融合梯度信息和邻域点云分布的 3D 线特征提取与配准

缪永伟^{1,2*}, 戴颖婷¹, 王海鹏¹, 刘复昌², 王金荣²

1. 浙江理工大学信息学院, 杭州 310018

2. 杭州师范大学信息科学与技术学院, 杭州 311121

* 通信作者. E-mail: ywmiao@zstu.edu.cn

收稿日期: 2021-01-11; 修回日期: 2021-03-21; 接受日期: 2021-05-05; 网络出版日期: 2021-12-17

国家自然科学基金 (批准号: 61972458) 和浙江省自然科学基金 (批准号: LY20F020017) 资助项目

摘要 针对传统点云场景重建中由于场景区域缺乏纹理、场景物体遮挡等导致重建结果不准确的局限性, 借助场景包含的几何特性和线结构信息, 有效利用 RGB-D 数据的梯度信息和邻域点云分布信息, 本文提出了一种针对点云场景的 3D 线特征提取、匹配和配准方法. 首先, 通过场景 RGB 图的梯度信息确定梯度方向相似的 3D 直线段支持域. 然后, 借助场景深度图呈现的邻域几何特性, 根据当前点与其邻近点之间的分布关系判断点云中的直线型边界线点云和直线型折边线点云并拟合得到场景 3D 线特征. 其次, 以点云场景提取的 3D 线特征为轴线, 构造等厚度同轴圆柱区域作为 3D 线段支持域, 并统计支持域中各层圆柱壁内点云梯度作为描述符实现 3D 线特征匹配. 最后, 利用基于线-线的迭代最近线配准算法, 迭代计算得到帧间场景的旋转与平移, 从而实现帧间点云场景的高效配准. 实验结果表明, 与已有的点云场景线特征提取及点特征配准方法比较, 本文直接基于线特征的方法配准效率高、计算量小、配准精准, 方法具有较强的鲁棒性.

关键词 点云场景, 3D 线特征, 线特征提取, 线特征配准, 三维重建

1 引言

在 3D 计算机视觉和三维重建中, 离散点云数据由于其数据获取容易、能有效表示三维形状或复杂场景、无需维护拓扑连接关系等优点得到广泛应用^[1]. 根据获得的场景点云数据 (包括 RGB-D 数据), 结合 SLAM 重建框架^[2], 能够重建出场景的 3D 数字化模型.

传统的点云场景重建主要基于点特征提取和配准实现, 如 Newcombe 等^[3]提出了一种仅使用深度图实现室内场景实时重建的方法 KinectFusion, 但是由于该方法使用固定体积的 TSDF (truncated

引用格式: 缪永伟, 戴颖婷, 王海鹏, 等. 融合梯度信息和邻域点云分布的 3D 线特征提取与配准. 中国科学: 信息科学, 2021, 51: 2069–2088, doi: 10.1360/SSI-2021-0009
Miao Y W, Dai Y T, Wang H P, et al. Extraction and registration of 3D lines by fusing gradient information and neighboring point cloud distribution (in Chinese). Sci Sin Inform, 2021, 51: 2069–2088, doi: 10.1360/SSI-2021-0009

signed distance fields) 模型表示场景, 其仅适合固定大小的室内小场景重建, 同时由于缺少回环检测和全局优化, 容易导致误差累积而造成重建场景的漂移. Whelan 等^[4]提出了一种适合于大中型场景三维重建的方法 Kintinuous, 该方法在传统 KinectFusion 方法基础上进行扩展, 把固定体积的 TSDF 模型随着相机运动进行移动, 从而能够实现大范围的室内场景重建. Whelan 等^[5]进一步提出利用 Surfel 模型实时重建室内房间大小场景的方法 ElasticFusion, 该方法克服了 Kintinuous 中不支持来回扫描同一片区域并进行数据融合的缺点. Nießner 等^[6]提出了一种利用体素哈希技术节省显存空间的实时重建方法 InfiniTAM. 利用 TSDF 模型表示, 该方法在重建场景时仅对场景表面划分网格进行表面数据的实时更新, 而不是将整个空间都划分成网格, 从而提高了重建方法的效率. Dai 等^[7]提出了一种能够有效减少优化变量数目的实时重建方法 BundleFusion, 该方法在场景建图时将全部帧划分成等大小的块, 先对块内进行局部位姿优化, 再对所有块进行全局位姿优化以加速优化过程. 上述方法通常需要提取 2D 图像中的离散特征点, 并通过前后帧图像之间的特征点匹配建立帧间对应关系, 从而进行特征跟踪或配准对齐. 然而, 由于 2D 图像提取的特征点数量通常规模较大, 其特征描述子计算和帧间匹配比较耗时; 另一方面, 如果图像质量较低、场景受到遮挡或者区域纹理较少, 通常难以提取出足够的场景有效特征点, 从而导致帧间匹配失效、位姿估计错误, 使得场景重建不准确.

为了克服这些困难, 基于线特征的 SLAM 重建得到研究者的普遍关注, 如 Zhang 等^[8]提出了一种以直线为特征的视觉 SLAM 重建方法 Line-SLAM, 该方法使用 PnL (perspective-n-line)^[9]进行相机位姿估计, 并将线特征重投影误差定义为 2D 线段两个端点到预测 3D 直线的距离, 再利用基于图优化的 SLAM 框架构建三维线状环境地图. Zhou 等^[10]提出的 Struct-Line-SLAM 则使用建筑结构线作为定位和建图的特征. 与其他线特征不同, 建筑结构线编码全局方向信息, 其随着时间的推移限制了相机的航向, 同时消除了累积的方向误差, 从而减少了位置漂移. Pumarola 等^[11]提出了一种点线结合的重建方法 PL-SLAM, 该方法在传统 SLAM 重建基础上加入了线特征提升其在弱纹理环境中的运行性能, 并通过 2D 线段端点逼近预测的 3D 直线投影点以最小化直线的重投影误差来估计相机位姿. Li 等^[12]提出了一种基于结构线特征的方法 Struct-PL-SLAM, 该方法使用结构线编码全局定位信息, 并利用有效的几何约束以消除建图时累积的方向误差, 从而有助于提高 SLAM 重建的精度. 然而, 由于传统基于线特征的 SLAM 方法仅适用于单目或双目图像数据而并不适用于 RGB-D 扫描数据, 同时上述方法使用的线特征提取及匹配均是基于 2D 图像的线特征提取及匹配, 并利用 PnL 的各种变形估计相机位姿得到 3D 线特征, 这些方法中并没有涉及到场景 3D 线特征的直接提取与匹配, 也未考虑 3D 线特征的直接配准. 不同于激光点云的线特征提取^[13~15], 场景 RGB-D 扫描数据往往由于点云精度低、法线杂乱无序、点云分布较稀疏等原因, 难以使用 RANSAC (random sample consensus)^[16]优化扫描点法线方向并拟合相同法线方向的离散点为直线, 从而导致传统适用于激光点云数据的特征提取方法并不适用于场景 RGB-D 数据; 且传统 RGB-D 场景重建中通常采用线段端点、线段中点或者两条匹配线段对应点作为输入并采用基于点的配准策略^[17,18], 因而亟待研究者提出针对点云场景 (或 RGB-D 数据) 的 3D 线特征描述及鲁棒提取, 同时需要相应考虑点云场景中 3D 线特征的直接匹配和配准方法.

为了有效利用点云数据所包含的梯度信息和邻域点云分布信息, 本文提出了一种针对场景 RGB-D 数据 (或者点云场景) 的 3D 线特征提取、匹配和配准方法. 其主要贡献如下: (1) 提出一种点云场景中 3D 线特征的鲁棒提取方法, 该方法融合场景数据的 RGB 梯度信息和点云分布特性以有效提取场景的 3D 线特征; (2) 提出一种结合 3D 线段几何属性和邻域点分布的圆柱型线段支持域构建方法, 并基于 3D 线段支持域提出相应匹配算法以实现 3D 线特征的直接匹配; (3) 提出一种直接通过 3D 线特征所在直线计算线特征间距离的基于线-线的迭代最近线 (line-to-line iterative closest line, line-to-line

图 1 (网络版彩图) 本文方法流程

Figure 1 (Color online) Flow diagram of the proposed method

ICL) 配准算法, 以实现点云场景线特征的高效配准.

其余章节安排如下: 第 2 节给出了本文方法的流程; 第 3 节阐述了点云场景 3D 线特征的鲁棒提取; 第 4 节介绍了点云场景 3D 线特征的匹配与配准; 第 5 节进行了实验讨论与结果分析; 第 6 节对论文进行了总结与展望.

2 本文方法流程

在传统基于点特征的场景重建中, 考虑到因场景区域缺乏纹理、场景内物体遮挡而导致重建结果不准确的局限性, 借助点云场景包含的几何特性和线结构信息, 本文提出了一种针对点云场景的 3D 线特征提取方法, 并基于 3D 线特征的圆柱型支持域及特征描述符进行 3D 线特征的匹配, 最终利用基于线-线的迭代最近线配准算法实现场景点云数据的高效配准. 本文方法流程如图 1 所示, 以下各节将对方法的各主要环节进行详细介绍.

3 点云场景 3D 线特征提取

为了高效提取点云场景中含有的大量 3D 线特征, 避免传统的基于图像匹配方法出现的几何信息缺失^[19,20]、基于 2D 与 3D 投影方法导致的投影误差较大^[21,22] 以及基于点云方法中由于平面分割导致线特征提取计算量过大^[14,23] 等局限性, 本文考虑采用 2D 信息与 3D 信息相融合的策略提取点云场景的 3D 线特征. 一般地说, 场景的二维 RGB 图像通常包含两类 2D 线特征^[24]: (1) 灰度边缘线, 是指图像中局部不连续的特征或者灰度值发生突变的位置, 即梯度最大的位置; (2) 纹理线, 是指含有相同纹理的相邻区域在不同尺度下的粗糙度突变或相邻区域的方向纹理之间的方向性突变. 相比较而

言, 三维点云场景中通常包含两类 3D 线特征^[13,14]: (1) 边界线, 是指表面轮廓 (facade outlines)、遮挡轮廓 (occluding contours)、跳高线 (height jump lines) 等各类平面或曲面的轮廓边缘; (2) 折边线, 是指平面相交线 (plane intersection lines)、尖锐特征线 (sharp feature lines)、折线 (breaklines) 等不同平面或曲面之间的相交线. 需要注意的是, 这里的特征线可以是直线也可以是曲线, 而本文的 3D 线特征特指直线段.

针对输入的场景 RGB-D 数据 (或点云数据), 本文提出的 3D 线特征提取和检测方法包含两个步骤, 即基于 2D 梯度信息的边缘检测预处理以及基于 3D 几何信息的边界线点云与折边线点云判断.

3.1 场景边缘检测预处理

在点云场景 3D 线特征提取中, 采样点的法向起到重要作用. 在法向计算中, 可以通过物体/场景表面进行局部平面拟合, 并通过计算局部平面法线得到采样点法线^[25]. 由于考虑到同一 3D 直线点集中所有采样点的法线方向应近似相同, 从而在 3D 线特征提取时需要找到方向接近且共线的点并拟合以提取场景的 3D 线特征. 然而, 受限于获取的点云数据精度, 导致 3D 直线上的采样点法线方向缺少一致性, 即相邻采样点其法向偏差较大, 难以在点云中直接判断为直线点集. 与无序、不规则分布的散乱点云不同, RGB 图像像素点呈均匀规则分布, 其梯度信息计算通常较精确, 在线特征提取中能准确地获取梯度方向相似的直线拟合区域^[26], 因而本文采用二维 RGB 图像的梯度信息找到对应的直线拟合区域并辅助 3D 线特征的提取.

由于场景 RGB 图像中的物体边缘通常是连续的, 受 EDLines 方法^[27]的启发, 本文使用边缘检测算法确定直线拟合区域. 该方法以 RGB 图像的梯度局部峰值点为锚点 (anchor points), 这些锚点将对应于场景边缘可能会覆盖的点, 在场景边缘检测时通过从一个锚点开始加入相邻点并绘制其整条边缘, 从而有效提取干净、连续的边缘像素链. 具体地说, 首先利用 Sobel 梯度算子计算出场景数据 RGB 图像的梯度信息: 水平梯度 G_x 和垂直梯度 G_y , 其每个像素点的梯度值为 $G = \sqrt{G_x^2 + G_y^2}$. 如果水平梯度较大, 即 $G_x > G_y$, 则假定垂直边缘穿过该像素; 否则, 假定是水平边缘穿过该像素, 而对于水平边缘穿过的像素点, 与其上下相邻点的梯度值进行比较; 对于垂直边缘穿过的像素点, 则与其左右相邻点的梯度值进行比较. 比较时, 如果该像素点的梯度值大于其两相邻像素点的梯度值一定的阈值 (该阈值控制两个锚点之间间隔的像素点数目, 实验中将阈值设为 0), 则该像素点被标记为一个锚点. 通过改变阈值可以调整锚点的数量, 阈值越低, 锚点越密集; 阈值越高, 锚点越稀疏. 然后, 通过在锚点之间绘制边来连接锚点生成完整的场景边缘像素链. 需要注意的是, 锚点只是边缘开始的位置, 并不意味着每个锚点都包含在最终的边缘中, 因而需要根据穿过锚点的边缘方向检查 3 个近邻像素点, 并选择梯度值最大的像素点作为下一个连接点. 一般来说, 如果一条水平边缘穿过锚点, 则将锚点向左或者向右移动 (一般先左后右) 寻找下一个连接点, 此时只考虑当前点左侧或者右侧的 3 个最近邻点; 如果一条垂直边缘穿过锚点, 则将锚点向上或者向下移动 (一般先上后下) 寻找下一个连接点, 此时只考虑当前点上方或者下方的 3 个最近邻点, 然后选择梯度值最大的近邻点作为下一个连接点, 并继续该移动过程直到当前像素点超出边缘区域 (即当前像素的梯度值为零), 或者遇到先前检测过的边缘点时停止. 最后, 根据连续边缘点的梯度方向将场景边缘像素链分割成若干具有相似梯度方向且位置相邻的像素点组成的连通区域, 各区域内边缘点梯度方向变化不超过阈值 $\tau = \pi/8 = 22.5^\circ$. 本文方法将各连通域作为 2D 直线拟合区域, 以此找到对应的 3D 直线拟合区域并生成具有相同直线编号的场景子点云, 后续将利用场景子点云的直线编号分别提取场景直线型边界线点云或场景直线型折边线点云.

图 2 (网络版彩图) 边界线点

Figure 2 (Color online) Sampling points of boundary lines

3.2 场景直线型边界线点云提取

作为点云场景边界线, 由于其一侧通常不包含点云数据点, 而另一侧则为点云数据点密集的三维平面或曲面, 因而可以利用点云数据点及其 K -邻域内的离散点分布均匀程度来判断是否为场景边界点. 这里可以通过判断当前点与其邻域点连线向量之间的夹角直观地度量该均匀程度. 如果当前点与其邻域点连线向量之间的最大夹角超过阈值, 当前点即为边界点, 反之则是内部点. 该阈值大小的设置将决定场景边界线点云提取的效果, 阈值越大, 则获取的边界线形状愈尖锐, 所得到的采样点数目越少; 反之则获取的边界线形状愈平缓, 所得到的采样点数目越多. 为了得到较完整的场景直线型边界线点云, 经实验将阈值设为 45° 是合适的选择. 在提取场景边界线点云过程中, 本文利用最小二乘法通过当前点 P 及其 K -邻域点拟合出 P 点的局部切平面 S , 并将当前点 P 及其 K -邻域点投影到切平面上; 然后, 根据 P 点与其 K -近邻点连线向量之间的最大夹角识别场景边界点, 即连线向量之间的最大夹角大于阈值时, 本文认为当前点是边界点, 如图 2 所示.

3.3 场景直线型折边线点云提取

一般地说, 场景点云数据中仅包含平面点、弧面点、折边点 3 种类型的数据点^[13,14]. 其中平面点和弧面点用于上述的场景直线型边界线点云判断, 而折边点则是指两平面或曲面的交线上的点. 这 3 类采样点的点云法线分布通常呈现一定规律性, 我们可以利用当前点及其 K -邻域内的离散点法向偏差情况判断采样点的分类. 对于平面采样点, 其 K -邻域内的邻域点法向基本相同, 即两两采样点的法向偏差近似为零; 对于弧面点, 其 K -邻域内的离散点法线方向偏差基本相同; 而对于折边点, 由于其法线方向同时垂直于两相交平面或曲面, 当两相交面靠近折边线时其采样点法线方向通常会发生骤变, 从而导致出现两个明显的法线方向偏差峰值. 因而, 在对当前点及其 K -邻域内邻域点集合的两两采样点法向偏差进行直方统计时, 平面点没有峰值, 弧面点有唯一的峰值, 而折边点则有两个峰值.

在场景直线型折边线点云提取中, 本文首先利用 PCA^[28] 计算当前点 P 及其 K -邻域点的法线方向, 并从中找到与当前点法线 N_P 夹角最大的邻近点法线 $N_{\max}$. 从而垂直平面法线方向可计算为 $N_V = N_P \times N_{\max}$, 相应得到与局部平面、局部弧面或者折边线垂直且过当前点 P 的垂直平面 V . 然后, 将当前点 P 及其 K -邻域点的法线投影到垂直平面 V 上, 并根据这些法线投影的分布情况识别场景点云的折边点. 如图 3 所示, 首先以当前点 P 为坐标原点, 其法线方向 N_P 为 u 轴, 建立局部坐标系 uw ; 然后将其所有 K -邻域点的法线均投影到局部坐标系 uw 所在的平面 V 上, 并把所有邻域

图 3 (网络版彩图) 折边线点. (a) 平面点, (b) 弧面点和 (c) 折边点

Figure 3 (Color online) Sampling points of folded edge lines. (a) A point on a plane; (b) a point on a curved surface; (c) a point on a folded edge

点法线投影均平移至局部坐标系原点 P , 再计算两两邻域采样点的法向偏差; 最后, 对所有邻域采样点的法向偏差进行直方统计, 并根据直方图中的峰值数目判断当前点 P 的类型. 为了便于找到直方图中的峰值, 这里采用拉伸直方图方法使得其峰值更为突出, 即对每个直方图元素 H_i 进行归一化, 并计算直方均值 $\bar{H}$ 和标准差 σ , 并判断 $H'_i = H_i - \bar{H}$ 与 σ 的大小, 若 $H'_i > \sigma$ 则认为 H_i 为一个峰值. 一般地说, 当前点 P 为折边点是指其邻域采样点的法向偏差直方图能找到两个峰值, 但由于受实际点云获取精度的影响, 两张平面相交形成的可能并不是简单折边线, 而是具有一定宽度的微小折边弧面, 从而导致其邻域点的法向偏差直方图可能出现多个距离相近的峰值, 故对最终的点云类型判断结果造成影响. 为了提高点云类型判断结果的准确性, 本文需要对当前采样点及其邻域点的法向偏差直方图进行邻近峰值聚集后, 再判断其峰值数量. 在法向偏差直方图中, 首先沿横坐标方向找到之前判定的第一个峰值元素 H_i 并设置容差 τ 以控制聚集的范围 (实验中 τ 设为 5). 然后, 从下标 $i + \tau$ 向下标 i 方向检查这个聚集中是否存在其他峰值元素, 如果存在则将该聚集中的所有峰值累加并作为新的峰

值 H_i , 再继续沿着横坐标方向找到下一个峰值元素并检查聚集, 重复上述操作直到检查完直方图内所有的聚集为止. 最后, 利用直方均值 $\bar{H}$ 判断聚集后的峰值数量, 并判断当前点的归属. 此时 $\bar{H}$ 表示聚集后的新均值, 若 $H_i > \bar{H}$ 或 $\alpha = H_i - \bar{H} < 0.2H_i$, 则认为 H_i 为一个峰值. 如果直方图中找到两个峰值, 则当前点 P 被确定为折边点.

3.4 场景 3D 线特征提取

场景 3D 线特征提取的目的是找到场景点云数据中的直线型边界线点云和直线型折边线点云, 并将其分别拟合成直线段作为场景 3D 线特征. 经过前述步骤可以得到带有直线编号的场景边界线点云和折边线点云. 然而, 由于光照投影、反射或折射等影响, 实际场景点云数据中往往混杂了距离实际场景边界线或折边线较近的阴影线点云^[29], 这些阴影线点云的存在将干扰后续的场景直线段拟合和 3D 线特征匹配, 因此这里需要通过对边界线点云或折边线点云进行聚类, 并根据类之间质心的欧氏距离将场景边界线点云、折边线点云和场景阴影线点云分开. 由于场景的线段簇点云分布形状并不规则, 本文使用基于密度的 DBSCAN (density-based spatial clustering of application with noise) 聚类方法^[28]实现线段簇点云聚类, 同时也可以滤除噪声点及离群点.

在场景 3D 线特征提取中, 本文首先利用 DBSCAN 聚类方法对具有同一直线编号的边界线点云或折边线点云进行聚类; 通过比较同一直线编号中每一个子聚类质心和场景点云总质心 z 值的大小, 确定子聚类的深度远近. 如果子聚类质心在总质心前方, 则认为当前子聚类属于真实的场景边界线点云或场景折边线点云; 如果子聚类质心在总质心后方, 则认为当前子聚类属于其他混入其中的场景阴影线点云而被直接丢弃. 然后, 在剩余聚类中取采样点数最多的聚类并作为初始的线特征拟合点集, 计算该拟合点集的 PCA 主方向^[25]和均值点, 从而得到一条过均值点的主方向直线; 同时计算其余聚类中的各采样点与该主方向直线之间的垂直距离, 并剔除与该主方向直线距离过大的点, 再将剩余采样点加入到线特征拟合点集中并更新该线特征拟合点集的 PCA 主方向和均值点以修正主方向直线; 最后将参与拟合主方向直线的点均投影到该直线上, 并找到线段的端点以便从直线中截取出线段作为场景的 3D 线特征.

4 场景 3D 线特征的匹配及配准

受场景 2D 线描述符 LBD (line band descriptor) 表示^[30]的启发, Burns 等^[26]引入的线支持域为 2D 图像空间中像素梯度方向相似的 2D 直线拟合区域, 而 2D 线描述符 LBD 则扩大了线支持域的范围, 其通常选取线特征周围以 2D 线特征为水平中心线的矩形局部区域作为 2D 线段支持域, 并将 2D 线段支持域划分为一组等宽度的条带, 然后统计条带内水平和垂直正负 4 个方向的像素梯度形成描述矩阵, 并通过 2D 线特征周围的梯度纹理变化描述 2D 线特征的相似性. 然而, 为了更好地描述场景 3D 线特征周围的点云分布, 本文需要考虑场景的 3D 线段支持域及线特征描述符的构建.

4.1 3D 线段支持域及描述符的构建

在支持域构建中, 我们以 3D 线特征为轴线向外分层扩展, 其每层圆柱壁厚度相同的同轴圆柱区域作为 3D 线段支持域. 如图 4 所示, 首先以提取得到的场景 3D 线特征 l 为轴线, 计算参与拟合当前线特征的场景点云中的点到轴线的距离, 并取最大距离为单层圆柱壁半径 r , 然后向外逐层扩展一个围绕 3D 线特征的等厚度同轴圆柱区域作为 3D 线段支持域, 支持域圆柱体半径为 $n \cdot r$, 本文实验中 n 取为 5.

图 4 (网络版彩图) 3D 线段支持域

Figure 4 (Color online) The supporting region of a 3D linear segment

在构建线特征描述符时, 以围绕 3D 线特征的等厚度圆柱壁内点云的 3D 梯度作为 3D 线段支持域的点云纹理. 一般来说, 场景点云的 3D 梯度描述采样点 K - 邻域内点云的分布及其法线方向变化 (反映了采样点邻域点云分布最密集的方向), 因而本文通过统计圆柱壁内所有采样点在 6 个方向上的 3D 梯度投影描述当前圆柱壁内的点云纹理, 并由内向外拼接每层圆柱壁的 6 个统计值形成描述符矩阵, 并用该描述符矩阵描述整个圆柱型 3D 线段支持域内的点云纹理. 具体地说, 为了构建 3D 线特征描述符, 首先要求出支持域各层圆柱壁内采样点的 3D 梯度信息. 基于场景点云的 3D SIFT 描述符中^[31], 对应 2D 像素梯度定义为像素点邻域灰度梯度直方图统计的峰值, 而 3D 点云梯度则定义为当前点 K - 邻域点到邻域中心点分布的角度直方图峰值, 场景点云 $P(x, y, z)$ 的 K - 邻域点到邻域中心点分布的距离和相对角度可计算为

$$\begin{cases} m(x, y, z) = \sqrt{L_x^2 + L_y^2 + L_z^2}, \\ \theta(x, y, z) = \arctan(L_y/L_x), \\ \phi(x, y, z) = \arctan\left(L_z/\sqrt{L_x^2 + L_y^2}\right), \end{cases}$$

其中, 三维点云的尺度空间 $L(x, y, z, \sigma)$ 定义为由一个尺度变化的高斯 (Gauss) 核函数 $G(x, y, z, \sigma)$ 与输入点云 $P(x, y, z)$ 卷积得到, 即 $L(x, y, z, \sigma) = G(x, y, z, \sigma) * P(x, y, z)$, 其中 $\sigma = \sqrt[3]{2}$ 是尺度空间因子, 高斯核函数 $G(x, y, z, \sigma) = \frac{1}{(\sqrt{2\pi}\sigma)^3} e^{-(x^2+y^2+z^2)/2\sigma^2}$; L_x, L_y, L_z 分别是在尺度空间下点 $P(x, y, z)$ 与邻域点的差分变量在 x, y, z 轴上的投影分量, θ 和 ϕ 分别为点 $P(x, y, z)$ 与上述邻域点的方位角和俯仰角.

进一步, 本文考虑到点云法线在点云特征刻画中的重要性, 在特征描述中结合点云法线分布将更全面地表达场景点云的几何特征和邻域点云空间分布特性, 故在计算场景点云的 3D 梯度时本文利用梯度方向和法线方向的角度偏差刻画当前点邻域内法线的变化. 设当前点 $P(x, y, z)$ 的邻域点集为 $\{P_k(x_k, y_k, z_k)\}$, 其中心点为 $P_c(x_c, y_c, z_c)$, 则结合法向特征的场景邻域点云 $P_k(x_k, y_k, z_k)$ 分布的距离

和相对角度可计算为

$$\begin{cases} m(x_k, y_k, z_k) = \sqrt{(x_k - x_c)^2 + (y_k - y_c)^2 + (z_k - z_c)^2}, \\ \theta(x_k, y_k, z_k) = \arctan[(y_k - y_c)/(z_k - z_c)], \\ \phi(x_k, y_k, z_k) = \arctan[(z_k - z_c)/\sqrt{(x_k - x_c)^2 + (y_k - y_c)^2}], \\ \delta(x_k, y_k, z_k) = \arccos \left[\frac{\mathbf{PP}_k \cdot \mathbf{N}}{\|\mathbf{PP}_k\| \cdot \|\mathbf{N}\|} \right], \end{cases}$$

其中梯度大小 $m(x_k, y_k, z_k)$ 为当前点邻域点集的中心点 $\mathbf{P}_c(x_c, y_c, z_c)$ 到邻域点 $\mathbf{P}_k(x_k, y_k, z_k)$ 之间的欧氏距离, θ 和 ϕ 分别为中心点 $\mathbf{P}_c(x_c, y_c, z_c)$ 与邻域点 $\mathbf{P}_k(x_k, y_k, z_k)$ 的方位角和俯仰角, δ 为当前点 $\mathbf{P}(x, y, z)$ 的法线 $\mathbf{N}$ 和向量 $\mathbf{PP}_k$ 的角度偏差. 为了使得点云 3D 梯度的方向沿着其 K -邻域内邻域点分布最密集且法向偏差最大的方向, 本文按照四元组 $(m, \theta, \phi, \delta)$ 统计当前点所有 K -邻域点相对于邻域中心点相对角度的直方图, 其直方图的峰值方向为当前点 $\mathbf{P}$ 的 3D 梯度方向, 其对应的幅值为当前点 $\mathbf{P}$ 的 3D 梯度大小.

结合场景点云法向特征计算得到点云 3D 梯度后, 本文将局部坐标系旋转使其 x 轴指向 3D 线特征方向, y 轴指向参与拟合 3D 线特征的点集拟合得到的切平面中与 x 轴垂直的方向, z 轴指向参与拟合 3D 线特征的点集拟合得到的切平面法线方向. 在该局部坐标系下, 首先计算 3D 线段支持域内场景点云 3D 梯度 $g = (g_x, g_y, g_z)$; 然后在 3D 线段支持域内第 i 层圆柱壁中分别累计该层中 6 个方向上的 3D 梯度幅值:

$$V_i^1 = \lambda \sum_{g_x > 0} g_x, V_i^2 = \lambda \sum_{g_x < 0} g_x, V_i^3 = \lambda \sum_{g_y > 0} g_y, V_i^4 = \lambda \sum_{g_y < 0} g_y, V_i^5 = \lambda \sum_{g_z > 0} g_z, V_i^6 = \lambda \sum_{g_z < 0} g_z,$$

其中权值 $\lambda = e^{-d_i^2/[2(n \cdot r)^2]}$, d_i 为第 i 层圆柱壁外侧到线特征中心轴线的距离, $n \cdot r$ 是圆柱型 3D 线段支持域的半径, 权值 λ 的引进将逐渐降低线段支持域内远离 3D 线段的外层圆柱壁内点云梯度的重要性. 最终, 3D 线段支持域内各圆柱壁的描述子计算为 $\text{CD}_i = (V_i^1, V_i^2, V_i^3, V_i^4, V_i^5, V_i^6)$. 本文提出的 3D 线特征圆柱型描述符 LCD (line cylindrical descriptor) 定义为 $\text{LCD} = (\text{CD}_1, \text{CD}_2, \dots, \text{CD}_n)$, 其计算如算法 1 所示.

4.2 基于 3D 线特征描述的配准

根据对点云场景每帧数据分别提取得到的 3D 线特征, 我们可以计算其相应的线特征描述符, 进一步可以确定两帧点云数据中线特征描述符距离最小的 3D 线特征匹配对, 并考虑两帧场景点云数据之间的旋转和平移, 从而实现场景点云数据的配准. 具体地说, 对齐两帧点云数据 (源帧点集 $P = \{\mathbf{p}_i\}$ 和目标帧点集 $Q = \{\mathbf{q}_i\}$) 时需要寻求刚体变换 $(\mathbf{R}, \mathbf{t})$, 使得其对点集 P 进行变换后能够与点集 Q 很好地重合. 自 Besl 与 McKay^[32] 提出点对点的迭代最近点 (iterative closest point, ICP) 算法以来, ICP 算法及其变种^[33, 34] 已成为广泛采用的三维点云配准方法. 在点对点的 ICP 配准中, 实现点云数据配准时使用了点对点的距离误差度量, 并通过计算给定的对应匹配点对 $(\mathbf{p}_i, \mathbf{q}_i)$, 寻找能够使点对点的距离误差总和目标函数最小化的刚体变换, 即最小化 $E = \sum_i \|\mathbf{R} \cdot \mathbf{p}_i + \mathbf{t} - \mathbf{q}_i\|^2$. 迭代最近线 ICL^[17, 18] 具有与 ICP 相同的逻辑, 线耦合的条件是迭代对应线之间的距离. 但是现有的迭代最近线方法通常借助于线段对应点, 将对应线之间的距离简化为对应点之间的距离来计算, 如果存在线段分段或遮挡则可能找不到对应点, 导致计算出的线段之间的距离误差过大. 在线特征提取与描述的基础上, 本文利用 3D 线特征所在的直线计算线段之间的距离, 提出了基于线-线的迭代最近线配准算法计算点云场景 3D 线集之间的旋转和平移.

Algorithm 1 Intrinsic descriptor calculation algorithm for 3D lines

Input: 3D line l and its fitting point set $S = \{P_k\}$;

- 1: Taking the 3D line l as a central axis, calculate the distance from its fitting points $\{P_k\}$ to the central axis, in which the maximum distance can be determined as the radius of single-layer cylindrical area, so that an n -layer coaxial cylinder support region with the iso-thick cylindrical area can be built (here n is taken as 5);
- 2: Taking the direction of input 3D line as x axis and the vertical direction on the fitting plane of point set S as y axis, we can find the corresponding approximate plane of the real point cloud surface which is contained in the cylinder area;
- 3: Calculate four corner vertices of the approximate plane through two endpoints of the input 3D line and the cylindrical radius, and then find the irregular quadrilateral 2D region which can be obtained by projecting the 3D cylinder support region in RGB channel;
- 4: Establish an index table of 2D points that correspond to 3D points and reverse 3D points in the cylinder area using 2D points, so that the adjacent sampling points can be divided into different layer cylindrical areas according to the distance with its 3D point to the central axis l ;
- 5: Lookup the neighboring points of 3D points using the above index table and calculate 3D gradient for each 3D point; then rotate the local coordinate system according to the input 3D line and calculate its projection $g = (g_x, g_y, g_z)$ of 3D gradient on the cylinder support region of the 3D line l ;
- 6: The descriptor of single-layer cylindrical area can be computed as the accumulated gradient component values of each cylinder area in 6 directions $g_x > 0, g_x < 0, g_y > 0, g_y < 0, g_z > 0, g_z < 0$, and the intrinsic descriptor of 3D line l can be determined by combining these single-layer descriptors;

Output: The intrinsic descriptor $LCD = (CD_1, CD_2, \dots, CD_n)$ for the 3D line l .

在点云场景的配准中, 为了简单起见, 利用 3D 线特征上的任意一点 (实验中, 我们使用拟合线特征点集的均值点) 和线特征的方向向量来表示线特征所在的直线. 给定点云场景两帧线集, 其中源帧线集表示为 $L^p = \{L_i^p | L_i^p = (p_i, v_i^p)\}$, $p_i(x_i^p, y_i^p, z_i^p)$ 是线上一点, $v_i^p = (l_i^p, m_i^p, n_i^p)$ 为 3D 线特征的单位方向向量; 目标帧线集表示为 $L^q = \{L_j^q | L_j^q = (q_j, v_j^q)\}$, 其中 $q_j(x_j^q, y_j^q, z_j^q)$ 是线上一点, $v_j^q = (l_j^q, m_j^q, n_j^q)$ 为 3D 线特征的单位方向向量. 该两帧线集中一组对应 3D 线之间的距离计算为 $d(L_i^p, L_j^q) = |p_i q_j \cdot s_{i,j}|$, 其中 $p_i q_j$ 为异面直线 L_i^p 和 L_j^q 分别任取一点的连线向量, $s_{i,j} = v_i^p \times v_j^q$ 为异面直线 L_i^p 与 L_j^q 公垂线的单位方向向量, 向量 $p_i q_j$ 在公垂线方向 $s_{i,j}$ 上的投影即为异面直线 L_i^p 与 L_j^q 之间的距离.

现在用 $T = (R, t)$ 表示线集 L^p 和 L^q 的刚性变换, 其中 R 和 t 分别表示旋转矩阵和平移向量. TL^p 表示刚性变换作用于线集 L^p , 其包含 $Tp_i = R \cdot p_i + t$ 和 $Tv_i^p = R \cdot v_i^p$, 则目标函数 E_{line} 可以定义为两帧对应线集之间的距离平方和加上对应线集之间的方向向量角度度量, 即

$$E_{\text{line}} = \sum_{i=1}^n d(TL_i^p, L_i^q) = \sum_{i=1}^n \{[(q_i - R \cdot p_i - t)(R \cdot v_i^p \times v_i^q)]^2 + (1 - (v_i^q)^T \cdot R \cdot v_i^p)\}.$$

点云场景 3D 特征线集之间配准的目标是寻求最佳 R 和 t 使得 E_{line} 最小.

类似于 ICP 方法, 这里首先需要通过两帧线集的方向向量协方差矩阵计算旋转矩阵 R 如下: 计算协方差矩阵 $C = \frac{1}{n} \sum_{i=1}^n [v_i^p \cdot (v_i^q)^T - \bar{v}^p \cdot (\bar{v}^q)^T]$, 其中 $\bar{v}^p = \frac{1}{n} \sum_{i=1}^n v_i^p$ 和 $\bar{v}^q = \frac{1}{n} \sum_{i=1}^n v_i^q$ 分别是源帧线集 L^p 方向向量的均值和目标帧线集 L^q 方向向量的均值, 旋转矩阵 R 可以通过协方差矩阵 C 的奇异值分解 SVD 和四元数表示得到, 这里使用 SVD 分解 $C = U\Lambda V^T$, 得到 $R = VU^T$. 计算得到旋转矩阵 R 后, 利用 R 对源帧线集进行旋转, 并计算旋转后的源帧线集与目标帧线集之间的平均距离作为平移向量 $t = \sum_{i=1}^n w_i (q_i - R \cdot p_i) \cdot (R \cdot v_i^p \times v_i^q)$, 其中 w_i 是权重, 即当前匹配对线特征长度和所有参与配准的线特征匹配对总长度的比值. 最后对源帧线集 L^p 进行旋转和平移变换, 得到更新后的新线集 L^p_* . 重复上述操作, 计算已更新的源帧线集 L^p_* 和目标帧线集 L^q 中所有对应直线对的最小化目标函数 E_{line} , 并更新旋转矩阵 R 和平移向量 t . 需要指出的是, 本文认为两帧线集距离最小同时其

角度偏差达到最小时取得最佳旋转矩阵 $\mathbf{R}$ 和平移向量 $\mathbf{t}$, 所以当前迭代的目标函数 E_{line} 大于前一次迭代的目标函数 E_{line} 时迭代停止, 如算法 2 所示.

Algorithm 2 Line-to-line iterative closest line (line-to-line ICL) registration algorithm

---

Input: The matched 3D line sets $\{L_i^p\}$ and $\{L_i^q\}$ which are extracted from the source frame and the target frame of scene data respectively;

- 1: **while** E_{line} of current iteration is not more than that of previous iteration **do**
- 2:   Compute the average direction vectors for the extracted 3D lines $\{L_i^p\}$ and $\{L_i^q\}$ respectively, that is, $\bar{\mathbf{v}}^p = 1/n \sum_{i=1}^n \mathbf{v}_i^p$, $\bar{\mathbf{v}}^q = 1/n \sum_{i=1}^n \mathbf{v}_i^q$;
- 3:   Calculate the covariance matrix for direction vectors of the extracted lines, that is, $\mathbf{C} = 1/n \sum_{i=1}^n [\mathbf{v}_i^p \cdot (\mathbf{v}_i^q)^T - \bar{\mathbf{v}}^p \cdot (\bar{\mathbf{v}}^q)^T]$;
- 4:   Decompose the covariance matrix using the singular value decomposition (SVD) method, that is, $\mathbf{C} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T$, in which $\mathbf{\Lambda}$ is a diagonal matrix and the rotation matrix can be computed as $\mathbf{R} = \mathbf{V}\mathbf{U}^T$;
- 5:   After the rotation transformation, the average difference between the aligned 3D lines can be computed as a translation vector, that is, $\mathbf{t} = \sum_{i=1}^n w_i (\mathbf{q}_i - \mathbf{R} \cdot \mathbf{p}_i) \cdot (\mathbf{R} \cdot \mathbf{v}_i^p \times \mathbf{v}_i^q)$;
- 6:   Calculate the target function value E_{line} , and update the extracted 3D lines from source frame scene data using the obtained rotation matrix $\mathbf{R}$ and translation vector $\mathbf{t}$;
- 7: **end while**

Output: The rotation matrix $\mathbf{R}$ and translation vector $\mathbf{t}$ between the matched 3D line sets.

---

5 实验结果及分析

本文实验平台为 Intel Core i7, GTX1660ti, 6 GB GPU, CUDA 10.2, 在 64 位 Win 10 系统下利用 Visual Studio 2017 环境开发. 本文提出的融合梯度信息和点云分布特性的场景 3D 线特征提取方法适用于普通精度的场景点云, 该方法直接利用场景结构线的几何特性抽取场景线特征, 有效避免了由于点云精度过低导致的场景平面分割欠准确的缺陷; 同时, 利用点云梯度信息指导直线拟合, 从而避免了线特征提取中直接利用法线判断直线时由于精度限制而导致提取失效的局限性. 本文实验分别利用室内场景 RGB-D 数据集 Matterport3D^[35] 和扩展的 ICL-NUIM 合成数据集^[36] 以验证方法的有效性.

5.1 点云场景 3D 线特征的提取

为了验证场景 3D 线特征提取方法的有效性, 本文在室内场景 RGB-D 数据集 Matterport3D^[35] 上对其进行了实验测试. 针对 3 个包含结构线较多的点云场景 (如书架、楼梯、百叶窗场景), 分别利用本文方法提取场景中的大量 3D 线特征, 并和其他方法进行比较. 图 5 分别给出了书架场景、楼梯场景和百叶窗场景 RGB-D 数据的 3D 线特征检测结果. 从中可以看到, 利用本文提出的 3D 线特征提取方法不仅能够提取大的场景框架结构 (如书架的框架、楼梯连续台阶和百叶窗的框架结构等), 而且能够提取小的场景细节结构 (如书架上摆放的小物体、楼梯的细栏杆和百叶窗的大量扇叶等), 各场景中的大量 3D 线特征都得到了有效的提取. 表 1 列出了本文方法对不同点云场景进行 3D 线特征提取的统计数据, 从中可以看到, 场景线特征数目大大少于点云数目, 同时由于省去了场景点云分割的时间消耗, 本文的 3D 线特征提取高效, 对书架场景、楼梯场景和百叶窗场景 RGB-D 数据的 3D 线特征提取分别约需 0.40, 0.37, 0.37 s, 其仅为 Lu 等^[15] 方法耗时的 1/10.

这里我们将本文提出的 3D 线特征提取方法与 Lu 等^[15] 的线特征提取方法进行比较. 对于普通扫描精度的室内场景点云数据来说, Lu 等^[15] 的方法往往由于场景点云法向的杂乱而导致直线拟合的失败, 从而使得场景中的线特征提取效果大大受限获取的场景点云数据的精度. 此外, Lu 等^[15] 的

图 5 (网络版彩图) 不同场景 3D 线特征提取

Figure 5 (Color online) 3D line extraction for different RGB-D scenes. (a) RGB images of different scenes; (b) depth images of corresponding scenes; (c) extraction results by using Lu et al.'s method^[15]; (d) our extraction results

表 1 不同场景的 3D 线特征提取统计

Table 1 Data statistics of 3D line extraction for different RGB-D scenes

RGB-D scenes	#Sampling points	Lu et al. ^[15]		Our method	
		#Extracted lines	Extraction time (s)	#Extracted lines	Extraction time (s)
Scene1 (bookshelf)	1207666	544	4.09	722	0.40
Scene2 (stairs)	1185903	369	3.91	1388	0.37
Scene3 (persiennes)	1309428	855	3.81	1049	0.37

方法提取效果还取决于场景面片分割的准确性, 对于复杂的普通精度室内场景点云数据, 往往会由于难以正确提取场景中的面片信息或忽略了部分小面片而造成场景线条提取的缺失, 同时复杂场景中的面片信息提取也是一个耗时的过程. 相反, 本文提出的 3D 线特征提取方法则兼顾了输入数据的 RGB 梯度信息和深度数据的几何信息, 其中 RGB 梯度信息提高了直线拟合区域的判断准确性, 而深度信息则避免了由于平面分割导致的时间消耗、点云精度不高造成的欠分割或过分割等引起的线特征提取错误, 对普通精度场景点云的大量线特征提取效果相对较好, 如图 5 所示.

为了验证场景 3D 线特征提取方法在弱光照条件下的适用性, 本文在扩展的 ICL-NUIM 合成数据集^[36]上进行了实验测试. 针对同一办公室室内场景, 图 6 分别给出了 4 种不同光照条件 — 普通照明 (static lighting)、全局照明 (global illumination)、局部照明 (local illumination)、手电筒照明 (flashlight) 下的场景 3D 线特征提取效果. 从中可以看到, 利用本文提出的方法在弱光照条件下仍具有较强的适用性. 表 2 给出了不同光照条件下场景 3D 线特征提取的数据统计, 虽然本文方法在不同光照条件下其提取到的线特征数目会由于场景明暗程度有所减少, 但本文方法提取的 3D 线特征基本能保持场景的框架完整性和局部细节, 而 Lu 等^[15]方法提取的场景线特征数目较少, 其在手电筒照明下由于光

图 6 (网络版彩图) 不同光照条件下场景 3D 线特征提取

Figure 6 (Color online) 3D line extraction for different lighting scenes, such as static lighting, global illumination, local illumination, and flashlight. (a) RGB images of different lighting scenes; (b) depth images of corresponding scenes; (c) extraction results by using Lu et al.'s method^[15]; (d) our extraction results of RGB-D scenes without depth noise; (e) our extraction results of RGB-D scenes with depth noise

表 2 不同光照条件下场景 3D 线特征提取统计

Table 2 Data statistics of 3D line extraction for different lighting scenes

Lighting scenes	#Sampling points	Lu et al. ^[15]		Our method	
		#Extracted lines	Extraction time (s)	#Extracted lines	Extraction time (s)
Static lighting	307200	201	1.49	874	0.39
Global illumination	307200	201	1.47	977	0.41
Local illumination	307200	201	1.41	757	0.38
Flashlight	307200	24	0.64	573	0.34

照过弱导致难以提取场景中的平面信息而使该方法失效。

为了验证 RGB-D 场景深度数据噪声误差对 3D 线特征提取效果的影响, 本文在扩展的 ICL-NUIM 合成数据集^[36]上进行了实验测试. 图 6(e) 分别给出了不同光照条件下具有深度数据噪声的场景 3D 线特征提取效果, 并和无深度数据噪声的场景线特征提取效果进行比较. 从中可以看到, 利用本文的 3D 线特征提取方法在带有深度噪声的场景点云中有一定的适用性, 但相较于光照影响, 场景深度误差带来的影响要更大一些. 本文方法在不同光照条件下对具有深度噪声的场景点云中提取的线特征数目分别为 745, 829, 626, 502, 其能基本保持场景线特征的完整性, 但部分细节会有少量扭曲和缺失.

5.2 点云场景 3D 线特征的配准

为了验证本文提出的场景 3D 线特征配准方法的有效性, 本文在 BundleFusion^[7]提供的室内场

图 7 (网络版彩图) 书桌场景的 3D 线特征配准

Figure 7 (Color online) Matching and registration of 3D lines for indoor desk scene. (a) RGB images of desk scene; (b) point clouds of desk scene; (c) matching of the extracted 3D lines; (d) our registration result of 3D lines

图 8 (网络版彩图) 房间角落场景的 3D 线特征配准

Figure 8 (Color online) Matching and registration of 3D lines for room corner scene. (a) RGB images of room corner scene; (b) point clouds of room corner scene; (c) matching of the extracted 3D lines; (d) our registration result of 3D lines

景 RGB-D 数据集中分别选取了 Office3 中书桌场景和房间角落场景, 并在同一场景中分别选择间隔 15 帧的具有位移偏差的两帧数据进行场景 3D 线特征配准实验. 图 7 展示了书桌场景 3D 线特征的配准结果, 从中可以看到, 书桌场景的 3D 线特征不仅包括如中间 2 条墙角线、柜门边框线、电脑边框线等场景框架结构, 还包括如墙上地图或其他图画的边框线, 利用本文提出的 3D 线特征配准方法能够有效配准这些线特征并能准确对齐两帧点云, 从而避免了场景内部的微小细节纹理 (如地图图案形成的短线、桌面杂乱小物体形成的短线等) 对场景配准带来的误差影响. 图 8 给出了房间角落场景 3D 线特征的配准结果, 从中可以看到, 场景中大量的屋顶、墙角、门等的 3D 特征线均能得到有效的匹配和配准, 其不仅包含场景框架结构特征线的配准而且含有场景的细节结构 3D 特征线的配准.

图 9 (网络版彩图) 点特征基础上点云场景不同配准方法的比较

Figure 9 (Color online) Comparisons of different point registration methods for point cloud scenes. (a) Different point cloud scenes; (b) matching results of 3D points extracted from different scenes; (c) registration results by using Point-to-Point^[32]; (d) registration results by using Point-to-Plane^[33]; (e) registration results by using symmetric ICP^[34]

表 3 不同场景的 3D 点配准统计^{a)}

Table 3 Data statistics of 3D point registration for different RGB-D scenes

Scenes	#Source points	#Target points	#Regist. points	Point-to-Point ^[32]			Point-to-Plane ^[33]			Symmetric ICP ^[34]		
				Iterat.	Regist.	Total	Iterat.	Regist.	Total	Iterat.	Regist.	Total
				times	time (s)	time (s)	times	time (s)	time (s)	times	time (s)	time (s)
Desk	406	327	406	17	0.027	77.667	11	0.215	77.326	11	0.223	77.403
Cabinet	530	517	530	23	0.035	65.479	22	0.263	65.058	4	0.218	65.200
Corner	732	796	732	34	0.048	42.569	31	0.278	43.099	81	0.315	43.043

a) Iterat.: Iteration; Regist.: Registration.

5.3 点云场景不同配准方法的比较

为了验证本文提出的 3D 线特征配准方法比传统点特征配准方法以及其他线特征配准方法更具鲁棒性和高效性, 本文在 BundleFusion^[7] 提供的室内场景 RGB-D 数据集 Office3 中选取书桌场景、柜子场景、房间角落场景并利用不同的配准方法 (如 Point-to-Point 方法^[32]、Point-to-Plane 方法^[33]、Symmetric ICP 方法^[34]) 进行场景 3D 点特征配准实验. 同时, 由于现有线特征配准方法大都将线配准问题转化为对应点配准进行处理, 故这里借助传统的点特征配准方法, 在场景 3D 线特征匹配的基础上使用参与拟合线特征的采样点作为点配准方法的输入, 并模拟线配准方法进行对比实验. 图 9 和 10 分别给出了不同场景的 3D 点配准和 3D 线配准的效果比较, 相应的统计数据见表 3

图 10 (网络版彩图) 线特征基础上点云场景不同配准方法的比较

Figure 10 (Color online) Comparisons of different line registration methods for point cloud scenes. (a) Matching results of 3D lines extracted from different scenes; (b) registration results by using Point-to-Point [32]; (c) registration results by using Point-to-Plane [33]; (d) registration results by using symmetric ICP [34]; (e) our registration results of 3D lines

表 4 不同场景的 3D 线配准统计 ^{a)}

Table 4 Data statistics of 3D line registration for different RGB-D scenes

Scenes	#Source#Target#Regist.			Point-to-Point [32]			Point-to-Plane [33]			Symmetric ICP [34]			Our method		
	lines	lines	lines	Iterat.	Regist.	Total	Iterat.	Regist.	Total	Iterat.	Regist.	Total	Iterat.	Regist.	Total
				times	time (s)	time (s)	times	time (s)	time (s)	times	time (s)	time (s)	times	time (s)	time (s)
Desk	154	161	35	26	6.226	6.859	17	4.067	4.700	10	3.425	4.058	6	1.379×10^{-3}	0.634
Cabinet	180	151	36	18	2.419	3.050	51	9.484	10.115	20	4.162	4.793	14	1.891×10^{-3}	0.632
Corner	155	193	61	11	0.866	1.493	14	1.481	2.108	13	1.381	2.008	13	1.873×10^{-3}	0.628

a) Iterat.: Iteration; Regist.: Registration.

和 4. 可以看出, 点云场景配准中基于线特征的配准比基于点的传统 ICP 配准方法更有效, 本文方法通过检测场景直线型边界线点和直线型折边线点从而能够提取场景中的大量框架结构线、纹理线条等; 而基于点的 ICP 配准中, 点特征的离散和有序性往往导致其配准结果易受特征点提取的误差影响 (如场景内部细小纹理、杂乱小物体等), 从而使得场景配准不准确. 如图 9 第 1 行书桌场景的点特征配准中利用 Point-to-Point 方法 [32] 与第 2 行书柜场景的点特征配准中利用 Point-to-Point 方法 [32] 和 Symmetric ICP 方法 [34] 得到的对齐结果存在扭曲或偏移, 这是由于点特征的离散性、无序性而将单帧场景点云的所有点特征及其重要性默认为相同, 而点配准方法仅考虑最小化点对点、点对面的误差度量, 书桌场景中桌面上的杂乱小物体、书柜场景中柜子上的点特征由于其误差度量不准确而干扰了整个场景的配准结果. 此外, 利用 Point-to-Plane 方法 [33] 和 Symmetric ICP 方法 [34] 时, 由于其缺少

点云场景结构信息而容易使配准迭代陷入局部最优,如图 9 中第 3 行房间角落场景利用 Symmetric ICP 方法^[34]和图 10 中第 2 行书柜场景利用 Point-to-Plane 方法^[33]和 Symmetric ICP 方法^[34]其配准效果偏移较明显.相反,在线特征基础上利用 Point-to-Point 方法的配准结果则明显优于在点特征基础上利用 Point-to-Point 方法的配准结果,如图 9(c)和 10(b)所示.

从表 3 和 4 可以看出,不同场景的线特征数目明显少于点特征数目,利用基于点特征的 ICP 配准策略比直接利用 3D 线特征进行配准更加耗时.另外,基于点特征的 ICP 配准效果取决于输入的特征点信息,虽然线特征数目少于点特征数目,但线特征包含了所有参与拟合的大量采样点信息;同时,由于场景线特征能体现出场景的大量结构信息,从而有效避免了配准过程中由于出现错位对齐或陷入局部最优而导致的配准错误.需要指出的是,表 3 和 4 中各种不同的点配准方法的迭代终止条件为:设置前后两次迭代的变换矩阵的最大容差为 $1\text{E}-10$,即前一个变换矩阵和当前变换矩阵的差异小于阈值 $1\text{E}-10$ 时认为方法已收敛;或前后两次迭代的欧氏距离均值的最大容差为 0.001,即迭代后点对的均方误差和小于阈值 0.001 时认为方法已收敛.以房间角落场景为例,Point-to-Point 方法^[32]需配准迭代 11 次耗时 0.866 s, Point-to-Plane 方法^[33]需配准迭代 14 次耗时 1.481 s, Symmetric ICP 方法^[34]则需配准迭代 13 次耗时 1.381 s,而直接基于线特征配准的本文方法配准迭代 13 次仅耗时 1.873×10^{-3} s,本文方法由于待配准的线特征数目少从而大大提高了场景配准的效率.

6 结束语

本文提出了一种融合场景数据的 RGB 梯度信息和邻域点云分布特性的 3D 线特征提取与配准方法.该方法由于结合了场景 RGB 梯度信息,在 3D 线特征提取中可以有效降低直线拟合对点云的精度要求;同时该方法直接利用点云分布特性提取场景中的主要结构线,从而避免了由于面片误分割(如欠分割或过分割)导致的提取误差.进一步,在线特征提取的基础上,本文方法利用 3D 线特征的线段支持域及线特征描述符实现了线特征匹配,以找到场景配准所需要的对应线对.最终,利用基于线-线的迭代最近线配准方法实现了场景点云数据的配准.实验结果表明,与已有的点云场景线特征提取及点特征匹配与配准方法比较,本文方法配准效率高、计算量小、配准精准,方法具有较强的鲁棒性.

然而,本文工作仅考虑了场景点云数据的配准问题,未来将考虑如何在 3D 线特征的基础上实现局部到全局的位姿优化策略,如何解决 3D 线特征加入到重建系统中面临的参数化问题等,从而实现基于 3D 线特征的场景点云高效重建.

参考文献

- 1 Miao Y W, Xiao C X. Geometric Processing and Shape Modeling of 3D Point-Sampled Models. Beijing: Science Press, 2014 [繆永伟,肖春霞.三维点采样模型的几何处理和形状造型.北京:科学出版社,2014]
- 2 Mur-Artal R, Tardös J D. ORB-SLAM2: an open-source SLAM system for monocular, stereo, and RGB-D cameras. IEEE Trans Robot, 2017, 33: 1255–1262
- 3 Newcombe R A, Izadi S, Hilliges O, et al. KinectFusion: real-time dense surface mapping and tracking. In: Proceedings of IEEE International Symposium on Mixed and Augmented Reality (ISMAR), Basel, 2011. 127–136
- 4 Whelan T, Johannsson H, Kaess M, et al. Robust real-time visual odometry for dense RGB-D mapping. In: Proceedings of IEEE International Conference on Robotics and Automation (ICRA), Karlsruhe, 2013. 5724–5731
- 5 Whelan T, Salas-Moreno R F, Glocker B, et al. ElasticFusion: real-time dense SLAM and light source estimation. Int J Robot Res, 2016, 35: 1697–1716
- 6 Nießner M, Zollhöfer M, Izadi S, et al. Real-time 3D reconstruction at scale using voxel hashing. ACM Trans Graph, 2013, 32: 1–11

- 7 Dai A, Nießner M, Zollhöfer M, et al. BundleFusion: real-time globally consistent 3D reconstruction using on-the-fly surface reintegration. *ACM Trans Graph*, 2017, 36: 1
- 8 Zhang G, Lee J H, Lim J, et al. Building a 3-D line-based map using stereo SLAM. *IEEE Trans Robot*, 2015, 31: 1364–1377
- 9 Xu C, Zhang L, Cheng L, et al. Pose estimation from line correspondences: a complete analysis and a series of solutions. *IEEE Trans Pattern Anal Mach Intell*, 2017, 39: 1209–1222
- 10 Zhou H, Zou D, Pei L, et al. StructSLAM: visual SLAM with building structure lines. *IEEE Trans Veh Technol*, 2015, 64: 1364–1375
- 11 Pumarola A, Vakhitov A, Agudo A, et al. PL-SLAM: real-time monocular visual SLAM with points and lines. In: *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, Singapore, 2017. 4503–4508
- 12 Li H, Yao J, Bazin J C, et al. A monocular SLAM system leveraging structural regularity in Manhattan world. In: *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, Brisbane, 2018. 2518–2525
- 13 Ni H, Lin X, Ning X, et al. Edge detection and feature line tracing in 3D-point clouds by analyzing geometric properties of neighborhoods. *Remote Sens*, 2016, 8: 710–718
- 14 Lin Y, Wang C, Chen B, et al. Facet segmentation-based line segment extraction for large-scale point clouds. *IEEE Trans Geosci Remote Sens*, 2017, 55: 4839–4854
- 15 Lu X, Liu Y, Li K. Fast 3D line segment detection from unorganized point cloud. 2019. ArXiv:1901.02532
- 16 Fischler M A, Bolles R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun ACM*, 1981, 24: 381–395
- 17 Kamgar-Parsi B. Algorithms for matching 3D line sets. *IEEE Trans Pattern Anal Machine Intell*, 2004, 26: 582–593
- 18 Alshawa M. ICL: iterative closest line—a novel point cloud registration algorithm based on linear features. In: *Proceedings of ISPRS 2nd Summer School in Ljubljana*, Ekscentar, 2007. 53–59
- 19 Hofer M, Maurer M, Bischof H. Efficient 3D scene abstraction using line segments. *Comput Vision Image Understanding*, 2017, 157: 167–178
- 20 Langlois P A, Boulch A, Marlet R. Surface reconstruction from 3D line segments. In: *Proceedings of IEEE International Conference on 3D Vision*, Quebec, 2019. 553–563
- 21 Brown M, Windridge D, Guillemaut J Y. A generalisable framework for saliency-based line segment detection. *Pattern Recogn*, 2015, 48: 3993–4011
- 22 Lin Y, Wang C, Cheng J, et al. Line segment extraction for large scale unorganized point clouds. *ISPRS J Photogrammetry Remote Sens*, 2015, 102: 172–183
- 23 Borges P, Zlot R, Bosse M, et al. Vision-based localization using an edge map extracted from 3D laser range data. In: *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, Anchorage, 2010. 4902–4909
- 24 Rosenfeld A, Thurston M. Edge and curve detection for visual scene analysis. *IEEE Trans Comput*, 1971, 100: 562–569
- 25 Vidal R, Ma Y, Sastry S. Generalized principal component analysis (GPCA). *IEEE Trans Pattern Anal Machine Intell*, 2005, 27: 1945–1959
- 26 Burns J B, Hanson A R, Riseman E M. Extracting straight lines. *IEEE Trans Pattern Anal Mach Intell*, 1986, 8: 425–455
- 27 Akinlar C, Topal C. EDLines: a real-time line segment detector with a false detection control. *Pattern Recogn Lett*, 2011, 32: 1633–1642
- 28 Lee J G, Han J, Whang K Y. Trajectory clustering: a partition-and-group framework. In: *Proceedings of the ACM SIGMOD International Conference on Management of Data*, Beijing, 2007. 593–604
- 29 Steder B, Rusu R B, Konolige K, et al. Point feature extraction on 3D range scans taking into account object boundaries. In: *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, Shanghai, 2011. 2601–2608
- 30 Zhang L, Koch R. Structure and motion from line correspondences: representation, projection, initialization and sparse bundle adjustment. *J Visual Commun Image Representation*, 2014, 25: 904–915
- 31 Scovanner P, Ali S, Shah M. A 3-dimensional SIFT descriptor and its application to action recognition. In: *Proceedings of the 15th ACM International Conference on Multimedia*, Augsburg, 2007. 357–360
- 32 Besl P J, McKay N D. Method for registration of 3-D shapes. In: *Proceedings of SPIE*, 1992. 1611: 586–606
- 33 Chen Y, Medioni G. Object modelling by registration of multiple range images. *Image Vision Computing*, 1992, 10:

145–155

- 34 Rusinkiewicz S. A symmetric objective function for ICP. *ACM Trans Graph*, 2019, 38: 1–7
- 35 Chang A, Dai A, Funkhouser T, et al. Matterport3D: learning from RGB-D data in indoor environments. In: *Proceedings of International Conference on 3D Vision*, Qingdao, 2017. 667–676
- 36 Park S, Schöps T, Pollefeys M. Illumination change robustness in direct visual SLAM. In: *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, Singapore, 2017. 4523–4530

Extraction and registration of 3D lines by fusing gradient information and neighboring point cloud distribution

Yongwei MIAO^{1,2*}, Yingting DAI¹, Haipeng WANG¹, Fuchang LIU² & Jinrong WANG²

1. *College of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China;*

2. *College of Information Science and Technology, Hangzhou Normal University, Hangzhou 311121, China*

* Corresponding author. E-mail: ywmiao@zstu.edu.cn

Abstract The traditional point-based scene reconstruction scheme always leads to inaccurate 3D reconstruction results due to its lack of scene texture and object occlusions. With the help of geometric characteristics and line features of 3D point cloud scenes, a novel extraction, matching, and registration method of 3D line features is presented by fusing RGB gradients and the spatial distribution of neighboring point clouds. Firstly, the supporting region of 3D linear segments with similar gradient directions can be determined by the gradients of an RGB scene image. Secondly, owing to the neighboring geometric features contained in the depth map, the point set of boundary lines and folded edge lines can also be extracted according to the spatial distribution of the current point and its neighboring points. The 3D line features can thus be fitted by these extracted point sets respectively. Thirdly, taking each extracted 3D line feature as a central axis, the supporting region can be constructed as the co-axial cylinder with the same thickness of each cylindrical layer. According to the 3D gradient statistics of point cloud within each cylindrical layer, we can calculate an intrinsic descriptor for each 3D line, which can be adopted to match 3D line features between scene frames. Finally, a novel line-to-line iterative closest line registration algorithm is introduced to compute the rotation and translation transformations, and the inter-frame point cloud scene data can be registered efficiently. The experimental results illustrate that the proposed line feature-based scheme has high registration efficiency, less computational cost, accurate registration, and strong robustness when compared with the existing line feature extraction and point-based registration methods.

Keywords point cloud scenes, 3D lines, line feature extraction, line registration, 3D reconstruction

Yongwei MIAO was born in 1971. He received his Ph.D. degree in computer graphics from the State Key Laboratory of CAD and CG, Zhejiang University in 2007. He is a professor at Zhejiang Sci-Tech University, and also a visiting professor of Hangzhou Normal University. His research interests include computer graphics, digital geometry processing, 3D computer vision, machine learning and deep learning.

Yingting DAI was born in 1991. She is a master candidate at Zhejiang Sci-Tech University. Her research interests include 3D computer vision and SLAM.

Fuchang LIU was born in 1982. He received his Ph.D. degree in computer science from Nanjing University of Science and Technology in 2009. He is an associate professor at Hangzhou Normal University. His research interests include computer graphics, scene understanding, and 3D computer vision.

Jinrong WANG was born in 1973. He received his Ph.D. degree in computer graphics from the State Key Laboratory of CAD and CG, Zhejiang University in 2012. He is an assistant professor at Hangzhou Normal University. His research interests include computer graphics, digital geometry processing, and 3D computer vision.