

清华 大学

综合 论文 训 练

题目：基于几何距离的场景分析与重构

系 别：交叉信息研究院

专 业：计算机科学与技术（计算机科学实验班）

姓 名：赵梓硕

指导教师：胡事民教授

2019 年 6 月 4 日

关于学位论文使用授权的说明

本人完全了解清华大学有关保留、使用学位论文的规定，即：学校有权保留学位论文的复印件，允许该论文被查阅和借阅；学校可以公布该论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存该论文。

(涉密的学位论文在解密后应遵守此规定)

签 名：赵梓硕 导师签名：胡秉氏 日 期：2019-6-5

中文摘要

本文利用欧氏空间中特征点之间的距离关系，设计了一种能在无先验知识的情况下将场景中运动的刚体进行聚类并计算其轨迹的算法，该算法能够有效地同时提升对室内运动场景进行分析的准确度和效率。

同时，本文提出了对一系列 MDS 算法的一种新评价方式“几何稳定性”，能够相比传统的基于点对间距离的评价标准更加真实地体现 MDS 对点集整体空间结构的还原精度，并证明了 cMDS 在一般情况下确实能够在输入距离存在独立高斯噪声时较为准确地还原点集的整体空间结构。并初步分析了基于有噪声稀疏距离图进行 MDS 的效果，讨论了在可能还原完整点集几何结构时稀疏度的理论极限，并提出了一种在更加稀疏情况下移除部分点以正确还原其余点几何结构的思路。这些结果为无线传感器网络定位、卫星定位等问题的精度和理论极限提供了理论依据。

之后，本文参考前人关于最小全局刚性生成子图的工作，以另一种角度提出了一种可生成欧氏空间中具有几何稳定性的最小代价连接图的算法，可以在边数与总代价均为线性的情况下对无线传感器网络进行可靠的定位。

关键词：计算几何；低秩近似；几何重建；压缩感知；组合优化

ABSTRACT

We designed an algorithm to cluster and trace moving rigid bodies in dynamic scenes via pairwise distances of landmarks in the Euclidean space without prior knowledge, which can effectively improve the accuracy and efficiency of indoor scene analysis in one shot.

We also proposed a new criterion *geometrical stability* for evaluation of MDS algorithms, which can more truthfully represent the fidelity of global structure instead of local pointwise distances, and proved the capability of the cMDS algorithm to reconstruct a generic structure under independent Gaussian noises, and made a preliminary analysis on the performance of MDS on sparse noisy distance graphs, and discussed about the theoretical limit of sparsity when the geometry can still be recovered. We also proposed an idea to remove some points and recover the geometric structure of the remaining majority of points when the graph is sparser than the limit, which renders a rigorous theoretical guarantee for the precision and theoretical bounds for wireless sensor network localization and satellite positioning.

Furthermore, we looked into previous work about minimum-cost global rigid spanning subgraph, and proposed an algorithm to construct a minimum-cost geometric-stable spanning graph in the Euclidean space, which enabled reliable wireless sensor network localization with linear numbers of edges and total cost.

Keywords: Computational Geometry; Low-Rank Approximation; Geometrical Reconstruction; Compressive Sensing ; Combinatorial Optimization

目 录

第 1 章 引言	1
1.1 课题背景	1
1.2 相关工作	1
1.2.1 欧氏嵌入与矩阵填充	1
1.2.2 欧氏距离矩阵的优化	4
1.2.3 随机图模型与图论理论	4
1.3 本文贡献	6
1.4 论文组织	6
第 2 章 ClusterSLAM 中基于几何距离的特征点聚类	8
2.1 ClusterSLAM 简介	8
2.2 面向 SLAM 的特征点聚类算法	9
2.2.1 基于特征点相似性的聚类	9
2.2.2 特征点相似性的量化方式	10
2.2.3 基准点策略：对于非完全观测的解决方案	11
2.3 算法分析与应用	12
2.3.1 算法分析	12
2.3.2 在场景分析中的应用	13
2.4 小结	14
第 3 章 对 MDS 算法的分析与探索	17
3.1 几何稳定的 MDS 算法	17
3.2 cMDS 的几何稳定性分析	18
3.2.1 对噪声 2-范数的估计	18
3.2.2 对真实 Gram 矩阵特征值与谱间隔的估计	19
3.2.3 对特征向量和特征值误差的估计	20
3.3 对于有噪声非完全观测情况的初步分析	22

3.3.1 实验探索	23
3.4 基于随机稀疏图的 MDS 之可行性	24
3.4.1 无向图的刚性性质	26
3.4.2 弱 MDS：对于 $\Theta(1/n)$ -稀疏之情况的重构可能性	26
3.4.3 高维情形	29
3.5 小结	30
第 4 章 面向几何稳定性的网格优化设计与传感器定位	32
4.1 概述	32
4.2 无线传感器定位模型	33
4.3 基于基准点框架的网格构造	35
4.3.1 贪心的均匀采样方案	35
4.3.2 全局刚性的证明	39
4.3.3 二维几何稳定性的证明	39
4.4 对比分析	43
第 5 章 结论	45
插图索引	46
公式索引	48
参考文献	51
致 谢	53
声 明	54
附录 A 外文资料的调研阅读报告	55

主要符号对照表

a.a.s.	渐进几乎必然, $\lim_{n \rightarrow \infty} \Pr[\cdot] = 1$
$neg(n)$	可忽略的, $o(1/poly(n))$
$a(n) \leq_* b(n)$	$\Pr[a(n) \leq b(n)] = neg(n)$
$a \sim b$	$a = (1 + o(1))b$
$a \lesssim b$	$a \leq (1 + o(1))b$
SLAM	同时定位与建图
MDS	Multi-Dimensional Scaling
cMDS	Classic Multi-Dimensional Scaling
SDP	半定规划, SemiDefinite Programming
NaN	无定义
log	指 \ln , 自然对数
\mathcal{S}^n	n 维对称矩阵的集合
\mathcal{E}^n	n 维平方欧氏距离矩阵的集合
\mathcal{K}_+^n	n 维对于 $\mathbf{1}_n^T$ 条件正定矩阵的集合
$\ \cdot\ _1$	矩阵的 1 范数, 所有特征值绝对值之和
$\ \cdot\ _2$	矩阵的 2 范数, 即谱范数
$\ \cdot\ _F$	矩阵的 Frobenius 范数
$\ \cdot\ $	向量的 2 范数, 即模长
I_n	n 阶单位矩阵
$\mathbf{1}_n$ 或 $\mathbf{1}$	所有元素皆为 1 的 n 维向量
$\mathbf{1}_{n \times n}$	所有元素皆为 1 的 $n \times n$ 矩阵
$B(X, r)$	与 X 中一点距离不超过 r 的所有点之集合
∂X	点集 X 的边界

第1章 引言

1.1 课题背景

在小毕设 ClusterSLAM^[1] 的工作中，我们利用同一刚体上两点距离保持恒定的几何性质，结合对双目视觉误差的估计，利用 Complete-Linkage Clustering 的聚类算法和 Bundle Adjustment 的优化算法，实现了对室内、室外场景中运动物体的识别和特征点分类。其中，对于室内场景达到了 91.54% 的准确率，相比基于语义识别的深度学习方法提升了 21.94%。对于室外场景的准确率为 94.15%，虽略低于深度学习算法的 96.70%，但节约了深度学习模型训练的巨大时间、算力、数据成本。

ClusterSLAM 算法对室内动态场景分析效果的显著提升，说明了在基于深度学习的图像处理领域日益热门的今天，基于几何模型的三维场景分析依然有着相当的潜力和探索空间。如果能充分利用欧氏空间本身具有的几何性质，并将其直接运用于模型中，则能节省让机器从零开始训练的大量成本。本研究将进一步从理论上探究基于几何约束对三维场景进行分析与重构的可能性和潜力，以此探索机器视觉、机器人学等领域的一条新道路，同时探索刚性和几何稳定性理论在传感器定位等领域的应用。

1.2 相关工作

1.2.1 欧氏嵌入与矩阵填充

1.2.1.1 MDS 概述

欧氏嵌入 (Euclidean Embedding) 是数据可视化、数据降维领域的一个重要工具：已知 n 个数据点间的相似度，将这些点嵌入到一个 r 维的欧氏空间，并通过嵌入后的几何距离尽可能保持这些点间的相似性信息。由于在机器学习等领域，将特征表示为欧氏空间中的向量，并利用其间距离进行优化的思路具有极为广泛的应用，例如自然语言处理 (NLP) 中 word2vec^[2] 等关于单词嵌入 (Word Embedding) 的算法，计算机视觉中用于人脸识别的 EigenFace^[3] 算法等等。当

输入数据的特征表现为向量空间中的点后，通过主成分分析（PCA）等算法，即可将数据压缩到较低维度的子空间，以此提取出数据中的主要特征，并将其与噪声分离，以对数据中的特征进行更加准确的分析。

在经典的 MDS^[4] 问题中，两个数据点 p_i, p_j 之间的相似度显式地表示为距离 d_{ij} ，而目标即是将 p_1, p_2, \dots, p_n 嵌入到 \mathbb{R}^k 中，使得 $\|p_i - p_j\| \approx d_{ij}$ 。

总体上，MDS 的应用领域分为两类：分析与重构。前者是已知一些数据的相似性，将本并非定义在欧氏空间的数据点嵌入到欧氏空间中，以便于利用欧氏空间的优良性质进行可视化或进一步分析^[5]；而后者是根据原本就在欧氏空间中的点之间的距离，还原出它们原本的空间结构。本课题对 MDS 的研究侧重于后者。

通过点对间距离重构点集的空间结构的思想，在多个领域有着极为广泛的应用。很多时候，获得一个场景的全局坐标系并直接求得各点较为准确的坐标并非容易。在涉及多点定位的领域^[6]，如室内多点定位^[7]、GPS 定位中，根据处于不同位置的多个传感器，最容易获得的信息即是传感器间的距离，而根据三角形的稳定性，已知这些传感器间两两的距离，它们之间相对的空间位置亦可得到确定，再根据已知的部分传感器的信息，即可求出所有传感器的位置。已有的关于无线传感器网络定位的文献对此问题亦有一些研究^[8]，对其在实际应用中的效果有较多的分析，但缺少基于几何、图论等角度的理论分析。

此外，在信息科学之外的一些领域，MDS 也有独特的应用，例如在计算化学中，根据分子间的键长推断分子的空间结构^[9]。

1.2.1.2 现有的 MDS 算法

最为原始的算法为常规 MDS (cMDS)^[4]，可以在 n 个点间全部 $\frac{n(n-1)}{2}$ 对距离都准确求得的情况下，准确恢复它们的空间位置，算法为：

令 $H_{ij} = I_n - \frac{1}{n}\mathbf{1}_{n \times n}$ ，根据平方距离矩阵

$$D_{ij} = d_{ij}^2, \quad (1-1)$$

可以直接得出 $\{p_i\}$ 的 Gram 矩阵：

$$G = -\frac{1}{2}HDH, \quad (1-2)$$

之后对 G 进行谱分解得 $G = U^T \Lambda U$, 并令 $K = \sqrt{\Lambda_+}U$, K 即为所求。

通过主成分分析的工具, 经典的 MDS 算法 (cMDS) 亦能对 d_{ij} 带有一定噪声的情况下获得较为理想的结果, 但是当输入数据存在一些错误 (outlier) 时, 则需要更加鲁棒的算法 (RobustMDS)^[10]。在 RobustMDS 的求解中, 一个常用的思想是利用凸松弛的方式, 在矩阵的秩 (0 范数) 难以直接进行优化的情况下, 对其 1 范数进行最小化。因为一些类型 RobustMDS 问题被证明为 NP-hard^[10], 相关成果多为从不同角度入手的局部优化或启发式算法^[11]。

在实际问题中, 常常出现部分点对间距离无法获取的情况, 因为 cMDS 和 RobustMDS 算法需要完整的距离矩阵作为输入, 对于数据缺失的情况, 则需要矩阵填充 (Matrix Completion) 的算法将缺失的数据补全。直观上, 用于主成分分析的矩阵在除去噪声后并非满秩, 而对于大部分秩为 r 的 $n \times n$ 随机对称矩阵, 只需根据 $O(nr \log^2 n)$ 个元素, 即能以大概率将其准确还原^[12]。由此, 通过随机准确观测 $O(nr \log^2 n)$ 对点间的距离, 亦能以大概率准确还原这 n 个点的相对位置^[13]。根据实验, 若点对距离存在一定噪声, 亦能得到满意的结果^[13], 但暂缺少相关的理论证明。此外, 对于一个既存在错误也存在缺失, 但其他数据无噪声的低秩矩阵, 在一定条件下亦可以准确还原^[14]。

事实上, “错误” 比 “缺失” 对还原的影响更大, 因为若能在存在一定数量的错误的情况下正确还原一个矩阵, 那么在存在同样数量缺失的情况下, 只须在缺失处填充任意数据, 即可将其归约为存在错误的还原问题。在无噪声的情况下, 现有算法可以容许常数比例的随机错误或大部分的缺失; 在同时存在噪声和错误情况下, RobustMDS 的多种模型已被证明是 NP-hard, 但对于同时存在噪声和缺失时的情况, 仍无充分的理论结果。

此外, 虽然当 $\{d_{ij}\}$ 得到准确还原时, $\{p_i\}$ 的相对位置也确实得到准确还原, 但现有的 MDS 研究并未证明当 $\{d_{ij}\}$ 较为接近真实值时, 数据点间的结构是否亦接近真实结构。事实上, 根据 Johnson-Lindenstrauss 引理, 对于任意 k 维欧氏空间中的 n 个点, 可以通过随机投影的方式, 将其映射到一个 $O(\log n / \epsilon^2)$ 维欧氏空间, 使任两点间距离的相对误差不超过 ϵ 。换言之, 尤其是在高维 ($k \gg \log n$)

空间中，即便能相对较好地还原出 G ，点对距离的微小误差仍可导致它们空间结构的巨大改变。因此，从整体的角度，普通的 MDS 算法能否较好地还原空间结构仍需要进一步分析。

1.2.2 欧氏距离矩阵的优化

对于 n 个点，并非对于任何给定的 $\frac{n(n-1)}{2}$ 组点对距离，都一定存在一个欧氏空间中的嵌入满足该约束。一般地，我们设所有存在欧式嵌入的 $n \times n$ 平方距离矩阵的集合记为 \mathcal{E}^n 。一般地， \mathcal{E}^n 为对角元素均为 0，且根据 (1-2) 所对应的 G 为半正定矩阵的所有 n 阶对称方阵之集合。由于所有半正定矩阵构成的集合是一个凸锥，而所有对角元素均为 0 的方阵构成的集合亦是一个凸锥，所以 \mathcal{E}^n 是一个凸锥。一般地，我们有^[11]

$$D \in \mathcal{E}^n \iff \text{diag}(D) = 0 \wedge -D \in \mathcal{K}_+^n \quad (1-3)$$

其中 \mathcal{K}_+^n 为一系列条件正定阵构成的凸锥：

$$\mathcal{K}_+^n = \{A \in \mathcal{S}^n \mid v^T A v \geq 0, \forall v \in \mathbf{1}^\perp\}$$

根据 Alfakih 等人的工作^[15]，当我们需要对 $D \in \mathcal{E}^n$ 做优化时，可以首先将 D 的定义域设为对角元素为 0 的对称方阵，而将 $G = -\frac{1}{2}H D H \succeq 0$ 作为一个约束条件，从而将其转化为一个半定规划问题。

1.2.3 随机图模型与图论理论

在社会网络等领域的研究中，随机图模型是一个常用的工具。通过按照某种特定的规则在大量顶点之间随机连边，即可生成一个相应的随机图。虽然随机图在局部上随机，但当顶点数 n 趋向于无穷时，其整体的统计性质往往收敛于特定的模式，这些特定的模式即可用于描述社会关系中虽个体行为无法预测，但整体行为有迹可循的现象。

1.2.3.1 Erdos-Renyi 图

Erdos-Renyi 图是随机图模型中最简单的一种，在图 $G(n, p)$ 中，对于每对顶点 v, v' ，无向边 (v, v') 存在的概率为 p 且互相独立。当 $n \rightarrow \infty$ 时，其连通性几乎必然地具有以下性质:^[16]

- 当 $p < (1 - \epsilon)/n$ 时，每个连通分支包含不超过 $O(\log n)$ 个点；
- 当 $p = 1/n$ 时，最大的连通分支包含 $\Theta(n^{2/3})$ 个点；
- 当 $p > (1 - \epsilon)/n$ 时，最大的连通分支包含 $\Theta(n)$ 个点，其他每个分支不超过 $O(\log n)$ 个点；
- 当 $p < (1 - \epsilon) \log n / n$ 时，全图不连通；
- 当 $p > (1 + \epsilon) \log n / n$ 时，全图连通。

虽然 Erdos-Renyi 图的一些统计特性与实际的社会网络有一定区别，但因为其模型较为简洁，又足以解释现实情境中出现的现象，此模型依然在网络科学等领域有着广泛的应用。在本课题中，我们使用此模型来研究稀疏 MDS 问题中不同稀疏度的邻接矩阵对空间结构重构可能的影响。

1.2.3.2 图的刚性

给定一个图 $G(V; E)$ ，我们试图在 k 维空间 \mathbb{R}^k 中找到 V 的一个嵌入，使得对于 E 中的每条边 $e = (v, v')$ ，我们约束 $\|v - v'\| = d_e$ 。如果这样的嵌入在全等意义下至多只有有限种（从而局部只有一种），则称 G 具有 k 维刚性，而若至多只有一种，则称 G 具有 k 维全局刚性。当嵌入中的点处于一般位置^①时，全局刚性是图 G 的性质，而不取决于具体的边长。^[17]

在后文中，我们提到一个图的刚性时，皆指其在一般位置意义下的刚性。

关于 2 维刚性，Maxwell-Laman 定理给出， W 是一个 2 维刚性图当且仅当其包含一个 Laman 子图^[18]。文献^[19] 进一步指出了一个图是 2 维全局刚性图的若干个充分条件：

定理 1.1 (Jackson): 下列结论成立：

1. 所有 6-连通图是 2 维全局刚性图。
2. 对于 $p = (\log n + 3 \log \log n + \omega(1)) / n$ ，Erdos-Renyi 图 $G(n, p)$ 渐近几乎必然^②为 2 维全局刚性图。

① 即不满足任何非平凡的多项式方程。

② 简称 a.a.s.。此外，约定后文中对于 $G(n, p)$ “a.a.s. 地拥有某性质”时，亦可直接简称其“拥有某性质”

而对于更高维的刚性，则并无类似的直接判定方式，与 Maxwell-Laman 定理类似的基于子图边数的条件只具有必要性而无充分性。不过，对于 3 维的情况，我们仍可以利用基于张力方法，在多项式时间内对一个图的全局刚性进行大概率的判定。^[17]

当且仅当稀疏 MDS 问题对应的图具有刚性时，给定此图所有边对应的点对距离，才在理论上有可能还原出这些点的空间位置。我们通过随机图的刚性性质，来从理论上分析根据几何距离对一个空间结构进行重构所需要的最少边数。

1.3 本文贡献

本课题分为三个部分，场景分析、场景重构和网格设计。

在场景分析中，提出了一种基于特征点间距离稳定性的刚体聚类算法，能对动态场景中不同运动轨迹的刚体进行识别，该算法能在理论上保证在输入数据较为准确的前提下能得出正确的聚类；在 ClusterSLAM 的实际运用中，使室内场景分析的准确率相比现有的深度学习算法有了较大提升，且运行效率亦比现有算法更高。

在场景重构中，对于 MDS 算法提出了“几何稳定性”一种新的评价方式，从整体结构的角度分析对点云还原的准确度，证明了 cMDS 算法的几何稳定性，讨论了在随机图模型中对于输入数据存在缺失情况的重构算法与理论可能性，并根据 Erdos-Renyi 理论，提出了在随机稀疏图中剔除部分不可靠点以保证整体还原基本准确的方案。

在网格设计中，基于几何稳定 MDS 的理论，设计了一种在点云间构造稀疏图结构的算法，能够用尽可能少的边数与总长，有效而稳定地建立足以描述和重建点云刚性结构的一系列几何约束。

1.4 论文组织

在第 2 章中，简单介绍前期工作 ClusterSLAM 与本人所作贡献，并初步从理论上分析所用基于几何距离对特征点进行聚类的算法的合理性，探讨算法的性能、寻找全局最优解的难度与对于不完整观测现象的解决方式。

而不致引起误解。事实上，对于随机图的任何描述都是概率意义上的：因为 $G(n, p)$ 可能为任何一个图，所以其不可能绝对必然地拥有任何非平凡的性质。

在第 3 章中，深入分析各种类型的 MDS 算法在几何结构重构中的效果，首先证明了普通的 cMDS 在仅有噪声情况下的有效性，其中定量分析了距离矩阵的各个特征值与噪声的 2-范数大小对重构的影响，从而对其重构误差得出了一个概率意义上的理论上界。之后试图提出一种能在同时存在噪声和数据缺失的情况下，能较为准确还原出几何结构的 MDS 算法，并分析所需数据量的极限。

在第 4 章中，分析了刚性理论中一个图具有刚性性质的必要与充分条件，并根据已有工作中对最小代价全局刚性图的研究，且针对该算法的缺陷，提出了一种对无线传感器网络进行稀疏连接，在理论上保证几何稳定性的前提下最小化连接代价，从而实现低成本可靠定位的方案。

第 2 章 ClusterSLAM 中基于几何距离的特征点聚类

在本章中，我们提出了一种通过基于点对间几何距离对动态场景中相对运动刚体进行聚类的算法，并介绍了在本人前期工作 ClusterSLAM 中该算法的应用，以及其在无须语义知识的情况下对相对运动刚体进行分割的有效性，从而说明了几何距离模型在场景分析中的意义。

2.1 ClusterSLAM 简介

对周围环境的理解与建模的技术，是现代导航系统的基础。随着自动驾驶技术的迅猛发展，同时定位与建图（SLAM）技术成为了当今计算机视觉领域的一大热点。SLAM 技术的核心即是，将一个机器人放在未知环境中，按照自主或既定的轨迹运动，并使其通过传感器获得的数据对所在环境进行建模，同时确定其自身的运动轨迹。然而，对于场景中存在的动态物体，由于在不同时间点会测得不同的位置，传统的 SLAM 技术就会面临困难。

在当今主流的动态场景 SLAM 算法中，由于现实场景中大部分物体都会保持静止，因此常见的做法是通过鲁棒性算法将动态物体剔除，之后利用地面坐标系对自身运动轨迹进行 SLAM。在 ClusterSLAM 的工作中，我们假设场景中所有物体都是刚体，并利用刚体上任意两点间几何距离保持不变的特性，对场景中已由前端提取出的特征点进行聚类，并通过光束法平差（Bundle Adjustment）的算法在聚类基础上进行优化，通过聚类-优化的交替迭代，求出双目视频中自身与场景中运动物体的轨迹。

本人在该项目中的主要贡献为：

- 利用同一刚体内两点距离保持恒定的性质，设计了描述特征点间相似性量度的矩阵；
- 注意到刚体性质的不可传递性，运用全连接层次聚类的算法对特征点进行聚类，获得了相比其他聚类算法准确度更高且效率更高的运行表现。

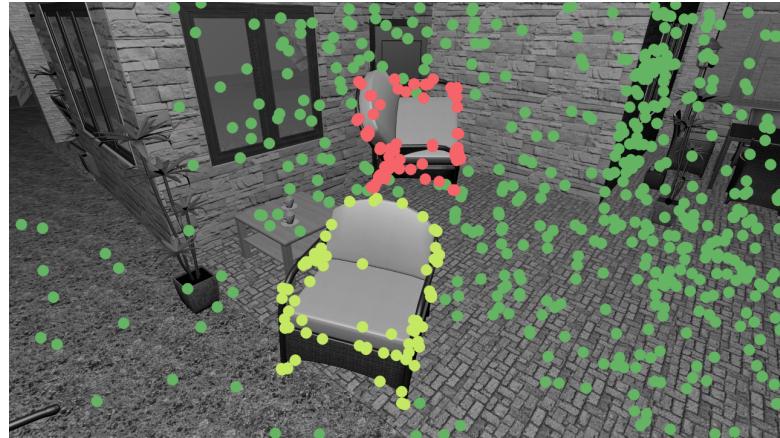


图 2.1 ClusterSLAM 聚类效果

2.2 面向 SLAM 的特征点聚类算法

本节讨论在动态场景的 SLAM 问题中，根据由双目视觉得出的特征点三维坐标，通过点对之间几何距离在各帧中的变化大小，将场景中的运动特征点按照刚性运动组进行聚类的算法。该算法能在不需要语义分割的情况下，基于几何信息高效地对存在相对运动的不同刚体进行聚类。

2.2.1 基于特征点相似性的聚类

首先，我们考虑三维坐标完全精确，且所有点在每一帧全部被观测的理想情况：假设一个刚体 X 上各特征点在 t_1 帧的坐标分别为 $x_i^{t_1}$ ，而在 t_2 帧的坐标分别为 $x_i^{t_2}$ ，那么令 $X^t = (x_1^t, \dots, x_m^t)$ ，则存在一个旋转变换 $R \in SO(3)$ 和平移变换 $T \in \mathbb{R}^3$ ，使得 $X^{t_2} = RX^{t_1} + T$ 。

在 X^{t_1}, X^{t_2} 均给定的情况下，我们可以通过经典的 ICP 算法^[20] 求出变换 (R, T) ，然而在此问题中， X 包含哪些点本身就是我们所求的问题，因此相比求出 R ，求出 X 的组成对聚类意义更大。在实际场景中，只要刚体 X 与 X' 存在相对运动，则对应的变换 $(R, T) \neq (R', T')$ 。从直觉上看，虽然直接求出一个刚体包含的点不容易，但是我们可以相对容易地验证一组点是否属于同一个刚体。

然而，点集的刚性不满足传递性：考虑将两个刚体的某一点连接在一起成为一个“双截棍”，则连接点与两个刚体上的任意点都满足刚体性质，但这两个刚体可以围绕连接点旋转，因此它们的并集不是同一个刚体。因此，要确定一组点组成一个刚体，我们需要其中每两点满足刚体性质^①，因此一些假设点与点

^① 严格地说，只要这些点之间存在一个全局刚性图，该图中每对点满足刚体性质，即可推出这些点中每

间的不相似性满足三角不等式（或者点与点的等价满足传递性）的聚类算法对此问题不适用。

考虑经典的 Kruskal 算法：在其最小生成树未生成完毕前，会生成一系列森林，而每一步会将距离最小的两棵树连接到一起。注意到此算法中，每次连接两个树所依据的是两棵树内点的**最小距离**，即两棵树只要存在一对点相似，这两棵树即视为相似。我们若将此算法进行一个修改：将两类进行合并时，是依据这两类间的**最大距离**，那么便能更加适应 ClusterSLAM 问题的结构。

2.2.2 特征点相似性的量化方式

由于在不知刚体组成时， (R, T) 难以由未聚类的点直接求得，而将特征点进行聚类并不需要得出它们具体的运动——只要知道两组点的运动是否“相同”即可。因此，我们可以只基于点对的距离来进行聚类，这样亦可以降低聚类过程的运算量。当然，在聚类完成后，我们依然可以根据 ICP 等算法对点的坐标进行优化。

首先我们易知，在运动的场景中，同一刚体上两点间的距离始终保持不变，因此，最简单的想法是将整个视频中两点三维距离的方差作为这两点不相似性的量度。粗略地看，假设由于观测误差导致的点对距离不确定性“上限”为 δ ，则当两点间距离的方差小于某一阈值 δ 时，则应视为这两点属于同一刚体。同时，对于两个不同的刚体，其应当至少存在一对点，其距离方差大于 δ ，即它们相对运动的尺度应大于观测噪声，否则我们便难以确定两点间的距离变化是来自相对运动还是来自观测误差。（事实上，在 ClusterSLAM 中，越远离相机的点观测误差越大，而具体的聚类也考虑了各处不同的观测误差的因素，但这里暂不过多讨论。）

因此，特征点聚类的问题可以描述为：将空间中的运动特征点不重不漏地分为尽可能少的类，使每类中任两点的不相似性不超过 δ 。然而，由于该聚类问题等价于 *clique-cover* 问题，计算其全局最优解是 NP-hard 的^[21]。因此，我们使用一种局部最优的算法，该算法可以在 $O(n^2 \log n)$ 时间内完成，且经实验验证有较好的表现：

1. 首先将每个点 p_i 分为单独一类 c_i ，两类间不相似性 $d(\cdot, \cdot)$ 设为对应点在各帧中几何距离的方差。

两点之间都满足刚性性质。由于存在此刚性图的条件严格强于存在一棵生成树，因此 Kruskal 算法不适用。我们将在3.4.1节和第4章详细讨论与图的刚性相关的问题。

2. 选取不相似性最小的两类 c_i, c_j , 若 $d(c_i, c_j) < \delta$, 将其合并为一类 c_k , 否则算法结束。
3. 对于其他类 $c_l \notin \{c_i, c_j, c_k\}$, 定义 $d(c_k, c_l) = \max\{d(c_i, l), d(c_j, l)\}$, 并移除 c_i, c_j 。
4. 跳转到第 2 步。

那么当上述算法结束时, 我们有

$$\max_{c_i} \max_{\{p, p'\} \subseteq c_i} d(p, p') < \delta \quad (2-1)$$

$$\min_{c_i \neq c_j} \max_{p_i \in c_i, p_j \in c_j} d(p_i, p_j) \geq \delta \quad (2-2)$$

即同一类任两点不相似性不超过 δ , 而任意不同两类中能分别取出一点, 它们的不相似性至少为 δ 。

其中, 对于两点无同时出现导致 $d(p_i, p_j)$ 无定义 (记为 NaN) 的情况, 我们可以定义 $\max\{NaN, x\} = x(x \in \mathbb{R})$, $\max\{NaN, NaN\} = NaN$ 。

2.2.3 基准点策略: 对于非完全观测的解决方案

在实际场景, 尤其是室外场景中, 我们往往难以在全部帧中观察到同一个特征点, 尤其是在一个点仅在较少的帧中出现时, 它与其他点距离的方差往往不具有代表性, 因此, 我们可以认为对出现较多的点包含的信息更为可靠, 并将它们作为基准点。

在使用基准点策略的聚类中, 我们首先只考虑基准点和它们之间的距离, 并将它们分为若干类。之后, 我们再将每一类的基准点通过 ICP 的算法在各帧建立局部坐标系, 并分析其他点在各个局部坐标系的运动, 并将其分在相对运动最小的类中。在实际实验中, 当使用恰当的阈值时, 基准点策略可以同时提升聚类算法的准确度与效率。

在室内场景中, 因为各个物体在画面中占的比例较大, 且场景本身移动较小, 因而各个物体上往往存在相当部分的点能有较多的观测, 此方案对聚类效果提升显著; 然而对于室外场景, 会存在部分物体 (如远处的车辆) 特征点过少而无一属于基准点的情况, 此时即会导致相应的刚体无法被识别。

因此，此方案主要运用于室内场景中。

2.3 算法分析与应用

本节首先分析聚类算法的效率，进一步，给出其在动态场景分析中的应用，并通过实验结果说明其有效性。

2.3.1 算法分析

首先考虑不取基准点的情况。假设共有 t 帧、 n 个点，则求得全部点对距离的时间复杂度为 $O(n^2t)$ 。

如果简单地用矩阵来存储类与类间的不相似性，那么在第 t 次合并时，第 2 步需要遍历矩阵 $\frac{(n+1-t)(n-t)}{2}$ 个元素，而第 3 步需要修改 $O(n)$ 个元素，每次修改的时间复杂度为 $O(1)$ ； n 个点进行合并的轮数可达 $\Theta(n)$ ，故总时间复杂度为 $\Theta(n^3)$ 。

若通过一个堆结构维护不相似性矩阵的最小元素，则寻找最小值和插入、删除值的时间复杂度均为 $O(\log n)$ 。因为第 2 步只须寻找一个最小值，而第 3 步需要对 $O(n)$ 个值进行操作，因此每次合并的时间复杂度为 $O(n \log n)$ ； n 个点进行合并的次数不超过 $n - 1$ 次，因此聚类时间复杂度为 $O(n^2 \log n)$ ，总时间复杂度为 $O(n^2(\log n + t))$ 。

对于部分 $d(c_i, c_j)$ 未定义的情况，在堆存储时可以将其存储为 $+\infty$ 以免在第 2 步被取到，而在第 3 步求最大值时，只须特判 $d(c_i, l), d(c_j, l)$ 中是否存在 $+\infty$ 即可。

事实上，对于全连接层次聚类问题，时间复杂度为 $O(n^2)$ 的 CLINK 算法^[22] 亦可以得出一个局部最优的聚类方案。然而，此方法得出的聚类结果依赖于输入点的顺序，在大部分问题中表现差于每次取具有最小不相似性的两类进行合并的算法，且不支持部分不相似性值无法获取的情况，因此我们最终选择上述算法。

对于取基准点情况，假设共有 n 个点，其中 m 个基准点，且总共被分为 c 类，则对基准点聚类的时间复杂度为 $O(m^2 \log m)$ ，对其余点归类的时间复杂度为 $O(nt)$ ，因此总时间复杂度为 $O(m^2 \log m + nt(n + c)) = O(m^2 \log m + n^2t)$ ，在实际应用中一般取适当阈值使 $m \approx 0.2n$ ，因此确实可以节省部分运行时间。

在 ClusterSLAM 中，我们同时尝试了谱聚类 (Spectral Clustering) 和

AP(Affinity Propagation) 的聚类方法，但另两种算法的运行效率均低于上述算法（上述算法：6.3ms，AP：13ms，谱聚类：180ms）。与之同时，AP 算法准确率明显偏低，谱聚类准确率在噪声较大时与我们的算法持平，但运行效率明显偏低。值得注意的是，在特征点坐标误差足够小的时候，我们的算法能够得到完全精确的聚类，但另两种算法不能。

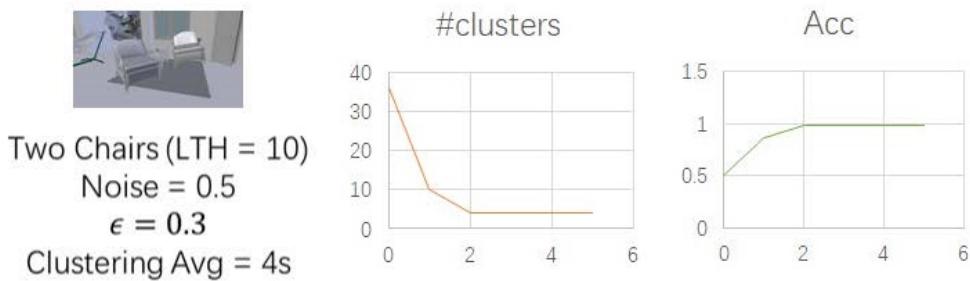


图 2.2 聚类-优化进程示意图

2.3.2 在场景分析中的应用

在 ClusterSLAM 中，我们采取了聚类-优化交替进行的迭代方案。由于根据双目视觉计算特征点的位置的算法本身存在一定误差，因此对原始数据进行的聚类难以保证其准确性。考虑到同一组特征点在同一个运动的局部坐标系内保持静止，我们可以基于该先验知识对各特征点的空间坐标进行优化，而优化后更加准确的坐标则能让我们对其进行更加准确的聚类，我们即可由此循环进行，从而最终同时得出准确度更高的聚类方案和特征点坐标。

在该聚类算法的实际应用中，聚类的几何容差阈值 δ 的取定，往往直接影响聚类结果。一般地，较低的 δ 易造成过分割，而较高的 δ 易造成欠分割。由于在聚类-优化的迭代方案中，过分割会将同一运动组分为多个运动轨迹相似的多组进行优化，当去除部分噪声后，后续的聚类中仍可被合并为同一类；但若将两个不同运动组误合为一类，则各点坐标将被错误地“优化”，在后续迭代中亦难以恢复。因此，我们倾向于选择相对较小的 δ ，在前几轮进行过分割，而随着优化的进行，聚类数目逐步减少，直至正确（如图2.2所示）。

由于现有的数据集缺少室内动态场景的数据，我们在测试中采取合成一系列动态场景，对其进行渲染，利用现有的前端提取特征点并进行追踪，再使用不同的算法分别计算其准确率。测试结果显示，采取了本节聚类算法的 Cluster-

SLAM 对室内场景达到了 91.54% 的准确率，相比现有的基于深度学习的语义分割算法 69.60% 的准确率提升了近 22 个百分点；对室外场景的准确率为 94.15%，略低于语义分割算法的 96.70%，但有着能够节约大量训练数据和算力成本的优势（如图2.4所示）。

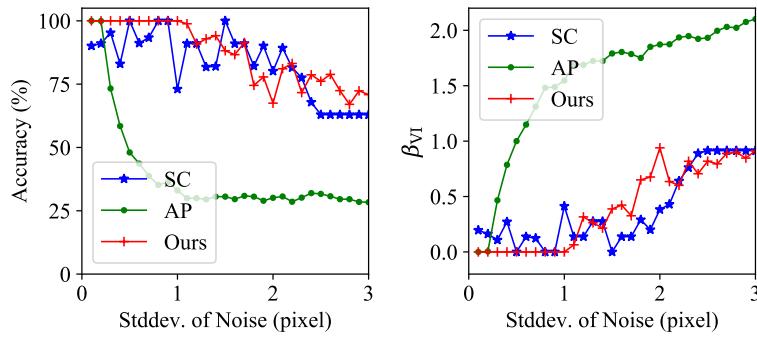


图 2.3 三种聚类算法对于实际数据的聚类效果

Table 1. Quantitative comparison on synthetic sequences. All prediction time of Mask R-CNN [?] is excluded in this table.

	Indoor Sequences							Outdoor Sequences								
	$\log \chi^2$	RMSE(m)	ATE	R.RPE	TRPE	Acc. (%)	β_{VI}	Time(s)	$\log \chi^2$	RMSE(m)	ATE	R.RPE	TRPE	Acc. (%)	β_{VI}	Time(s)
Full BA	9.61	2.10	0.53/-	0.48/-	0.87/-	52.73	1.19	5.45	10.92	14.39	12.94/-	0.73/-	42.55/-	81.39	0.84	6.00
P. DCS	12.80	1.53	0.63/1.61	0.49/1.82	0.99/2.60	56.05	1.18	3.83	13.85	11.22	9.36/-	0.73/-	37.61/-	80.22	0.82	1.86
Sem. SEG	9.31	0.84	0.31/0.51	0.12/0.87	0.49/0.95	69.60	1.19	3.82*	8.55	2.69	1.65/3.09	0.18/0.32	2.34/ 8.11	96.70	0.24	5.28*
Ours w/o U	7.88	1.21	0.35/0.34	0.15/0.37	0.53/0.57	65.60	0.96	9.60	8.56	2.48	1.86/5.13	0.02 /0.40	3.18/12.32	86.51	0.64	6.96
Ours w/o I	8.65	1.05	0.15/0.31	0.05/0.57	0.21/0.55	76.16	1.06	9.47	9.70	9.56	2.12/3.44	0.47/0.20	4.47/9.94	81.83	0.57	5.84
Ours	7.15	0.44	0.01/0.12	0.01/0.29	0.02/0.22	91.54	0.40	11.15	6.52	0.63	0.53/3.37	0.02/0.18	1.10/8.65	94.15	0.27	6.14

图 2.4 ClusterSLAM 中对各种算法在合成数据集上运行效果的比较

2.4 小结

在本节中，我们介绍了一种基于几何距离对动态场景中的运动特征点进行聚类的算法，从理论上分析了其正确性和效率。进一步，通过实验说明了 Cluster-SLAM 算法中，将该聚类算法与传统 SLAM 中的优化算法进行结合，可以在不需要深度学习的模型训练之情况下，对动态场景中的特征点进行分类，进而不依赖语义知识而依然有效地识别不同的运动物体（分别采用了 SUNCG 与 CARLA 合成数据集，效果如图2.5与图2.6所示）。尤其是在室内场景中，ClusterSLAM 算法的准确率相比深度学习算法还获得了较大提升，这充分体现了几何信息对于解决场景分析问题的价值。

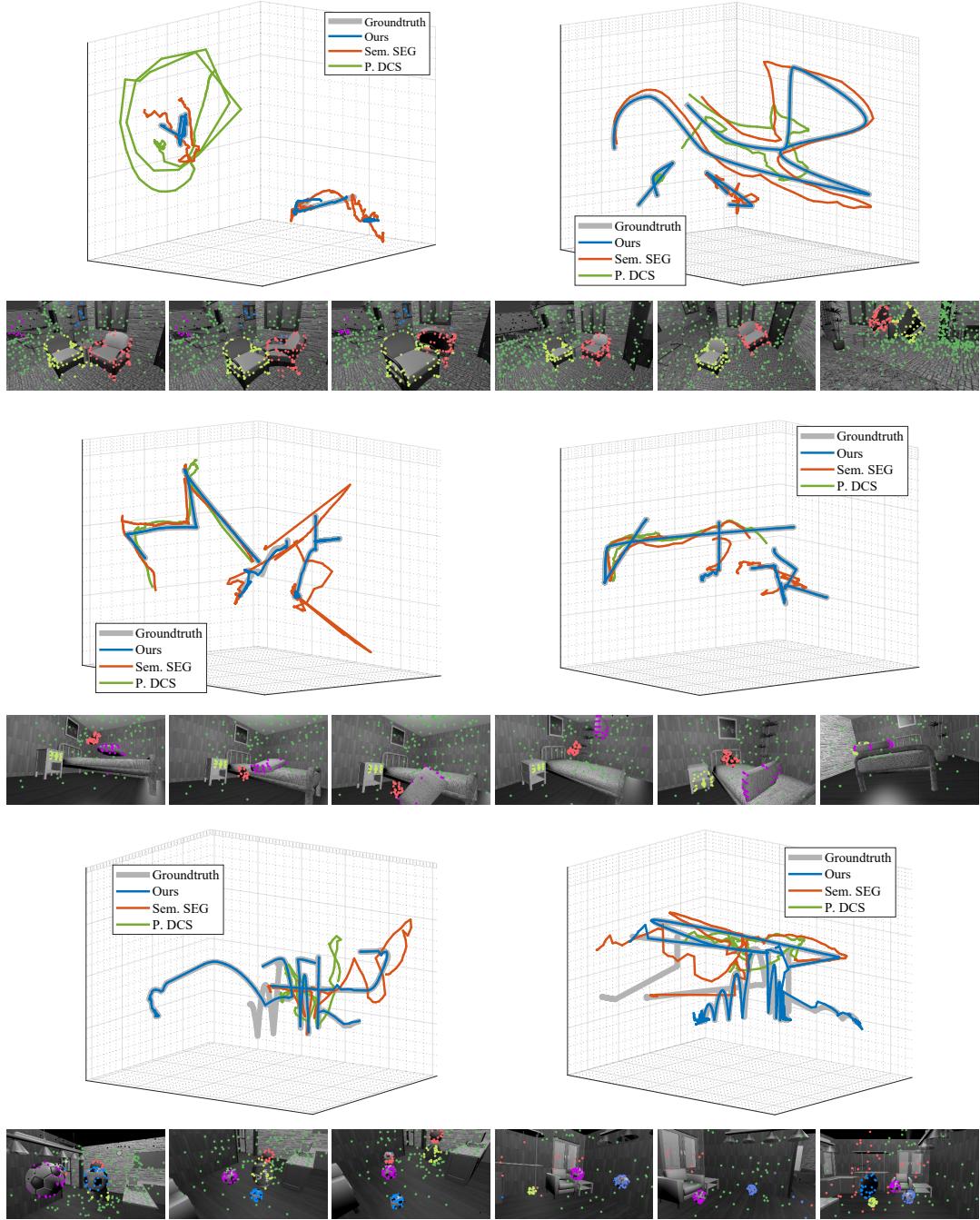


图 2.5 ClusterSLAM 中室内动态场景的分析效果

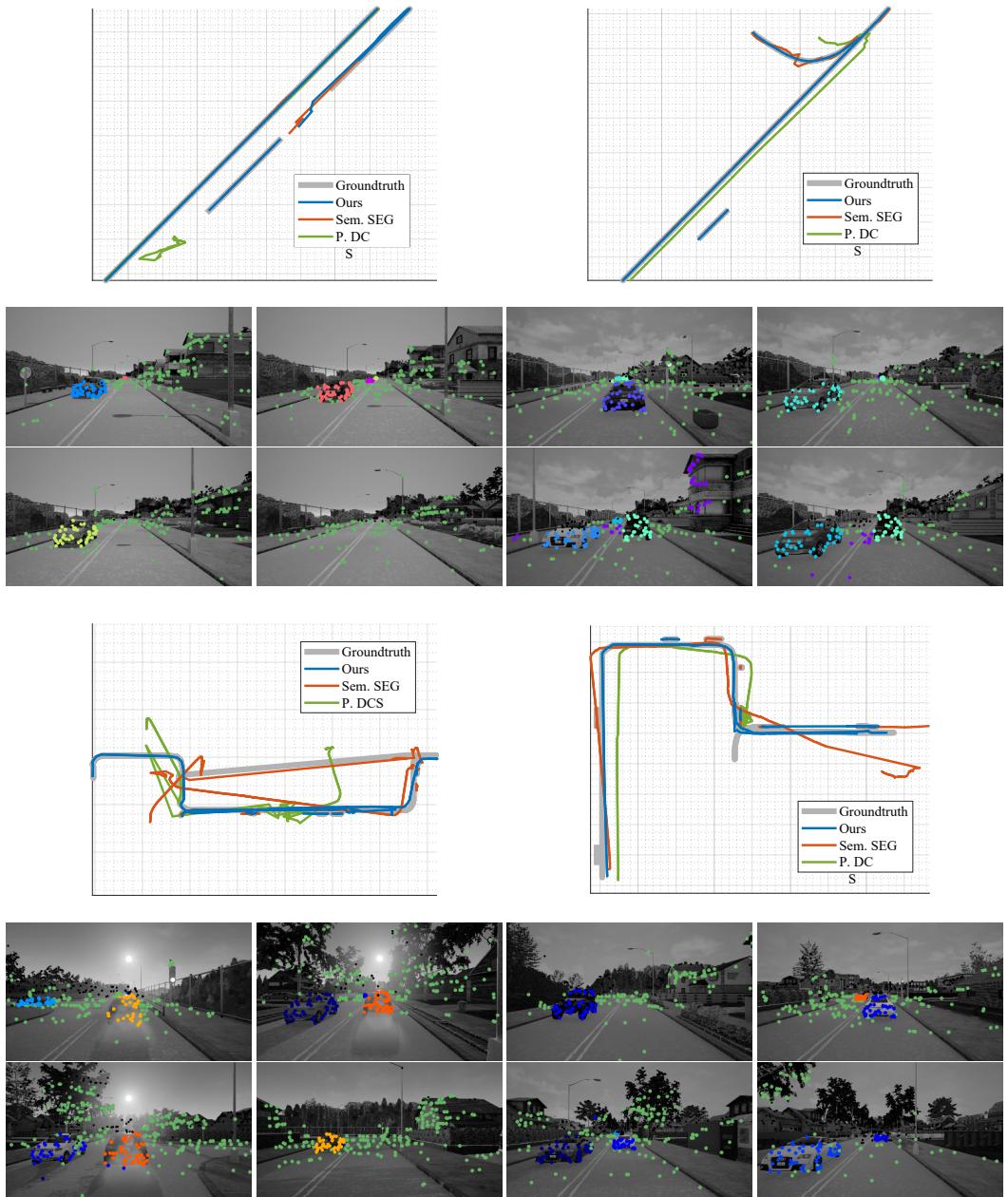


图 2.6 ClusterSLAM 中室外动态场景的分析效果

第 3 章 对 MDS 算法的分析与探索

在本章中，我们分析了经典的 cMDS 算法，并证明了其在输入距离矩阵存在一定误差的情况下仍能稳定地恢复点云的整体几何结构之性质。同时，我们从理论上探讨了基于稀疏几何距离约束对点云进行重构的理论极限，并进行了一系列的实验验证。

3.1 几何稳定的 MDS 算法

在考虑各种 MDS 算法的性能时，前人的工作常以重构后的点云中点与点距离与真实值的差异来评价其准确度，然而，点对间距离的准确度并不直接代表整体结构的准确度。由于同一个几何结构在刚体运动的变换下等价，为讨论两个点云结构的相似性，我们需要用一个刚体变换将其尽可能重合，之后分析对应点间的距离。

因此，我们从整体几何结构的角度，给出了对 MDS 准确度的一个新评价方式---几何稳定性：

定义 3.1： 如果对于输入坐标矩阵 $P \in \mathbb{R}^{k \times n}$, $\|P\|_{2,\infty} \leq 1$ ^①，依据随机变量 X ，以 s 的概率对其中点对的距离进行独立同分布的伯努利随机采样，加上 $\delta \sim \delta(X)$ 的绝对误差，并以概率 c 将其值替换为任意错误值 $E(P, X, \cdot)$ 所得的输入 Q ，求得的 $K = f(Q)$ 满足：

$$\sup_E \lim_{n \rightarrow \infty} \Pr_{P \sim \mathcal{P}; X \sim \mathcal{X}} [\exists S \in \mathbb{E}(k) \text{ s.t. } \|K - SP\|_F \leq \epsilon \sqrt{n}] > 0.99 \quad (3-1)$$

则称 f 为一个 $[k, s, c, \delta, \epsilon]$ -几何稳定的 MDS 算法（简写为 $[k, s, c, \delta, \epsilon]$ -MDS）。其中 $\mathbb{E}(k)$ 为 k 维欧氏空间的对称群。

当 $\lim_{n \rightarrow \infty} \epsilon = 0$, $\delta \sim N(0, \sigma)$ 时，可简写为 $[k, s, c, \sigma]$ 。当 $\sigma = \Theta(1)$ 时，简写为 $[k, s, c]$ 。

根据现有的工作，我们有：

^① 即每个点都在单位球内。

定理 3.1: 存在 $[k, \Theta(1), \Theta(1), 0, 0]$ -MDS。^[14]

定理 3.2: 存在 $[k, O(k \log^2 n/n), 0, 0, 0]$ -MDS。^[13]

此外，根据1.2.1节关于错误和缺失的讨论，我们有

定理 3.3: 若存在 $[k, s, c, \delta, \epsilon]$ -MDS，那么也存在 $[k, (1 - c)s, 0, \delta, \epsilon]$ -MDS。

3.2 cMDS 的几何稳定性分析

首先我们定义 \mathcal{P} 为在各个特征方向上投影的方差固定的一系列 P 的集合。我们通过谱分解的稳定性，证明了

定理 3.4: 对于谱间隔不为 0 (参见第3.2.3节) 的 \mathcal{P} 和充分大的 $n > n_0(\mathcal{P}, \sigma)$ ，在将矩阵的每个元素均减去 σ^2 后，cMDS 是 $\left[3, 1, 0, \sigma, \frac{w(\mathcal{P})\sigma}{\sqrt{n}}\right]$ -MDS，其中 $n_0(\mathcal{P}, \sigma)$ 仅取决于 \mathcal{P} 和 σ ， $w(\mathcal{P})$ 仅取决于 \mathcal{P} 。

3.2.1 对噪声 2-范数的估计

在3.1节的模型中，考虑 $f(n, k) = 1$ ， $\delta \sim N(0, \sigma^2)$ 的情况。我们假设将加上噪声得到的距离视为真实距离，还原后的矩阵为 \tilde{D} ，并假设当 $\sigma < 1$ ，并定义噪声矩阵为 $N = \tilde{D} - D$ ，那么 $N_{ij} = 2d_{ij}\delta_{ij} + \delta_{ij}^2$ ，且 D, \tilde{D}, N 均为对称矩阵。首先定义：

定义 3.2: 定义 $f(n)$ 是可忽略的，如果对于任何多项式 $g(n)$ ，有 $\lim_{n \rightarrow \infty} f(n)g(n) = 0$ 。记作 $f(n) = neg(n)$ 。

那么我们有：

定理 3.5: 存在常数 $C_2 > 0$ ，使得对于充分大的 n ，有

$$\Pr \left[\|N - \sigma^2 \mathbf{1}_{n \times n}\|_2 > C_2 \sigma \sqrt{n} \right] = neg(n) \quad (3-2)$$

为此，我们首先给出正态分布尾部概率的估计：

引理 3.1: 对于满足正态分布的随机变量 $X \sim N(\mu, \sigma^2)$ ，有

$$\Pr [|X - \mu| > t\sigma] \leq \sqrt{\frac{2}{\pi}} \cdot \frac{\exp(-t^2/2)}{t}. \quad (3-3)$$

据此可得：

引理 3.2: 存在正常数 C_1 , 使得

$$\Pr \left[\max |N_{ij}| > C_1 \sigma \sqrt{\ln n} \right] = neg(n) \quad (3-4)$$

由此, 又因为高斯分布以及其尾部的 1,2 阶矩均有限, 若将 \tilde{D} 中大于 $C_1^2 \sigma^2 \ln n$ 的元素截断至 $C_1^2 \sigma^2 \ln n$, 这对所有均值、方差的影响亦是可忽略的。

注意到 $E[N_{ij}] = E[\delta_{ij}^2] = \sigma^2$, $Var[N_{ij}] = E[(N_{ij} - \sigma^2)^2] = 4d_{ij}^2 \sigma^2 + 2\sigma^4 \leq 16\sigma^2 + 2\sigma^4$, 根据 Bernstein 不等式^[23] (令 $M = C_1 \sigma \sqrt{\ln n}$; 对于固定的 j , 令 $X_i = N_{ij} - \sigma^2$) :

引理 3.3 (Bernstein 不等式): 设 X_1, \dots, X_n 为期望为 0 的独立随机变量。若对任何 i 都有 $|X_i| \leq M$, 那么对于任何正数 t ,

$$\Pr \left[\sum_{i=1}^n X_i > t \right] \leq \exp \left(-\frac{\frac{1}{2}t^2}{\sum_i E[X_i^2] + \frac{1}{3}Mt} \right) \quad (3-5)$$

可知对于任意 $t > 0$, 对于固定的第 j 列,

$$\Pr \left[\left| \sum_i (N_{ij} - \sigma^2) \right| > t \right] \leq 2 \exp \left(-\frac{\frac{1}{2}t^2}{(16\sigma^2 + 2\sigma^4)n + \frac{1}{3}C_1\sigma\sqrt{\ln n}} \right) + neg(n) \quad (3-6)$$

再根据圆盘定理, 即可推出定理3.5。

3.2.2 对真实 Gram 矩阵特征值与谱间隔的估计

各个非零特征向量对应的特征值即为点云在该方向上坐标的平方和, 在模型一定的情况下, 与 n 成正比。因此, 对于固定的 \mathcal{P} , G 的谱间隔 (这里假设所有非零特征值不同, 且多重零特征值只算一次) 为

$$\gamma_0 = (1 + o(1))(g(\mathcal{P}) \cdot n) \quad (3-7)$$

且其最大，最小特征值分别为：

$$\lambda_1 = (1 + o(1))(h(\mathcal{P}) \cdot n) \quad (3-8)$$

$$\lambda_k = (1 + o(1))(j(\mathcal{P}) \cdot n) \quad (3-9)$$

3.2.3 对特征向量和特征值误差的估计

对于对称矩阵 $A = G$ 和 $E = H(N - \sigma^2 \mathbf{1}_{n \times n})H$, 定义 $\tilde{A} = A + E$, 那么对于 $G = U^T \Lambda U$, $U = (u_1, \dots, u_k)^T$, $\Lambda = \text{Diag}(\lambda_1, \dots, \lambda_k, 0, \dots, 0)$, $\lambda_1 > \dots > \lambda_k > 0$, 假设第 i 个点还原后的坐标为 \tilde{p}_i , $\tilde{P} = (\tilde{p}_1, \dots, \tilde{p}_n)$, 那么有：

$$\tilde{P}_{i,\cdot} = \sqrt{\lambda_i} \cdot u_i \quad (3-10)$$

于是, 这里我们记 $\delta x := x - \tilde{x}$,

$$\begin{aligned} \delta P_{i,\cdot} &\sim \delta \sqrt{\lambda_i} \cdot u_i + \sqrt{\lambda_i} \cdot \delta u_i \\ &\sim \frac{\delta \lambda_i}{2\sqrt{\lambda_i}} u_i + \sqrt{\lambda_i} \cdot \delta u_i \end{aligned} \quad (3-11)$$

根据 Weyl 定理, 有:

$$|\lambda_i - \tilde{\lambda}_i| \leq \|E\|_2 \leq_* C_2 \sigma \sqrt{n} \quad (3-12)$$

由^[24]的(7)可得, 设 u_i, \tilde{u}_i 分别为 A 和 \tilde{A} 的第 i 大特征向量, 且 γ_0 为 G 的谱间隔, 那么

引理 3.4: 若 $\gamma_0 > 2\|E\|_2$, 则

$$\max_i \|u_i - \tilde{u}_i\| \leq \frac{2\sqrt{2}\|E\|_2}{\gamma_0} \quad (3-13)$$

因为 $\|H\|_2 = 1$, 所以根据定理3.5,

$$\|E\|_2 \leq \|N - \sigma^2 \mathbf{1}_{n \times n}\| \leq_* C_2 \sigma \sqrt{n} \quad (3-14)$$

其中 $a(n) \leq_* b(n)$ 表示 $\Pr[a(n) \leq b(n)] = 1 - neg(n)$ 。^① 结合 (3-7) 可知,

定理 3.6: 对于

$$n > \Theta(1) \frac{C_2^2 \sigma^2}{g^2(\mathcal{P})} \quad (3-15)$$

有

$$\max_i \|u_i - \tilde{u}_i\| \leq_* (1 + o(1)) \frac{2\sqrt{2}C_2}{g(\mathcal{P})} \cdot \frac{\sigma}{\sqrt{n}} \quad (3-16)$$

再记 $a \lesssim_* b$, 当 $a \leq_* (1 + o(1))b$ 。根据 (3-11)、(3-12) 和定理3.6得, 对于足够大 (满足3-15) 的 n , 有

$$\begin{aligned} \|\delta P\|_F^2 &\lesssim_* \sum_{i=1}^k \left(\frac{C_2 \sigma \sqrt{n}}{2\sqrt{\lambda_i}} u_i + \sqrt{\lambda_i} \frac{2\sqrt{2}C_2}{g(\mathcal{P})} \cdot \frac{\sigma}{\sqrt{n}} \right)^2 \\ &\leq 2C_2^2 \sigma^2 \left[\sum_{i=1}^k \left(\frac{\sqrt{n}}{2\sqrt{\lambda_i}} \right)^2 + \sum_{i=1}^k \left(\sqrt{\lambda_i} \frac{2\sqrt{2}}{g(\mathcal{P})\sqrt{n}} \right)^2 \right] \\ &\leq 2C_2^2 \sigma^2 k \left(\frac{n}{4\lambda_k} + \frac{8\lambda_1}{g^2(\mathcal{P})n} \right) \\ &\sim 2C_2^2 \sigma^2 k \left(\frac{1}{4j(\mathcal{P})} + \frac{8h(\mathcal{P})}{g^2(\mathcal{P})} \right) \end{aligned} \quad (3-17)$$

于是定理3.4得证。

注意到 h, j, g 分别表示点云在各个主方向上投影方差的最大值、最小值和最小间距。因此, 当每个方向上的散布宽度都不太小, 且不太接近之时, (3-17)

^① 这是一个比 a.a.s. 更强的条件, 它保证了 $poly(n)$ 个如此事件的交仍以 $1 - neg(n)$ 概率成立, 而后者只保证常数个 a.a.s. 事件的交仍 a.a.s..

能够对 cMDS 还原结果的整体结构误差给出一个较好的上界。此外，根据 (3-16) 和 (3-17)，我们还可以得到：

推论 3.1： 只要 $\|\tilde{D} - E[\tilde{D}]\|_2 \le_* \Delta(n) = o(n)$ ，则 $cMDS(\tilde{D} - (E[\tilde{D}] - D))$ 是 $[3, 1, 0] - MDS$ 。

之所以要求 $j(\mathcal{P})$ 不能过小，是因为当 $\lambda \rightarrow 0$ 时， $\sqrt{\lambda}$ 随 λ 的变化极为敏感。对于实际应用中的此类情况，我们可以直接将其视为一个较低维度的形体即可。

之所以要求 $g(\mathcal{P})$ 不能过小，是因为在此节证明中，变换 $S = (R; T)$ 是通过将重构前后的点云进行主成分分析，并将特征值从大到小排序，并将对应特征向量对齐的方式实现的。当 $g(\mathcal{P})$ 很小时，这种方法将难以在存在噪声影响的情况下对各个特征向量进行准确的匹配，从而难以对点云进行准确的对应，因此无法据此得到嵌入误差的估计。然而，实际上 3.1 中定义的误差是对所有可能的变换 S 求最小值，因此可能仍存在一个较好的上界。这个问题有待进一步研究。

3.3 对于有噪声非完全观测情况的初步分析

在实际应用中，很多情况下，并非每一对点的距离都能得到观测。在 3.1 节中已提到，对于 D 完全准确的情况，根据低秩特性，我们只需其一部分的值即可对其进行还原。在此节中，我们可以首先将其视为一个优化问题：已知输入距离矩阵 $\tilde{D} \in \mathcal{S}^n$ 和观测矩阵 $W \in \mathcal{S}^n, W_{ij} \in \{0, 1\}$ ，求一个 k 维的距离矩阵 $\hat{D} \in \mathcal{E}^n$ ，使得在

$$L_0 = \|W \circ (\hat{D} - \tilde{D})\|_F, \quad (3-18)$$

$$s.t. \quad \hat{D} \in \mathcal{E}^n \quad (3-19)$$

$$\text{rank}(H\hat{D}H) \leq k \quad (3-20)$$

中， L_0 尽可能小。

此优化问题的困难之处在于，(3-20) 是一个非凸的约束条件。在相关领域的文献中，常常可以用凸松弛^[12] 的方式，利用矩阵的 1-范数作为惩罚项，以对结果进行稀疏化，即化为如下的优化问题 (*)：

$$\text{minimize} \quad L_1 = \|W \circ (\hat{D} - \tilde{D})\|_F^2 + \mu \|H\hat{D}H\|_1 \quad (3-21)$$

$$s.t. \quad \hat{D} \in \mathcal{E}^n \quad (3-22)$$

根据^[15]，由所有 $n \times n$ 的距离矩阵组成的集合 \mathcal{E}^n 是凸集，因此这是一个凸优化问题。注意到 (3-22) 并不直接保证 \hat{D} 是一个不超过 k 维的距离矩阵，但我们可以对其进行 k 维的 cMDS，将其化为一个 k 维距离矩阵。

事实上，根据^[15]，如果 \hat{D} 的误差足够小，那么即使不加上 1-范数惩罚项，仍然能得到较好的结果，我们首先通过一系列实验进行验证。

3.3.1 实验探索

3.3.1.1 实验模型

探索这个问题时，我们首先用一系列实验来进行探究。

首先我们令 P 为心形线（这样 $\lambda_1 = \lambda_2$ ）上的 $n = 50$ 个点，即

$$P_i = (r_i \cos \theta_i, r_i \sin \theta_i) \quad (3-23)$$

其中 $\theta_i = i \cdot \frac{2\pi}{n}$, $r_i = 1 - \sin(\theta_i)$ 。

之后我们用 Erdos-Renyi 图模型生成 W ，即对于每组 $i < j$, $W_{ij} = W_{ji}$ 独立地以概率 p 为 1，以概率 $(1-p)$ 为 0。然后对于观测的每对点间距离，加上均值为 0、标准差为 $\sigma = 0.05$ 的高斯噪声，并用以上算法尝试进行重构，观察其效果。

3.3.1.2 初步实验结果

首先进行不加惩罚项 ($\mu = 0$) 的优化。实验发现，对于 $p = 0.50$ ，足以进行较为准确的重构（如图3.1）；对于 $p = 0.23$ ，能够基本正确地恢复出心形线结构（如图3.2）。注意到在多次实验中曲线的朝向不一致，这说明了对于不同的噪声， $\tilde{\lambda}_1$ 和 $\tilde{\lambda}_2$ 对应的特征向量完全不同，但依然能恢复出正确的形状。而对于 $p = 0.16$ ，则形状出现了严重的失真（如图3.3）。

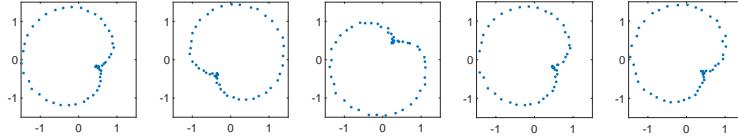


图 3.1 $p = 0.50, \mu = 0$ 时的重构效果

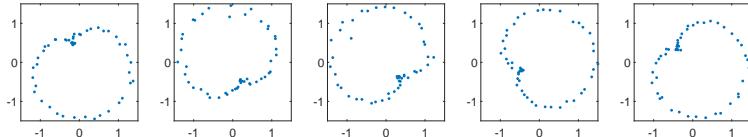


图 3.2 $p = 0.23, \mu = 0$ 时的重构效果

之后，我们对 $p = 0.16$ 的情况进行带有 l_1 惩罚项的优化，结果如图3.4、3.5、3.6所示。我们观察到，当 μ 较大时形状失真会变得更加严重，而当 μ 较小时，其表现与不加惩罚项时无显著差别，因此， l_1 惩罚项并不能明显降低重构所需数据量的下限。

我们再对 $p = 0.23$ 的情况进行带有 l_1 惩罚项的优化，探索当数据足够的情况下，是否能提升重构的效果。如图3.7、3.8、3.9可见，此正则化项亦无法明显提升重构表现。由此可见，影响此问题中重构准确程度的主要因素并非模型维度，而是输入数据信息量的局限性。因此，我们需要对 p 对重构效果的影响进行进一步分析。

3.4 基于随机稀疏图的 MDS 之可行性

首先我们从理论上分析无噪声之情况。由^[13]知，对于非退化情形的点云，当 $p = \Omega(k \log^2 n / n)$ 时，已有能对其结构进行准确重构的算法。显然当 $p = 0$ 时，我们无法进行任何重构，因此我们试图探究一个临界值，即对于 Erdos-Renyi 随机图模型而言，当 p 多大时，我们理论上可能对结构进行重构？

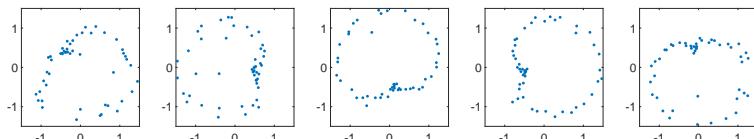


图 3.3 $p = 0.16, \mu = 0$ 时的重构效果

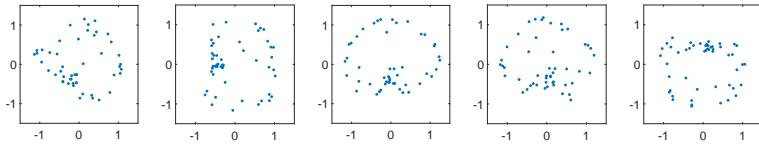


图 3.4 $p = 0.16, \mu = 1$ 时的重构效果

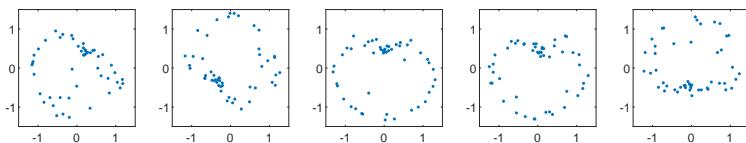


图 3.5 $p = 0.16, \mu = 0.2$ 时的重构效果

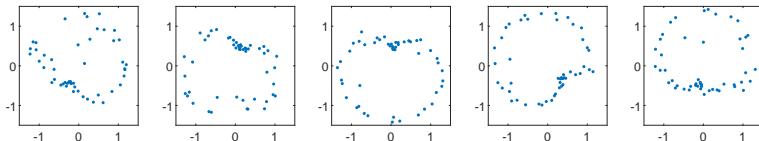


图 3.6 $p = 0.16, \mu = 0.05$ 时的重构效果

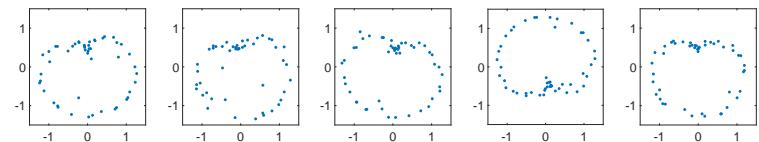


图 3.7 $p = 0.23, \mu = 1$ 时的重构效果

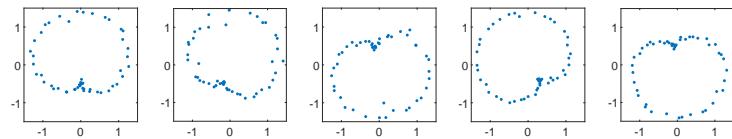


图 3.8 $p = 0.23, \mu = 0.2$ 时的重构效果

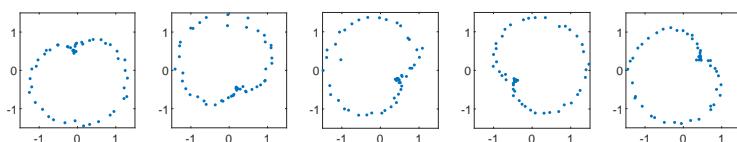


图 3.9 $p = 0.23, \mu = 0.05$ 时的重构效果

3.4.1 无向图的刚性性质

一个较为容易得到的结论是， p 不能小于 $\Omega(\log n/n)$ 。根据 Erdos-Renyi 模型^[16]，当 $p < (1 - \epsilon)\log n/n$ 时，对于足够大的 n ， W ^①以大概率不是连通图。因此可以将点云分为不连通的两部分，且其间没有任何边约束。在这种情况下，我们是不可能对整体结构进行准确的重构的。

对于 $k \in \{2, 3\}$ 的情况，现有的算法和不可能性结论之间存在一个 $\log n$ 的距离。因此，对于 $\log n/n \ll p \ll \log^2 n/n$ 的情形，还有待进一步的分析。

3.4.1.1 2 维全局刚性图

事实上，即便输入数据完全精确，重构的必要条件是 (W, D) 具有**刚性**，从物理意义上，即是在 n 个点之间按照图 W 连上刚性杆后，其结构无法进行局部变形。一个比刚性更加严格的条件是**全局刚性**，即给定这些刚性杆的长度，这些点只有至多一种可能的空间结构。

根据^{[25][17]}可知，在 2 维和 3 维的一般位置意义下，全局刚性是图 W 的性质而与 G 无关，因此我们可以称图 W 具有全局刚性，当且仅当对任一个一般位置的 D ， (W, D) 具有全局刚性。

根据定理1.1.2，在二维空间中，对于无噪声情况，理论上只须 $p = \Theta(\log n/n)$ 即可根据 $W = G(n, p)$ 还原出 n 个点的空间结构，因此^[13] 的算法可能仍有提升空间。

3.4.2 弱 MDS：对于 $\Theta(1/n)$ -稀疏之情况的重构可能性

我们首先考虑 2 维的情况。

在 Erdos-Renyi 图模型中，虽然在 $p = o(\log n/n)$ 时 W 非连通，从而其不具有刚性，但另一个方面，只要 $p > (1 + \epsilon)/n$ 时， $G(n, p)$ 就拥有一个大小为 $\Theta(n)$ 的连通分支^[16]。因此，对于 $p = \Omega(1/n)$ ，我们依然可能选择其中一部分的点，并根据它们来还原点云的大致结构。根据^[26]，当 $p = c/n$ 且 $c > 3.59$ 时， W 存在一个顶点数为 $\Theta(n)$ （取决于 c 的值且不小于 $0.74n$ ）的刚性子图。因此若能剔除不属于该子图的点，而对其中大部分“较好”的点进行 MDS，就可能在 p 更小的情况下对点云 P 进行部分重构，而不影响对其整体结构的识别。

^① 本文中讨论的图均为无向图，且符号上不区分一个图和它的邻接矩阵。

由^[26]可知，对于足够大的 c ， $G(n, c/n)$ 的 $(3+2)$ -core^① 占全图的比例是 c 的函数，且其刚性子图包含 $(3+2)$ -core 中 $(1 - o(n))$ 比例的点。由于 $(3+2)$ -core 可以在 $O(n^3)$ ^② 时间内求得，只要有 $p = \Omega(1/n)$ ，理论上仍可以在剔除至多常数比例的点后，对剩余的 $\Theta(n)$ 个点进行 MDS；因为其中 $(1 - o(n))$ 的点均在刚性子图上。

然而， $(3+2)$ -core 的结构并不一定能被正确地还原。考虑到其中度为 2 的顶点，将其关于两邻点连线作镜像对称，所得的框架仍满足约束，因此，一个图能够在 2 维空间中被正确还原（具有全局刚性）的一个必要条件是其每个点的度数均不小于 3，所以在 2 维弱 MDS 的实际应用中，从最简单的角度考虑，我们可以筛选出 W 的 d -core，其中 $d \geq 3$ 。

由于理论上判定一个图的全局刚性是一个强 NP-hard 问题^[25]，因此我们难以精确地找到点集的最大全局刚性子集，而定理1.1.1 中的 6-连通之充分条件既难以判定也明显过强（蕴涵 $d > 6$ ，而 $p > d/n$ ，据数值估计在 $n > 1000$ 时才可能对比定理1.1.1 更稀疏的情况进行求解）。直觉上看，全局刚性是比刚性更加严格的条件，因此取 $d > 3$ 更利于较好地筛选出较为可靠的点。经过我们的实验，在实际重构中（由于此节主要考虑图本身的刚性，对于距离矩阵我们取了较小的误差），取 $d = 4$ 确实能够去除一些不可靠点，如图3.10和3.11所示。其中每列代表一个不同的随机图，而每行代表不同的算法：

- 第一行：直接对全部点根据 (*) 进行优化，之后进行 cMDS。（对照）
- 第二行：对全部点根据进行优化，之后进行 cMDS，但只取 d -core 输出。
- 第三行：对全部点进行优化，取 d -core 进行 cMDS。
- 第四行：取 d -core 进行优化，之后进行 cMDS。

根据观察，第二、三种方案普遍能够过滤原图中部分不可靠点，但第四种在一些情况下效果相反比第一种更差。这说明不可靠点亦包含部分有益的信息，因而不应完全排除。因此，我们可选用第二种或第三种方案作为弱 MDS 的算法。

① 首先求出图 G 的最小度不小于 3 的极大子图 (3 -core)，然后递归并上与当前子图至少 2 个顶点相邻的点，直至无法继续进行操作时，形成的子图即为 G 的 $(3+2)$ -core。

② 猜想可以通过堆结构优化至 $O(n^2 \log n)$ （或用其他算法优化到更低的复杂度），但由于我们并不使用 $(3+2)$ -core，本文不作探讨。

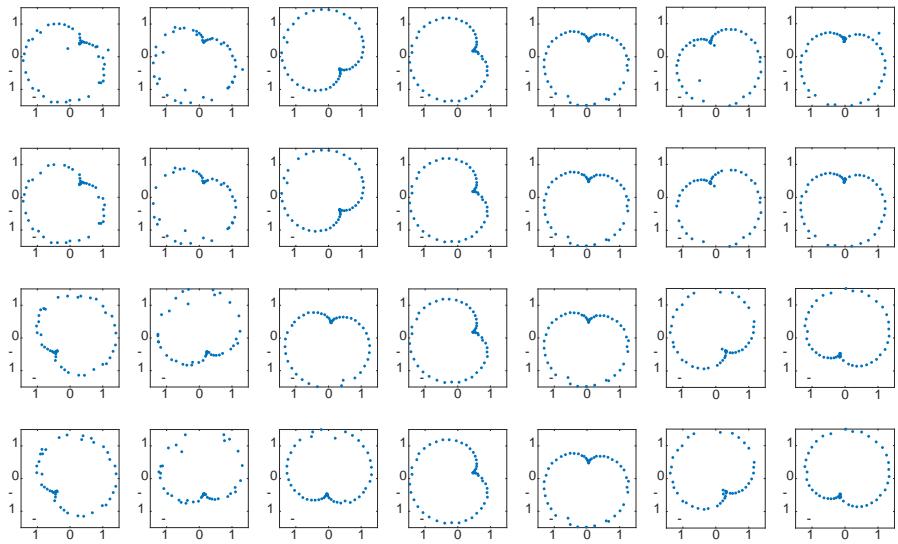


图 3.10 $n = 50, p = 0.15, \sigma = 0.005$ 时的弱 MDS 效果对比

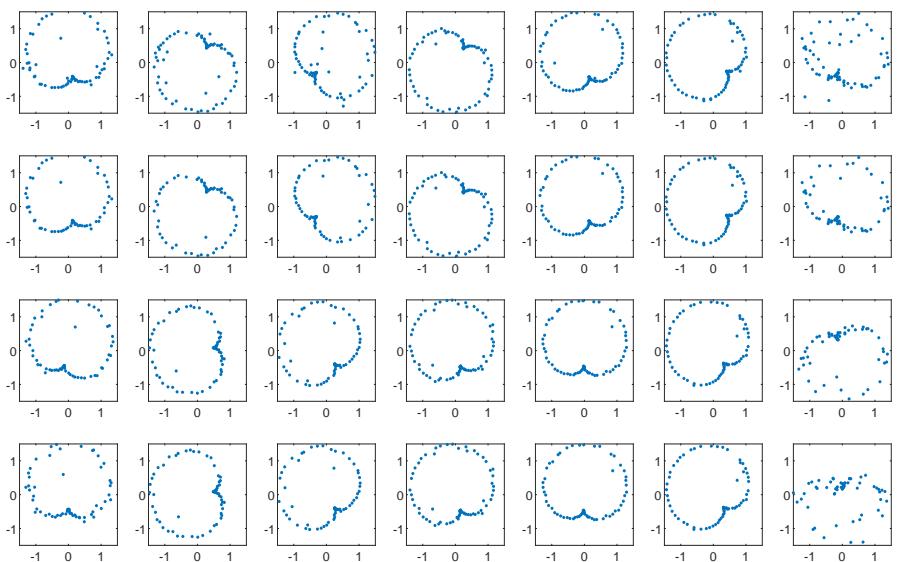


图 3.11 $n = 70, p = 0.10, \sigma = 0.005$ 时的弱 MDS 效果对比

3.4.3 高维情形

除前文图示中的二维情形中，前文中的各种 MDS 算法，都可自然地应用于高维情形。虽然对于高维刚性的理论分析仍为一开放问题，但实验结果显示，在需要更多边数的前提下，前文中的 MDS 算法在高维情形下仍有较为可靠的表现。下面的图例仅展示三维点云重构之效果。

在此节的实验中，我们令 $n = 50$, $\sigma = 0.03$ ，且点集 $P = \{p_i(x_i, y_i, z_i)\}$ 构成三维空间中的一条底面投影为椭圆的螺线：

$$x_i = 2 \cos(0.3i)$$

$$y_i = \sin(0.3i)$$

$$z_i = 0.1i$$

对于 $p = 1.00$ ，如图3.12所示，cMDS 算法能够较为准确地恢复螺旋线结构，仅能看到极小的失真；当 $p = 0.30$ 时，如图3.13所示，3.3节的算法亦能基本恢复出螺旋线的结构。

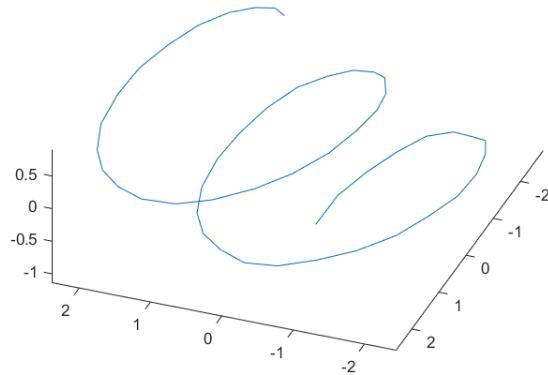


图 3.12 三维 cMDS 效果 ($n = 50, p = 1.00$)

当 $p = 0.22$ 时，3.3节的算法所恢复出的螺旋线已出现严重的失真（如图 3.14 左）。然而，若利用3.4.2节的策略，令 $d = 6$ ，通过该节提出的两种弱 MDS 策略，则能对 P 的 6-core（共 46 个点，占全部点的 92%）进行基本正确的重构，其效果与3.3节在 $p = 0.30$ 时的情况相当（如图 3.14 中、右）。

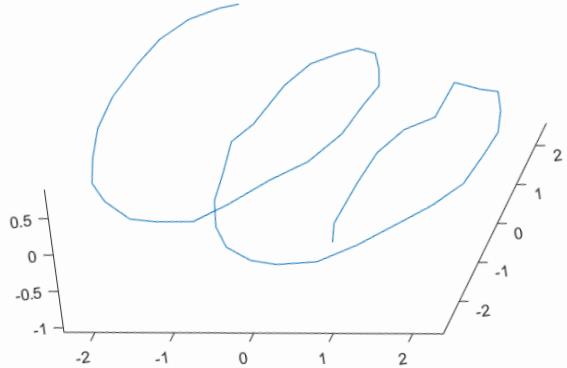


图 3.13 三维稀疏 MDS 效果 ($n = 50, p = 0.30$)

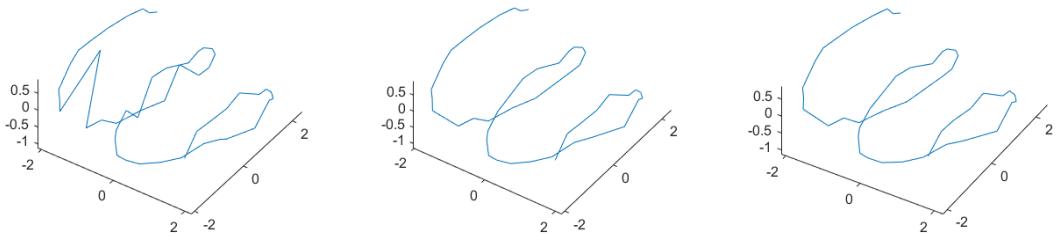


图 3.14 三维弱 MDS 效果 ($n = 50, p = 0.22, d = 6$)

3.5 小结

在本章中，我们提出了**几何稳定性**这一定量描述点云重构准确度的概念，并证明了当边长观测存在噪声但完全的情况下，cMDS 算法在常数维度下的几何稳定性，并利用二维刚性理论和 Erdos-Renyi 图理论中 Giant Component 的现象，提出了一种在距离观测随机而稀疏的情况下，选取点云中大部分可靠点并还原其相对空间位置的算法。前者从理论上证明了 cMDS 算法在恢复空间结构意义上的可靠性，而后者指出了一种在随机图边数低于整体刚性的理论极限的情况下，仍能恢复点云所蕴含的几何结构的可能性。

然而，和 2 维刚性不同的是，对 3 维刚性的刻画在本质上更为复杂，因为 Laman 形式的条件仅为 3 维刚性的一个必要非充分条件，而 3 维刚性的充分条件需要利用平衡张力矩阵的工具^{[25][17]}。因此虽然 3 维全局刚性在一般位置意义下仅依赖于邻接矩阵，从而是一个图的性质，但现有的工具还无法直接从图的意义上对 3 维刚性进行准确的判定。

根据实验模拟，在 3 维空间中进行重构除需要略多的边之外，和 2 维情况

无本质差别，因此本人猜想使一个 Erdos-Renyi 图具有 3 维全局刚性和含有 $\Theta(n)$ 个点的 3 维全局刚性所需要的 p 仍分别为 $\Theta(\log n/n)$ 和 $\Theta(1/n)$ 。由于 3 维全局刚性蕴涵 2 维全局刚性，所以该界的必要性自然成立。然而，其充分性仍有待进一步探索和证明。

第4章 面向几何稳定性的网格优化设计与传感器定位

在第3章中，我们讨论了 W 为完全图与随机稀疏图的情形，分析了保证一个点云结构具有全局刚性所需的刚性约束数目的最小值；然而，考虑在无线传感器网络、几何建模以至建筑结构设计等领域的应用中，我们并不必被动地观测点与点之间的边约束，而可以主动在这些点之间加上一些边使其具有刚性。

在这种情况下，第3章的结论提供了一个最基本的保证：只要随机连上一定数量的边，所得结构几乎必定具有刚性性质。但在实际应用中，为考虑经济性，我们往往希望通过特定的算法来构造一个“较好的”图 W ，或对已有的图做某种程度的优化，例如减少边数，或减少总长度等等，以节约材料成本。因此，本章将从组合优化的角度，试图解决此类问题。

4.1 概述

在第3章中，我们探讨了完全图和 Erdos-Renyi 随机图的刚性性质，证明了完全图的几何稳定性，并对基于 Erdos-Renyi 图模型的稀疏距离矩阵进行了欧氏嵌入的尝试，说明了只须随机询问 n 个点间 $\Theta(n \log n)$ 或 $\Theta(n)$ 条边，理论上即可恢复这些点全部/大部分的几何结构。相比基于随机图的重构，在一些实际问题中我们更关注“最优图结构”，即在满足一定刚性性质的前提下，尽可能地减小所得图的“成本”。为解决此问题，我们需要研究保证一个图具有刚性性质所需的“最低要求”。

对于一个刚性图所需的最少边数，根据 Maxwell-Laman 定理^[17]：

引理 4.1 (Maxwell-Laman): 一个图 W , $|V(W)| \geq k$ 具有 k 维刚性的必要条件是其存在一个生成子图 L ，使得：

- $|E(L)| = k|V(L)| - \frac{k(k+1)}{2}$;
- $\forall W \subseteq V(L) \text{ s.t. } |W| \geq k, |E(L_W)| \leq k|W| - \frac{k(k+1)}{2}$ 。

其中对于 $k = 2$ 为充要条件，但对 $k \geq 3$ 不充分。

我们知道由 $n(\geq k)$ 个点构成的 k 维刚性图至少含有 $kn - \frac{k(k+1)}{2}$ 条边。又根据^[25]，有条件

引理 4.2 (Connelly): 图 W 在 k 维空间中具有一般位置^① 意义下的全局刚性的必要条件是：

- W 是 $(k+1)$ -连通图；
- $\forall e \in E(W), W \setminus \{e\}$ 是刚性图。

因此，全局刚性图至少含有 $kn - \frac{k(k+1)}{2} + 1$ 条边。

在考虑具有某种性质的“最小”生成子图时，除考虑最小边数之外，我们还常常需要优化其最小边权和，以满足实际问题中最小成本的需求，如最小生成树问题、旅行商问题等。Jordán 等人^[27] 还探讨了 k 维欧氏空间中点集的最小全局刚性生成子图问题，并对每个 k 得出了一个常数近似比的近似算法，其中对 $k = 2, 3$ 的近似比分别为 2 和 $4 + o(1)$ 。

然而，^[27] 中给出的全局刚性子图并未给出相应的定位算法，即虽然给定图 W 中每条边对应的距离约束，确实存在唯一的满足此约束之结构，但根据这些稀疏约束求出每个点的相对位置仍需通过一系列的矩阵填充的算法（时间复杂度较大），且在边长存在误差时，其几何稳定性并无理论保证：若每个点只与其附近小范围的点相连，则其误差的累积效应可能使整个几何结构的误差远大于局部边长误差。

事实上，该文献中对于一般的 k 的解决方式正是先找到关于 n 个点近似最短的 Hamilton 回路 p_{H_1}, \dots, p_{H_n} ，再将 Hamilton 回路上所有距离不超过 k 的点相连。这里我们考虑点集 $P = \{p_t = (\cos \frac{2\pi t}{n}, \sin \frac{2\pi t}{n}) \mid t \in [n]\}$ 和 $Q = (1 + nr/k)P$ ，则 P 与 Q 对应边长度之差仅为 $O(r)$ ，但它们的 Hausdorff, Bottleneck, 2-Wasserstein 距离均为 nr/k 。这说明当图 W 的图论直径^② 较大时，基于局部点对距离对整体结构的重构确实会因累积误差而变得不准确。因此，我们考虑适当放宽“最小成本”的条件，在保证其几何稳定性的前提下，构造成本尽可能小的全局刚性图。

4.2 无线传感器定位模型

在室内多点定位等问题中，我们往往需要通过多个传感器之间的距离关系来对整体环境进行重构。第3章的内容证明了当 n 个传感器两两相连时，根据其间的 $\frac{n(n-1)}{2}$ 组距离即可稳定地求出这些传感器之间的相对位置。然而在实际应用

① 定义见1.2.3.2节，后文将直接简称为“ k 维全局刚性”。

② 即从一点到另一点最短路径边数的最大值。

中，建立信道存在功耗、频谱等成本，因此在网络结构足以稳定得出传感器间的几何关系的前提下，尽可能降低通讯成本，亦是一个需要研究的问题。上一节中关于最小全局刚性子图的分析说明了根据^[27] 的算法得出的图结构虽具有全局刚性，但在距离数据存在一定误差时，可能导致对整体结构的还原出现较大失真。因此，为解决最优定位网络问题，我们需要构造一个最小成本的**几何稳定的全局刚性图**，即：

定义 4.1(最小成本定位图): 对于 k 维空间中的 n 个点 $P = p_{[n]}$ 和 P 的一个子集 P_- ，试求图 W s.t. $V(W) = P$ 和可能的其他标记 L ，使得根据 W 中所有边在 P 中的长度加上 $\delta \sim N^{|E(W)|}(0, \sigma^2)$ 的绝对误差得到的输入 $Q = \{\|x_i - x_j\|\}_{E(W)} + \delta$ ，存在一个算法 f ，使得 $K = f(Q, L)$ 满足

$$\lim_{n \rightarrow \infty} \Pr_{P \sim \mathcal{P}; \delta \sim N^{|E(W)|}(0, \sigma^2)} [\exists S \in \mathbb{E}(k) \text{ s.t. } \|(K - SP)|_{P_-}\|_F \leq \epsilon \sqrt{n}] > 0.99 \quad (4-1)$$

其中 $(K - SP)|_{P_-}$ 表示 P_- 对应的列构成的子矩阵。且对于给定的成本函数 $c : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ s.t. $c(x) = z(x) + b$ (其中 z 为连续不减函数， $z(0) = 0$)，使

$$C = \sum_{e(p_i, p_j) \in W} (z(\|p_i - p_j\|_2) + b) \quad (4-2)$$

最小。

因为 W 的选取和 P 的具体分布有关，我们假设 $\|p_i - p_j\|_2$ 的排序是已知的，但询问具体值则需要 $z(\|p_i - p_j\|_2) + b$ 的代价。

在此模型中，我们假设每建立一条信道需要固定的成本（如频谱等），我们记作 b ；而距离越远时需要越大的发射功率，将其带来的额外成本记作 $z(\cdot)$ 。当 $b = 0, z(x) = x$ 时目标即类似^[27] 中的最小总长问题，而 $z(x) = 0$ 时即为上一节中的最小边数问题。

4.3 基于基准点框架的网格构造

4.3.1 贪心的均匀采样方案

直观上，为使局部误差的积累不致引起整体结构的不稳定，我们需要一些连接非邻近点的边，使得整体结构能够得到一定的约束，起到“骨架”的作用。由定理3.4知完全图具有几何稳定性，因此结合2.2.3节的思路，我们可以首先在 P 中选取 $\rho(n) = \omega(1)$ （待定，计算总成本时再做优化^[28]）个代表点组成点集 R ，将其中任两点相连接，则 R 已具有几何稳定性。之后，对于 $P \setminus R$ 中的点，则根据其与 R 中点的距离关系来推断它们的位置。

为使总边数在最优值的常数比例之内，即不超过 $O(n)$ ，需要 $\rho(n) = O(\sqrt{n})$ 。为使其余点附近不太大的邻域内都存在 R 中的点，我们希望 R 的分布尽可能均匀。

在这里，我们总假设 $n \geq \rho(n) \geq k + 2$ 。

事实上，我们定义：

定义 4.2： 点集 Y 关于距离度量 $\|\cdot\|$ 是 ϵ -稀疏的，当且仅当 $\forall y, y' \in Y, \|y - y'\| < \epsilon \implies y = y'$ 。

定义 4.3： 点集 Y 是 X 关于距离度量 $\|\cdot\|$ 的一个 ϵ -覆盖，当且仅当 $\forall x \in X, \exists y \in Y, s.t. \|x - y\| \leq \epsilon$ 。

定义 4.4： 对于点集 Y ，定义 Y 的 t -膨胀为 $B(Y, t) = \{v | \exists y \in Y, \|y - v\| \leq t\}$ 。

那么用贪心的方式，对 P 进行最远点优先采样，则得到的 R 是 P 的一个 e -稀疏- e -覆盖，又称 e -网^[29]。对于 e -网，据定义有：

命题 4.1： 若点集 Y 是 X 的一个 ϵ -网（记为 $Y \in \mathcal{E}_X(\epsilon)$ ），则：

1. $X \subseteq B(Y, \epsilon)$ ；
2. $B(Y, \epsilon/2)$ 为两两不交的球之并。

根据命题4.1，对于 k 维空间，设 k 维单位球的体积为 U_k ，记点集 S 的体积为 $V(S)$ 。对 $Y \in \mathcal{E}_X(\epsilon)$ ，由于 $B(Y, \epsilon/2) \subseteq B(X, \epsilon/2)$ ，我们有：

$$|Y| U_k (\epsilon/2)^k \leq V(B(X, \epsilon/2)) \quad (4-3)$$

Algorithm 1 最远点优先采样

```

1:  $d[1 : n] = +\infty;$ 
2: for  $t = 1 : \rho(n)$  do
3:   选取使  $d[i]$  最大的  $i$  (之一)
4:    $e = d[i]$ 
5:    $r_t = p_i$ 
6:   for  $j = 1 : n$  do
7:      $d[j] = \min(d[j], \|p_i - p_j\|)$ 
8:   输出  $R = \{r_{[\rho(n)]}\}, e$ 
```

注：虽然 P 中点的坐标与 $\|p_i - p_j\|$ 均未知，但问题定义中我们已知 $\{\|p_i - p_j\|\}$ 的序关系，因此第 7 步可以正确运行。

设 $Diam(X)$ 为 X 的直径，即 $Diam(X) = \sup_{x, x' \in X} \|x - x'\|$ ，那么有：

$$\forall x_0 \in X, \quad B(X, \epsilon/2) \subseteq B(x_0, Diam(X) + \epsilon/2) \quad (4-4)$$

于是，

$$V(B(X, \epsilon/2)) \leq U_k(Diam(X) + \epsilon/2)^k \quad (4-5)$$

据 (4-3) 与 (4-5) 可得：

定理 4.1： 若 $Y \in \mathcal{E}_X(\epsilon)$ ，则：

$$\epsilon \leq \frac{2Diam(X)}{\sqrt[k]{|Y|} - 1} \quad (4-6)$$

我们记 $\mathcal{D} = Diam(X)$ ，由定理4.1可得，对于 P 中的每一个点 p ， R 中都存在一个与之距离不超过 e 的点 r_p ，其中

$$e = \frac{2\mathcal{D}}{\sqrt[k]{|R|} - 1} = O(\mathcal{D}\rho^{-\frac{1}{k}}(n)) \quad (4-7)$$

一般地，有：

命题 4.2：若 $B(R, e)$ 连通，则 $\forall p \in P, l \leq |R|$,

$$|\{r \in R : \|p - r\| \leq (2l - 1)e\}| \geq l \quad (4-8)$$

事实上，我们只须作 $B(R, e)$ 关于 R 的 Voronoi 划分^[30]，设 $r \in R$ 对应的面片为 V_r ，且 $p \in V_{r_0}$ ，那么有

$$B(r, e/2) \subseteq V_r \subseteq B(r, e) \quad (4-9)$$

再作一条连续闭曲线 ω ，使 ω 是 $B(R, e/2)$ 的 $e/2$ -覆盖（只需按算法 1 作可 $e/2$ -覆盖 $B(R, e/2)$ 的离散点集，再取其任意一条 Hamilton 回路即可）。则对于任意 $r \in R$ ， ω 必与 $B(r, e/2)$ 相交，从而与 V_r 相交。通过 ω 经过每一个 V_r 的顺序，可诱导一条经过 R 中所有点（每个点至少一次）的一条回路 C ，由广度优先搜索可知， C 中与 r_0 距离不超过 l 的顶点不少于 l 个。

对于其中每个顶点 r_i ，考虑其在 C 中到 r_{i_0} 的最短路径，其中每条边的两顶点 r, r' 对应的面片相邻，于是 $B(r, e) \cap B(r', e) \neq \emptyset$ ，故 $\|r - r'\| \leq 2e$ ，因此 $\|p - r_i\| \leq \|p - r_{i_0}\| + \|r_i - r_{i_0}\| \leq (2l - 1)e$ 。

故命题 4.2 得证。即只要 $B(R, e)$ 连通，则 R 中距 p 第 l 近的点与 p 距离不超过 $(2l - 1)e$ 。

据此，我们若将 R 中任两点相连，同时将 $P \setminus R$ 中每个点与 R 中任意 $k + 1$ 个距离不超过 $(2k + 1)e$ 的点相连，那么总成本为

$$\begin{aligned} C &= O(\rho^2(n)(b + z(\mathcal{D})) + (k + 1)n(b + z((2k + 1)e))) \\ &= O(z(\mathcal{D})\rho^2(n) + kbn + knz(O(k\mathcal{D})\rho^{-1/k}(n))) \end{aligned} \quad (4-10)$$

考虑最短总长的目标，即令 $b = 0, z(x) = x$ ，并对于固定的 k ，

$$C = O(\mathcal{D} \cdot (\rho^2(n) + n\rho^{-1/k}(n))) \quad (4-11)$$

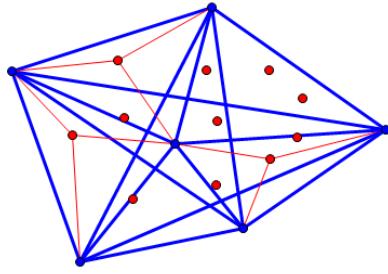


图 4.1 网格构造示意图

为使 C 无穷大的阶最小，我们使 $\rho^2(n)$ 与 $n\rho^{-1/k}(n)$ 同阶^[28]，此时 $\rho(n) = \Theta(n^{\frac{k}{2k+1}})$ 。对于充分大的 n ，

$$C \leq \Lambda_k \mathcal{D} n^{\frac{2k}{2k+1}}. \quad (4-12)$$

其中 Λ_k 是仅依赖于维数 k 的常数。

而对于 b 为大于 0 的常数的情况，在 (4-11) 中，对于固定的 k ，由于 $\lim_{x \rightarrow \infty} z(x) = z(0) = 0$ ，只要 $\rho(n) = \omega(1)$ ，就有

$$knz(O(k\mathcal{D})\rho^{-1/k}(n)) = o(n) \quad (4-13)$$

从而，

$$C = O(z(\mathcal{D})\rho^2(n) + bn) \quad (4-14)$$

于是只要 $z(\mathcal{D}) = o(n)$ ，就可以令 $\rho(n) = \omega(1)$ 且 $\rho(n) = O\left(\sqrt{\frac{n}{z(\mathcal{D})}}\right)$ ，使得

$$C = O(bn) \quad (4-15)$$

注意到 (4-15) 的结果和最优解至多相差常数倍，因为 n 个点的刚性图至少需要 $\Omega(n)$ 条边。事实上，在 $n \rightarrow \infty$ 时，这个近似比可任意接近 $\frac{k+1}{k}$ 。

我们将在下两节证明该网格的全局刚性与 $k = 2$ 时的几何稳定性。

4.3.2 全局刚性的证明

在4.1节中，我们指出了一个图具有全局刚性的一系列必要条件，以此推出了全局刚性图的一系列“下界”性质。然而，为保证构造一个满足全局刚性的图，我们需要全局刚性的一些充分条件。

Jordán 等人在^[27] 中指出了全局刚性图的一些递归性质，并以此递归地构造了最小边长和不超过最优值常数倍的全局刚性图，我们受其启发，完成了4.3.1节中的构造，并证明其全局刚性：

引理 4.3 (^[27]): 对于任意 k 维空间和 $m \in \mathbb{N}_+$, m 阶完全图 K_m 是全局刚性的。

引理 4.4 (^[27]): 对于任意 k 维空间和 $|G| \geq k + 2$, G 是全局刚性的，当且仅当它能由 K_{k+2} 通过一系列的以下操作生成：

- A 操作: 连接任两个原不相邻的顶点。
- H 操作: 取边 $e(v, v') \in E(G)$, 新建一顶点 w , 取 $V_+ \subseteq V(G) \setminus \{v, v'\}$, $|V_+| = k - 1$, 删除边 e , 并将 $V_+ \cup \{v, v'\}$ 中的每个顶点分别与 w 相连。

于是，可将4.3.1节中的操作视为，先作一完全图 $K_{\rho(n)}$, 在之后每轮里，我们首先加入一个顶点，将其与 $k + 1$ 个已有点相连并删除 $K_{\rho(n)}$ 中的一条边 (H 操作)，再将被删的边加入图中 (A 操作)，直到 P 中全部点均已加入。因此，由于 $\rho(n) \geq k + 1$, 4.3.1节所得的网格确为 k 维全局刚性的。

4.3.3 二维几何稳定性的证明

我们首先考虑，当 P 足够稠密地覆盖一个“实心”区域 \mathcal{L} , 而点 p 处于 \mathcal{L} 内且远离边缘时，我们可以在 p 附近找到 R 中的 3 个点，并根据它们来稳定地确定 p 的位置。事实上，我们有定理：

定理 4.2: 对 $k = 2$, $p \in P$, 若 P 是 $B(p, e + \delta)$ 的一个 δ -覆盖，则

$$|R \cap B(p, 2e + \delta)| \geq 3, \quad (4-16)$$

且 $\exists r_i, r_j \in (R \cap B(p, 2e + \delta))$, 使得

$$\angle r_i p r_j \in \left[\frac{\pi}{2} - \phi, \frac{\pi}{2} + \phi \right], \quad (4-17)$$

其中

$$\phi = \arccos \frac{1}{2(1 + \frac{\delta}{e})}. \quad (4-18)$$

证明 因 P 是 $B(p, e + \delta)$ 的一个 δ -覆盖, 故 P 也是圆周 $\partial B(p, e + \delta)$ 的一个 δ -覆盖, 故对于 $\partial B(p, e + \delta)$ 上任一点 u , 存在对应的 $p' \in P$ 使 $\|u - p'\| \leq \delta$ 。又因 R 是 P 的一个 δ -覆盖, 所以 $\exists r' \in R$ 使 $\|p' - r'\| \leq e$, 从而有 $\|u - r'\| \leq e + \delta$ 。因此, R 是 $\partial B(p, e + \delta)$ 的一个 $(e + \delta)$ -覆盖。

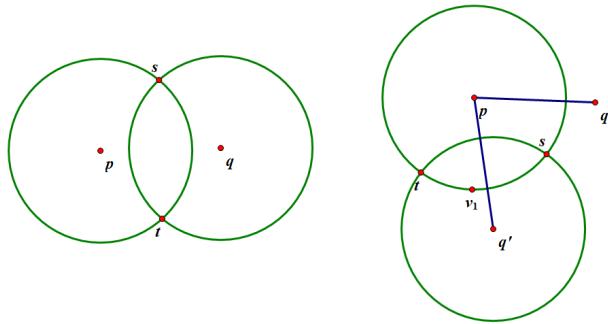


图 4.2 定理4.2证明图例

考慮

$$Q = \{q \in R \mid B(q, e + \delta) \cap \partial B(p, e + \delta) \neq \emptyset\} \quad (4-19)$$

由于 R 是 e -稀疏的, 对于任意 $q \in Q$, 有 $\|p - q\| \geq e$ 。设 $\partial B(p, e + \delta)$ 与 $\partial B(q, e + \delta)$ 交于点 s, t , 则 s, t 均在 $p - q$ 的中垂线上。由于 $\frac{\|p - q\|}{2} \geq \frac{e}{2}$, 可得 s, t 之间的弧 $A_q = \partial B(p, e + \delta) \cap B(q, e + \delta)$ 的圆心角为

$$\psi_q = 2 \arccos\left(\frac{\|p - q\|}{2(e + \delta)}\right) \leq 2\phi. \quad (4-20)$$

由于 R 是圆周 $\partial B(p, e + \delta)$ 的 $(e + \delta)$ -覆盖, 故 Q 也是。于是

$$\sum_{q \in Q} \psi_q \geq 2\pi. \quad (4-21)$$

又因 $\phi \in (0, \pi/2)$, 得

$$|Q| \geq 3. \quad (4-22)$$

(4-16) 得证。

任取 $q_0 \in Q$, 过 p 作 $p - q_0$ 的垂线, 与 $\partial B(p, e + \delta)$ 交于两点 v_1, v_2 。由于

$$\{A_q | q \in Q\} = \partial B(p, e + \delta), \quad (4-23)$$

必 $\exists q'$ 使得 $v_1 \in A_{q'}$ 。已证 $A_{q'}$ 为 q' 两侧圆心角各不超过 ϕ 的弧之并, 因此 q' 与 v_1 之间弧的圆心角不超过 ϕ , 故令 $r_i = q_0$, $r_j = q'$, (4-17) 得证。 ■

之后我们作 $\angle r_i p r_j$ 对应优弧的平分点 l , 亦存在 $r_k \in Q$ 使得 $l \in A_{r_k}$, 那么同理有 $\angle r_k p l \leq \phi$ 。因此, 对于 p 关于直线 $r_i r_j$ 的对称点 p' , 必有 $\|p - r_k\| \leq \|p' - r_k\|$ 。

于是, 我们将根据 r_i, r_j, r_k 的位置和 $\|r_i - p\|, \|r_j - p\|$ 来推断 p 的位置。一般地, 分别以 r_i, r_j 为中心, $d_i = \|r_i - p\|, d_j = \|r_j - p\|$ 为半径作圆, 取两圆交点中距离 r_k 较近的一个, 记为 $\tilde{p} = \tilde{p}(r_i, d_i, r_j, d_j, r_k)$ 。我们有:

定理 4.3: 对于仅依赖于 ϕ 的常数 $\lambda_\phi = \frac{1}{\cos \phi} = 2(1 + \frac{\delta}{e})$,

$$\frac{\|\partial \tilde{p}\|}{\partial d_i} \leq \lambda_\phi \quad (4-24)$$

$$\frac{\|\partial \tilde{p}\|}{\partial d_j} \leq \lambda_\phi \quad (4-25)$$

且

$$\left\| \frac{\partial(\tilde{p})}{\partial(r_i)} \right\| \leq \lambda_\phi \quad (4-26)$$

$$\left\| \frac{\partial(\tilde{p})}{\partial(r_j)} \right\| \leq \lambda_\phi \quad (4-27)$$

其中 $\frac{\partial(y)}{\partial(x)}$ 表示 Jacobi 矩阵。

证明 我们只须证明 (4-24) 和 (4-26)。

首先证明(4-24)。固定 r_i, r_j 和 d_j , 根据余弦定理, 在 $|d_i - \|r_i - r_j\|| < d_j$ 或 $\alpha \in (0, \pi)$ 的条件下, 已知标量 α 和 d_i 中的任一个, 都可唯一确定另一个, 且在 \tilde{p} 与 r_k 在直线 $r_i r_j$ 同侧的约束下, \tilde{p} 的位置可以唯一确定。

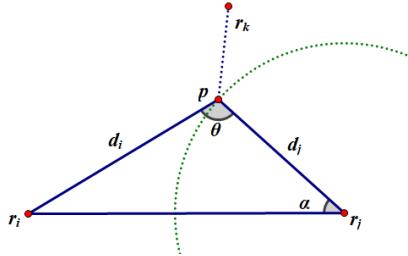


图 4.3 定理4.3证明图例

将 d_i 视为 α 的函数, 则有

$$\frac{dd_i}{d\alpha} = d_j \sin \alpha \quad (4-28)$$

于是根据反函数定理, 有

$$\frac{d\alpha}{dd_i} = \frac{1}{d_j \sin \alpha} \quad (4-29)$$

又因 d_j 固定时, p 只能在 $\partial B(r_j, d_j)$ 上运动, 所以

$$\frac{\|\partial p\|}{\partial \alpha} = d_j \quad (4-30)$$

$$\frac{\|\partial p\|}{\partial d_i} = \frac{1}{\sin \alpha}. \quad (4-31)$$

由于 $|\theta - \frac{\pi}{2}| \leq \phi$, 令 $\lambda_\phi = \frac{1}{\cos \phi} = 2(1 + \frac{\delta}{e})$, (4-24)得证。

再考虑(4-26)。固定 p , 根据三角不等式, 当 r_i 扰动不超过 δ 时, d_i 的变化不超过 δ 。因此, 若移动 \tilde{p} 使 d_i 仍满足约束, 根据(4-24), \tilde{p} 的扰动不超过 $(1 + o(1))\delta \lambda_\phi$ 。故(4-26)得证。 ■

根据 3.2 节的结论，我们知道 R 中的点满足几何稳定性，而根据定理 4.2 和定理 4.3，可得对于 P “内部远离边界”的点，亦满足几何稳定性。

对于任意固定的 k ，亦可用类似的方式证明，可以在每个点邻域内找到 $k+1$ 个 R 中一般位置的点，用同样的方式几何稳定地重构 P 。此处暂不作具体证明。

4.4 对比分析

在 3-4 章，本文主要分析了在根据欧氏空间离散点集中根据点对间距离重构点集的整体结构问题需求下，所需的几何约束模型与重构效果之间的关系。我们首先证明了当一般位置的 n 个点以完全图相连时，全局刚性与几何稳定性都有较好的保证，但其边数和总长存在较大的冗余，因此，在无线传感器网络等实际应用中，我们可以将其进行一些简化。

算法	维数 k	边数	总长	全局刚性	几何稳定性	建图复杂度	重构复杂度(2D)
全连接	任意	$\Theta(n^2)$	$O(n^2 \mathcal{D})$	是	是	$O(n^2)$	$O(n^3)$
Erdos-Renyi	2, 猜想任意	$\Theta(c_1(k)n \log n)$	$O(c_1(k)n \log n \mathcal{D})$	全部点, a.a.s.	未知	$O(n^2)$	SDP
Erdos-Renyi (弱)	2, 猜想任意	$\Theta(c_2(k)n)$	$O(c_2(k)n \mathcal{D})$	$\Theta(n)$ 个点, a.a.s.	未知	$O(n^2)$	SDP
最小全局刚性图	任意	$\Theta(kn)$	$O(k \cdot \text{OPT}, \Omega(c_3(k)n^{\frac{k-1}{k}} \mathcal{D})$	是	否	$O(n \log n)$	SDP
本节方案	任意	$\Theta(kn)$	$O(k^2 n^{\frac{2k}{2k+1}} \mathcal{D})$	是	非边缘点	$O(n \log n)$	$O(n^{6/5})$

图 4.4 多种图结构的性能对比

最为简单的想法是随机地在 n 个点间连边。根据 Erdos-Renyi 图模型的结论，至少需要 $\Omega(n \log n)$ 条边才能保证图以大概率连通，且在 2 维情况下， $\Theta(n \log n)$ 条边亦足以使图具有全局刚性。如果允许选择性移除部分点以使整体结构能得到还原，边数还可降低至 $\Theta(n)$ 。由于 k 维全局刚性图至少含有 $\Omega(kn)$ 条边，此边数与理论极限仅相隔一个常数。然而，对于基于 Erdos-Renyi 图的 MDS 和弱 MDS，我们并没有一个直接的算法对其进行求解，而需要将其转化为一个较为复杂的半定规划问题。虽然该问题为凸优化问题，可求得其全局最优解，但其时间复杂度较高。据实验，对 $n = 75$ 的情况，使用 CVX 进行求解已需要数小时时间。

在一些实际问题，如无线通信、建筑结构设计中，我们除关注所需的边数之外，每条边的长度亦和实际成本密切相关。在文献^[27] 中，作者提出了在给定欧氏空间中 n 个点坐标，求得其小总边长刚性图的近似解。然而，在输入边长数据

存在误差时，基于 Erdos-Renyi 图和最小全局刚性图的重构均无几何稳定性的理论保证，因此重构结果可能对输入误差敏感。尤其是最小全局刚性图的设计中，最小总长的目标使得其所有边都倾向于仅连接邻近的点，因此虽然其理论上有全局刚性的保证，但在多点定位问题中会有严重的累积误差问题，在建筑结构设计中也易因小的局部变形导致整体结构出现不稳定的现象。此外，当前亦没有基于最小全局刚性图进行重构的直接算法，因此其运算复杂度同样较高。

而本节提出的定位网格方案，在保证全局刚性、非边缘点的几何稳定性的同时，存在一个直接求解的重构算法，且其边数、总长和计算复杂度都相比完全图 MDS 有极大提升，亦明显优于基于 Erdos-Renyi 图的算法。相比最小全局刚性图算法，亦只有总长相对较长，但在全局几何稳定性的要求下，这个次线性的总长成本已较为令人满意。

第 5 章 结论

本文中，我们以 ClusterSLAM 中基于特征点间几何距离的聚类算法作为背景，从理论上分析了距离约束作为描述点集结构的特征的有效性，并探究了在保证点集整体结构能够被有效还原的前提下，尽可能除去冗余几何约束的方案。

对于物理世界的描述，几何特征往往最为直接而本质，但对于当今计算机视觉领域的大部分问题，主流的解决方案是利用机器学习、深度学习等工具进行基于图像的识别和处理。这是因为在实际应用中，基于点阵数据的输入图像虽然包含场景的几何特征，但容易受到光照、相机分辨率等多个因素的影响，从而包含一定程度的噪声甚至错误^[31]，而当今计算几何等领域的研究往往缺乏对输入噪声稳定性的理论探讨，因此在计算机视觉领域应用有限。在本文中，我们基于传统的**刚性**理论提出了更强的**几何稳定性**理论，刻画了在点对距离输入存在一定误差时恢复点集整体空间结构的可能性，并提出了一种针对输入点集自适应地生成代价尽可能小的几何稳定连接的算法。基于本文中的理论与算法，我们将可以更加有效地将几何工具运用在计算机视觉问题，以及无线网络定位、几何特征提取、建筑结构设计等其他实际问题之中。

当然，本文提出的理论与算法也存在一些值得改进之处，如：

- 当前的几何稳定性理论仅针对独立而随机的噪声，对于更加一般的噪声，或者存在错误的情况的处理（鲁棒性），还有待进一步研究。
- 由于基于特征向量的点集对齐方案之局限性，几何稳定性的理论保证会对一些特殊的点集分布失效。有待寻找更加有效的数学工具来进行更加细致的分析。
- 第 4 章的算法设计还相对较为粗糙，可能可以进一步优化。同时，对于几何稳定结构的代价下界还有待探究。

在今后的工作中，我们希望能够通过更深刻的数学工具来解决以上问题，同时继续将几何模型与机器学习、优化等工具相结合，以扬长避短，寻求一条能够更加高效、准确、可靠地解决计算机视觉及其他领域相关问题的新道路。

插图索引

图 2.1 ClusterSLAM 聚类效果	9
图 2.2 聚类-优化进程示意图	13
图 2.3 三种聚类算法对于实际数据的聚类效果	14
图 2.4 ClusterSLAM 中对各种算法在合成数据集上运行效果的比较	14
图 2.5 ClusterSLAM 中室内动态场景的分析效果	15
图 2.6 ClusterSLAM 中室外动态场景的分析效果	16
图 3.1 $p = 0.50, \mu = 0$ 时的重构效果	24
图 3.2 $p = 0.23, \mu = 0$ 时的重构效果	24
图 3.3 $p = 0.16, \mu = 0$ 时的重构效果	24
图 3.4 $p = 0.16, \mu = 1$ 时的重构效果	25
图 3.5 $p = 0.16, \mu = 0.2$ 时的重构效果	25
图 3.6 $p = 0.16, \mu = 0.05$ 时的重构效果	25
图 3.7 $p = 0.23, \mu = 1$ 时的重构效果	25
图 3.8 $p = 0.23, \mu = 0.2$ 时的重构效果	25
图 3.9 $p = 0.23, \mu = 0.05$ 时的重构效果	25
图 3.10 $n = 50, p = 0.15, \sigma = 0.005$ 时的弱 MDS 效果对比	28
图 3.11 $n = 70, p = 0.10, \sigma = 0.005$ 时的弱 MDS 效果对比	28
图 3.12 三维 cMDS 效果 ($n = 50, p = 1.00$)	29
图 3.13 三维稀疏 MDS 效果 ($n = 50, p = 0.30$)	30
图 3.14 三维弱 MDS 效果 ($n = 50, p = 0.22, d = 6$)	30
图 4.1 网格构造示意图	38

图 4.2 定理 4.2 证明图例	40
图 4.3 定理 4.3 证明图例	42
图 4.4 多种图结构的性能对比.....	43

公式索引

公式 1-1	2
公式 1-2	3
公式 1-3	4
公式 2-1	11
公式 2-2	11
公式 3-1	17
公式 3-2	18
公式 3-3	18
公式 3-4	19
公式 3-5	19
公式 3-6	19
公式 3-7	19
公式 3-8	20
公式 3-9	20
公式 3-10	20
公式 3-11	20
公式 3-12	20
公式 3-13	20
公式 3-14	21
公式 3-15	21
公式 3-16	21

公式 3-17	21
公式 3-18	22
公式 3-19	22
公式 3-20	22
公式 3-21	23
公式 3-22	23
公式 3-23	23
公式 4-1	34
公式 4-2	34
公式 4-3	35
公式 4-4	36
公式 4-5	36
公式 4-6	36
公式 4-7	36
公式 4-8	37
公式 4-9	37
公式 4-10	37
公式 4-11	37
公式 4-12	38
公式 4-13	38
公式 4-14	38
公式 4-15	38
公式 4-16	39
公式 4-17	39

公式 4-18	40
公式 4-19	40
公式 4-20	40
公式 4-21	40
公式 4-22	41
公式 4-23	41
公式 4-24	41
公式 4-25	41
公式 4-26	41
公式 4-27	41
公式 4-28	42
公式 4-29	42
公式 4-30	42
公式 4-31	42

参考文献

- [1] Huang J, Yang S, Zhao Z, et al. ClusterSLAM: A SLAM backend for simultaneous rigid body clustering and motion estimation[Z]. 2019.
- [2] Yang L, Liu Z, Chua T S, et al. Topical word embeddings[C]//Twenty-ninth Aaai Conference on Artificial Intelligence. 2015.
- [3] Kshirsagar V P, Baviskar M R, Gaikwad M E. Face recognition using eigenfaces[C]//International Conference on Computer Research & Development. 2011.
- [4] Young G, Householder A S. Discussion of a set of points in terms of their mutual distances [J]. Psychometrika, 1938, 3(1): 19-22.
- [5] Pawliczek P, Dzwinel W, Yuen D A. Visual exploration of data by using multidimensional scaling on multicore cpu, gpu, and mpi cluster[J]. Concurrency & Computation Practice & Experience, 2014, 26(3): 662-682.
- [6] Shang Y, Rum W. Improved mds-based localization[C]//Joint Conference of the IEEE Computer & Communications Societies. 2004.
- [7] Shon M, Jo M, Choo H. An interactive cluster-based mds localization scheme for multimedia information in wireless sensor networks[J]. Computer Communications, 2012, 35(15): 1921-1929.
- [8] Qiao D. Solutions for wireless sensor network localization[Z]. 2012.
- [9] Trosset M. Applications of multidimensional scaling to molecular conformation: volume 29 [Z]. 1997.
- [10] Cayton L, Dasgupta S. Robust euclidean embedding[C]//International Conference on Machine Learning. 2006.
- [11] Zhou S, Xiu N, Qi H D. Robust euclidean embedding via edm optimization[Z]. 2018.
- [12] Gross D. Recovering low-rank matrices from few coefficients in any basis[J]. Information Theory, IEEE Transactions on, 2011, 57: 1548 - 1566.
- [13] Tasissa A, Lai R. Exact reconstruction of euclidean distance geometry problem using low-rank matrix completion[J]. IEEE Transactions on Information Theory, 2018, PP.
- [14] Zhang H, Lin Z, Zhang C. Completing low-rank matrices with corrupted samples from few coefficients in general basis[J]. IEEE Transactions on Information Theory, 2015, 62.
- [15] Alfakih A, K. Khandani A, Wolkowicz H. Solving euclidean distance matrix completion problems via semidefinite programming[J]. Computational Optimization and Applications, 1999, 12: 13-30.
- [16] Erds P, Rényi A. On the evolution of random graphs[J]. Bulletin de l' Institut International de Statistique, 1961, 38.

- [17] Gortler S J, Healy A D, Thurston D P. Characterizing generic global rigidity[J]. American Journal of Mathematics, 2010, 132(4): 897-939.
- [18] Laman G. On graphs and of plane skeletal structures[J]. Journal of Engineering Mathematics, 1970, 4: 331-340.
- [19] Jackson B, Servatius B, Servatius H. The 2-dimensional rigidity of certain families of graphs [J]. Journal of Graph Theory, 2007, 54: 154-166.
- [20] Zinsser T, Schmidt J, Niemann H. A refined icp algorithm for robust 3-d correspondence estimation[C]//International Conference on Image Processing. 2003.
- [21] Karp R M. Reducibility among combinatorial problems[Z]. 1972.
- [22] Defays D. An efficient algorithm for a complete link method[J]. Computer Journal, 1977, 20(4): 364-366.
- [23] Janson S. On concentration of probability[Z]. 1999.
- [24] Fan J, Wang W, Zhong Y. An ℓ_∞ eigenvector perturbation bound and its application to robust covariance estimation[J]. Journal of Machine Learning Research, 2016, 18.
- [25] Connelly R. Generic global rigidity[M]. 2005: 549-563
- [26] Prasad Kasiviswanathan S, Moore C, Theran L. The rigidity transition in random graphs[J]. Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms, 2010.
- [27] Jordán T, Mihálykó A. Minimum cost globally rigid subgraphs[J]. Egerváry Research Group on Combinatorial Optimization, 2019.
- [28] Li Jiawei. Private communication[Z]. 2017.
- [29] Heck D, Deussen O. Farthest-point optimized point sets with maximized minimum distance [C]//Acm Siggraph Symposium on High Performance Graphics. 2012.
- [30] Okabe A, Boots B, Sugihara K. Spatial tessellations: Concepts and applications of voronoi diagrams[M]. 2000
- [31] Mao Jiayuan. Private communication[Z]. 2019.

致 谢

衷心感谢导师胡事民教授在大学四年里的精心指导。他的言传身教不仅帮助我提升了学习、科研能力，还让我懂得了许多做人的道理，使我终生受益。感激组里杨晟、黄家晖学长的指点，让我对科研这项工作有了更深的理解。

感谢清华大学交叉信息院、计算机系、数学系丰富的选修课程，使我得以在本科四年打下坚实而广泛的知识基础，同时能在与同学们的交流切磋中获得宝贵的科研灵感，让我能更好地克服学术之路上的艰难险阻。

在美国德克萨斯大学奥斯汀分校(UT Austin)的科研期间，感谢 Qixing Huang 助理教授和 Chandrajit Bajaj 教授的指导与鞭策，让我培养了身为一个科研工作者所必需的坚毅与专注，也让我在异国他乡迈出了科研的第一步。

在毕业论文开题期间，感谢 ETH Zurich 的 Zhiwu Huang 博士后为我推荐了一些文献，让我对相关的研究方向有了更多的认识与理解。

在自大三上以来的科研生活中，还感谢邓丛悦同学，让我了解了计算几何的一些领域，并在我迷茫之时支持我坚持理想，怀着探索广阔天地、寻求学术之美的初心不懈前行。愿与你共勉。

声 明

本人郑重声明：所呈交的学位论文，是本人在导师指导下，独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本学位论文的研究成果不包含任何他人享有著作权的内容。对本论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明。

签 名：赵梓硕 日 期：2019.6.4.

附录 A 外文资料的调研阅读报告

Euclidean embedding

The concept of recovering relative positions (in Euclidean space) from their pairwise distances, known as multi-dimensional scaling or Euclidean embedding, dates back to 1938[19]. For a Euclidean distance matrix D of n points, we compute the spectral decomposition $U\Lambda U^T = B = -\frac{1}{2}HDH$ in which $H = I - \frac{1}{n}\mathbf{1}\mathbf{1}^T$ [19], and it has been proven that for arbitrary symmetric D , it can be an Euclidean distance matrix if and only if B is positive semidefinite and its rank is the minimum dimension of the space in which the points can be embedded with consistent pairwise distances described in D [6]. Therefore, we can do approximate Euclidean embedding by principal component analysis as the Classical MDS does[2].

We can be interested by the required number of dimensions to approximate the pairwise distance, and as the Johnson–Lindenstrauss lemma shows, as long as the n points are originally from a higher dimensional space, we can map them into a $\Omega(\epsilon^{-2} \log n)$ -dimensional Euclidean space by *random projection*[7], regardless of the dimension of the original space. This result is good and bad: on the one hand, it shows that a relatively low-dimensional space is sufficient to contain points from an arbitrary high dimensional spaces preserving the pairwise distance information; on the other hand, it can imply the difficulty to robustly recover the original high-dimensional structure from pairwise distances alone, as even “squashing” the structure into a lower-dimensional space does not necessarily result in a significant distortion on pairwise distances.

Although our work is expected to recover the **original** structure from pairwise distances, most existing papers only desire a **consistent** embedding that preserves the pairwise distance information from the data, no matter if it resembles the high-dimensional structure—it does not even care whether the high-dimensional structure actually exists or not. In natural language processing, word embedding[18] is a class of multiple methods that embed words and topics into a latent vector space based on their affinity: it does not need an explicit D to approximate, but with the information how two objects are close to each other. The paradigm of Euclidean embedding also has application in data science, especially about social network.[12][9]

The methods above introduce a variety of ways to do Euclidean embedding, against small noises. However, the robust Euclidean embedding[2] is needed when there are a few outliers. In the common field of PCA, the robust PCA[17] provides a series of ways to recover a low-dimensional matrix against outliers, which is widely applied in latent semantic analysis in NLP[11]. However, a corrupted entry in D will diffuse to multiple entries in B , so robust PCA does not directly imply robust Euclidean embedding. It has also been proven that the 1-dimensional robust Euclidean embedding with l_1 -norm is NP-hard[2], while the case for other low dimensions is still left open. [2] has used the trace heuristics and outperformed the common MDS for image distance embedding, but its performance was still unable to compete with the nearest-neighbor algorithm. A more recent work[20] looks into it as an optimization problem, and [13] deals with it via the max-correntropy criterion.

Geometry-aware distance measure

When doing optimization, which is also the essential procedure of machine learning, we must have a target function to optimize, which can be the “distance” between two objects. For two singletons, many norms or geodesics can be a well-defined one. However, when going to point sets or distributions, things can be more complicated.

In the scope of information theory, the KL divergence[10] describes the distance of two distributions as the number of extra bits needed in optimal coding for mistaking one distribution for another, which

is asymmetric. The total variation distance[1], in the other perspective, is defined by the minimum total “mass” that is needed to be transported in one distribution to coincide with the other.

In the two formulations above, the latent space for distribution has discrete topology: any pair of different points have identical weight. If we define the cost for unit mass corresponding to the starting and ending point for the transportation, it is the theory of optimal transport[16]. Particularly, the Wasserstein distance can be used for GAN to make a geometry-aware loss function for improved image generation and discrimination.

Another application for Wasserstein distance is topological data analysis. The persistence diagram is a multiset of points on an extended plane that depicts the topological features of an implicit function, and it is proven to have the stability w.r.t. the implicit function in the bottleneck distance[3], i.e. ∞ -Wasserstein distance. A recent work proposed a way to do gradient descent by synchronize two persistence diagrams, i.e. match corresponding points into a vector space[14] for distance computation, and [5] has given a more mathematical view about the geometry and topology in the space of persistence diagrams.

Non-metric distance measurement

It is not the main topic in this research, but in certain scenarios it does occur. Non-metric property generally exists in non-transitive settings, as earlier studied in ClusterSLAM for rigid body detection.

When doing clustering with a non-metric distance, the complete-linkage clustering is preferred as it requires each pair of points satisfy the clustering criterion[15]. However, as the choice of merging element can be arbitrary, there can be multiple satisfying results. Our method to heuristically merge the smallest distance works well in ClusterSLAM with a time complexity of $O(n^2 \log n)$, while the lower bound for possible complexity is $O(n^2)$. The CLINK algorithm [4] has $O(n^2)$ complexity but typically generates a bad result which is also dependent on the order of input. However, it has proposed a insight that the complete-linkage clustering is equivalent to finding a locally minimum ultrametric relation superior to the distance matrix.

By equivalence to clique-cover problem, finding the minimum possible number of clusters for complete-linkage clustering is generally NP-hard.[8] However, for other criterions for optimal clustering or under some conditions, the problem remains open.

References

- [1] Martin Arjovsky, Soumith Chintala, and LSon Bottou. Wasserstein gan. 2017.
- [2] Lawrence Cayton and Sanjoy Dasgupta. Robust euclidean embedding. In *International Conference on Machine Learning*, 2006.
- [3] David Cohensteiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. In *Acm Symposium on Computational Geometry*, 2005.
- [4] D. Defays. An efficient algorithm for a complete link method. *Computer Journal*, 20(4):364–366, 1977.
- [5] Vincent Divol and ThSo Lacombe. Understanding the topology and the geometry of the persistence diagram space via optimal partial transport. 01 2019.
- [6] F. R. Hodson, P. H. A. Sneath, and J. E. Doran. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*, 53(3/4):325–338, 1966.
- [7] W. B. Johnson and J. Lindenstrauss. Extensions of lipschitz mappings into a hilbert space. *Contemporary Mathematics*, 26, 1984.
- [8] Richard M. Karp. Reducibility among combinatorial problems. 1972.
- [9] Mohammad Khoshneshin and W. Nick Street. Collaborative filtering via euclidean embedding. In *Acm Conference on Recommender Systems*, 2010.

- [10] S. Kullback and R. A. Leibler. On information and sufficiency. *Annals of Mathematical Statistics*, 22(1):79–86, 1951.
- [11] Thomas K Landauer. *Latent Semantic Analysis*. 2010.
- [12] Wentao Li, Min Gao, Wenge Rong, Junhao Wen, and Tong Dou. Social recommendation using euclidean embedding. In *International Joint Conference on Neural Networks*, 2017.
- [13] Fotios D. Mandanas and Constantine L. Kotropoulos. Robust multidimensional scaling using a maximum correntropy criterion. *IEEE Transactions on Signal Processing*, 65(4):919–932, 2016.
- [14] Adrien Poulenard, Primoz Skraba, and Maks Ovsjanikov. Topological function optimization for continuous shape matching. In *Computer Graphics Forum*, pages 13–25, 2018.
- [15] R. R. Sokal and C.D. Michener. A statistical method of evaluating systematic relationships. *University of Kansas Sci Bull*, 28:1409–1438, 1958.
- [16] Cedric Villani. Topics in optimal transportation. *Ams Graduate Studies in Mathematics*, page 370, 2003.
- [17] Huan Xu, Constantine Caramanis, and Sujay Sanghavi. Robust pca via outlier pursuit. *IEEE Transactions on Information Theory*, 58(5):3047–3064, 2012.
- [18] Liu Yang, Zhiyuan Liu, Tat Seng Chua, and Maosong Sun. Topical word embeddings. In *Twenty-ninth Aaai Conference on Artificial Intelligence*, 2015.
- [19] Gale Young and A. S. Householder. Discussion of a set of points in terms of their mutual distances. *Psychometrika*, 3(1):19–22, 1938.
- [20] Shenglong Zhou, Naihua Xiu, and Hou-Duo Qi. Robust euclidean embedding via edm optimization. 03 2018.

综合论文训练记录表

学生姓名	赵梓硕	学号	2015011255	班级	计科 50
论文题目	基于几何距离的场景分析与重构				
主要内容以及进度安排	<p>在前阶段工作 Cluster SLAM 中，基于空间距离不确定性的刚体识别与聚类算法，让动态场景的分析准确率在室内条件下获得了较大提升。本课题将进一步从理论上探究基于几何约束对三维场景进行分析、重构的可能性与潜力，以探索机器视觉、机器人学的一种新思路。</p> <p>此项目将首先调研传统的 Robust Euclidean Embedding 问题，并分析其在存在数据错误与缺失时的表现，并结合 NLP 领域 Word Embedding 等算法进行思考；此外，还将研究分布等拓扑空间中的距离量度，了解 Optimal Transport 理论，并将其运用于分析、重构之中。</p>				
	<p style="text-align: right;">指导教师签字: <u>胡事迹</u></p> <p style="text-align: right;">考核组组长签字: <u>张维峰</u></p> <p style="text-align: right;">2019 年 3 月 27 日</p>				
中期考核意见	<p>进度良好！</p> <p style="text-align: right;">考核组组长签字: <u>张维峰</u></p> <p style="text-align: right;">2019 年 4 月 27 日</p>				

指导教师评语	<p>场景分析与重构是计算机图形学的热点研究问题，怎样突破传统方法的关键在于几何距离的精度聚类、均差重构等，并用于几何网格稳定性的优化设计。论文写作规范，成果有创新，同意参加论文答辩。</p> <p>指导教师签字：胡翠红</p> <p>2019年6月12日</p>
评阅教师评语	<p>本文有以下贡献：设计了无先验知识的运动场景的刚体聚类；MPS算法的几何稳定性定义；最小代价连接图算法。本文创新性强，有理论证明的依据，是一篇合格的本科毕业论文</p> <p>评阅教师签字：吴文波</p> <p>2019年6月15日</p>
答辩小组评语	<p>答辩展示，结构清晰，逻辑合理，展示了不同层面的三项工作成果，并有理论证明支持。建议加强研究背景介绍和理论成果在实际应用中的实现举例。</p> <p>答辩小组组长签字：高鸣宇</p> <p>2019年6月16日</p>

总成绩：A
教学负责人签字：姚期智

2019年6月20日