论 文

结构化场景重建中一种融合直线信息的平面拟合方法

傅康平^{①*}, 王伟^①, 胡立华^②, 胡占义^①^① 中国科学院自动化研究所模式识别国家重点实验室, 北京 100190^② 太原科技大学计算机学院, 太原 030024

* 通信作者. E-mail: kpfu@nlpr.ia.ac.cn

收稿日期: 2014-08-28; 接受日期: 2014-10-23; 网络出版日期: 2015-02-12

国家自然科学基金 (批准号: 61333015, 61273280) 资助项目

摘要 结构化场景是三维重建中非常重要的一类场景. 本文分析了结构化场景的特点, 并针对该类场景下多视立体点云的平面拟合问题, 提出了一种融合直线信息的改进 PEaRL 算法. 本文方法首先利用三维直线模型和三维点云模型生成候选平面. 直线信息的引入使候选平面的数量得以降低, 同时其可靠性得以提升. 然后, 利用直线对点云的邻接关系进行约束, 提高了平面相交处的平面拟合准确性. 在能量优化过程中, 本文方法对 PEaRL 算法所使用能量函数的平滑项和标记惩罚项进行了改进, 以更好地适应多视立体点云的不规则性. 在模拟数据和真实数据集上的对比实验表明, 本文方法有较高的运行效率, 且可获得比 PEaRL 等多模型拟合算法更可靠的平面拟合结果.

关键词 结构化场景 平面拟合 模型拟合 能量优化 MVS PEaRL 三维直线

1 引言

在基于图像的室外场景三维重建中, 由于各种干扰因素 (如弱纹理区域、透视畸变、光照变化等) 的影响, 传统利用像素匹配或扩散的多视立体 (multiple view stereo, MVS) 算法 (如 PMVS^[1], DAISY^[2] 等) 通常只能获得不完整的重建结果. 特别在结构化的城市建筑场景中, 由于更多弱纹理、倾斜表面等区域的存在, 此问题尤为严重.

为了克服上述问题, 利用特定场景模型 (如 Manhattan-world 场景模型^[3]) 或分段平滑假设的重建算法^[4~7] 近年来得到深入研究和广泛应用. 相对而言, 此类算法不但可获得完整的重建结果, 更有效地表达场景核心信息, 而且由其获得的平面结构模型往往也便于存储和传输. 在实际应用中, 此类算法为了推断完整的场景结构, 增强对复杂结构场景的适应性, 通常需要确定场景结构中尽可能多的主平面或方向. 因此, 获取场景的平面信息具有非常重要的研究意义.

如图 1 所示, 在 PMVS 算法所获得的三维点云模型中, 弱纹理区域 (如墙面等)、重复纹理区域及非朗伯 (Lambertian) 平面 (如玻璃等) 无法得以重建; 而在已重建的主要结构中, 墙体及窗户边缘的直线结构往往得到了较好的保留, 这也为这些区域空间信息的进一步推断提供了可靠的约束或先验. 因此, 在结构化场景的三维重建中, 如何结合直线信息, 在 MVS 方法所获得的点云模型基础上高效地获取场景平面信息是本文所要讨论的内容.

引用格式: 傅康平, 王伟, 胡立华, 等. 结构化场景重建中一种融合直线信息的平面拟合方法. 中国科学: 信息科学, 2015, 45: 889–902, doi: 10.1360/N112014-00150

图 1 PMVS 算法^[1]获得的准稠密点云示例

Figure 1 Examples of quasi-dense point models obtained by PMVS^[1]. (a) Point model of the Tsinghua math building data set; (b) details of the point model in (a); (c) point model of the Valbonne data set

现有针对三维点云进行平面拟合的诸多方法中, 利用空间点聚类或区域生长的方法^[8]通常对点云的质量要求较高(如激光扫描技术获取的规则稠密点云)。然而, MVS 方法得到的点云是非规则的准稠密点云。相对于激光扫描数据, 其中不但存在大量的缺失信息, 而且也存在更多的误差及外点。因此, 在此类点云模型上直接通过空间点聚类或区域生长得到的平面信息一般并不可靠。基于多模型拟合(如 Multi-RANSAC^[9], J-linkage^[10]等)的方法虽然在某些条件下可获得较好的拟合结果, 但由于其对于拟合模型(如平面)的数量、相关阈值等因素较为敏感, 所以该类方法在复杂场景中的适应性通常较差。

最近, Isack 和 Boykov^[11]提出了基于能量优化的多模型拟合 PEaRL 方法。该方法将传统的多模型拟合问题转换为全局能量优化中的多标记问题, 利用 Graph-Cut 的方法进行求解^[12~14]。通过在能量函数中融入模型之间的平滑性约束, 该方法克服了传统多模型拟合方法孤立地对单个模型进行拟合而导致的 inconsistency 问题。然而, 此方法由于对初始候选模型的数量与质量、空间点邻域关系的可靠性等具有较高的依赖, 实际应用中不但计算复杂度高, 而且难以获得较好的拟合结果。

Pham 等^[15]新近提出了基于 PEaRL 框架的 RCMSA 方法。该方法利用“随机聚类模型”来提高初始模型的生成质量, 并将模型生成与数据拟合两个关键步骤融入了统一的模拟退火框架以提高模型拟合的准确率。但该方法无法利用更多的先验信息, 因此在处理复杂场景的三维点云平面拟合问题时, 仍然无法获得令人满意的结果。

为了克服这些问题, 本文以 PEaRL 能量优化框架为基础, 将结构化场景中的直线¹⁾信息融入到该框架的各个关键步骤中。引入的直线信息不仅可以通过缩减候选平面的数量及增强候选平面的可靠性来提高平面优化环节的可靠性及速度, 而且可以对空间点间不可靠的邻接关系进行检测与剔除, 从而提升整体算法的稳定性。此外, 本文以自适应的方式对能量函数中平面数量的惩罚进行了改进, 使得算法具有更高的适应性。在模拟数据及真实数据集上的实验表明, 本文算法可以更快地获得更准确的平面拟合结果。

2 PEaRL 算法

MVS 点云的平面拟合问题可以视为数据的多模型拟合问题, 相关的方法(如 J-linkage^[16], PEaRL, RCMSA 等)通常首先根据数据的分布产生模型假设, 然后利用不同的方法对模型进行调整以获得可靠的多模型拟合结果。由于本文算法为 PEaRL 算法的改进, 因而此处对其主要流程先作一简要介绍。

1) 如无特别说明, 本文中提到的直线均指线段。

PEaRL 算法基于能量优化的方法, 其核心为式 (1) 所示的能量函数. 其中, p, q 代表原始数据; L_p 代表数据 p 的模型标记; N 为邻接关系; $\delta(\cdot)$ 为指示函数 (当条件成立时值为 1, 否则值为 0); $|L_L|$ 为当前候选标记集中非空标记的个数. 该算法主要流程分为三步: 候选模型生成 (proposal), 能量优化 (expand) 和模型调整 (re-estimate labels).

$$E(L) = \sum_p \|p - L_p\| + \lambda \cdot \sum_{(p,q) \in N} w_{pq} \cdot \delta(L_p \neq L_q) + \beta \cdot |L_L|. \quad (1)$$

以三维点的平面拟合为例. 在 PEaRL 算法中, 候选模型 (平面) 通过随机采样的方式生成, 而数据之间的邻接关系, 则通过对三维点进行 Delaunay 三角化获得. 在获得候选平面集及数据点的邻接关系后, PEaRL 算法利用全局能量优化的方法对数据进行平面划分, 并根据优化结果重新构造候选平面. 为了在优化过程中剔除多余的平面, 能量函数中引入了针对非空候选平面数量的惩罚项 $\beta \cdot |L_L|$.

以上过程通过有限次的迭代, 候选平面数量不断缩减, 每个候选平面的可靠性不断增强, 最终获得全局最优的平面拟合结果.

PEaRL 算法使用的是一种迭代的计算方式, 而每次迭代都会利用既有的数据内点对模型进行调整, 因此, 该算法在实验中表现出了良好的准确度. 但实际应用中, 为了保证算法的可靠性, 候选平面应尽可能涵盖所有真实平面. 因而, 初始候选平面的数量往往需要远大于真实平面数量, 从而算法的计算复杂度一般很高. 此外, 随机生成的候选平面的可靠性通常也难以得到保证. 如文献 [15] 所分析, 完全随机采样产生的候选模型, 非常容易被可能存在的外点及误差所影响, 因而, 算法很容易收敛到局部最优解. 而在确定数据之间的邻接关系时, Delaunay 三角化通常会使平面交汇处点的划分不够准确, 从而影响平面优化的精度.

针对 PEaRL 算法存在的不足以及结构化场景 MVS 点云平面拟合问题的特点, 本文提出一种将三维直线信息融合入 PEaRL 框架的改进方法.

3 针对结构化场景 MVS 点云平面拟合的一种改进 PEaRL 算法

结构化场景中, 直线元素普遍存在. 而利用现有的直线重建方法, 也可方便地获得三维直线模型. 本文将三维直线信息融入 PEaRL 算法框架中, 以期提高候选平面及数据点邻接关系的可靠性, 进而提升平面拟合结果的准确度. 本节首先分析结构化场景的特点并提出结构化场景假设, 之后介绍改进的 PEaRL 算法的各个关键步骤.

3.1 结构化场景假设

结构化场景 (如建筑物、街道等) 中存在大量的平面元素, 而平面中又往往存在大量的直线元素. 此处, 我们首先分析结构化场景的特点, 并提出如下结构化场景假设:

假设 1 对于场景中的某一条直线, 其位置有两种可能: 位于某个平面内部或位于某个平面边缘 (即两个平面相交处).

对于该假设中的第一种情况, 直线的位置如图 2(a) 所示; 而对于第二种情况, 又可细分为图 2(b) 和图 2(c) 两种. 图 2(b) 中, 相交的两个平面均被重建出来; 图 2(c) 所示的情况中, 由于可能存在的遮挡等因素的影响, 相交的两个平面中, 另一平面不可见.

在基于图像的三维重建结果中, 结构化场景中的绝大多数直线均满足此假设.

假设 2 任意平面中至少存在一条直线.

图 2 直线与平面的位置关系

Figure 2 Relationship between lines and planes. (a) An example of a line lying in a plane; (b) an example of a line lying at the intersection of two planes; (c) an example of a line lying at the intersection of two planes while only one of them is visible

该假设要求平面中至少一条边为直线或平面内部存在直线. 该假设条件相对严苛, 但建筑物、城市街道等结构化场景, 大都满足该假设.

3.2 三维直线模型获取

三维直线模型的获取有多种方法, 其中, 文献 [17,18] 的方法可根据已标定图像直接重建三维直线模型; 而方法 [19] 则根据既有 MVS 结果快速获得可靠的直线匹配, 进而获得三维直线模型. 相对而言, 由于方法 [19] 直接基于 MVS 点云模型进行直线匹配的求取, 因此, 不但准确度较高, 而且所获直线模型与点云模型的一致性较好. 因此, 本文真实数据实验中所使用的三维直线模型均采用方法 [19] 获得. 在下文中, 我们将三维点集记为 P_t , 而将三维直线集记为 L_n .

3.3 候选平面生成

基于结构化场景假设, 我们提出一种候选平面的生成方法. 该方法结合空间直线及其邻域中的空间点, 获得相对可靠的候选平面.

给定空间直线 L_{AB} (端点为 A, B , 长度为 $|L_{AB}|$), 我们将以 L_{AB} 为中心的一个圆柱形区域视为 L_{AB} 的邻域. 本文方法通过考察 L_{AB} 邻域内点的分布来生成候选平面. 为便于表示, 我们首先对坐标系进行旋转和平移, 建立以 A 点为原点, 且 B 点位于 z 轴正半轴的柱面坐标系, 如图 3 所示. 此时, 空间中任一点 X_θ 的坐标可表示为式 (2) 所示形式.

$$\begin{cases} x = \rho \cdot \cos\theta, \\ y = \rho \cdot \sin\theta, \quad \rho \in [0, +\infty), \theta \in [0, 2\pi), \\ z = z. \end{cases} \quad (2)$$

直线 L_{AB} 可定义一个半平面族. 我们利用角度扫描的方法, 衡量该半平面族内每个半平面 P_α 与其邻域中空间点的一致性. 空间点 X_θ 与半平面 P_α 的距离可以直接用欧氏距离求得. 考虑到半平面自身的方向性, 该距离可表示为

$$d(X_\theta, P_\alpha) = \begin{cases} \rho \cdot |\sin(\theta - \alpha)|, & |\theta - \alpha| < \pi, \\ \rho, & \text{其他.} \end{cases} \quad (3)$$

图 3 基于直线 L_{AB} 建立的柱面坐标系Figure 3 Cylindrical coordinate system based on line L_{AB}

考虑到场景的复杂性, L_{AB} 的邻域内很可能存在属于其他平面的空间点. 而距离 L_{AB} 越近的点, 其分布情况越能可靠地反映 L_{AB} 所在平面的信息. 基于此, 我们为邻域内的每个空间点设置权值, 该权值随点与直线 L_{AB} 距离的增大而衰减, 如式 (4).

$$D(X_\theta, P_\alpha) = e^{-\rho^2/\mu^2} \cdot d(X_\theta, P_\alpha). \quad (4)$$

考虑到直线 L_{AB} 为有限长度, 因此只考虑 z 坐标位于 $[0, |L_{AB}|]$ 范围内的点. 综上所述, 点 X_θ 与半平面 P_α 之间的代价函数可用式 (5) 与式 (6) 表示.

$$C(X_\theta, P_\alpha) = e^{-\rho^2/\mu^2} \cdot \rho \cdot |\sin(\theta - \alpha)| \cdot \Delta(X_\theta, P_\alpha), \quad (5)$$

$$\Delta(X_\theta, P_\alpha) = \delta(|\theta - \alpha| < \pi) \cdot \delta(0 \leq X_z \leq |L_{AB}|). \quad (6)$$

定义点集 $N(P_\alpha) = \{X | \Delta(X, P_\alpha) > 0\}$, 半平面 P_α 的整体代价函数可记为式 (7) 所示形式.

$$C(P_\alpha) = \begin{cases} \frac{1}{|N(P_\alpha)|} \cdot \sum_{X \in N(P_\alpha)} C(X, P_\alpha), & |N(P_\alpha)| > 0, \\ +\infty, & \text{其他.} \end{cases} \quad (7)$$

图 4(a) 给出了图 2(b) 中示例直线的邻域点. 对该直线进行平面扫描所获得的半平面族的代价函数曲线 $C(P_\alpha)$ 如图 4(b) 所示. 该曲线中, 最小的两个极小值对应的半平面如图 4(c) 所示. 可以看到, 此代价函数最小的两个极小值与真实平面位置相对应.

根据前文中结构化场景的两个假设, 每条直线至多对应两个平面. 因此可从平面族的代价函数中选取两个最小的极小值, 对应生成两个候选平面. 而为了增强方法的鲁棒性, 我们在实际计算中, 求取代价函数的所有极小值, 并由每个极小值计算获得一个候选平面.

图 4 直线的半平面扫描示例

Figure 4 An example of half-plane sweeping. (a) The line shown in Figure 2(b) and its neighborhood; (b) curve of the cost function in plane sweeping; (c) planes corresponding to the two smallest cost values

3.4 邻接关系确定

对于几何模型拟合问题, PEaRL 算法采用了 Delaunay 三角化来确定数据点之间的相邻关系. 而在三维点的平面拟合中, 该方法通常会产生较多的错误邻接关系: 如位于不同平面中且深度变化较大的两个数据点, 虽然实际并不相邻, 但是很可能被错误地确定为具有相邻关系. 因此, 本文根据当前直线模型所提供的约束, 检测并剔除错误的邻接关系, 从而提高了平面优化环节的可靠性, 尤其是提高了平面边界处的拟合精度.

本文方法生成的候选平面相对可靠, 因此我们直接利用候选平面信息对 Delaunay 三角化获得的邻接关系进行处理. 对于 Delaunay 三角化获得的边集 $\text{Edge} = \{e_{(X,Y)} | X, Y \in \text{Pt}\}$, 若点 X 与点 Y 位于同一直线的邻域内且分属于不同候选平面, 则去除该边. 最终获得的边集记为 $\overline{\text{Edge}}$.

通过直线对空间点的邻接关系进行约束, 边界处的拟合结果的精度可以得到有效提升. 同时, 场景中深度变化较小处的平面结构也可被有效地拟合出来.

3.5 能量函数设计

本文方法所使用的能量函数仍沿用 PEaRL 算法的基本形式, 但对其中的平滑项及平面数量惩罚项进行了修改.

在文献 [11] 中, 将 PEaRL 算法应用于数据点拟合时, 使用了固定的边权 ($w_{pq} = 1$). 正如前述, 由于 MVS 点云的不完整性, 相应区域的边界数据点在 Delaunay 三角化过程中将被连接起来以构成邻域系统, 如图 1(c) 中所示区域的数据点连接; 然而, 这些实际并不连续的数据点通常会对能量函数中平滑项约束的可靠性产生较大的影响, 最终导致平面拟合结果中存在较大的错误. 基于此, 本文将 W_{XY} 定义为式 (8) 所示形式.

$$W_{XY} = e^{-\|X-Y\|}. \quad (8)$$

而对于平面数量惩罚项, 原 PEaRL 算法以非空标记集数目为依据, 对所有标记进行统一惩罚. 此举旨在保证整体拟合准确度的同时, 尽量减少模型数目. 但由于该惩罚项并不区分不同模型中内点的数目, 因此模型数目约束的作用相对不明显. 基于此, 我们采用文献 [20] 中的方法, 将非空模型的内点数引入到惩罚项中, 将惩罚项改为 $\sum_{L_i \in L} e^{-|L_i|}$ 的形式. 其中, L 为标记集, $|L_i|$ 为模型 L_i 的内点数. 此种形式更鼓励削减内点数较少的模型, 保留内点数较多的模型, 从而使最终的结果更加稳定.

图 5 真实数据集图片样张

Figure 5 Sample images of the real data sets. (a) VALBONNE; (b) WADHAM; (c) MATH-Tsinghua math building; (d) SCHOOL-Tsinghua school

综上所述, 本文新的能量函数如式 (9) 所示.

$$E(L) = \sum_X \|X - L_X\| + \lambda \cdot \sum_{(X,Y) \in \overline{\text{Edge}}} e^{-\|X-Y\|} \cdot \delta(L_X \neq L_Y) + \beta \cdot \sum_{L_i \in L} e^{-|L_i|}. \quad (9)$$

3.6 整体算法

本文方法的整体算法如算法 1 所示. 在获取到候选平面及邻接关系后, 本算法沿用 PEaRL 的迭代优化框架: 利用 α -expansion 方法对定义的能量函数进行优化, 优化完成后对模型进行重估并重新优化, 两个步骤迭代至能量值不再下降.

算法 1 针对结构化场景 MVS 点云平面拟合的一种改进 PEaRL 算法

输入

三维点集 P_t ,
三维直线集 L_n ;

输出

点集的平面标号 Label;

```

1: 候选平面集  $\text{Plane} \leftarrow \emptyset$ ;
2: 对点集  $P_t$  的元素进行 Delaunay 三角化, 得到边集  $\text{Edge}$ ;
3: for each  $L_{AB}$  in  $L_n$  do
4:   获取  $L_{AB}$  的邻域点集  $P_{tN}$ ;
5:   计算  $L_{AB}$  对应平面族的  $C(P_\alpha)$  曲线;
6:   获取  $C(P_\alpha)$  曲线的极小值, 并将其对应的平面加入  $\text{Plane}$  集合;
7:   删除  $\text{Edge}$  边集中两端点均属于  $P_{tN}$ , 且两端点位于不同候选平面的边;
8: end for
9:  $\overline{\text{Edge}} \leftarrow \text{Edge}$ ;
10:  $\text{Label} \leftarrow \text{PEaRL-mod}(\text{Plane}, \overline{\text{Edge}})$ ; // PEaRL-mod 为使用式 (9) 所示能量函数的 PEaRL 优化框架
11: return Label;
```

4 实验结果及分析

为了验证本文算法的可行性, 我们分别采用模拟数据与真实数据对算法进行了测试. 其中, 真实数据集的图片样张如图 5 所示, 其中前两组为标准数据, 后两组为自有数据. 同时, 为了验证本文算法的性能, 我们在模拟数据实验中将本文算法与 PEaRL²⁾, RCMSA³⁾ 及 J-linkage⁴⁾ 进行了定量比较, 并在

2) <http://vision.csd.uwo.ca/code/>.

3) <http://cs.adelaide.edu.au/trung/home.php>.

4) <http://www.diegm.uniud.it/fusiello/demo/jlk/>.

真实数据集上对比了 4 种方法的平面拟合结果及时间开销. 用于对比的各方法均取自作者所公开的源码. 本节所有实验的硬件环境为配置有 Intel-i7 CPU, 16 G 内存的个人计算机, 软件环境为 Matlab 2012.

4.1 参数选择

在本文实验中, 所有数据集均采用相同的参数配置. 实验中发现, 对于能量函数中平滑项权重 λ , 其在区间 $[0.2, 0.8]$ 取值对结果影响并不大; 而对于惩罚模型复杂度的标记项权重 β , 将其设置为 $[10, 25]$ 之间任意值时均获得相同的结果. 因而, 实验中将两者分别设置为 0.5 与 15. 对于候选平面选取时的参数 μ , 本文则采用以下统计方法进行确定.

本方法引入的参数 μ 与半平面扫描时的直线邻域大小直接相关. 在实验中, 我们发现该参数与空间点之间的平均最小距离 $\bar{d}_{\min}$ 存在一定相关性. 即可通过 $\mu = n \cdot \bar{d}_{\min}$ 确定参数 μ 的取值. 直观上, 当 n 取值过小时, 邻域内的点无法准确反映直线周围平面的分布情况; 而当 n 取值过大时, 又会引入直线所在平面之外的其他平面结构, 进而影响平面信息的准确性. 针对此问题, 我们对不同 n 取值下所产生的候选平面质量进行评价. 为了更好地适应非规则点云模型, 本节所有实验均在真实数据集上进行.

给定 n 的取值, 我们在每个数据集中随机选取 10 条直线 L_{sample}^i , 并分别手工标记每条直线两侧的邻域点集. 对于当前直线 L_{sample}^i , 其两侧邻域平面由手工标记的邻域点集通过 RANSAC 的方法拟合获得, 记为 P_j^i , 而基于直线 L_{sample}^i , 通过前文方法进行平面扫描所获得的平面记为 $\tilde{P}_k^i$. 给定直线 L_{sample}^i , 其候选平面误差由式 (10) 计算获得.

$$E(L_{\text{sample}}^i) = \begin{cases} \frac{1}{|\tilde{P}^i|} \sum_k \min(\text{angle}(\tilde{P}_k^i, P_j^i)), & |\tilde{P}^i| > 0, \\ \pi, & \text{其他}, \end{cases} \quad (10)$$

其中 $\text{angle}(\cdot)$ 表示两个平面的夹角. 而给定 n 的候选平面平均误差则由 $E(n) = \frac{1}{10} \sum_i E(L_{\text{sample}}^i)$ 计算获得.

图 6 给出的是 n 在不同取值下 $([2, 100])$ 的候选平面平均误差曲线. 从曲线中可以观察到, 当 n 取值在 $[10, 30]$ 区间内时, 几组数据的误差均较低. 而 n 的取值越大, 平面扫描的计算量越大. 综合两个因素, 我们将 n 的取值定为 10, 即 μ 取值 $10 \cdot \bar{d}_{\min}$. 需要注意的是, 该实验现象也符合准稠密点云的特点. 如图 1 所示, 直线结构附近的纹理相对丰富, 其邻域内重建所得的空间点相对密集, 因此可通过空间点之间的平均最小距离来确定直线邻域, 进而确保产生较高质量的候选平面.

4.2 合成数据实验

为定量对比不同方法的平面拟合精度, 我们合成了一个简易的房屋模型. 该模型如图 7(a) 图所示, 共有 9 个面, 17 条直线. 除顶部 4 个面外, 我们为每个面随机生成具有高斯噪声的 100 个点. 顶部 4 个面上, 我们共生成 300 个具有高斯噪声的点. 生成点的同时, 对每条直线的端点位置也添加同样分布的高斯噪声. 我们共生成均值为 0, 标准差 $\sigma \in \{0, 0.005, 0.01, 0.015, 0.02\}$ 5 组数据. 其中 $\sigma = 0.02$ 的模型俯视图如图 7(b) 所示.

我们在这些数据上分别运行本文算法, PEaRL, RCMSA 和 J-linkage. 这 4 种方法的最终结果均为数据点的标号信息, 我们将每个数据点的真正所属平面与最终的标号信息进行比较, 并统计其中错误标记的点的个数. 4 种方法结果中的错误点数统计数据如图 7(c) 所示.

图 6 候选平面平均误差曲线 $E(n)$ Figure 6 Average error $E(n)$ of candidate planes under different thresholds

图 7 合成数据模型及合成数据实验对比

Figure 7 The synthesized data and error comparison of different methods. (a) The synthesized house model; (b) the synthesized model under noise level of 0.02; (c) error curves of different methods at different noise levels

从实验结果中我们可以发现, 在所有噪声水平下, 本文方法的错误点数均最少. RCMSA 方法采用了与 PEaRL 相似的框架, 但 RCMSA 方法通过改进候选模型的产生方式获取更为可靠的初始候选模型. 从实验结果来看, 当数据无噪声时, RCMSA 表现出比 PEaRL 更好的性能. 但随着噪声的引入, RCMSA 的错误点数反而高于 PEaRL, 因此其鲁棒性不如 PEaRL 算法.

对于 J-linkage 方法, 由于该方法依赖随机生成的候选模型, 因此其错误数与噪声水平并没有显著的对立关系. 但该方法的整体误差要高于 PEaRL 算法及本文算法.

为了模拟由点云获取直线的方法 (如方法 [19]) 所面临的直线模型的误差问题, 本实验所使用的直线模型添加了与点模型同样水平的噪声. 通过实验结果可以观察到, 由于本方法在模型调整时采用了与 PEaRL 算法相同的迭代优化方式, 因此对直线误差同样有较好的鲁棒性.

4.3 标准数据集实验

为了对比真实数据上的平面拟合效果, 我们选取了 VALBONNE 和 WADHAM 两组标准数据集⁵⁾,

5) <http://www.robots.ox.ac.uk/vgg/data/data-mview.html>.

图 8 VALBONNE 及 WADHAM 数据集的点云及直线模型

Figure 8 Point and line models of VALBONNE and WADHAM

表 1 标准数据集实验统计

Table 1 Statistics of experiments on the two public data sets

	VALBONNE (7865 points)		WADHAM (28799 points)	
	Plane number	Time (min)	Plane number	Time (min)
Our method	15	0.6	16	2.8
PEaRL	7	7.5	13	21.1
RCMSA	5	0.4	8	2.7
J-linkage	7	30.1	84	97.5

并分别运行了本文算法、PEaRL 算法、RCMSA 算法和 J-linkage 算法。

图 8 给出了利用 PMVS⁶⁾ 得到的 VALBONNE, WADHAM 两个场景的三维点云, 以及在此基础上, 利用文献 [19] 中方法得到的直线模型. 图 9 分别为 VALBONNE 场景、WADHAM 场景的平面拟合结果. 在每组结果中, 从左至右分别为本文算法、PEaRL 算法、RCMSA 算法和 J-linkage 算法; 其中, 不同颜色标注的点代表不同的平面标号. 表 1 给出了两组实验的平面数、运行时间等统计数据.

从实验结果中可以发现, 由于本文方法在平面拟合过程中添加了基于直线的约束信息, 因此可以更好地区分场景中距离较近的平面 (如 VALBONNE 数据中正门两侧立柱式结构以及 WADHAM 数据中屋顶的立柱式结构). 而对于数据中的噪声点, 本文方法也可以进行有效区分 (如 VALBONNE 数据右侧的噪声数据). 此外, 本文方法可以生成更加可靠的候选平面, 因此整个算法所需要处理的候选平面数目远少于 PEaRL 算法. 所以, 本文算法的效率也明显优于 PEaRL 算法.

方法 [19] 可基于点云模型快速获得直线模型, 在此组数据集的实验中, 直线建模所需的时间小于 1 分钟. 考虑此部分时间开销后, 本方法的整体时间开销仍然优于 PEaRL 和 J-linkage, 与 RCMSA 算法相当. 值得注意的是, 不可避免的直线误匹配使得直线模型中存在明显的错误直线 (如地面上的错误直线). 由错误直线所产生的错误候选平面可在后续优化过程中得以去除, 因此最终结果的准确性并没有受此影响.

4.4 自有数据集实验

如上文所述, 本方法可用于对已有 MVS 点云模型直接进行平面建模. 图 10 展示的是通过 PMVS 算法获得的“清华数学楼”和“清华学堂”两组数据集的点云模型, 以及以点云模型为基础获得的直线模型. 两场景均为建筑物场景, 符合结构化场景假设. 同时, 该场景并不满足 Manhattan-world 场景模型.

6) <http://www.di.ens.fr/pmvs/>.

图 9 标准数据集实验结果对比

Figure 9 Comparison on the standard data sets. On each row, the results are generated by our method, PEaRL, RCMSA and J-linkage respectively

图 10 “清华学堂”及“清华数学楼”数据集的点云及直线模型

Figure 10 Point and line models of MATH and SCHOOL

图 11 自有数据集实验结果对比

Figure 11 Comparison on our own data sets. On each row, the results are generated by our method, PEaRL, RCMSA and J-linkage respectively

图 11 给出了 4 种方法在“清华数学楼”数据集上的对比. 表 2 为该节实验的相关统计数据.

从两组结果中可以观察到, 本文方法可以将点云模型中的主要平面成功提取出来. 两个模型中均存在多个非正交平面, 且这些平面的面积相对较小. 对于传统的多模型拟合方法, 由于没有关于场景的额外信息, 因此很难将这些平面提取出来; 而本方法由于利用了直线信息, 因此可以获得更为准确的结果.

值得注意的是, 通过图 10 清华数学楼的模型我们可以观察到, 该场景下部结构中存在大量曲线元素, 而直线元素不够丰富. 但在图 11 所示的结果中, 该部分的平面信息仍可被准确提取出来. 如结构化场景假设二所述, 这些平面中只要存在至少一条直线 (边界处或内部直线纹理均可), 便可利用本文方法恢复其平面信息.

在时间开销方面, 由于该组数据规模较大, J-linkage 算法所耗时间明显多于其他方法. 方法 [19] 在

表 2 自有数据集实验统计

Table 2 Statistics of experiments on our own two data sets

	MATH (69836 points)		SCHOOL (54530 points)	
	Plane number	Time (min)	Plane number	Time (min)
Our method	16	11.8	16	7.3
PEaRL	7	68.1	4	35.5
RCMSA	8	19.5	8	13.0
J-linkage	16	283.0	36	197.0

此组数据上进行直线重建的时间小于 10 分钟, 考虑此时间开销后, 本方法总时间开销仍然优于 PEaRL 算法, 而与 RCMSA 算法相当.

5 结论

针对结构化场景 MVS 点云模型的平面拟合问题, 本文提出了一种基于 PEaRL 框架的改进方法. 该方法首先利用场景的直线信息生成比较可靠的候选平面, 并根据生成的候选平面约束空间点的相邻关系. 之后, 本文方法改进了 PEaRL 算法的能量函数, 并通过能量优化获得最终的场景平面信息.

直线信息的引入为点云的平面拟合提供了良好的初始模型及有效约束. 实验表明, 在结构化场景 MVS 点云模型的平面拟合问题上, 本文方法表现出了优于其他多模型拟合方法的准确度. 同时本方法由于得到的初始平面比较可靠, 计算速度较 PEaRL 算法也有明显提升.

参考文献

1 Furukawa Y, Ponce J. Accurate, dense, and robust multiview stereopsis. *IEEE Trans Pattern Anal Mach Intell*, 2010, 32: 1362–1376

2 Tola E, Strecha C, Fua P. Efficient large-scale multi-view stereo for ultra high-resolution image sets. *Mach Vision Appl*, 2012, 23: 903–920

3 Furukawa Y, Curless B, Seitz S M, et al. Manhattan-world stereo. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Miami, 2009. 1422–1429

4 Hong L, Chen G. Segment-based stereo matching using graph cuts. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Washington, 2004. I-74–I-81

5 Zitnick C L, Kang S B. Stereo for image-based rendering using image over-segmentation. *Int J Comput Vision*, 2007, 75: 49–65

6 Wang Z F, Zheng Z G. A region based stereo matching algorithm using cooperative optimization. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Anchorage, 2008. 1–8

7 Wang W, Gao W, Hu Z Y. Dense 3D scene reconstruction based on semantic constraint and graph cuts. *Sci Sin Inform*, 2014, 44: 774–792 [王伟, 高伟, 胡占义. 基于语义约束与 Graph Cuts 的稠密三维场景重建. *中国科学: 信息科学*, 2014, 44: 774–792]

8 Nguyen A, Le B. 3D point cloud segmentation: a survey. In: *Proceedings of IEEE Conference on Robotics, Automation and Mechatronics*, Manila, 2013. 225–230

9 Zuliani M, Kenney C S, Manjunath B S. The multiransac algorithm and its application to detect planar homographies. In: *Proceedings of IEEE International Conference on Image Processing*, Genova, 2005. III-153-6

10 Toldo R, Fusiello A. Image-consistent patches from unstructured points with J-linkage. *Image Vision Comput*, 2013, 31: 756–770

11 Isack H, Boykov Y. Energy-based geometric multi-model fitting. *Int J Comput Vision*, 2012, 97: 123–147

- 12 Boykov Y, Veksler O, Zabih R. Efficient approximate energy minimization via graph cuts. *IEEE Trans Pattern Anal Mach Intell*, 2001, 20: 1222–1239
- 13 Kolmogorov V, Zabih R. What energy functions can be minimized via graph cuts? *IEEE Trans Pattern Anal Mach Intell*, 2004, 26: 147–159
- 14 Boykov Y, Kolmogorov V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Trans Pattern Anal Mach Intell*, 2004, 26: 1124–1137
- 15 Pham T, Chin T J, Yu J, et al. The random cluster model for robust geometric fitting. *IEEE Trans Pattern Anal Mach Intell*, 2014, 36: 1658–1671
- 16 Toldo R, Fusiello A. Robust multiple structures estimation with J-linkage. In: *Proceedings of European Conference on Computer Vision*, Marseille, 2008. 537–547
- 17 Bay H, Ferrari V, Gool L V. Wide-baseline stereo matching with line segments. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, San Diego, 2005. 329–336
- 18 Werner T, Zisserman A. New techniques for automated architectural reconstruction from photographs. In: *Proceedings of European Conference on Computer Vision*, London, 2002. 541–555
- 19 Fu K P, Shen S H, Hu Z Y. Line matching across views based on multiple view stereo. *Acta Autom Sin*, 2014, 40: 1680–1689 [傅康平, 申抒含, 胡占义. 一种基于多视立体视觉的多视图直线匹配方法. *自动化学报*, 2014, 40: 1680–1689]
- 20 Delong A, Osokin A, Isack H, et al. Fast approximate energy minimization with label costs. *Int J Comput Vision*, 2012, 96: 1–27

A plane fitting approach using line information to reconstruct structured scenes

FU KangPing^{1*}, WANG Wei¹, HU LiHua² & HU ZhanYi¹

1 *National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China;*

2 *School of Computer Science and Technology, Taiyuan University of Science and Technology, Taiyuan 030024, China*

*E-mail: kpfu@nlpr.ia.ac.cn

Abstract Structured scenes are an important element of 3D reconstruction. Characteristics of structured scenes are analyzed, based on which a modified PEaRL algorithm is proposed to perform plane fitting on points provided by multiple view stereo algorithms. In this method, candidate planes are generated from both 3D lines and 3D points. The number of candidate planes is reduced using information provided by the 3D lines; furthermore, the reliability of the planes is improved. Subsequently, 3D lines are employed to obtain a more reliable neighborhood system, which leads to more accurate plane fitting results at plane intersections. For energy optimization, the smoothness term and label cost term are improved to better manage the irregularity of multi-view stereo (MVS) point clouds. Experiments on synthesized and actual data show that the proposed method provides satisfactory time efficiency, and obtains more reliable plane fitting results than other multi-model fitting methods, including PEaRL.

Keywords structured scene, plane fitting, model fitting, energy optimization, MVS, PEaRL, 3D lines

FU KangPing was born in 1987. He received his B.S. degree in Computer Science from the University of Science and Technology of China in 2009, and is currently a Ph.D. candidate at the Institute of Automation, Chinese Academy of Sciences. His current research interests include computer vision and pattern recognition.

WANG Wei received his Master's degree in 2011 from Southwest JiaoTong University. Currently, he is a Ph.D. candidate at the Institute of Automation, Chinese Academy of Sciences. His research interests include computer vision and machine learning.

HU LiHua was born in 1982. She received her M.S. degree in computer application technology from Taiyuan University of Science and Technology in 2006. Since 2009, she has been a lecturer at the School of Computer Science and Technology, Taiyuan University of Science and Technology. Her research interests include computer vision, pattern recognition, and reverse engineering.

HU ZhanYi was born in 1961. He received his B.S. degree in automation from the North China University of Technology in 1985, and his Ph.D. degree (Docteur d'Etat) in computer vision from the University of Liege, Belgium, in 1993. He is now a Professor at the Institute of Automation, Chinese Academy of Sciences.