ij}^I   e_{ij}^I \in \mathcal{E}^I\}$			random $\{n_{ij}^I   e_{ij}^I \in \mathcal{E}^I\}$								
				1st trail			2nd trail			3rd trail		
	$n_{ij}$	$r_{ij}$	$\tilde{r}$	$n_{ij}$	$r_{ij}$	$\tilde{r}$	$n_{ij}$	$r_{ij}$	$\tilde{r}$	$n_{ij}$	$r_{ij}$	$\tilde{r}$
ALM	6012	0.57	0.83	21	1.96	0.83	104	2.85	0.83	235	0.38	0.83
ELS	2152	0.81	0.51	324	0.99	0.51	899	0.49	0.51	90	2.94	0.51
MDR	6477	0.33	0.82	641	1.35	3.23	44	1.66	0.80	151	1.12	0.85
MND	7472	2.65	0.51	530	1.68	0.51	65	1.93	0.51	191	1.57	0.51
NYC	3158	0.07	1.00	229	2.84	1.00	110	1.22	1.00	163	1.37	1.00
PDP	3680	0.83	0.90	652	0.40	0.90	77	2.58	0.90	145	1.37	0.90
PIC	5345	2.86	1.66	36	1.23	1.64	30	1.65	1.65	29	2.07	1.66
ROF	5973	1.08	1.65	62	2.59	1.51	834	2.61	1.65	88	0.87	1.64
TOL	5846	1.46	2.39	37	0.92	2.38	42	1.15	2.39	1352	2.57	2.39
TFG	5630	1.38	3.30	37	1.08	3.58	32	0.59	7.76	32	2.14	7.79
USQ	3808	0.99	4.40	160	2.19	4.41	33	1.29	4.46	121	0.97	4.40
VNC	7419	0.81	1.01	389	1.07	1.02	24	2.60	1.01	36	2.55	1.02
YKM	7720	2.23	1.57	235	1.05	1.57	599	0.17	1.57	26	0.61	1.57
Data	$\arg \max\{n_{ij}^O   e_{ij}^O \in \mathcal{E}^O\}$			random $\{n_{ij}^O   e_{ij}^O \in \mathcal{E}^O\}$								
				1st trail			2nd trail			3rd trial		
	$n_{ij}$	$r_{ij}$	$\tilde{r}$	$n_{ij}$	$r_{ij}$	$\tilde{r}$	$n_{ij}$	$r_{ij}$	$\tilde{r}$	$n_{ij}$	$r_{ij}$	$\tilde{r}$
ALM	3024	5.90	0.82	16	42.20	0.84	32	17.11	0.83	114	3.80	0.82
ELS	2013	27.38	0.51	154	3.01	0.50	181	7.25	0.51	56	5.49	0.50
MDR	3545	3.78	0.92	35	10.42	3.25	38	4.58	3.22	25	29.17	0.86
MND	7107	3.25	0.51	71	3.53	0.51	181	3.28	0.51	140	5.08	0.51
NYC	2894	4.13	0.99	70	5.80	1.00	58	23.50	0.98	104	10.50	1.01
PDP	1482	5.74	0.89	103	34.07	0.90	18	27.61	0.89	24	19.43	0.90
PIC	4942	4.32	1.69	20	38.13	1.66	18	21.12	1.70	21	11.66	1.64
ROF	5131	8.43	1.51	146	6.02	3.37	34	49.09	3.34	294	8.85	3.23
TOL	6120	5.36	2.45	42	7.34	2.40	83	5.56	2.37	21	8.20	2.36
TFG	3875	5.77	8.22	34	4.13	7.86	34	9.91	7.76	32	18.61	8.26
USQ	2427	3.37	4.39	61	16.86	4.38	73	5.88	4.43	95	5.42	4.38
VNC	5439	3.51	1.02	156	3.56	1.02	124	3.15	1.02	133	31.04	1.01
YKM	5644	9.42	1.57	200	6.93	1.57	25	19.15	1.58	18	18.38	1.57

在此初始图像对敏感度实验中,将初始图像对分别置为特征匹配对数最多的外极几何内值/外值边连接的两幅图像,以及分别从 $\mathcal{E}^I$ 与 $\mathcal{E}^O$ 随机选择的三条外极几何边连接的两幅图像。然后分别以选出的不同图像对作为初始种子,利用本章方法进行旋转平均化操作。

由表 2-3 可知: 1) 对于大多数测试数据,本章方法对初始图像对的选取并不敏感,即使连接该图像对的外极几何边是外值边,或者采用随机的方式进行初始图像对的选取; 2) 尽管对初始图像对的选取不敏感,更加精确的初始图像对的选取更易于获得更精确的绝对旋转估计。因此,与增量式 SfM 不同,本章方法对初始种子的选取与构建不太敏感,这受益于相对于 SfM,旋转平均化涉及的优化问题更为轻量化(更少的待估计参数量)以及本章提出的诸如 WSSNS 以及 WL/GO+RRA 等关键技术。

### 2.3.4 对比实验

在本节中,对本章提出的增量式旋转平均化方法在测量外值滤除策略以及参数估计精度、效率等方面与其它旋转平均化方法进行了对比实验。

在对本章对比实验进行介绍之前,由于下文中还会提到,在此先对其中的一个对比方法,基于加权随机生成树(Weighed Random Spanning Tree, WRST)的旋转平均化方法,的原理进行简要介绍。WRST 方法主要包括两部分: 加权随机树生成与基于 RANSAC 的旋转平均化,下面分别进行介绍。

对于旋转平均化问题,用于 RANSAC 的最小集合为从外极几何图上获取的生成树的边集。因此,WRST 方法首先要采用随机的方式进行树的生成,采用的思想为广度优先搜索。具体来说,将外极几何图的第一个顶点作为树生成过程的起点。假设当前已被与未被生成树覆盖的顶点集合分别记作 $\mathcal{V}_1$ 与 $\mathcal{V}_2$ , $\mathcal{V}_1$ 与 $\mathcal{V}_2$ 之间的边集记作 $\mathcal{E}_{12}$ 。从 $\mathcal{E}_{12}$ 中随机选择一条边并将该边连接的 $\mathcal{V}_2$ 中的顶点由 $\mathcal{V}_2$ 移至 $\mathcal{V}_1$ 。迭代进行上述随机选择过程,直至 $\mathcal{V}_2$ 成为空集。另外,由于在 RANSAC 流程中希望随机获取的最小集合中至少有一次全为内值,WRST 中采用加权的方式进行随机树的生成。具体来说,假设 $\mathcal{E}_{12}$ 中每条边被随机选中的概率 $P(e_{ij})$ 与其上的特征匹配对数成正比:

$$P(e_{ij}) = \frac{n_{ij}}{\sum_{\substack{v_m \in \mathcal{V}_1, v_n \in \mathcal{V}_2 \\ e_{mn} \in \mathcal{E}_{12}}} n_{mn}}$$

在通过随机树生成获取最小集合之后，将其引入基于 RANSAC 的绝对旋转估计过程。各绝对旋转的初始化方式为首先获取根节点与其他各节点之间的路径，然后通过对路径上经过边的相对旋转进行链式计算。在此基础上，可以通过下式计算各随机生成树对应的支撑边集：

$$\mathcal{E}_t^S = \left\{ d_\theta^R(\mathbf{R}_{ij}, \mathbf{R}_j^t \mathbf{R}_i^{tT}) < \theta_{th}^R \right\}$$

for  $v_i, v_j \in \mathcal{V}, e_{ij} \in \mathcal{E}$

其中， $\{\mathbf{R}_i^t\}$ 为根据第 $t$ 棵随机生成树获取的绝对旋转集合初值， $S$ 表示支撑。通常情况下，在 RANSAC 过程中，会随机生成若干最小集合，其中最优的最小集合为对应支撑集合元素数量最多的最小集合。WRST 中随机生成的最小集合数量为 100。另外，WRST 在选取最优的最小集合时也不仅是考虑支撑集合的元素数量 $|\mathcal{E}_t^S|$ ，而是再一次把边上的特征匹配数纳入考虑，即下式最大的最小集合为最优最小集合：

$$t^* = \arg \max \sum_{e_{ij}^t \in \mathcal{E}_t^S} n_{ij}^t$$

其中， $\{\mathcal{E}_t^S\}$ 为各次迭代中生成的随机生成树的边集。然后，绝对旋转可通过下式进行优化：

$$\{\mathbf{R}_i^*\} = \arg \min \sum_{\substack{v_i, v_j \in \mathcal{V} \\ e_{ij}^{t^*} \in \mathcal{E}_{t^*}^S}} \left( \omega_{ij}^{t^*} d_\theta^R(\mathbf{R}_{ij}, \mathbf{R}_j \mathbf{R}_i^{t^*T}) \right)^2$$

其中， $\omega_{ij}^{t^*} = n_{ij}^{t^*} \cos d_\theta^R(\mathbf{R}_{ij}, \mathbf{R}_j \mathbf{R}_i^{t^*T})$ 。

### (1) 测量外值滤除策略对比实验

首先在测量外值滤除策略方面对本章方法与两个有代表性的基于测量外值滤除的旋转平均化方法，WRST<sup>[29]</sup>以及 OMSTs<sup>[31]</sup>，进行对比，对比结果如表 2-4 所示。另外，为分析不同滤除策略的有效性，原始外极几何图的边集 $\mathcal{E}$ 也参与各评测指标的计算。需要注意的是， $\mathcal{E}$ 的召回率永远是100%，这是因为其包含所有的相对旋转测量值，当然也会包含所有的相对旋转测量内值。

表 2-4 相对旋转测量外值滤除策略对比实验结果

Data	WRST <sup>[29]</sup>			OMSTs <sup>[31]</sup>			Our IRA			Original EG		
	Precision	Recall	F-Score	Precision	Recall	F-Score	Precision	Recall	F-Score	Precision	Recall	F-Score
ALM	76%	27%	39%	69%	6%	11%	84%	95%	89%	53%	100%	69%
ELS	89%	35%	51%	81%	11%	19%	93%	96%	94%	51%	100%	67%
MDR	79%	26%	51%	67%	25%	36%	85%	90%	88%	29%	100%	45%
MND	91%	57%	70%	80%	8%	15%	94%	97%	95%	67%	100%	80%
NYC	80%	31%	44%	71%	16%	26%	85%	93%	89%	42%	100%	59%
PDP	90%	59%	71%	74%	9%	16%	91%	97%	93%	65%	100%	78%
PIC	73%	16%	26%	61%	11%	18%	85%	91%	88%	38%	100%	55%
ROF	80%	29%	43%	71%	21%	33%	86%	91%	89%	50%	100%	67%
TOL	75%	44%	55%	71%	25%	37%	83%	90%	86%	54%	100%	70%
TFG	71%	13%	22%	65%	9%	15%	87%	92%	89%	50%	100%	67%
USQ	62%	20%	30%	55%	34%	42%	75%	84%	79%	45%	100%	62%
VNC	86%	32%	47%	67%	8%	14%	91%	96%	93%	54%	100%	70%
YKM	82%	36%	50%	72%	20%	31%	85%	93%	89%	53%	100%	70%

由表 2-4 可知，所有的相对旋转外值滤除策略得到的内值边集的精度均高于原始外极几何图的边集，这在一定程度上说明了各方法的有效性。然而，尽管 WRST 与 OMSTs 方法的精度高于原始外极几何图，其与本章的增量式旋转平均化方法在外值滤除方面仍有明显差距。另外，WRST 与 OMSTs 方法的召回率相当低，这说明对于这两种方法，许多相对旋转测量内值未成功探测。

如表 2-4 所示，1DSfM 数据集中各测试数据的相对旋转测量值内值比例在 50% 左右，因此，对于 WRST 方法来说，在有限次随机生成树的产生过程中，产生一个边全为相对旋转测量内值的生成树是几乎不太可能的，对于规模较大的数据更是如此。由于在 WRST 方法中，通过链式计算得到的绝对旋转会用于相对旋转测量外值滤除，而一条外值边就会导致在链式计算的过程中经过这条边之后的所有绝对旋转的错误估计。因此，WRST 的外值滤除精度可以较高，但是其召回率会比较低。对于 OMSTs 方法，一方面，尽管特征匹配对数更多的边上的相对旋转测量值可能会更加准确，但是仅考虑特征匹配对数这个因素的话会不可避免的引入外值边；另一方面，由于 OMSTs 方法仅在整個外极几何图中选取若干最大生成树，毫无疑问其召回率会很低。对于本章方法，由于其在增量式绝对旋转计算的过程中采用与增量式 SfM 类似的方式，迭代式地判断各外极几何边是否属于内值边，因此该方法可以对内外值边有效区分。本章方法在不同测试数据上的精度、召回率、以及 F-Score 均在 90% 左右。

## (2) 参数估计精度、效率对比实验

对于旋转平均化中绝对旋转估计精度对比实验，本章方法与

Weiszfeld<sup>[25]</sup>, DISCO<sup>[26]</sup>, IRLS-GM<sup>[7]</sup>, IRLS- $\ell_{\frac{1}{2}}$ <sup>[27]</sup>, WRST<sup>[29]</sup>, OMSTs<sup>[31]</sup>等方法进行了对比。其中, Weiszfeld, DISCO, 以及 IRLS- $\ell_{\frac{1}{2}}$ 的结果由文献[27]中摘录, IRLS-GM 的结果通过执行作者提供的开源代码获取, WRST 与 OMSTs 的结果通过对这两种方法复现获取。另外, 为再一次验证本章方法在相对旋转测量外值滤除方面的有效性, 在此将由本章方法获取的相对旋转测量内值集合导入 IRLS- $\ell_{\frac{1}{2}}$ 中, 并将此方法记为 IRLS- $\ell_{\frac{1}{2}} \setminus w$  IRA。本章方法以及上述各对比方法的旋转估计精度对比结果如表 2-5 所示。

表 2-5 旋转平均化方法精度对比实验结果, 单位: 角度

Data	Weiszfeld <sup>[25]</sup>	DISCO <sup>[26]</sup>	IRLS-GM <sup>[7]</sup>	IRLS- $\ell_{\frac{1}{2}}$ <sup>[27]</sup>	WRST <sup>[29]</sup>	OMSTs <sup>[31]</sup>	Our IRA	IRLS- $\ell_{\frac{1}{2}} \setminus w$ IRA
ALM	3.57	4.21	2.12	2.14	2.11	1.26	0.83	1.23
ELS	1.66	1.82	1.02	1.15	1.32	0.75	0.51	0.52
MDR	4.37	6.64	2.75	3.08	35.38	1.12	0.85	1.02
MND	0.92	1.03	0.77	0.71	1.03	0.68	0.51	0.55
NYC	2.43	2.59	1.43	1.40	4.51	1.30	1.00	1.11
PDP	3.35	5.25	2.19	2.62	1.48	1.73	0.90	1.30
PIC	7.65	14.66	2.38	3.12	14.40	1.41	1.67	1.63
ROF	2.11	13.69	1.59	1.70	10.55	1.85	1.51	1.48
TOL	2.73	2.89	2.55	2.45	4.08	2.10	2.45	2.45
TFG	13.20	91.62	1.85	2.03	13.25	2.63	3.30	3.22
USQ	13.54	7.50	4.34	4.97	15.39	3.83	4.40	4.22
VNC	5.14	14.57	4.47	4.64	3.63	3.30	1.02	1.06
YKM	2.73	2.34	1.71	1.62	2.90	1.55	1.57	1.44

由表 2-5 可知, 本章方法整体上来讲取得了最优的绝对旋转估计精度, 这验证了本章中提出的这一新型增量式旋转平均化流程的有效性以及其对相对旋转测量外值的鲁棒性。方法 IRLS- $\ell_{\frac{1}{2}} \setminus w$  IRA 较好的结果验证了本章方法在相对旋转测量外值滤除方面的有效性。尽管通过本章方法大部分相对旋转测量外值可以有效滤除, 由表 2-4 可知, 通过本章方法获取的内值边集仍存在约 10% 的外值边, 这是导致 IRLS- $\ell_{\frac{1}{2}} \setminus w$  IRA 方法与本章方法仍有微小差距的原因。

另外, 在此还进行了本章方法与方法 Weiszfeld, DISCO, 以及 IRLS- $\ell_{\frac{1}{2}}$ 的效率对比, 结果如表 2-6 所示。上述三个对比方法的运行时间同样是由文献[27]中摘录。尽管没有在本地运行上述三种方法的程序, 文献[27]中所采用的 CPU 主频为 2.67GHz, 性能与本文 CPU 类似, 因此用[27]报导的运行时间进行效率对比是合理的。由表可知, 本章方法在效率方面相对于三种对比方法优势明显。尽管本章方法执

行全局优化的次数更多,其所采用的更加简单的损失函数以及更加精确的初始化使其收敛速度更快,运行时间更短。

表 2-6 旋转平均化方法效率对比实验结果,单位:秒

Data	Weiszfeld <sup>[25]</sup>	DISCO <sup>[26]</sup>	IRLS- $\ell_{\frac{3}{2}}$ <sup>[27]</sup>	Our IRA
ALM	55	3917	37	16
ELS	21	470	4	3
MDR	30	560	6	3
MND	10	1608	12	8
NYC	63	446	4	3
PDP	17	583	5	4
PIC	1635	15604	776	165
ROF	121	1559	25	12
TOL	34	479	4	3
TFG	5128	43616	1725	620
USQ	221	466	12	4
VNC	222	4085	49	22
YKM	16	641	4	5

## 2.4 本章小结

本章提出了一种简单却有效的旋转平均化流程,增量式旋转平均化。该新型流程有着与增量式 SfM 类似的工作流程,如初始种子选取与构建,下一最优视图选取,局部与全局优化,以及重计算与重优化等。因此,该流程也像增量式 SfM 一样对测量外值较为鲁棒。另外,对于特定的旋转平均化任务,本章中还提出了若干关键技术,以进一步推进绝对旋转估计效果。在实验验证环节,通过参数设置实验,对关键技术的消融实验,以及对比实验,验证了本章方法的有效性。

### 3. 层级式旋转平均化

#### 3.1 引言

在上一章的主体内容与实验部分, 分别对提出的增量式旋转平均化 (Incremental Rotation Averaging, IRA) 与基于加权随机生成树 (Weighed Random Spanning Tree, WRST) 的旋转平均化方法进行了介绍。上述两种方法有两个共同的特点: 1) 都是在对绝对旋转估计的同时滤除相对旋转测量外值, 对外值较为鲁棒; 2) 都是在整个外极几何图上进行绝对旋转估计。

然而, 这两种方法各自均存在一些不足。对于 IRA, 1) 一方面来讲, 尽管其中迭代进行局部/全局优化与重旋转平均化, 随着输入的外极几何图规模的增大, 其漂移问题会不可避免地越发严重; 另一方面来讲, 尽管 IRA 中涉及的参数量远小于增量式 SfM, 由于其仍遵循增量式参数估计与迭代全局优化流程, 效率是 IRA 的另一个问题。上述两问题会使得 IRA 在大规模场景中拓展性较差。2) 正如 Wilson 等人<sup>[40]</sup>从理论上分析的那样, 旋转平均化问题的困难程度主要受到两个因素的影响, 分别是噪声程度和图结构。然而, IRA 主要关注外极几何边上的噪声程度, 而忽略了另外一个关键因素, 外极几何图的图结构。对于 WRST, 由于基于 RANSAC 的参数估计方法的结果对于最小集合的大小以及其中的内值比例十分敏感, 而 WRST 的最小集合为整个外极几何图上的随机生成树对应的边集。因此, WRST 方法在实际中几乎不可能得到边集均为内值的整个外极几何图上的随机生成树, 当面对大规模图像集合时更是如此。

针对上述问题, 为提升 IRA 与 WRST 方法在大规模与高噪声情形下的鲁棒性, 根据 Wilson 等人<sup>[40]</sup>的建议, 本章将在 SfM<sup>[51~52]</sup>中广泛采用的分而治之的思想引入旋转平均化流程之中, 提出了层级式的旋转平均化流程。并且在该层级式旋转平均化流程之中分别引入 IRA 以及 WRST 两种旋转平均化方法, 进而提出了两种新型的层级式旋转平均化方法: 层级增量式旋转平均化 (IRA++) 以及层级 RANSAC 旋转平均化 (Hierarchical RANSAC-based Rotation Averaging, HRRRA) 方法。具体来说, 首先通过受限社区发现 (constrained community

detection) 算法<sup>[53]</sup>将输入的原型外极几何图聚类为若干子图。然后, 对于每个类内子图, 采用某种旋转平均化方法(对于 IRA++采用 IRA, 对于 HRRR 采用 WRST), 计算其中各顶点相对于各类内子图局部坐标系的绝对旋转。接下来, 通过两两子图之间的外极几何边集, 利用基于投票的单旋转平均化 (single rotation averaging) 方法估计两两类内子图局部坐标系之间的相对旋转。在此基础上, 再对类间子图采用某种旋转平均化方法 (IRA 或 WRST), 估计各类内子图局部坐标系的绝对旋转 (即局部坐标系相对于全局坐标系的旋转)。最后, 基于上述结果, 各类内子图的各项点的绝对旋转先对齐至全局坐标系下, 然后进行全局优化以获取最终的旋转平均化结果。

本章工作的主要贡献包括以下两点: 1) 为提升现有旋转平均化方法在大规模与高噪声情形下的鲁棒性, 提出了一种层级式旋转平均化流程; 2) 将现有的两种旋转平均化方法, IRA 与 WRST, 紧密融合到了提出的新型层级式旋转平均化之中, 以分布式解决底层类内子图与顶层类间子图的旋转平均化问题。

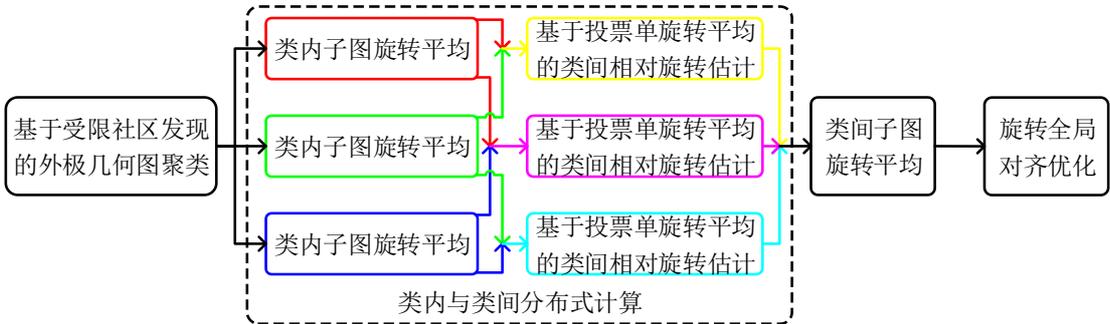


图 3-1 层级式旋转平均化流程图

### 3.2 方法介绍

本章提出的层级式旋转平均化流程的输入与 IRA 相同, 也是匹配图像对之间的相对旋转测量值以及 SIFT<sup>[15]</sup>特征匹配对数, 记作  $\{R_{ij}, n_{ij} | e_{ij} \in \mathcal{E}\}$ , 输出为每个相机优化后的绝对旋转:  $\{R_i^* | v_i \in \mathcal{V}\}$ 。图 3-1 展示了本章提出的层级式旋转平均化方法的基本流程: 1) 在原始外极几何图上执行基于受限社区发现的外极几何图聚类算法, 得到若干子图; 2) 对每个类内子图采用某种旋转平均化方法 (对于 IRA++采用 IRA, 对于 HRRR 采用 WRST) 估计相机在类内局部坐

标系下的绝对旋转；3) 采用基于投票的单旋转平均化方法求解两两子图局部坐标系之间的相对旋转；4) 在类间子图上执行某种旋转平均化方法 (IRA 或 WRST)，估计各子图局部坐标系相对于全局坐标系的绝对旋转；5) 采用全局旋转的对齐和优化，首先将原始外极几何图中的所有绝对旋转对齐至一个统一的全局坐标系下，然后对其进行全局优化，得到最终的旋转平均化结果。上述五个层级式旋转平均化步骤分别详述如下。

### 3.2.1 基于受限社区发现的外极几何图聚类

为了应对现有旋转平均化方法中的拓展性和鲁棒性问题，并将图结构纳入考量范围，在此将原始外极几何图  $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$  聚类为若干更小的子图，以实现类内与类间相对于原始外极几何图更紧密的连接性。考虑到可扩展性的问题，需要对每个聚类子图的大小进行限制，此处通过受限外极几何图聚类算法<sup>[53]</sup>实现。经过聚类的外极几何子图记作  $\{\mathcal{G}_p = \{\mathcal{V}_p, \mathcal{E}_p\}\}$ 。此外，需要注意，除了外极几何子图内部有边集  $\{\mathcal{E}_p\}$  外，外极几何子图之间也存在边集  $\{\mathcal{E}_{pq}\}$ ，即  $\{\mathcal{E}_p\} \cup \{\mathcal{E}_{pq}\} = \mathcal{E}$ 。另外，如果把每个外极几何子图  $\mathcal{G}_p$  看作一个顶点，将其记作  $v_p \in \mathcal{V}_c$ ，把每两个外极几何子图之间的边集  $\mathcal{E}_{pq}$  看作一条边，将其记作  $e_{pq} \in \mathcal{E}_c$ ，则可以得到一个聚类层级上的外极几何图  $\mathcal{G}_c = \{\mathcal{V}_c, \mathcal{E}_c\}$ ，其中  $\mathcal{V}_c$  和  $\mathcal{E}_c$  分别代表该图的顶点集与边集。为了区别相机层级上的外极几何子图  $\{\mathcal{G}_p = \{\mathcal{V}_p, \mathcal{E}_p\}\}$  和聚类层级上的外极几何子图  $\mathcal{G}_c = \{\mathcal{V}_c, \mathcal{E}_c\}$ ，将其分别称为类内与类间子图。

### 3.2.2 类内子图旋转平均

在获取类内与类间子图后，首先对各类内子图  $\mathcal{G}_p$  采用某种旋转平均化方法 (对于 IRA++ 采用 IRA，对于 HRRR 采用 WRST)，获取其中各顶点在其局部坐标系下的绝对旋转，记作  $\{\mathbf{R}_m^p \mid v_m^p \in \mathcal{V}_p\}$ 。将子图  $\mathcal{G}_p$  边集  $\mathcal{E}_p$  上的相对旋转测量值与特征匹配对数，记作  $\{\mathbf{R}_{mn}^p, n_{mn}^p \mid e_{mn}^p \in \mathcal{E}_p\}$ ，输入到某种旋转平均化方法中以实现类内旋转平均化。在此之后，为将上述结果对齐到全局坐标系下，并对其进行全局优化，还需要进行两个操作。首先是需要求得每对类内子图局部

坐标系之间的相对旋转，记作 $\{\mathbf{R}_{pq}^c | e_{pq}^c \in \mathcal{E}_c\}$ ；其次是需要估计各类内子图局部坐标系相对于全局坐标系的绝对旋转，记作 $\{\mathbf{R}_p^c | v_p^c \in \mathcal{V}_c\}$ 。这两步操作的具体流程分别在 3.2.3 节与 3.2.4 节进行描述。

### 3.2.3 基于投票单旋转平均的类间相对旋转估计

为了估计两个类内子图 $\mathcal{G}_p$ 和 $\mathcal{G}_q$ 在聚类层级上的相对旋转 $\mathbf{R}_{pq}^c$ ，首先要获得它们在相机层级上的相对旋转测量值集合，记作 $\{\mathbf{R}_{mn}^{pq} | e_{mn}^{pq} \in \mathcal{E}_{pq}\}$ ，其中 $\mathbf{R}_{mn}^{pq}$ 代表第 $p$ 个类内子图中第 $m$ 个相机与第 $q$ 个类内子图中第 $n$ 个相机之间的相对旋转测量值。随后，利用 $\{\mathbf{R}_m^p | v_m^p \in \mathcal{V}_p\}$ 和 $\{\mathbf{R}_n^q | v_n^q \in \mathcal{V}_q\}$ 中的绝对旋转，可通过下式对 $\{\mathbf{R}_{mn}^{pq}\}$ 中的相对旋转测量值进行变换用于描述聚类层级上的相对旋转：

$$\hat{\mathbf{R}}_{mn}^{pq} = \mathbf{R}_n^{q*T} \mathbf{R}_{mn}^{pq} \mathbf{R}_m^{p*}$$

注意，本章中每个类内子图的局部坐标系均建立在其第一个顶点上。

理论上讲，由于 $\{\hat{\mathbf{R}}_{mn}^{pq} | e_{mn}^{pq} \in \mathcal{E}_{pq}\}$ 中的各个旋转矩阵均表示同一个聚类层级上的相对旋转，（ $\mathcal{G}_p$ 与 $\mathcal{G}_q$ 局部坐标系之间的相对旋转），它们的取值应该互相保持一致。然而，由于 $\{\mathbf{R}_m^p\}$ ， $\{\mathbf{R}_n^q\}$ ，以及 $\{\mathbf{R}_{mn}^{pq}\}$ 之中均存在不可避免的参数估计误差，上述现象在实际情况下不可能发生。因此，此处的类间相对旋转估计问题可以归结为单旋转平均化问题，即从一个单一旋转的若干估计中通过平均化的方式获取最优旋转<sup>[54]</sup>。本章提出一种基于投票策略的单旋转平均化方法以解决该问题，具体描述如下。

首先从集合 $\{\hat{\mathbf{R}}_{mn}^{pq}\}$ 中采用加权随机的方式选取一个旋转值，并将其记为 $\hat{\mathbf{R}}_{mn}^{pq \dagger}$ ，与 WRST 中的加权随机树的生成方式类似，某条边 $e_{mn}^{pq}$ 的选取概率与其上特征匹配对数 $n_{mn}^{pq}$ 成正比。然后，可通过下式获取 $\hat{\mathbf{R}}_{mn}^{pq \dagger}$ 的支撑边集：

$$\begin{aligned} \mathcal{E}_{pq}^S &= \left\{ d_{\theta}^R \left( \hat{\mathbf{R}}_{ij}^{pq}, \hat{\mathbf{R}}_{mn}^{pq \dagger} \right) < \theta_{th}^R \right\} \\ &\text{for } v_i^p \in \mathcal{V}_p, v_j^q \in \mathcal{V}_q, e_{ij}^{pq} \in \mathcal{E}_{pq} \end{aligned}$$

重复迭代 100 次，然后同样与 WRST 方法类似，将其中最优的最小集合（在此为某随机选取的单个旋转矩阵 $\hat{\mathbf{R}}_{mn}^{pq \dagger}$ ）确定为其支撑边集上特征匹配对数加和最大的最小集合，具体方式如下式所示：

$$\widehat{\mathbf{R}}_{mn}^{pq \dagger*} = \arg \max \sum_{\substack{v_m^p \in \mathcal{V}_p, v_n^q \in \mathcal{V}_q \\ e_{mn}^{pq} \in \mathcal{E}_{pq}^{\mathcal{S}}}} n_{mn}^{pq}$$

然后，将 $\mathbf{R}_{pq}^c$ 初始化为 $\widehat{\mathbf{R}}_{mn}^{pq \dagger*}$ ，并通过下式对其进行优化：

$$\mathbf{R}_{pq}^{c*} = \arg \min \sum_{\substack{v_m^p \in \mathcal{V}_p, v_n^q \in \mathcal{V}_q \\ e_{mn}^{pq*} \in \mathcal{E}_{pq}^{\mathcal{S}*}}} \left( \omega_{mn}^{pq*} d_{\theta}^R(\widehat{\mathbf{R}}_{mn}^{pq*}, \mathbf{R}_{pq}^c) \right)^2$$

其中， $\mathcal{E}_{pq}^{\mathcal{S}*}$ 为 $\widehat{\mathbf{R}}_{mn}^{pq \dagger*}$ 的支撑边集， $\omega_{mn}^{pq*} = n_{mn}^{pq*} \cos d_{\theta}^R(\widehat{\mathbf{R}}_{mn}^{pq*}, \widehat{\mathbf{R}}_{mn}^{pq \dagger*})$ 。

### 3.2.4 类间子图旋转平均

通过类间相对旋转估计，可获取用于执行类间子图 $\mathcal{G}_c$ 旋转平均化的输入，记为 $\{\mathbf{R}_{pq}^{c*}, |\mathcal{E}_{pq}^{\mathcal{S}*}| | e_{pq}^c \in \mathcal{E}_c\}$ 。值得注意的是，对类间子图旋转平均化而言，聚类层级各边 $e_{pq}^c$ 的支撑集越大时越容易获取可靠的类间相对旋转估计。因此，在此采用 $|\mathcal{E}_{pq}^{\mathcal{S}*}|$ 对各聚类层级边进行加权。之后，再次执行某种旋转平均化算法（IRA 或 WRST）可对 $\{\mathbf{R}_p^{c*}\}$ 进行估计。

### 3.2.5 旋转全局对齐优化

在通过类内与类间子图旋转平均化方法求得 $\{\mathbf{R}_m^p | v_m^p \in \mathcal{V}_p\}$ 与 $\{\mathbf{R}_p^c | v_p^c \in \mathcal{V}_c\}$ 后，原始外极几何图中所有相机的绝对位姿均可以根据以下公式对齐至全局坐标系下：

$$\mathbf{R}_i^{\dagger} = \mathbf{R}_m^p \mathbf{R}_p^c$$

随后，原始外极几何图中的内值边集 $\mathcal{E}^{\mathbb{I}}$ 可以根据 $\{\mathbf{R}_{ij} | e_{ij} \in \mathcal{E}\}$ 中的相对旋转测量值与借助 $\{\mathbf{R}_i^{\dagger} | v_i \in \mathcal{V}\}$ 获取的相对旋转反算值 $\{\mathbf{R}_j^{\dagger} \mathbf{R}_i^{\dagger T}\}$ 之间的旋转角度误差进行判断：

$$\mathcal{E}^{\mathbb{I}} = \left\{ d_{\theta}^R(\mathbf{R}_{ij}, \mathbf{R}_j^{\dagger} \mathbf{R}_i^{\dagger T}) < \theta_{th}^R \right\}$$

for  $v_i, v_j \in \mathcal{V}, e_{ij} \in \mathcal{E}$

然后，对原始外极几何图上的旋转进行全局优化，得到本方法最终的旋转平均化结果，优化公式如下：

$$\{\mathbf{R}_i^*\} = \arg \min \sum_{\substack{v_i, v_j \in \mathcal{V} \\ e_{ij}^{\parallel} \in \mathcal{E}^{\parallel}}} \left( \omega_{ij}^{\parallel} d_{\theta}^R(\mathbf{R}_{ij}^{\parallel}, \mathbf{R}_j \mathbf{R}_i^{\top}) \right)^2$$

其中，内边权重  $\omega_{ij}^{\parallel} = n_{ij}^{\parallel} \cos d_{\theta}^R(\mathbf{R}_{ij}^{\parallel}, \mathbf{R}_j^{\dagger} \mathbf{R}_i^{\dagger \top})$  用于在优化期间对集合  $\mathcal{E}^{\parallel}$  提供的约束进行加权。

### 3.3 实验验证

本节对本章中提出的层级式旋转平均化流程进行了较为详尽的验证。首先比较了层级式旋转平均化流程中产生的类内、类间子图与原始外极几何图的图结构的连接紧密度。然后对 HRRR 与 IRA++ 两个基于层级式旋转平均化流程的旋转平均化方法分别进行了评测。对于 HRRR 的评测，包括不同最大聚类大小的影响，消融实验，精度对比实验等；对于 IRA++，包括外值滤除、分而治之策略的对比实验，以及效率、精度的对比实验等。具体实验过程详述如下。

#### 3.3.1 图结构对比实验

正如 Wilson 等人<sup>[40]</sup>宣称的那样，图结构是影响旋转平均化问题难度的两个关键因素之一。因此，在此对在 IRA 以及 WRST 中使用的原始外极几何图  $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$  的图结构与在 IRA++ 以及 HRRR 中使用的类内与类间子图  $\{\mathcal{G}_p = \{\mathcal{V}_p, \mathcal{E}_p\}\}$  与  $\mathcal{G}_c = \{\mathcal{V}_c, \mathcal{E}_c\}$  的图结构的连接紧密度采用 Wilson 等人<sup>[40]</sup>提出的图结构紧密度指标进行比较。该图结构指标定义如下： $\frac{\lambda_2(\mathbf{L}(\mathcal{G}))}{|\mathcal{V}|}$ ，其中  $\mathbf{L}(\mathcal{G}) = \mathbf{D}(\mathcal{G}) - \mathbf{A}(\mathcal{G})$  为  $\mathcal{G}$  的图拉普拉斯， $\mathbf{D}(\mathcal{G})$  与  $\mathbf{A}(\mathcal{G})$  分别为  $\mathcal{G}$  的度矩阵与邻接矩阵， $\lambda_2(\mathbf{L}(\mathcal{G}))$  为矩阵  $\mathbf{L}(\mathcal{G})$  的第二小特征值，称为代数连通度。为比较在不同方法中使用的外极几何图的图结构，在此对 1DSfM 数据集中各测试数据的  $\mathcal{G}$ ， $\mathcal{G}_c$  的图结构指标  $\frac{\lambda_2(\mathbf{L}(\mathcal{G}))}{|\mathcal{V}|}$ ， $\frac{\lambda_2(\mathbf{L}(\mathcal{G}_c))}{|\mathcal{V}_c|}$  以及  $\{\mathcal{G}_p\}$  中全部子图的图结构指标均值  $\frac{\sum_{p=1}^{v_c} \frac{\lambda_2(\mathbf{L}(\mathcal{G}_p))}{|\mathcal{V}_p|}}{|\mathcal{V}_c|}$  进行了计算，结果如表 3-1 所示。对所有类内子图计算图结构指标均值的原因在于每个测试数据都含有多个类内子图，且不同测试数据子图个数不一致。

表 3-1 IRA/WRST 与 IRA++/HRRR 方法图结构对比实验结果

Method	Graph structure indicator	ALM	ELS	GDM	MDR	MND	NYC	PDP
IRA/WRST	$\frac{\lambda_2(\mathcal{L}(\mathcal{G}))}{ \mathcal{V} }$	1.32e-3	3.75e-3	1.14e-3	1.81e-3	2.09e-3	2.47e-3	1.15e-3
IRA++/HRRR	$\frac{\lambda_2(\mathcal{L}(\mathcal{G}_c))}{ \mathcal{V}_c }$	1.00	1.00	5.22e-1	1.00	1.00	1.00	1.00
	$\frac{\sum_{p=1}^{ \mathcal{V}_c } (\lambda_2(\mathcal{L}(\mathcal{G}_p)) /  \mathcal{V}_p )}{ \mathcal{V}_c }$	1.13e-2	2.20e-2	1.48e-2	7.89e-3	1.24e-2	5.77e-3	4.33e-2
Method	Graph structure indicator	PIC	ROF	TOL	TFG	USQ	VNC	YKM
IRA/WRST	$\frac{\lambda_2(\mathcal{L}(\mathcal{G}))}{ \mathcal{V} }$	2.33e-4	7.83e-4	1.67e-3	5.92e-5	4.06e-4	5.84e-4	1.61e-3
IRA++/HRRR	$\frac{\lambda_2(\mathcal{L}(\mathcal{G}_c))}{ \mathcal{V}_c }$	2.43e-1	4.82e-1	7.14e-1	3.48e-2	2.63e-1	4.00e-1	1.00
	$\frac{\sum_{p=1}^{ \mathcal{V}_c } (\lambda_2(\mathcal{L}(\mathcal{G}_p)) /  \mathcal{V}_p )}{ \mathcal{V}_c }$	1.01e-2	8.40e-3	1.24e-2	5.38e-3	7.50e-3	1.14e-2	5.95e-3

由表 3-1 可知，类内与类间子图的图结构指标在全部测试数据上均大于原始外极几何图，这说明通过进行受限外极几何图聚类，本章中用到的类内与类间子图的连接性相较原始外极几何图均更加紧密。这说明本章中提出的层级式旋转平均化流程需要求解的旋转平均化问题更为简单，验证了本章提出的面向旋转平均化问题的分而治之策略的有效性。

### 3.3.2 HRRR 验证实验

#### (1) 参数设置评测

对于 HRRR 算法，在此主要讨论了其中受限社区发现算法中的最大聚类大小参数对最终旋转平均化结果的影响。此处该参数的取值集合为 $\{1, 50, 100, 200, \infty\}$ 。其中，该参数取值为1时表示原始外极几何图中的每个顶点构成一个类，即 WRST 方法；该参数取值为 $\infty$ 时表示在进行图聚类时不对最大聚类大小进行限制。该参数设置实验结果如表 3-2 所示，其中 $\tilde{r}$ 表示参数估计中值误差。

表 3-2 HRRR 最大聚类大小设置评测结果， $\tilde{r}$ 的单位：角度

Data	Constrained community detection <sup>[53]</sup> with different max cluster sizes									
	1		50		100		200		$\infty$	
	$ \mathcal{V}_c $	$\tilde{r}$	$ \mathcal{V}_c $	$\tilde{r}$	$ \mathcal{V}_c $	$\tilde{r}$	$ \mathcal{V}_c $	$\tilde{r}$	$ \mathcal{V}_c $	$\tilde{r}$
ALM	627	2.11	14	1.12	7	1.03	6	1.09	3	1.11
ELS	247	1.32	6	0.60	5	0.59	4	0.63	2	0.60
GDM	742	21.26	18	4.01	11	4.04	6	4.17	4	3.91
MDR	394	35.38	11	2.66	6	2.54	3	2.24	2	1.47
MND	474	1.03	12	0.61	7	0.62	4	0.60	3	0.53
NYC	376	4.51	11	1.28	5	1.24	4	1.29	3	1.32
PDP	354	1.48	9	0.94	6	0.92	5	0.98	4	0.95
PIC	2508	14.45	51	5.77	29	4.87	18	4.53	4	11.53
ROF	1134	10.55	27	2.63	12	2.48	7	2.52	4	2.27
TOL	508	4.08	14	1.98	7	2.05	5	2.27	4	2.47
TFG	5433	13.25	113	5.29	58	4.88	32	4.91	3	5.94
USQ	930	15.35	25	3.83	13	3.77	7	3.75	5	5.35
VNC	918	3.63	23	1.85	13	1.84	9	1.87	4	1.86
YKM	458	2.90	11	1.63	7	1.57	5	1.59	3	1.72

由表 3-2 可知：1) 聚类数 $|\mathcal{V}_c|$ 与顶点数 $|\mathcal{V}|$ 正相关，与最大聚类

大小负相关；2) 将 HRRR 方法的最大聚类大小设置为适中的范围时 ( $\{50,100,200\}$ ) 其结果整体上优于将其设置为极端值 ( $\{1,\infty\}$ )，而且当其值在合适的范围内变化时，该参数对最终的旋转平均化结果影响不明显。因此，对于 HRRR 方法来说，其中是否包含 (受限) 外极几何图聚类操作比较关键，而用于进行该操作的参数最大聚类大小具体的取值并不十分重要，因此本章将 HRRR 方法中该参数的值设为 100。另外，在 4 组测试数据 (GDM, MDR, MND, 以及 ROF) 中，参数最大聚类大小取值为  $\infty$  时结果最优。一个可能的原因是由于图聚类结果受图结构影响明显，与不受限方法相比，受限外极几何图聚类在上述测试数据上可能使得类内与类间连接性变弱，进而使得旋转平均化结果变差。

**(2) 消融实验**

接下来，对 HRRR 方法进行了消融实验，以验证其有效性，各消融实验情形简述如下。

表 3-3 消融实验对比结果，单位：角度    表 3-4 精度对比实验结果，单位：角度

Data	w/o weig.	w/o clus.	w/o cons.	w/o opti.	HRRR	Data	Weiszfeld <sup>[25]</sup>	IRLS-GM <sup>[7]</sup>	IRLS- $\ell_{1/2}$ <sup>[27]</sup>	OMSTs <sup>[31]</sup>	HRRR
ALM	1.21	2.11	1.11	2.34	1.03	ALM	3.57	2.12	2.14	1.26	1.03
ELS	0.97	1.32	0.60	2.23	0.59	ELS	1.66	1.02	1.15	0.75	0.59
GDM	21.18	21.26	3.91	13.36	4.04	GDM	—	39.15	28.20	45.15	4.04
MDR	9.31	35.38	1.47	6.98	2.54	MDR	4.37	2.75	3.08	1.12	2.54
MND	0.74	1.03	0.53	1.77	0.62	MND	0.92	0.77	0.71	0.68	0.62
NYC	2.45	4.51	1.32	3.09	1.24	NYC	2.43	1.43	1.40	1.30	1.24
PDP	1.24	1.48	0.95	2.22	0.92	PDP	3.35	2.19	2.62	1.73	0.92
PIC	19.44	14.45	11.53	10.33	4.87	PIC	7.65	2.38	3.12	1.41	4.87
ROF	5.97	10.55	2.27	6.01	2.48	ROF	2.11	1.59	1.70	1.85	2.48
TOL	2.17	4.08	2.47	3.40	2.05	TOL	2.73	2.55	2.45	2.10	2.05
TFG	17.54	13.25	5.94	15.62	4.88	TFG	13.20	1.85	2.03	2.63	4.88
USQ	5.31	15.35	5.35	7.51	3.77	USQ	13.54	4.34	4.97	3.83	3.77
VNC	2.50	3.63	1.86	4.14	1.84	VNC	5.14	4.47	4.64	3.30	1.84
YKM	1.95	2.90	1.72	2.87	1.57	YKM	2.73	1.71	1.62	1.55	1.57

**不含加权 (w/o weig.):** 在这种情形下，旋转平均化的整个过程中所有外极几何边同等对待，即在生成随机树，获取最优的最小集合，以及进行绝对旋转优化时均不考虑边上的特征匹配对数或相对旋转反算误差。

**不含聚类 (w/o clus.):** 在这种情形下，不再进行基于社区发现的外极几何图聚类，此种情形等同于 WRST 方法。

**不含受限 (w/o cons.):** 在这种情形下，外极几何图采用不受限的方法进行图聚类，即不设定最大聚类大小。

**不含优化 (w/o opti.):** 在这种情形下，仅将绝对旋转对齐至全局坐标系而不进行全局优化。

上述消融实验结果如表 3-3 所示。由表可知，与原始 HRRa 方法相比，各消融实验情形的精度均明显下降，因此，上述消融实验验证了 HRRa 中的边加权、（受限）图聚类、以及全局优化等操作的有效性。

### (3) 精度对比实验

对于精度对比实验，在此对 HRRa 与 Weiszfeld<sup>[25]</sup>，IRLS-GM<sup>[7]</sup>，IRLS- $\ell_{\frac{1}{2}}$ <sup>[27]</sup>，以及 OMSTs<sup>[31]</sup>等 4 个方法进行对比，实验结果如表 3-4 所示。由表可知，HRRa 在所有对比方法中取得了整体最优的表现：在所有 14 个测试数据上有 9 个测试数据取得了最优精度，2 个测试数据取得了次优精度。尽管在 3 个相对规模较大的测试数据（PIC，ROF，以及 TFG）上并未取得较好结果，导致这种情况的原因可能并不仅是规模因素，还有可能是数据的其他特点，例如外极几何图的噪声程度与连接紧密度<sup>[40]</sup>等

## 3.3.3 IRA++验证实验

### (1) 外值滤除对比实验

在此对本文第 2 章中的 IRA 方法与本章 IRA++方法的相对旋转测量外值的滤除效果进行对比实验。首先获取两方法估计的相对旋转测量内值集合，然后采用 1.4.3 节中给出的评测方法进行评测，结果如表 3-5 所示，IRA 与 IRA++方法中精度、召回率、以及 F-score 较大者在表中进行了加粗。

表 3-5 IRA++外值滤除对比实验结果

Data	Original EG			IRA			IRA++		
	Precision	Recall	F-score	Precision	Recall	F-score	Precision	Recall	F-score
ALM	52.63%	100.00%	68.96%	84.22%	<b>94.83%</b>	89.21%	<b>84.70%</b>	94.47%	<b>89.32%</b>
ELS	<b>50.86%</b>	100.00%	67.43%	92.54%	96.45%	94.46%	<b>93.56%</b>	<b>96.46%</b>	<b>94.99%</b>
GDM	24.66%	100.00%	39.57%	53.24%	<b>82.41%</b>	64.68%	<b>61.99%</b>	81.40%	<b>70.38%</b>
MDR	29.33%	100.00%	45.35%	<b>85.25%</b>	<b>90.49%</b>	<b>87.79%</b>	85.15%	88.39%	86.74%
MND	66.94%	100.00%	80.19%	93.75%	96.89%	95.30%	<b>94.07%</b>	<b>97.05%</b>	<b>95.54%</b>
NYC	41.93%	100.00%	59.09%	85.10%	<b>93.46%</b>	89.09%	<b>85.99%</b>	93.26%	<b>89.48%</b>
PDP	64.56%	100.00%	78.46%	90.53%	<b>96.63%</b>	93.48%	<b>91.01%</b>	96.32%	<b>93.59%</b>
PIC	38.38%	100.00%	55.47%	<b>84.60%</b>	<b>90.83%</b>	<b>87.61%</b>	84.41%	85.94%	85.17%
ROF	50.27%	100.00%	66.91%	86.29%	<b>91.38%</b>	<b>88.76%</b>	<b>87.01%</b>	90.03%	88.49%
TOL	53.92%	100.00%	70.07%	83.02%	89.77%	86.26%	<b>83.79%</b>	<b>90.63%</b>	<b>87.08%</b>
TFG	49.88%	100.00%	66.56%	<b>86.68%</b>	<b>92.48%</b>	<b>89.49%</b>	85.95%	88.06%	87.00%
USQ	44.52%	100.00%	61.62%	75.18%	<b>83.60%</b>	79.17%	<b>76.44%</b>	82.29%	<b>79.26%</b>
VNC	54.12%	100.00%	70.23%	90.89%	<b>95.64%</b>	<b>93.20%</b>	<b>91.00%</b>	94.77%	92.85%
YKM	53.48%	100.00%	69.69%	85.13%	<b>93.43%</b>	89.09%	<b>85.99%</b>	92.61%	<b>89.18%</b>

由表 3-5 可知，在全部测试数据上，相比于原始外极几何图边集，

IRA 与 IRA++获得的内值边集在精度与 F-score 两个指标上取得了明显提升，这证实了两种方法在相对旋转测量外值滤除方面的有效性。另外，通过对 IRA 与 IRA++比较可知，IRA++在精度与 F-score 方面存在优势，但是在召回率方面略有差距。需要注意的是，对于原始外极几何边集精度最差的 GDM 数据，相对于 IRA，IRA++方法在精度与 F-score 方面分别取得了8.75%与5.70%的领先。尽管 IRA++方法未在三个指标上取得全部领先，在 IRA++方法取得领先的两个指标中，F-score 由于兼顾精度与召回率，是一个更加综合的指标，而在参数估计问题中，精度也比召回率指标相对更为重要一些。因此，整体上讲 IRA++在相对旋转测量外值滤除方面表现优于 IRA。

## (2) 分治策略对比实验

为验证本文采用的分治策略的有效性，在此进行了若干对比实验，包括采用不同聚类算法（社区发现算法<sup>[53]</sup>与归一化割算法<sup>[55]</sup>）的对比实验以及采用不同最大聚类大小（{50,100,200,∞}）的受限社区发现算法的对比实验。实验结果如表 3-6 所示，该表中包含聚类数，估计中值误差，以及总体运行时间，其中对不同对比方法中估计中值误差与总体运行时间的最优结果进行了加粗处理。

表 3-6 IRA++分治策略对比实验结果，其中 $|V_c|$ ， $\bar{r}$ ，以及 $T_\Sigma$ 分别表示含聚类数，估计中值误差，以及总体运行时间

Data	Constrained community detection <sup>[53]</sup> with different max cluster sizes												Normalized cut <sup>[55]</sup>		
	50			100			200			∞					
	$ V_c $	$\bar{r}$	$T_\Sigma$	$ V_c $	$\bar{r}$	$T_\Sigma$	$ V_c $	$\bar{r}$	$T_\Sigma$	$ V_c $	$\bar{r}$	$T_\Sigma$	$ V_c $	$\bar{r}$	$T_\Sigma$
ALM	14	0.82°	9s	7	<b>0.80°</b>	9s	6	0.81°	19s	3	0.81°	21s	7	0.82°	12s
ELS	6	0.48°	4s	5	<b>0.46°</b>	4s	4	0.49°	3s	2	0.49°	4s	5	0.50°	2s
GDM	18	5.21°	9s	11	2.88°	7s	6	2.92°	11s	4	4.87°	13s	11	<b>2.35°</b>	8s
MDR	11	0.95°	4s	6	<b>0.83°</b>	4s	3	0.89°	6s	2	0.87°	6s	6	0.96°	4s
MND	12	0.50°	6s	7	0.50°	7s	4	<b>0.49°</b>	11s	3	0.50°	14s	7	0.54°	8s
NYC	11	1.08°	3s	5	<b>0.95°</b>	3s	4	0.99°	5s	3	1.00°	6s	5	1.07°	4s
PDP	9	0.84°	5s	6	<b>0.75°</b>	4s	5	0.83°	7s	4	0.85°	7s	6	0.90°	4s
PIC	51	1.94°	40s	29	1.70°	42s	18	1.80°	76s	4	<b>1.66°</b>	114s	29	1.83°	66s
ROF	27	1.33°	9s	12	1.24°	10s	7	<b>1.18°</b>	19s	4	1.43°	21s	12	1.46°	14s
TOL	14	2.08°	5s	7	1.33°	4s	5	1.31°	7s	4	<b>1.30°</b>	7s	7	2.36°	5s
TFG	113	1.96°	73s	58	<b>1.74°</b>	80s	32	1.77°	133s	3	<b>1.74°</b>	339s	58	2.91°	156s
USQ	25	4.37°	5s	13	<b>3.70°</b>	5s	7	3.90°	8s	5	3.85°	9s	13	4.86°	6s
VNC	23	1.00°	11s	13	<b>0.94°</b>	12s	9	0.97°	23s	4	1.00°	27s	13	1.11°	18s
YKM	11	1.46°	4s	7	<b>1.38°</b>	4s	5	1.43°	8s	3	1.47°	8s	7	1.41°	5s

由表可知，对于采用不同最大聚类大小的受限社区发现算法的对比实验，整体上讲，随着最大聚类大小的增大，聚类数变小，总体运行时间变长。另外，上述对比实验的估计精度差别并不明显，而当最大聚类大小设为100时精度最高。因此，本文中将该参数设为100。对

于不同聚类算法的对比实验,在此将本文采用的受限社区发现算法与另一种在大规模 SfM 中常用的聚类算法<sup>[52, 56]</sup>, 归一化割算法进行了对比。需要注意的是,在通过归一化割方法进行图聚类时,需要提前对聚类数进行设置。基于公平比较的考虑,在此将归一化割算法的聚类数设为与受限社区发现算法最大聚类大小设为100的聚类数相同。由表可知,基于受限社区发现算法的层级式旋转平均化方法(IRA++)精度与效率均优于基于归一化割算法的层级式旋转平均化方法,这证实了本文采用的分治策略的有效性。

### (3) 估计效率对比实验

在效率对比实验中,将 IRA++与 IRA, IRLS- $\ell_{\frac{1}{2}}^{[27]}$ , 以及一个本地实现的层级式版本的 IRLS- $\ell_{\frac{1}{2}}$ 方法(记作 HIRLS- $\ell_{\frac{1}{2}}$ )进行了对比。对于 HIRLS- $\ell_{\frac{1}{2}}$ 的实现,此处将 IRA++中的基于 IRA 的类内与类间子图旋转平均化替换为基于 IRLS- $\ell_{\frac{1}{2}}$ 的方法。上述四种对比方法在不同测试数据上的运行时间如表 3-7 所示,其中对各测试数据的最优结果进行了加粗处理。

表 3-7 IRA++估计效率对比实验结果

Method	ALM	ELS	GDM	MDR	MND	NYC	PDP
IRLS- $\ell_{\frac{1}{2}}^{[27]}$	39s	4s	16s	6s	12s	4s	5s
HIRLS- $\ell_{\frac{1}{2}}$	23s	<b>3s</b>	13s	7s	13s	7s	8s
IRA	16s	<b>3s</b>	<b>7s</b>	<b>3s</b>	8s	<b>3s</b>	<b>4s</b>
IRA++	<b>9s</b>	4s	<b>7s</b>	4s	<b>7s</b>	<b>3s</b>	<b>4s</b>
Method	PIC	ROF	TOL	TFG	USQ	VNC	YKM
IRLS- $\ell_{\frac{1}{2}}^{[27]}$	776s	25s	4s	1725s	12s	49s	<b>4s</b>
HIRLS- $\ell_{\frac{1}{2}}$	59s	23s	9s	132s	11s	25s	13s
IRA	165s	12s	<b>3s</b>	620s	<b>4s</b>	22s	5s
IRA++	<b>42s</b>	<b>10s</b>	4s	<b>80s</b>	5s	<b>12s</b>	<b>4s</b>

由表可知,在规模相对较小的测试数据上,各对比方法的效率相当。然而,随着测试数据规模的增长,基于层级式流程的旋转平均化方法(IRA++与 HIRLS- $\ell_{\frac{1}{2}}$ )的效率优势越发明显。需要注意的是,本章实验均在一台个人电脑上进行,而层级式旋转平均化流程中的两个关键步骤,类内子图旋转平均化(3.2.2 节)以及基于投票单旋转平均化的类间相对旋转估计(3.3.3 节),均可很轻易地进行并行化实现,进而实现更大的提速。因此,采用本章提出的层级式旋转平均化流程可以极大程度上缓解大规模旋转平均化效率较低问题。另外,由于

IRA 的效率高于 IRLS- $\ell_{\frac{1}{2}}$ , 相对于 HIRLS- $\ell_{\frac{1}{2}}$ , IRA++ 效率也相应更高。

#### (4) 估计精度对比实验

在精度对比实验中, 将 IRA++ 与 9 种代表性旋转平均化方法进行了对比, 包括 4 种基于鲁棒损失的方法 (IRLS-GM<sup>[7]</sup>, IRLS- $\ell_{\frac{1}{2}}$ <sup>[27]</sup>, MPLS<sup>[28]</sup>, 以及 HIRLS- $\ell_{\frac{1}{2}}$ ), 3 种基于外值滤除的方法 (OMSTs<sup>[31]</sup>, HRRRA, 以及 IRA), 以及 2 种基于深度学习的方法 (NeuRoRA<sup>[43]</sup> 以及 MSPRA<sup>[44]</sup>)。实验结果如表 3-8 所示, 其中对各测试数据精度最高的 4 种方法进行了高亮处理。

表 3-8 IRA++ 估计精度对比实验结果

Data	IRLS-GM <sup>[7]</sup>	IRLS- $\ell_{\frac{1}{2}}$ <sup>[27]</sup>	MPLS <sup>[28]</sup>	HIRLS- $\ell_{\frac{1}{2}}$	OMSTs <sup>[31]</sup>	HRRRA	IRA	NeuRoRA <sup>[43]</sup>	MSPRA <sup>[44]</sup>	IRA++
ALM	2.12°	2.14°	1.16°	1.10°	1.26°	1.03°	0.83°	1.16°	1.07°	0.80°
ELS	1.08°	1.15°	0.88°	0.98°	0.75°	0.59°	0.51°	0.64°	0.83°	0.46°
GDM	35.83°	28.20°	9.87°	3.36°	45.15°	4.04°	5.32°	2.94°	3.69°	2.88°
MDR	4.52°	3.08°	1.26°	1.21°	1.12°	2.54°	0.85°	1.13°	1.09°	0.83°
MND	0.77°	0.71°	0.51°	0.66°	0.68°	0.62°	0.51°	0.60°	0.50°	0.50°
NYC	1.43°	1.40°	1.24°	1.21°	1.30°	1.24°	1.00°	1.18°	1.12°	0.95°
PDP	2.16°	2.62°	1.93°	1.10°	1.73°	0.92°	0.90°	0.79°	0.76°	0.75°
PIC	4.14°	3.12°	1.81°	2.86°	1.41°	4.87°	1.67°	1.91°	1.80°	1.70°
ROF	1.62°	1.70°	1.37°	1.40°	1.85°	2.48°	1.51°	1.31°	1.19°	1.24°
TOL	2.59°	2.45°	2.20°	2.22°	2.10°	2.05°	2.45°	1.46°	1.25°	1.33°
TFG	1.94°	2.03°	—	1.81°	2.63°	4.88°	3.30°	2.25°	—	1.74°
USQ	4.93°	4.97°	3.48°	3.64°	3.83°	3.77°	4.40°	2.01°	1.85°	3.70°
VNC	4.87°	4.64°	2.83°	1.68°	3.30°	1.84°	1.02°	1.50°	1.10°	0.94°
YKM	1.70°	1.62°	1.45°	1.55°	1.55°	1.57°	1.57°	0.99°	0.91°	1.38°

由表可知, 在所有对比方法中, IRA++ 在绝对旋转估计精度方面取得了整体最优的结果。另外, 通过比较 IRA 与 IRA++ 方法可知, IRA++ 几乎在所有的测试数据上均实现了精度的明显提升, 尤其是在噪声程度最高的 GDM (5.32° vs. 2.88°), 以及数据规模最大的 TFG (3.30° vs. 1.74°) 上。上述结果证实了 IRA++ 相对于 IRA 方法在估计精度上的优势。

### 3.4 本章小结

针对现有旋转平均化方法大都是在整个外极几何图上进行绝对旋转估计, 在大规模与高噪声情形下的鲁棒性较差问题, 本章基于分治策略提出了一种新型旋转平均化流程: 层级式旋转平均化。基于本章流程, 可将原始的一个规模更大、难度更高的旋转平均化问题转化为多个规模更小、难度更低的子问题。在此基础上, 该流程可以分别求解各子问题并将求解结果融合以获取原始问题的解。另外, 本章将

IRA 与 WRST 方法融入到上述流程之中形成了 IRA++ 以及 HRRA 算法。在实验验证环节, 对本章提出的分治策略以及 IRA++ 与 HRRA 涉及到的关键技术与最终的绝对旋转估计结果等进行了评测, 验证了本章提出的层级式旋转平均化流程以及基于此流程构建的两种旋转平均化方法的有效性。



## 4. 增量式平移平均化

### 4.1 引言

本章通过将增量式 SfM 中增量式参数估计思想引入平移平均化问题，提出了一种新型的平移平均化方法，称为增量式平移平均化 (Incremental Translation Averaging, ITA)。与传统的平移平均化方法不同，本章方法不再同时估计所有相机的绝对位置，而是通过一种序列式方式对其进行准确、鲁棒估计。此外，为了提升增量式平移平均化结果的准确性和鲁棒性，本章方法中还融合了若干项受增量式 SfM 技术启发的关键技术，如：1) 初始四元组和下一最优视图的选取策略；2) 加权局部与全局优化以及重平移平均化。本章提出的这一简单、高效的平移平均化方法在不依赖于任何其它方法中必不可少的鲁棒操作的前提下在 1DSfM 数据集上达到了当前最优性能。值得注意的是，尽管本章提出的增量式平移平均化方法的基本思想与本文第二章中的增量式旋转平均化方法相似，但两者之间存在若干本质区别：1) 本章提出的增量式平移平均化方法是目前已知的首个采用增量式参数估计流程解决平移平均化问题的方法；2) 在本章方法中，为应对平移平均化过程中的一些特定问题，创新性地提出了包括基于四元组的初始种子选取，基于线性三焦张量算法的相机绝对位置计算，以及基于平方弦距离的相机绝对位置优化等特定操作。

本章工作的主要贡献主要包括以下三个方面：1) 提出了一种新型的增量式估计相机绝对位置的平移平均化方法，该方法具有与增量式 SfM 相同的准确度高和鲁棒性强的优点；2) 开发了若干受增量式 SfM 启发的关键技术，并将其融合到提出的平移平均化方法中，有效改善了相机绝对位置估计的结果；3) 受益于增量式参数估计流程的有效性，本章提出的增量式平移平均化方法较少依赖于其它方法中常用的鲁棒操作，使得整个流程更为简单、高效。

### 4.2 方法介绍

本章提出的增量式平移平均化方法的输入为匹配图像对之间根据已知的绝对旋转变换至全局坐标系下的相对平移，将其记作

$\{c_{ij}|e_{ij} \in \mathcal{E}\}$ 。输出为各相机优化后的绝对位置，记作 $\{c_i^*|v_i \in \mathcal{V}\}$ 。图 4-1 展示了本章提出的增量式平移平均化方法的基本流程：1) 首先，利用基于局部优化的初始四元组选取策略进行初始种子的选取和恢复；2) 其次，根据基于加权支撑集合的下一最优视图选择策略确定顶点的添加顺序；3) 在完成下一最优视图的选取和初始化后，执行加权局部或全局优化；4) 最后，在完成全局优化的基础上再进行一次重平移平均化，进一步改善估计结果的准确性和鲁棒性。其中，下一最优视图选取，加权局部或全局优化和重平移平均化等步骤需要迭代运行，直到所有相机均已完成定位或不再有相机可以继续定位为止。

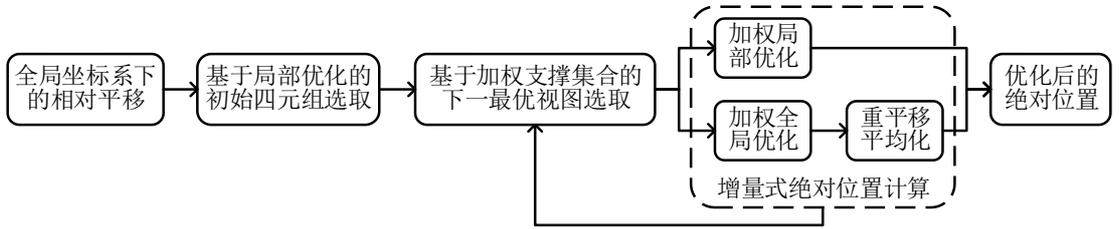


图 4-1 增量式平移平均化流程图

### 4.2.1 基于局部优化的初始四元组选取

对于本章提出的增量式平移平均化方法的初始种子选取与恢复，一种最为直接的解决方案是通过下式选取具有最小旋转回环偏差的相机对或相机三元组作为初始种子：

$$i^*, j^* = \arg \min \{d_{\theta}^R(\mathbf{I}, \mathbf{R}_{ij}^T \mathbf{R}_j \mathbf{R}_i^T) | v_i, v_j \in \mathcal{V}, e_{ij} \in \mathcal{E}\} \text{ 或}$$

$$i^*, j^*, k^* = \arg \min \left\{ \begin{array}{l} d_{\theta}^R(\mathbf{I}, \mathbf{R}_{ij}^T \mathbf{R}_j \mathbf{R}_i^T) \\ +d_{\theta}^R(\mathbf{I}, \mathbf{R}_{ik}^T \mathbf{R}_k \mathbf{R}_i^T) \\ +d_{\theta}^R(\mathbf{I}, \mathbf{R}_{jk}^T \mathbf{R}_k \mathbf{R}_j^T) \end{array} \middle| v_i, v_j, v_k \in \mathcal{V}, t_{ijk} \in \mathcal{T} \right\}$$

其中， $\mathcal{T}$ 表示由边集 $\mathcal{E}$ 中全部外极几何边所组成的三元组集合。对于一条外极几何边 $e_{ij}$ 而言，其上高精度的相对旋转 $\mathbf{R}_{ij}$ 并不能保证高精度的相对平移 $\mathbf{t}_{ij}$ 。因此，平移平均化过程中初始种子的选择应取决于其本身的位置回环偏差，而不是旋转回环偏差。此外，由于相对平移中尺度信息的丢失，相机三元组是相机定位中所需的最小配置，本章方法在相机三元组的基础上添加了一个相机用于其评估初始种子选择和恢复的有效性。因此，出于对鲁棒性问题的考虑，本章方法用相

机四元组代替相机三元组或图像对作为初始种子。

在介绍初始四元组选择策略之前，应当注意到，一个外极几何图中通常包含许多相机四元组，当顶点数量很大时上述现象更为显著。为使四元组的选取过程更加高效，本方法中只考虑旋转回环偏差最小的前 $n_1$ 条边所构成的四元组，被选中的四元组集合记作 $Q^* \in Q^*$ 。

对于 $Q^*$ 中的一个四元组 $q_{ijkl}$ ，为获得相机在局部坐标系 $\{v_i, v_j, v_k, v_l\}$ 中的绝对位置，需要在这一局部坐标系下进行如下基于四元组的优化：

$$\mathbf{c}_i^*, \mathbf{c}_j^*, \mathbf{c}_k^*, \mathbf{c}_l^* = \arg \min \sum_{\substack{v_i, v_j \in \mathcal{V}_{q_{ijkl}} \\ e_{ij} \in \mathcal{E}_{q_{ijkl}}} } d_{ch}^t \left( \mathbf{c}_{ij}, \frac{\mathbf{c}_j - \mathbf{c}_i}{\|\mathbf{c}_j - \mathbf{c}_i\|_2} \right)^2$$

其中， $\mathcal{V}_{q_{ijkl}}$ 与 $\mathcal{E}_{q_{ijkl}}$ 分别表示为四元组 $q_{ijkl}$ 构成的顶点与边集合。为求解上述非线性最小二乘问题，首先要对 $\{\mathbf{c}_i, \mathbf{c}_j, \mathbf{c}_k, \mathbf{c}_l\}$ 进行初始化。具体来说，首先将 $\mathbf{c}_i$ 与 $\mathbf{c}_j$ 分别初始化为 $\mathbf{0}$ 以及 $\mathbf{c}_{ij}$ ，然后可利用1.3.2节中提到的三焦张量算法对三元组 $t_{ijk}$ 与 $t_{ijl}$ 进行处理，以实现 $\mathbf{c}_k$ 以及 $\mathbf{c}_l$ 的初始化。

完成上述优化后，通过下式计算 $Q^*$ 中每个四元组的选取代价：

$$C_{ijkl} = \sum_{\substack{v_i, v_j \in \mathcal{V}_{q_{ijkl}} \\ e_{ij} \in \mathcal{E}_{q_{ijkl}}} } \omega_{ij} = \sum_{\substack{v_i, v_j \in \mathcal{V}_{q_{ijkl}} \\ e_{ij} \in \mathcal{E}_{q_{ijkl}}} } \mathbf{c}_{ij} \cdot \frac{\mathbf{c}_j^* - \mathbf{c}_i^*}{\|\mathbf{c}_j^* - \mathbf{c}_i^*\|_2}$$

然后，通过下式对初始四元组进行选取：

$$i^*, j^*, k^*, l^* = \arg \max \{C_{ijkl} | q_{ijkl} \in Q^*\}$$

在此基础上，对应于初始四元组的四个优化后的绝对位置 $\mathbf{c}_{i^*}^*, \mathbf{c}_{j^*}^*, \mathbf{c}_{k^*}^*, \mathbf{c}_{l^*}^*$ 构成了增量式平移平均化方法的初始种子。

## 4.2.2 基于加权支撑集合的下一最优视图选取

对于本章提出的增量式平移平均化方法，下一最优视图可以简单地选取为与当前已估计绝对位置的相机中连接边数最多的相机。然而， $\mathcal{E}$ 中的每条边 $e_{ij}$ 上有着不同的相对平移测量误差 $d_{\theta}^t \left( \mathbf{c}_{ij}, \frac{\mathbf{c}_j^* - \mathbf{c}_i^*}{\|\mathbf{c}_j^* - \mathbf{c}_i^*\|_2} \right)$ ，在

进行下一最优视图选取时不应将其同等对待。因此，为提升平移平均化方法的鲁棒性，本章方法提出了一种面向增量式平移平均化方法的基于加权支撑集合的下一最优视图选取策略。

在外极几何图中，将已进行绝对位置估计的顶点和未进行绝对位置估计的顶点分别记为 $v_i \in \mathcal{V}_1$ 和 $\mathcal{V}_2$ ，可知 $\mathcal{V}_1 \cup \mathcal{V}_2 = \mathcal{V}$ 。下一最优视图的选取是指从 $\mathcal{V}_2$ 选择一个合适的顶点，将其移至 $\mathcal{V}_1$ 中以使得增量式参数估计的流程可以更加鲁棒。为了提高选择下一最优视图的效率，本章方法只考虑 $\mathcal{V}_2$ 中与 $\mathcal{V}_1$ 中顶点共享最多边的前 $n_2$ 个顶点。被选中的顶点记作 $v_m \in \mathcal{V}_2^* \subseteq \mathcal{V}_2$ 。

对于 $\mathcal{V}_2^*$ 中的每个顶点 $v_m$ ，获取其与 $\mathcal{V}_1$ 中顶点之间的三元组集合并将其记作 $t_{ijm} \in \mathcal{T}_{1m}$ 。需要注意在此引入了相机三元组以消除绝对位置估计中的尺度不确定性。对于三元组集合 $\mathcal{T}_{1m}$ 中的每个相机三元组 $t_{ijm}$ ，顶点 $v_m$ 的绝对位置可以再次利用 1.3.2 节中介绍的三焦张量算法，在给出已估计的相机绝对位置 $\mathbf{c}_i$ 与 $\mathbf{c}_j$ ，以及相对位置测量值 $\mathbf{c}_{ij}$ ， $\mathbf{c}_{im}$ ， $\mathbf{c}_{jm}$ 的情况下进行计算。将得到的 $v_m$ 绝对位置的估计结果记为 $\mathbf{c}_m^{ij}$ 。理想情况下，位置集合 $\{\mathbf{c}_m^{ij} | t_{ijm} \in \mathcal{T}_{1m}\}$ 中的每个位置向量应该相等。但在实际情况下，受到相对平移的测量误差和绝对位置的估计误差的影响，这一现象不会发生。为了进行下一最优视图的选取，此处对 $\{\mathbf{c}_m^{ij}\}$ 中每个绝对位置的选取代价计算如下：

$$C_m^{ij} = \sum_{\substack{v_k \in \mathcal{V}_1 \\ e_{km} \in \mathcal{E}_{1m}}} \omega_{km}^{ij} = \sum_{\substack{v_k \in \mathcal{V}_1 \\ e_{km} \in \mathcal{E}_{1m}}} c_{km} \cdot \frac{c_k - c_m^{ij}}{\|c_k - c_m^{ij}\|_2}$$

其中， $\mathcal{E}_{1m}$ 表示 $\mathcal{V}_1$ 中的顶点与顶点 $v_m$ 之间的边集。随后，可通过下式对集合 $\{\mathbf{c}_m^{ij}\}$ 中的代表性绝对位置 $\mathbf{c}_m^{i^*j^*}$ 进行选取：

$$i^*, j^* = \arg \max \{C_m^{ij} | t_{ijm} \in \mathcal{T}_{1m}\}$$

最后，可通过下式对下一最优视图选取：

$$m^* = \arg \max \{C_m^{i^*j^*} | v_m \in \mathcal{V}_2^*\}$$

即本章提出的下一最优视图选取策略将下一最优视图选定为顶点 $v_{m^*}$ ，其绝对位置可初始化为 $\mathbf{c}_{m^*}^{i^*j^*}$ 。

由于本章中提出的面向增量式平移平均化的下一最优视图选取策略基于由重计算位置偏差进行加权的支撑集合，因而其对相对平移测量外值更加鲁棒。

### 4.2.3 加权局部/全局优化与重平移平均化

完成下一最优视图的选择后，顶点 $v_{m^*}$ 的绝对位置已初始化为 $\mathbf{c}_{m^*}^{i^*j^*}$ 。对当前估计的绝对位置进行局部或全局优化，可以使估计结果更加精确和鲁棒。考虑到效率问题，与增量式旋转平均化类似，在常规情况下仅进行局部优化，而只有在当前已估计的绝对位置的相机数量增长到一定的比率 $r$ 时，才进行全局优化。此处的局部优化和全局优化的执行也采用加权的方式进行。此外，为了应对增量式参数估计流程中的漂移问题，在每次全局优化完成后，还需对外极几何图中已估计绝对位置的子图进行重平移平均化。本章提出的面向增量式平移平均化的加权局部/全局优化与重平移平均化技术详述如下。

#### (1) 加权局部优化

在完成下一最优视图的选择和初始化的基础上，能够通过如下方式获取 $\mathcal{E}_{1m^*}$ 中的内值边集合 $\mathcal{E}_{1m^*}^{\parallel}$ ：

$$\mathcal{E}_{1m^*}^{\parallel} = \left\{ d_{\theta}^t \left( \mathbf{c}_{im^*}, \frac{\mathbf{c}_{m^*}^{i^*j^*} - \mathbf{c}_i}{\|\mathbf{c}_{m^*}^{i^*j^*} - \mathbf{c}_i\|_2} \right) < \theta_{th}^t \right\}$$

for  $v_i \in \mathcal{V}_1, e_{im^*} \in \mathcal{E}_{1m^*}$

之后，采用如下方式对绝对位置 $\mathbf{c}_{m^*}$ 进行优化：

$$\mathbf{c}_{m^*}^* = \arg \min_{\substack{v_i \in \mathcal{V}_1 \\ e_{im^*} \in \mathcal{E}_{1m^*}^{\parallel}}} \sum \left( \omega_{im^*}^{\parallel} d_{ch}^t \left( \mathbf{c}_{im^*}^{\parallel}, \frac{\mathbf{c}_{m^*} - \mathbf{c}_i}{\|\mathbf{c}_{m^*} - \mathbf{c}_i\|_2} \right) \right)^2$$

其中，各内值边对应的权重 $\omega_{im^*}^{\parallel} = \mathbf{c}_{im^*}^{\parallel} \cdot \frac{\mathbf{c}_{m^*}^{i^*j^*} - \mathbf{c}_i}{\|\mathbf{c}_{m^*}^{i^*j^*} - \mathbf{c}_i\|_2}$ 。

#### (2) 加权全局优化

加权全局优化与加权局部优化类似，首先通过下式从所有当前估计的绝对位置的边集合 $\mathcal{E}_1 \cup \mathcal{E}_{1m^*}$ 中获取内值边集合 $(\mathcal{E}_1 \cup \mathcal{E}_{1m^*})^{\parallel}$ ：

$$(\mathcal{E}_1 \cup \mathcal{E}_{1m^*})^{\parallel} = \left\{ d_{\theta}^t \left( \mathbf{c}_{ij}, \frac{\mathbf{c}_j - \mathbf{c}_i}{\|\mathbf{c}_j - \mathbf{c}_i\|_2} \right) < \theta_{th}^t \right\}$$

for  $v_i, v_j \in \mathcal{V}_1 \cup \{v_{m^*}\}, e_{ij} \in \mathcal{E}_1 \cup \mathcal{E}_{1m^*}$

其中， $\mathcal{E}_1$ 代表 $\mathcal{V}_1$ 中所有顶点之间的边构成的集合。之后，绝对位置集合 $\{\mathbf{c}_i | v_i \in \mathcal{V}_1 \cup \{v_{m^*}\}\}$ 通过如下方式进行优化：

$$\{\mathbf{c}_i^*\} = \arg \min \sum_{\substack{v_i, v_j \in \mathcal{V}_1 \cup \{v_m^*\} \\ e_{ij} \in (\mathcal{E}_1 \cup \mathcal{E}_{1m^*})^{\parallel}}} \left( \omega_{ij}^{\parallel} d_{ch}^t \left( \mathbf{c}_{ij}^{\parallel}, \frac{\mathbf{c}_j - \mathbf{c}_i}{\|\mathbf{c}_j - \mathbf{c}_i\|_2} \right) \right)^2$$

其中，各内值边对应得权重  $\omega_{ij}^{\parallel} = \mathbf{c}_{ij}^{\parallel} \cdot \frac{\mathbf{c}_j - \mathbf{c}_i}{\|\mathbf{c}_j - \mathbf{c}_i\|_2}$ 。

### (3) 重平移平均化

得到经过全局优化的绝对位置集合  $\{\mathbf{c}_i^*\}$  后，再利用加权全局优化中的方法分别重新计算内边集合以及对当前已估计的绝对位置集合进行再一次的全局优化。

## 4.3 实验验证

本节对本章中提出的增量式平移平均化方法进行了较为详尽的验证，包括参数设置评测，关键技术的消融实验，种子敏感度验证，以及关于相对平移测量外值滤除与绝对位置估计精度、效率的对比实验等。需要注意的是，对于平移平均化化问题，需要提供绝对旋转将相对平移测量值变换至全局坐标系下，而为了公平起见，由于目前绝大部分平移平均化方法<sup>[9, 23, 32, 34~36]</sup>均使用 IRLS-GM 算法<sup>[7]</sup>的绝对旋转估计结果，本章方法也是如此。

### 4.3.1 参数设置评测

本章提出的增量式旋转平均化方法中主要有 4 个预先定义的参数，包括：1) 在进行初始四元组选取时的预选边数  $n_1$ ，2) 在进行下一最优视图选取时的预选顶点数  $n_2$ ，3) 平移角度距离阈值  $\theta_{th}^t$ ，以及 4) 用于控制局部与全局优化操作切换的增长比率  $r$ 。其中，本章方法对  $n_1$  与  $n_2$  值的变化并不敏感，在实验中分别将其设为 100 与 10。然而，参数  $\theta_{th}^t$  与  $r$  的值会对最终的平移平均化结果有着显著影响。因此，此处先对上述两参数的设置进行评测。

在参数设置评测的过程中，固定其中一个参数值而改变另一个，并在各测试数据上进行平移平均化操作，评测结果如表 4-1 所示。其中，参数  $\theta_{th}^t$  的变化范围与步长分别为  $[2^\circ, 8^\circ]$  以及  $2^\circ$ ；参数  $r$  的变化范围与步长分别为  $[120\%, 180\%]$  以及  $20\%$ 。表 4-1 的左半部分与右半部

分分别为 $\theta_{th}^t$ 与 $r$ 的参数设置测评结果。另外需要注意，在对参数 $\theta_{th}^t$ 进行设置评测时， $r$ 的值固定为150%；在对参数 $r$ 进行设置评测时， $\theta_{th}^t$ 的值固定为5°。

表 4-1 平移角度距离阈值 $\theta_{th}^t$ 与增长比率 $r$ 的参数设置评测结果

Data	# reconstructed cameras   estimation error/m   elapsed time/s on different values of $\theta_{th}^t$ and $r$ .							
	$\theta_{th}^t = 2^\circ$	$\theta_{th}^t = 4^\circ$	$\theta_{th}^t = 6^\circ$	$\theta_{th}^t = 8^\circ$	$r = 120\%$	$r = 140\%$	$r = 160\%$	$r = 180\%$
ALM	535 0.6 11	545 0.5 15	581 0.5 17	582 0.5 18	543 0.5 18	578 0.5 18	543 0.5 17	580 0.5 15
ELS	220 1.8 1	227 1.5 2	229 1.4 2	230 1.3 3	229 1.2 3	230 1.3 2	230 1.3 2	221 1.3 2
MDR	309 7.2 2	333 6.0 3	335 8.2 4	341 7.4 5	340 7.1 4	337 6.4 3	338 6.5 3	337 7.1 4
MND	431 0.6 7	447 0.4 9	451 0.4 11	451 0.4 13	447 0.4 16	447 0.4 10	446 0.4 10	446 0.4 10
NYC	311 1.0 2	324 0.7 3	329 0.7 4	331 0.6 4	329 0.7 5	327 0.7 3	327 0.7 3	327 0.7 3
PDP	306 0.8 3	319 0.6 5	326 0.7 5	330 0.6 6	328 0.7 7	326 0.6 5	327 0.6 5	328 0.7 5
PIC	1945 3.1 192	2190 2.0 359	2267 1.4 378	2276 1.2 432	2280 1.7 385	2215 1.7 384	2227 1.8 380	2229 2.0 378
ROF	1019 1.9 21	1050 1.6 25	1063 2.8 30	1065 2.0 32	1059 2.2 36	1059 2.7 29	1057 2.7 28	1056 3.0 27
TOL	431 2.6 5	446 1.8 7	457 1.6 7	463 1.7 8	453 1.7 10	456 1.9 8	455 1.8 7	450 1.9 7
USQ	621 4.8 5	691 5.6 8	705 5.3 9	725 5.4 13	695 5.4 10	695 5.2 8	689 5.2 8	693 5.3 8
VNC	745 1.0 18	771 1.0 24	818 1.2 28	831 1.1 29	814 1.1 35	778 1.1 30	781 1.1 26	798 1.2 24
YKM	390 0.9 3	410 0.8 5	418 0.8 5	419 0.8 6	411 0.8 7	413 0.8 5	414 0.8 4	411 0.8 4

由表 4-1 可知，由于随着 $\theta_{th}^t$ 取值的增大，在绝对位置计算与优化过程中会引入更多的相对平移测量值，这会使得标定的相机数量与运行时间均增大。另外，对于某些测试数据，如 ALM, ELS, MND, NYC, PIC, 以及 YKM, 随着 $\theta_{th}^t$ 取值的增大，绝对位置估计的中值误差下降；然而，对于另外一些测试数据，如 MDR, PDP, ROF, TOL, USQ, 以及 VNC, 在 $\theta_{th}^t$ 的取值与绝对位置的估计精度之间并不存在明显的关系。因此，参数 $\theta_{th}^t$ 对于本章方法有着明显且不规则的影响。对于参数 $r$ ，由表可知，运行时间整体上随着 $r$ 取值的增大会下降，而标定的相机数量与绝对位置的估计精度均不会受参数 $r$ 取值的明显影响。因此，为确保对足够的相机进行重建并且在估计精度与效率之间达到一个较好的平衡，本章实验中将 $\theta_{th}^t$ 与 $r$ 的值分别设置为5°与150%。

另外，由表 4-1 可知，不管是对于 $\theta_{th}^t$ 还是 $r$ ，均不存在能够使全部测试数据取得最优平移平均化结果的全局统一的参数设置。然而，作者认为对于某一参数在所有测试数据上的全局统一的参数设置可以提升方法的鲁棒性。

### 4.3.2 消融实验

基于上述参数设置测评，在此通过若干消融实验验证本章方法提出的各关键技术的有效性，包括基于局部优化的初始四元组选取（Local Optimization-based Initial Quaternary Selection, LOIQS），基于

加权支撑集合的下一最优视图选取（Weighted Supporting Set-based NBV Selection, WSSNS），以及加权局部/全局优化（Weighted Local/Global Optimization, WL/GO）与重平移平均化（Re-Translation Averaging, RTA）等。各消融实验结果如表 4-2 所示。在进行结果分析前，先对表 4-2 中不同的消融实验的情形简介如下。

表 4-2 不含 LOIQS, WSSNS, 加权, WGO, WLO, 或 RTA 情形下的消融实验结果, 其中 w/o 表示不包含某关键技术

Data	# reconstructed cameras   estimation error/m   elapsed time/s on different ablation study situations							
	w/o LOIQS	w/o WSSNS	w/o weighting	w/o RTA	w/o WGO	w/o WLO	Our ITA	
ALM	575 0.5 21	536 0.4 12	578 0.5 20	574 0.5 17	584 0.7 13	576 1.1 7	575 0.5 17	
ELS	227 1.2 2	227 1.4 2	227 1.3 3	227 1.5 2	229 2.0 1	228 2.6 1	230 1.3 2	
MDR	335 6.4 4	273 4.6 3	334 7.8 5	334 7.1 3	336 7.4 1	332 9.7 1	339 7.0 3	
MND	452 0.4 11	440 0.4 11	447 0.4 10	451 0.5 8	447 0.7 4	447 1.4 3	446 0.4 10	
NYC	323 1.0 4	282 0.6 2	330 0.7 4	327 0.7 2	327 0.9 1	326 1.3 1	327 0.6 3	
PDP	324 0.7 5	256 0.5 4	324 0.7 6	324 0.7 4	322 0.9 1	326 1.4 1	326 0.6 5	
PIC	2208 1.7 359	1969 1.7 261	2230 1.9 393	2226 2.0 383	2227 2.8 352	2214 3.6 349	2199 1.8 381	
ROF	1060 3.2 27	836 3.3 17	1060 2.8 26	1054 3.1 22	1051 4.3 16	1046 7.0 15	1062 2.3 28	
TOL	455 1.8 7	431 1.7 4	451 1.8 7	455 2.6 5	453 3.0 2	453 5.7 2	453 1.8 7	
USQ	691 9.3 8	560 7.5 5	694 9.8 9	696 5.4 8	707 7.3 6	697 8.0 5	703 5.6 8	
VNC	807 1.2 28	664 0.8 19	777 1.2 29	795 1.4 23	793 1.9 17	814 3.5 14	783 1.0 27	
YKM	415 0.9 6	367 0.8 4	415 0.8 5	411 0.9 5	410 1.0 2	412 2.1 2	414 0.8 5	

**不含 LOIQS:** 在这种情形下, 不进行 LOIQS 操作, 而将初始种子置为具有最小旋转回环偏差的相机三元组 (参见 4.2.1 节)。

**不含 WSSNS:** 在这种情形下, 不进行 WSSNS 操作, 下一最优视图选取为与当前已估计的绝对位置集合中共享边数最多的相机。

**不含加权:** 在这种情形下, 相对平移误差在整个流程中均不再考虑, 即对外极几何图上的所有边同等对待。具体来说, 在此情形下将 4.2.3 节中的  $\omega_{im}^I$  与  $\omega_{ij}^I$  的值均置为 1。

**不含 RTA:** 在这种情形下, 在每次加权全局优化后不再进行重平移平均化操作。

**不含 WGO:** 在这种情形下, 在增量式绝对位置计算的过程中, 加权全局优化与重旋转平均化均不再进行, 即在整个过程中仅对最新估计的绝对位置进行加权局部优化。

**不含 WLO:** 在这种情形下, 在增量式绝对位置计算的过程中, 不再有优化操作, 即各新计算的绝对位置仅将其定为通过基于加权支撑集合的下一最优视图选取过程获取的初值。

由表 4-2 可知: 1) 除了不含 WSSNS 这一消融实验情形外, 重建相机数量在所有消融实验情形中变化不大, 这说明本章提出的面向增量式平移平均化的 WSSNS 操作对于选取合适的下一最优视图以

及定位足够数量的相机是一个必要的保证; 2) 在各消融实验情形下, 几乎所有的测试数据上的平移平均化的估计误差均有增大, 这证实了本章提出的增量式平移平均化方法中的各项关键技术 (如 LOIQS, WSSNS, WL/GO, 以及 RTA 等) 的有效性; 3) 对于运行时间这一指标来说, 相对于原始的增量式平均化方法, 在各消融实验情形下, 运行时间均有所下降, 需要注意的是, 由于涉及到的优化操作与计算负担依次下降, 原始增量式平移平均化方法, 不含重平移平均化, 不含加权全局优化, 以及不含加权局部优化这四种情形的运行时间依次下降。

### 4.3.3 种子敏感度验证

由于增量式 SfM 对于初始种子选取较为敏感, 为验证本章提出的增量式平移平均化方法对初始种子选取的敏感程度, 在此通过比较本章方法与四次随机的初始种子选择, 进行了初始种子敏感度实验验证。各次随机初始种子选择通过每次从定义于 4.2.1 节中的四元组集合  $Q^*$  中随机选取一个四元组作为增量式平移平均化方法的初始种子实现。该种子敏感度验证实验结果如表 4-3 所示, 其中角度选择代价

$$(\text{angular selection cost}) \text{ 定义为 } \sum_{v_i, v_j \in \mathcal{V}_{q_{ijkl}}} \cos^{-1} \omega_{ij}; \mathcal{V}_{q_{ijkl}}, \mathcal{E}_{q_{ijkl}}, e_{ij} \in \mathcal{E}_{q_{ijkl}}$$

$\omega_{ij}$  均定义于 4.2.1 节。

表 4-3 初始种子选取敏感度实验结果

Data	Angular selection cost <sup>o</sup> / # reconstructed cameras / estimation error/m / elapsed time/s									
	Random selection from the pre-selected quaterinty set: random $\{C_{ijkl}   q_{ijkl} \in Q^*\}$									
	1st trial		2nd trial		3rd trial		4th trial		LOIQS	
ALM	8.64	575 0.5 20	12.70	576 0.5 19	5.37	575 0.5 17	3.77	573 0.5 17	0.30	575 0.5 17
ELS	15.34	227 1.3 2	3.56	230 1.4 2	25.44	227 1.3 2	5.94	227 1.3 2	1.00	230 1.3 2
MDR	27.00	335 7.7 5	56.02	333 7.6 4	5.54	341 8.1 4	41.22	335 8.0 4	1.46	339 7.0 3
MND	10.47	447 0.4 10	2.54	447 0.4 10	2.75	447 0.4 10	9.59	448 0.4 9	0.47	446 0.4 10
NYC	36.95	328 0.8 4	63.27	326 0.8 5	8.82	328 0.7 5	12.20	327 0.7 5	0.69	327 0.6 3
PDP	26.34	322 0.7 6	28.86	327 0.7 5	1.54	324 0.7 5	4.29	320 0.7 5	0.09	326 0.6 5
PIC	2.73	2211 1.8 361	87.70	2226 2.6 361	91.65	2206 1.9 374	23.69	2223 2.1 370	1.09	2199 1.8 381
ROF	1.14	1056 2.5 26	11.62	1062 3.0 26	9.15	1061 3.5 25	39.90	1060 2.6 26	0.23	1062 2.3 28
TOL	6.11	455 1.8 7	33.90	451 1.9 7	1.44	457 1.7 7	6.03	457 1.8 8	0.14	453 1.8 7
USQ	15.62	696 7.0 7	39.45	704 8.9 8	39.41	703 6.8 8	41.20	694 9.2 8	1.48	703 5.6 8
VNC	4.04	812 1.2 29	9.68	816 1.2 29	49.11	812 1.2 27	5.29	805 1.1 28	0.44	783 1.0 27
YKM	5.35	412 0.8 5	7.35	414 0.8 5	7.19	415 0.9 5	8.28	412 0.8 6	0.38	414 0.8 5

由表 4-3 可知: 1) 对于大多数测试数据, 即使随机选取的四元组中的边集存在外极几何外值边 (参见表 4-3 中的角度选取代价), 本章方法在绝对位置估计精度、效率, 以及重建相机数量等方面对初

始相机四元组的选取均不敏感；2) 尽管最终的平移平均化结果对初始种子的选取不敏感，更加精确的相机四元组的选取倾向于获得整体上更精确的绝对位置估计结果。因此，与增量式 SfM 不同，本章方法对初始种子的选取与构建不太敏感，这受益于相对于 SfM，平移平均化涉及的优化问题更为轻量化（更少的待估计参数量）以及本章提出的诸如 WSSNS 以及 WL/GO+RTA 等面向平移平均化的关键技术。

### 4.3.4 对比实验

#### (1) 测量外值滤除策略对比实验

首先在测量外值滤除策略方面，对本章方法与两个有代表性的外极几何图外值边滤除方法（基于绝对旋转估计辅助的 IRA 以及在面向平移平均化的代表性相对平移测量值滤除方法 1DSfM<sup>[9]</sup>）进行了对比实验，实验结果如表 4-4 所示。另外，在此除了对精度、召回率、以及 F-Score 三个指标进行比较外，不同外值滤除方法保留下来的内值边数， $|\mathcal{E}_X^I|$ ，以及内值边数真值， $|\mathcal{E}_{GT}^I|$ ，也列于表 4-4 中。不同于其它三个指标的值越高表现越好的评价方式，在内值边保留数量这一指标上的评价方式为最好的方法 X\* 其保留的内值边数与内值边数真值最为接近，即：

$$X^* = \arg \min_{X \in \{IRA, 1DSfM, ITA\}} \left| |\mathcal{E}_X^I| - |\mathcal{E}_{GT}^I| \right|$$

所有对比方法在保留边数、精度、以及 F-Score 指标上的最优与次优结果，以及在召回率指标上的次优与再次优结果（由于最优结果永远是原始外极几何图的 100%，无参考意义）在表 4-4 中分别由单下划线与双下划线标出。

表 4-4 相对平移测量外值滤除策略对比实验结果

Data	# reserved edges   Precision/%   Recall/%   F-Score/% on different relative translation outlier filtering strategies.																
	$ \mathcal{E}_{GT}^I $	Original EG			IRA			1DSfM <sup>[9]</sup>			Our ITA						
ALM	50671	97015	<u>52.23</u>	<u>100.00</u>	68.62	57555	<u>70.91</u>	<u>80.46</u>	<u>75.38</u>	74109	<u>60.04</u>	<u>87.68</u>	<u>71.28</u>	55508	<u>82.06</u>	<u>89.88</u>	<u>85.79</u>
ELS	7472	20010	<u>37.34</u>	<u>100.00</u>	54.38	<u>10642</u>	<u>59.49</u>	<u>84.46</u>	<u>69.81</u>	13346	<u>51.04</u>	<u>90.34</u>	<u>65.22</u>	8003	<u>77.08</u>	<u>82.16</u>	<u>79.54</u>
MDR	7572	23594	<u>32.08</u>	<u>100.00</u>	48.58	<u>7400</u>	<u>74.21</u>	<u>71.99</u>	<u>73.09</u>	16505	<u>40.15</u>	<u>86.71</u>	<u>54.89</u>	<u>7270</u>	<u>64.12</u>	<u>61.37</u>	<u>62.72</u>
MND	31877	52340	<u>60.90</u>	<u>100.00</u>	75.70	36240	<u>78.94</u>	<u>89.66</u>	<u>83.96</u>	38714	<u>71.88</u>	<u>87.19</u>	<u>78.80</u>	<u>33176</u>	<u>87.31</u>	<u>90.84</u>	<u>89.04</u>
NYC	8646	20569	<u>42.03</u>	<u>100.00</u>	59.19	<u>9529</u>	<u>74.24</u>	<u>81.34</u>	<u>77.63</u>	14406	<u>51.32</u>	<u>84.96</u>	<u>63.99</u>	9313	<u>79.14</u>	<u>85.21</u>	<u>82.06</u>
PDP	14960	24676	<u>60.63</u>	<u>100.00</u>	75.49	<u>17020</u>	<u>79.82</u>	<u>90.72</u>	<u>84.92</u>	17528	<u>71.62</u>	<u>83.80</u>	<u>77.23</u>	15821	<u>88.14</u>	<u>93.19</u>	<u>90.59</u>
PIC	111893	309418	<u>36.16</u>	<u>100.00</u>	53.11	<u>128541</u>	<u>68.32</u>	<u>77.86</u>	<u>72.78</u>	185296	<u>50.72</u>	<u>81.81</u>	<u>62.62</u>	<u>111082</u>	<u>77.63</u>	<u>76.53</u>	<u>77.07</u>
ROF	38321	70081	<u>54.68</u>	<u>100.00</u>	70.70	<u>37370</u>	<u>79.74</u>	<u>77.64</u>	<u>78.68</u>	43341	<u>68.04</u>	<u>76.83</u>	<u>72.17</u>	39635	<u>88.68</u>	<u>91.69</u>	<u>90.16</u>
TOL	15390	23782	<u>64.71</u>	<u>100.00</u>	78.58	13913	<u>86.39</u>	<u>77.84</u>	<u>81.89</u>	17266	<u>75.18</u>	<u>84.04</u>	<u>79.36</u>	<u>16104</u>	<u>89.06</u>	<u>93.14</u>	<u>91.06</u>
USQ	5908	24627	<u>23.94</u>	<u>100.00</u>	38.63	<u>12573</u>	<u>40.40</u>	<u>83.55</u>	<u>54.46</u>	17401	<u>32.17</u>	<u>90.86</u>	<u>47.52</u>	7294	<u>52.85</u>	<u>62.84</u>	<u>57.41</u>
VNC	55478	103198	<u>53.76</u>	<u>100.00</u>	69.93	<u>58921</u>	<u>79.75</u>	<u>84.49</u>	<u>82.05</u>	75377	<u>64.60</u>	<u>87.54</u>	<u>74.34</u>	59873	<u>86.01</u>	<u>92.74</u>	<u>89.25</u>
YKM	16543	27691	<u>59.74</u>	<u>100.00</u>	74.80	16276	<u>82.40</u>	<u>80.96</u>	<u>81.68</u>	20630	<u>69.74</u>	<u>86.83</u>	<u>77.35</u>	<u>17601</u>	<u>87.38</u>	<u>92.96</u>	<u>90.08</u>

由表 4-4 可知：1) 本章方法在精度与 F-Score 这两项指标上在绝大多数的测试数据上取得了最佳结果（除 MDR 外）。由于平移平均化问题通常为一个已知量冗余的超定问题，相比于召回率，精度是更为重要的评价指标，而本章方法相对于原始外极几何图在各测试数据上均能实现30%左右的精度指标提升，这证实了本章方法在滤除相对平移测量外值方面的有效性以及相对于其他方法的优势；2) 尽管 IRA 在相对旋转测量外值滤除方面取得了不错的表现，对于相对平移测量外值滤除任务而言，相较于 ITA，IRA 仍有较为明显的差距（相对于原始外极几何图的精度指标提升分别为约30%以及约20%），这表明对于同一条外极几何边，其相对旋转与相对平移的内外值情况并不一致，进而限制了如 BATA<sup>[35]</sup>等基于绝对旋转辅助的平移平均化方法的表现；3) 对于 1DSfM<sup>[9]</sup>方法来说，其仅取得了相对于原始外极几何图10%左右的精度指标提升，然而由于该方法可以保留大量外极几何边，在召回率指标上表现较好，因此可以像 ShapFit/Kick<sup>[34]</sup>等方法一样，将 1DSfM 用作本章方法的预处理步骤。

## (2) 参数估计精度、效率对比实验

对于平移平均化中绝对位置估计精度、效率对比实验，在此将本章方法与若干有代表性的平移平均化方法进行了对比，包括一种基于特征轨迹的方法 SATA<sup>[23]</sup>，一种基于外极几何图优化的方法 VGO<sup>[24]</sup>，两种基于鲁棒优化的方法 SFSK<sup>[34]</sup>以及 BATA<sup>[35]</sup>，以及一种基于秩约束的方法 ACEM<sup>[38]</sup>。对比结果如表 4-5 所示，其中在重建相机数量，绝对位置估计精度、效率三个方面的最优与次优结果均通过单下划线与双下划线分别标出。需要注意的是，SFSK<sup>[34]</sup>方法在 YKM 数据上的结果以及 ACEM<sup>[38]</sup>在 PIC，ROF，以及 USQ 数据上的结果是缺失的；另外，SFSK<sup>[34]</sup>与 BATA<sup>[35]</sup>均未报导重建相机数量的情况。另外，与其它几乎所有的平移平均化方法<sup>[23~24, 32, 34~36, 38]</sup>一样，此处其它对比方法的运行时间直接从其原始文章中摘录。由于在此仅对平移平均化本身进行效率对比，因此表 4-5 中仅列举了不同方法的平移平均化本身的运行时间而忽略了如特征匹配、旋转平均化、最终的全局捆绑调整等操作的运行时间。尽管由于对比方法不开源的问题使得在同一平台下对所有方法进行效率对比几乎不太可能，为更直观地理解此处的效率对比情况，在此将各方法中使用的 CPU 主频同样列于表 4-5 中。

表 4-5 平移平均化重建相机数量与参数估计精度、效率对比实验结果

Data	# reconstructed cameras   estimation error/m   elapsed time/s on different translation averaging methods								
	SATA <sup>[23]</sup>	VGO <sup>[24]</sup>	SFSK <sup>[34]</sup>	BATA <sup>[35]</sup>	ACEM <sup>[38]</sup>	Our ITA	ITA w/ IRA	ITA w/ IRA*	
	2.30GHz	2.26GHz	2.60GHz	3.40GHz	3.60GHz	2.60GHz	2.60GHz	2.60GHz	
ALM	574 0.5 78	533 1.4 69	- 0.9 27	- 0.6 24	482 1.2 172	575 0.5 17	577 0.5 17	523 0.4 10	
ELS	223 2.5 37	203 3.7 19	- 1.9 4	- 1.5 2	211 6.1 87	230 1.3 2	227 1.1 2	224 1.0 2	
MDR	317 2.7 31	272 8.7 67	- 6.0 9	- 1.8 4	168 6.9 23	339 7.0 3	336 6.5 4	294 1.5 2	
MND	452 0.4 62	416 2.0 133	- 0.8 19	- 0.3 10	416 1.0 224	446 0.4 10	452 0.4 10	429 0.4 9	
NYC	338 0.8 38	294 2.8 71	- 1.4 11	- 0.6 4	277 2.2 67	327 0.6 3	328 0.7 4	304 0.4 3	
PDP	340 2.0 43	302 2.9 23	- 3.6 7	- 4.2 4	275 3.5 60	326 0.6 5	322 0.6 4	299 0.5 5	
PIC	2276 1.3 328	1928 5.2 544	- 1.2 464	- 1.0 114	-  -	2199 1.8 381	2231 1.5 402	1929 0.7 268	
ROF	1077 2.9 131	966 6.8 385	- 4.3 62	- 1.6 29	-  -	1062 2.3 28	1053 2.6 27	997 2.2 22	
TOL	465 1.9 52	409 9.3 145	- 2.3 18	- 2.2 7	414 5.0 121	453 1.8 7	452 2.0 7	422 1.9 5	
USQ	570 5.5 41	701 4.5 141	- 8.9 28	- 4.3 10	-  -	703 5.6 8	697 5.3 9	588 3.0 6	
VNC	842 2.7 117	771 6.7 185	- 1.9 74	- 1.9 30	674 4.2 273	783 1.0 27	778 1.2 26	720 0.8 17	
YKM	417 2.3 46	409 3.9 31	-  -	- 0.9 9	341 2.7 91	414 0.8 5	411 0.8 6	385 0.7 6	

由表 4-5 可知，本章提出的增量式平移平均化方法在参数估计精度与效率方面均取得了整体最优结果。在各对比方法之中，SATA<sup>[23]</sup>与 BATA<sup>[35]</sup>在精度方面取得了略低于本章方法的结果，然而它们为实现精度的提升，或者需要特征轨迹与局部捆绑调整（SATA<sup>[23]</sup>），或者依赖于基于 LUD<sup>[32]</sup>的初始化与基于旋转辅助的优化（BATA<sup>[35]</sup>）。若去除上述因素，SATA<sup>[23]</sup>与 BATA<sup>[35]</sup>方法的精度均会下降，而本章方法并不依赖上述任何因素。对于重建相机数量这一指标，本章方法仅略差于基于特征轨迹的方法 SATA<sup>[23]</sup>，这是由于该方法中通过引入特征轨迹的方式提供了远多于原始平移平均化问题的约束条件。

另外，为验证本章提出的增量式平移平均化方法对不同旋转平均化结果的适应性，在此将 IRLS-GM<sup>[7]</sup>的绝对旋转估计结果替换为 IRA，并将该方法记为 ITA w/ IRA，实验结果列于表 4-5 第 8 列。由表可知，ITA w/ IRA 的相机绝对位置的估计精度略高于原始 ITA（在全部测试数据上的精度误差均值：1.93m vs. 1.98m），这与第 2 章中旋转平均化精度对比结果一致。另外，对于原始 ITA 与 ITA w/ IRA，均将原始外极几何图作为平移平均化问题的输入，但 IRA 方法本身可以在估计绝对旋转的过程中可以给出各外极几何边的内外值标签。因此，对于 ITA w/ IRA，可以将通过 IRA 过程外值边滤除后的外极几何图作为输入，并将该方法记为 ITA w/ IRA\*，实验结果列于表 4-5 第 9 列。由表可知，对于绝对位置估计精度，ITA w/ IRA\*方法实现了明显的提升，这验证了本章提出的 ITA 方法在引入如外值滤除等鲁棒操作后可以取得更好的结果。但是，由于外值边滤除后的外极几何图中边的数量有所减少，ITA w/ IRA\*方法重建的相机数量以及运行时间也相应的减少与缩短。

## 4.4 本章小结

本章通过将增量式 SfM 中的增量式参数估计思想引入平移平均化问题，提出了一种称作增量式平移平均化的新型平移平均化方法。该方法结合了增量式参数估计流程精度高与鲁棒性强以及全局式运动平均化方法简单、高效的优点。在实验验证环节，对本章方法进行了全面评测以验证其有效性；另外，还将本章方法与若干代表性的平移平均化方法进行对比，验证了本章方法在精度、鲁棒性、简洁性、以及效率等方面的优势。



## 5. 总结与展望

### 5.1 总结

本文对基于图像的大规模场景三维重建中的运动平均化方法进行了系统性研究。研究工作主要包括如下三个方面：

1) 针对外极几何图中相对旋转测量外值过多，现有方法对其鲁棒性较差的问题，提出了基于增量式参数估计的增量式旋转平均化方法，可实现在滤除相对旋转测量外值的同时对绝对旋转进行精确、鲁棒估计；

2) 针对在大规模场景下增量式旋转平均化方法与基于 RANSAC 的旋转平均化方法分别存在的误差累积、拓展性差以及最小集尺寸过大、一致集获取困难等问题，将分而治之的思想引入上述两方法，提出了分布增量式旋转平均化与层级 RANSAC 旋转平均化方法，提升了算法的精度、效率、鲁棒性以及拓展性；

3) 针对平移平均化问题的特点与难点，在相机位置算子、流程关键技术、优化目标函数等方面进行了改进，将增量式旋转平均化方法拓展至平移平均化问题，实现了对该问题的精确、高效、简洁、鲁棒的求解。

### 5.2 展望

尽管本文对大规模运动平均化方法的高效性与鲁棒性问题进行了一些初步的探索并取得了一些结果，但仅通过本文工作远不能完全解决大规模运动平均化问题，后续可考虑从如下三个方面开展工作：

1) 将本文提出的层级式旋转平均化流程引入平移平均化问题，以构建层级式平移平均化流程，实现对大规模平移平均化问题的高效、鲁棒求解；

2) 本文提出的层级式旋转平均化流程中采用的基于受限社区发现的外极几何图聚类算法为现有方法，其作为外极几何图的预处理步骤，与旋转平均化问题相对独立。后续可考虑研究一种专门用于层级式运动平均化的图聚类方法，可在估计相机绝对位姿的同时对大规模外极几何图进行聚类；

3) 本文提出的增量式平移平均化方法采用线性三焦张量算法对相机绝对位置进行初始化,在相机做直线运动时其求解不确定性增大,这会导致相机绝对位置估计的误差累积现象。后续可以考虑分别研究增量式尺度平均化以及已知尺度的增量式平移平均化方法,以克服上述问题。

## 参考文献

- [1] N. Snavely, S. M. Seitz, and R. Szeliski, “Modeling the world from Internet photo collections,” *International Journal of Computer Vision*, vol. 80, no. 2, pp. 189–210, 2008.
- [2] Y. Furukawa and J. Ponce, “Accurate, Dense, and Robust Multiview Stereopsis,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 8, pp. 1362–1376, 2010.
- [3] M. Jancosek and T. Pajdla, “Multi-view reconstruction preserving weakly-supported surfaces,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011, pp. 3121–3128.
- [4] M. Waechter, N. Moehrle, and M. Goesele, “Let there be color! Large-scale texturing of 3D reconstructions,” in *European Conference on Computer Vision (ECCV)*, 2014, pp. 836–850.
- [5] C. Wu, “Towards linear-time incremental structure from motion,” in *International Conference on 3D Vision (3DV)*, 2013, pp. 127–134.
- [6] J. L. Schonberger and J. M. Frahm, “Structure-from-motion revisited,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 4104–4113.
- [7] A. Chatterjee and V. M. Govindu, “Efficient and robust large-scale rotation averaging,” in *IEEE International Conference on Computer Vision (ICCV)*, 2013, pp. 521–528.
- [8] P. Moulon, P. Monasse, and R. Marlet, “Global fusion of relative motions for robust, accurate and scalable structure from motion,” in *IEEE International Conference on Computer Vision (ICCV)*, 2013, pp. 3248–3255.
- [9] K. Wilson and N. Snavely, “Robust global translations with 1DSfM,” in *European Conference on Computer Vision (ECCV)*, 2014, pp. 61–75.
- [10] L. Zhu, S. Shen, X. Gao, and Z. Hu, “Large scale urban scene modeling from MVS meshes,” in *European Conference on Computer Vision (ECCV)*, 2018, pp. 640–655.
- [11] J. Han, M. Rong, H. Jiang, H. Liu, and S. Shen, “Vectorized indoor surface reconstruction from 3D point cloud with multistep 2D optimization,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 177, pp. 57–74, 2021.
- [12] J. Sun, Y. Xie, L. Chen, X. Zhou, and H. Bao, “NeuralRecon: Real-time coherent 3D reconstruction from monocular video,” in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021, pp. 15598–15607.
- [13] X. Gao, S. Shen, Y. Zhou, H. Cui, L. Zhu, and Z. Hu, “Ancient Chinese architecture 3D preservation by merging ground and aerial point clouds,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 143, pp. 72–84, 2018.
- [14] N. Snavely, I. Simon, M. Goesele, R. Szeliski, and S. M. Seitz, “Scene reconstruction and visualization from community photo collections,” *Proceedings of the IEEE*, vol. 98, no. 8, pp. 1370–1390, 2010.
- [15] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [16] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, “Speeded-up robust features (SURF),” *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.
- [17] D. Nister and H. Stewenius, “Scalable recognition with a vocabulary tree,” in *IEEE*

- Conference on Computer Vision and Pattern Recognition (CVPR), 2006, pp. 2161–2168.
- [18] M. Muja and D. G. Lowe, “Scalable nearest neighbor algorithms for high dimensional data,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 11, pp. 2227–2240, 2014.
- [19] R. Hartley, “In defense of the eight-point algorithm,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 6, pp. 580–593, 1997.
- [20] D. Nister, “An efficient solution to the five-point relative pose problem,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 6, pp. 756–770, 2004.
- [21] M. A. Fischler and R. C. Bolles, “Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography,” *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.
- [22] S. Agarwal, N. Snavely, S. M. Seitz, and R. Szeliski, “Bundle adjustment in the large,” in *European Conference on Computer Vision (ECCV)*, 2010, pp. 29–42.
- [23] Z. Cui and P. Tan, “Global structure-from-motion by similarity averaging,” in *IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 864–872.
- [24] C. Sweeney, T. Sattler, T. Hollerer, M. Turk, and M. Pollefeys, “Optimizing the viewing graph for structure-from-motion,” in *IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 801–809.
- [25] R. Hartley, K. Aftab, and J. Trunpf, “L1 rotation averaging using the Weiszfeld algorithm,” in *IEEE conference on computer vision and pattern recognition (CVPR)*, 2011, pp. 3041–3048.
- [26] D. Crandall, A. Owens, N. Snavely, and D. Huttenlocher, “SfM with MRFs: Discrete-continuous optimization for large-scale structure from-motion,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 12, pp. 2841–2853, 2013.
- [27] A. Chatterjee and V. M. Govindu, “Robust relative rotation averaging,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 958–972, 2018.
- [28] Y. Shi and G. Lerman, “Message passing least squares framework and its application to rotation synchronization,” in *International Conference on Machine Learning (ICML)*, 2020.
- [29] V. M. Govindu, “Robustness in motion averaging,” in *Asian conference on computer vision (ACCV)*, 2006, pp. 457–466.
- [30] C. Zach, M. Klopschitz, and M. Pollefeys, “Disambiguating visual relations using loop constraints,” in *IEEE conference on computer vision and pattern recognition (CVPR)*, 2010, pp. 1426–1433.
- [31] H. Cui, S. Shen, W. Gao, H. Liu, and Z. Wang, “Efficient and robust large-scale structure-from-motion via track selection and camera prioritization,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 156, pp. 202–214, 2019.
- [32] O. Ozyesil and A. Singer, “Robust camera location estimation by convex programming,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 2674–2683.
- [33] V. M. Govindu, “Combining two-view constraints for motion estimation,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2001, pp. 218–225.
- [34] T. Goldstein, P. Hand, C. Lee, V. Voroninski, and S. Soatto, “ShapeFit and ShapeKick for robust, scalable structure from motion,” in *European Conference on Computer Vision (ECCV)*,

2016, pp. 289–304.

- [35] B. Zhuang, L. Cheong, and G. H. Lee, “Baseline desensitizing in translation averaging,” in IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 4539–4547.
- [36] Z. Cui, N. Jiang, C. Tang, and P. Tan, “Linear global translation estimation with feature tracks,” in British Machine Vision Conference (BMVC), 2015, pp. 46.1–46.13.
- [37] N. Jiang, Z. Cui, and P. Tan, “A global linear method for camera pose registration,” in IEEE International Conference on Computer Vision (ICCV), 2013, pp. 481–488.
- [38] Y. Kasten, A. Geifman, M. Galun, and R. Basri, “Algebraic characterization of essential matrices and their averaging in multiview settings,” in IEEE/CVF International Conference on Computer Vision (ICCV), 2019, pp. 5894–5902.
- [39] A. Geifman, Y. Kasten, M. Galun, and R. Basri, “Averaging essential and fundamental matrices in collinear camera settings,” in IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 6020–6029.
- [40] K. Wilson, D. Bindel, and N. Snavely, “When is rotations averaging hard?” in European Conference on Computer Vision (ECCV), 2016, pp. 255–270.
- [41] A. P. Bustos, T. Chin, A. Eriksson, and I. Reid, “Visual SLAM: Why bundle adjust?” in International Conference on Robotics and Automation (ICRA), 2019, pp. 2385–2391.
- [42] C. K. Chng, A. Parra, T. J. Chin, and Y. Latif, “Monocular rotational odometry with incremental rotation averaging and loop closure,” in Digital Image Computing: Techniques and Applications (DICTA), 2020, pp. 1–8.
- [43] P. Purkait, T. J. Chin, and I. Reid, “NeuRoRA: Neural robust rotation averaging,” in European Conference on Computer Vision (ECCV), 2020, pp. 137–154.
- [44] L. Yang, H. Li, J. A. Rahim, Z. Cui, and P. Tan, “End-to-end rotation averaging with multi-source propagation,” in IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021, pp. 11774–11783.
- [45] T. Shen, S. Zhu, T. Fang, R. Zhang, and L. Quan, “Graph-based consistent matching for structure-from-motion,” in European Conference on Computer Vision (ECCV), 2016, pp. 139–155.
- [46] A. Eriksson, C. Olsson, F. Kahl, and T. Chin, “Rotation averaging and strong duality,” in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 127–135.
- [47] X. Wang, T. Xiao, and Y. Kasten, “A hybrid global structure from motion method for synchronously estimating global rotations and global translations,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 174, pp. 35–55, 2021.
- [48] Q. Dong, X. Gao, H. Cui, and Z. Hu, “Robust camera translation estimation via rank enforcement,” *IEEE Transactions on Cybernetics*, vol. 52, no. 2, pp. 862–872, 2022.
- [49] H. Cui, X. Gao, S. Shen, and Z. Hu, “HSfM: Hybrid structure-from-motion,” in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 2393–2402.
- [50] S. Haner and A. Heyden, “Covariance propagation and next best view planning for 3D reconstruction,” in European Conference on Computer Vision (ECCV), 2012, pp. 545–556.
- [51] H. Cui, S. Shen, X. Gao, and Z. Hu, “CSfM: Community-based structure from motion,” in

- IEEE International Conference on Image Processing (ICIP), 2017, pp. 4517–4521.
- [52] S. Zhu, R. Zhang, L. Zhou, T. Shen, T. Fang, P. Tan, and L. Quan, “Very large-scale global SfM by distributed motion averaging,” in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 4568–4577.
- [53] L. Zhou, Z. Luo, M. Zhen, T. Shen, S. Li, Z. Huang, T. Fang, and L. Quan, “Stochastic bundle adjustment for efficient and scalable 3D reconstruction,” in European Conference on Computer Vision (ECCV), 2020, pp. 364–379.
- [54] R. Hartley, J. Trumpf, Y. Dai, and H. Li, “Rotation averaging,” *International Journal of Computer Vision*, vol. 103, p. 267–305, 2013.
- [55] I. S. Dhillon, Y. Guan, and B. Kulis, “Weighted graph cuts without eigenvectors a multilevel approach,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 11, pp. 1944–1957, 2007.
- [56] B. Bhowmick, S. Patra, A. Chatterjee, V. M. Govindu, and S. Banerjee, “Divide and conquer: Efficient large-scale structure from motion using graph partitioning,” in Asian Conference on Computer Vision (ACCV), 2014, pp. 273–287.

## 致谢

行文至此，谨向从事博士后研究工作两年多以来关心、帮助过我的人致以最诚挚的感谢。

首先要感谢我的博士后合作导师解则晓教授。是解老师在十年前将我引入了三维视觉的大门，让我对这一理论与应用完美融合的领域产生了浓厚的兴趣，并最终确立为自己终身学习的方向与奋斗的目标。从事博士后研究工作以来，时常与解老师就研究方向、发展路线、职业规划等方面进行深入讨论。解老师知无不言，为我指引方向、确立目标、排忧解难。解老师作为师长，虽为吾师、更似长辈，让我时时刻刻都能感受到一位长辈对晚辈的关怀与温暖。

我还要感谢我的博士研究生导师胡占义研究员与申抒含研究员。虽然我已经博士毕业并离开了自动化所，但是仍与两位老师保持着联系。还记得在 2021 年毕业后第一次回所时胡老师对我的教诲。胡老师教导我带学生应当在质不在量，应当事必躬亲，与学生共同进步、共同提高，而不应当做甩手掌柜，当“老板”。申老师这边毕业之后更是一直保持着较为密切的联系，在师生关系维持的同时也建立了良好的合作关系。在过去两年多的时间里，申老师虽然事务缠身，仍不厌其烦地对我在研究方向、职业发展、论文撰写、项目申请等方面进行悉心指导。虽取得的成绩只是皮毛，但这些也与申老师的帮助最为密不可分。是两位老师的培养让我从一个刚入门的学生逐渐成长为一个具有一定的独立从事科研工作能力的研究者，师恩永不敢忘。

感谢中国海洋大学自动化及测控系的其他师长，褚东升教授、黎明教授、李庆忠教授、宋大雷教授、綦声波教授、杨荣荣老师、迟书凯老师等，学生时代就受到您们的教诲，工作以来又获得您在工作、生活等各方面的帮助，使我能够尽快适应新的人生阶段，完成由学生到老师，由研究生到科研工作者的身份的转变。感谢系里的“同辈”

老师，李坤乾老师、李崇老师、姚鹏老师、陈震老师等，诸位老师作为系里的新一代骨干力量，各方面都十分优秀，也在工作、生活上为我提供了诸多帮助。诸位老师让我看到了自身的差距与不足，在过去的两年多时间里，我一直以诸位老师为榜样，不敢有丝毫懈怠，以使自己不枉处于这样一个优秀的集体之中。

感谢北京科技大学刘红敏教授、樊彬教授，中国科学院自动化研究所董秋雷研究员、崔海楠老师，与诸位老师的合作拓展了我的视野，丰富了我的知识，让我更有信心地沿着科研道路继续前行。

感谢阿里巴巴达摩院 XR 实验室朱灵杰博士，学生时期就与灵杰建立了良好的合作关系并一直维持至今。灵杰敏捷的思维，开阔的视野，严谨的态度，规范的作风都对我影响颇深。

感谢和我一起工作、学习过的同学们，包括研究生：邵玮婧、叶旋、陶冬冬、于筱萱、李梦晗、马孝冬等，本科生：公劲夫、李悦、罗家铮、王欣宇、张秉钧、黄鹏、黄媛媛、李元戎、亓颖颖、孙英明、孙睿璟、王格格等，与诸位共同成长、共同进步，是我弥足珍贵的经验与财富。

感谢我的父母以及我的爱人吴焕芝对我的支持与理解，你们是我坚强的后盾和温暖的依靠，感谢我的儿子高铭泽，你是我快乐与灵感的最重要来源。

本文的完成离不开中国博士后科学基金会，国家自然科学基金委员会，山东省科技厅，中国科学院自动化研究所模式识别国家重点实验室，青岛市人力资源和社会保障局，以及中国海洋大学水利工程博士后流动站、工程学院、人事处、科技处等单位、部门的支持，在此一并感谢。

最后，感谢各位专家百忙之中前来参加我的博士后出站答辩。

2022 年 4 月于青岛

## 博士生期间发表的学术论文、专著

- [1] **X. Gao**, S. Shen, Y. Zhou, H. Cui, L. Zhu, and Z. Hu, “Ancient Chinese architecture 3D preservation by merging ground and aerial point clouds,” *ISPRS Journal of Photogrammetry Remote Sensing*, vol. 143, pp. 72–84, 2018. (SCI, IF: 8.979, h5-index: 82)
- [2] **X. Gao**, L. Hu, H. Cui, S. Shen, and Z. Hu, “Accurate and efficient ground-to-aerial model alignment,” *Pattern Recognition*, vol. 76, pp. 288–302, 2018. (SCI, IF: 7.740, CCF-B, h5-index: 99)
- [3] **X. Gao**, S. Shen, Z. Hu, and Z. Wang, “Ground and aerial meta-data integration for localization and reconstruction: A review,” *Pattern Recognition Letters*, vol. 127, pp. 202–214, 2019. (SCI, IF: 3.756, CCF-C, h5-index: 66)
- [4] **X. Gao**, H. Cui, L. Zhu, T. Shi, and S. Shen, “Multi-source data-based 3D digital preservation of large-scale ancient Chinese architecture: A case report,” *Virtual Reality & Intelligent Hardware*, vol. 1, no. 5, pp. 525–541, 2019.
- [5] H. Cui, **X. Gao**, S. Shen, and Z. Hu, “HSfM: Hybrid structure-from-motion,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 1212–1221. (EI, CCF-A, h5-index: 356)
- [6] L. Zhu, S. Shen, **X. Gao**, and Z. Hu, “Large scale urban scene modeling from MVS meshes,” in *European Conference on Computer Vision (ECCV)*, 2018, pp. 640–655. (EI, CCF-B, h5-index: 197)
- [7] H. Cui, S. Shen, **X. Gao**, and Z. Hu, “Batched incremental structure-from-motion,” in *International Conference on 3D Vision (3DV)*, 2017, pp. 205–214. (EI, CCF-C, h5-index: 44)
- [8] T. Shi, S. Shen, **X. Gao** and L. Zhu, “Visual localization using sparse semantic 3D map,” in *IEEE International Conference on Image Processing (ICIP)*, 2019, pp. 315–319. (EI, CCF-C, h5-index: 60)
- [9] H. Cui, S. Shen, **X. Gao**, and Z. Hu, “CSfM: Community-based structure from motion,” in *IEEE International Conference on Image Processing (ICIP)*, 2017, pp. 4517–4521. (EI, CCF-C, h5-index: 60)
- [10] Y. Zhou, S. Shen, **X. Gao**, and Z. Hu, “Accurate mesh-based alignment for ground and aerial multi-view stereo models,” in *IEEE International Conference on Image Processing (ICIP)*, 2017, pp. 2627–2631. (EI, CCF-C, h5-index: 60)



## 博士后期间发表的学术论文、专著

- [1] **X. Gao**, L. Zhu, Z. Xie, H. Liu, and S. Shen, “Incremental rotation averaging,” *International Journal of Computer Vision*, vol. 129, pp. 1202–1216, 2021. (SCI, IF: 7.410, CCF-A, h5-index: 72)
- [2] **X. Gao**, L. Zhu, H. Cui, Z. Xie, and S. Shen, “IRA++: Distributed incremental rotation averaging,” *IEEE Transactions on Circuits and Systems for Video Technology*, 2021. (SCI, IF: 4.685, CCF-B, h5-index: 77)
- [3] **X. Gao**, S. Shen, L. Zhu, T. Shi, Z. Wang, and Z. Hu, “Complete scene reconstruction by merging images and laser scans,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 10, pp. 3688–3701, 2020. (SCI, IF: 4.685, CCF-B, h5-index: 77)
- [4] **X. Gao**, J. Luo, K. Li, and Z. Xie, “Hierarchical RANSAC-based rotation averaging,” *IEEE Signal Processing Letters*, vol. 27, pp. 1874–1878, 2020. (SCI, IF: 3.109, CCF-C, h5-index: 65)
- [5] **X. Gao**, L. Zhu, H. Cui, Z. Hu, H. Liu and S. Shen, “Complete and accurate indoor scene capturing and reconstruction using a drone and a robot,” *IEEE Sensors Journal*, vol. 21, no. 10, pp. 11858–11869, 2021. (SCI, IF: 3.301, h5-index: 79)
- [6] Q. Dong, **X. Gao**, H. Cui, and Z. Hu, “Robust camera translation estimation via rank enforcement,” *IEEE Transactions on Cybernetics*, vol. 52, no. 2, pp. 862–872, 2022. (SCI, IF: 11.448, CCF-B, h5-index: 127)
- [7] J. Han, L. Zhu, **X. Gao**, Z. Hu, L. Zhou, H. Liu, and S. Shen, “Urban scene LOD vectorized modeling from photogrammetry meshes,” *IEEE Transactions on Image Processing*, vol. 30, pp. 7458–7471, 2021. (SCI, IF: 10.856, CCF-A, h5-index: 123)
- [8] L. Zhu, S. Shen, **X. Gao**, and Z. Hu, “Urban scene vectorized modeling based on contour deformation,” *ISPRS International Journal of Geo-Information*, vol. 9, no. 3, pp. 162, 2020. (SCI, IF: 2.899, h5-index: 47)
- [9] S. Chi, X. Ye, **X. Gao**, Z. Xie, and D. Tao, “Coded marker-based high-accuracy motion estimation (in Chinese),” *Optics and Precision Engineering*, vol. 29, no. 7, pp. 1720–1730, 2021. (Correspondence, EI, h5-index: 10)
- [10] Z. Xie, W. Shao, **X. Gao**, H. Gong, H. Wang, and Y. Jiao, “Discrete epipolar curve model-based underwater multiple line-structured light binocular measuring method (in Chinese),” *Chinese Journal of Lasers*, vol. 48, no. 13, pp. 1304003, 2021. (Correspondence, EI, h5-index: 9)



## 个人简历

1. 2008年9月至2012年6月，就读于中国海洋大学工程学院自动化专业，获工学学士学位；
2. 2012年9月至2015年6月，就读于中国海洋大学工程学院控制理论与控制工程专业，获工学硕士学位，导师解则晓教授；
3. 2015年9月至2019年6月，就读于中国科学院自动化研究所模式识别与智能系统专业，获工学博士学位，导师胡占义研究员、申抒含研究员；
4. 2019年8月至今，进入中国海洋大学工程学院水利工程流动站从事博士后研究工作，合作导师解则晓教授。

