

# PLP-SLAM: 基于点、线、面特征融合的视觉 SLAM 方法

李海丰<sup>1,2</sup>, 胡遵河<sup>1</sup>, 陈新伟<sup>2</sup>

(1. 中国民航大学计算机科学与技术学院, 天津 300300;

2. 福建省信息处理与智能控制重点实验室(闽江学院), 福建 福州 350121)

**摘要:** 基于点特征的视觉 SLAM (同时定位与地图构建) 算法存在计算量大、环境存储空间负荷高、定位误差较大的问题, 为此, 提出了一种基于点、线段、平面特征融合的视觉 SLAM 算法——PLP-SLAM. 在扩展卡尔曼滤波(EKF)框架下, 首先利用点特征估计机器人当前位姿, 然后构建了基于点、线、平面特征的观测模型, 最后建立了带平面约束的线段特征数据关联方法及系统状态更新模型, 并利用线段和平面特征描述环境信息. 在公开数据集上进行了实验, 结果表明, 本文 PLP-SLAM 算法能够成功完成 SLAM 任务, 平均定位误差为 2.3 m, 优于基于点特征的 SLAM 方法, 并通过基于不同特征的 SLAM 实验表明了本文提出的点、线、面特征融合的优越性.

**关键词:** 同时定位与地图构建; 点线面特征融合; 扩展卡尔曼滤波; 线段特征; 平面特征

**中图分类号:** TP242

**文献标识码:** A

**文章编号:** 1002-0446(2017)-02-0214-07

## PLP-SLAM: A Visual SLAM Method Based on Point-Line-Plane Feature Fusion

LI Haifeng<sup>1,2</sup>, HU Zunhe<sup>1</sup>, CHEN Xinwei<sup>2</sup>

(1. College of Computer Science and Technology, Civil Aviation University of China, Tianjin 300300, China;

2. Fujian Provincial Key Laboratory of Information Processing and Intelligent Control (Minjiang University), Fuzhou 350121, China)

**Abstract:** To reduce the computation and memory cost, and improve the localization accuracy of point-based visual simultaneous localization and mapping (SLAM) methods, a novel visual SLAM algorithm, named PLP-SLAM, is proposed by fusing the point, line segment and plane features. Under the EKF (extended Kalman filter) framework, the current robot pose is estimated using the point features firstly. Then, the observation models of point, line segment and plane are set up. Finally, the data association of line segments and state-updating models with plane constraints are formulated, and the unknown environment is described by line segment and plane feature map. The experiments are carried out on a public dataset, and the results show that the proposed PLP-SLAM can accomplish the SLAM task with a mean localization error of 2.3 m, which is better than the point-based SLAM method. Furthermore, the SLAM experiment results with different feature combinations illustrate the superiority of the proposed point-line-plane feature fusion.

**Keywords:** simultaneous localization and mapping; point-line-plane feature fusion; extended Kalman filter; line segment feature; plane feature

## 1 引言 (Introduction)

SLAM 是指利用机器人自身传感器在未知环境中同时实现自身定位与环境地图构建, 其主要思想是: 一方面, 依靠已创建的地图信息进行自定位; 另一方面, 根据定位结果对地图进行更新. SLAM 技术是移动机器人自主导航的核心和关键, 是机器人自主导航领域研究的热点问题<sup>[1]</sup>. 随着计算机视觉设备及方法的快速发展, 近年来视觉 SLAM 问题引起了学者的广泛关注<sup>[2-3]</sup>.

依据所采用方法的不同, SLAM 可分为基于概率框架的方法和基于图优化的方法. 基于概率框架的方法出现较早, 其典型代表是基于卡尔曼滤波器的 SLAM 算法<sup>[4-5]</sup>, 此类方法实现简单, 但存在计算量较大和误差累积的问题, 因而主要用于中小场景的 SLAM 应用, 近年出现的基于随机有限集的方法<sup>[6-7]</sup>非常具有代表性. 基于图优化的方法近年来成为研究热点, 最具代表性的算法包括 PTAM-SLAM (parallel tracking and mapping-

基金项目: 国家自然科学基金 (61305107, U1333109); 天津市应用基础与前沿技术研究计划重点项目 (14JCZDJC32500); 中央高校基本科研业务费 (3122016B006); 福建省信息处理与智能控制重点实验室开放课题 (MJUKF201732); 福建省科技厅引导性课题 (2015H0031).

通信作者: 陈新伟, chenxw\_mju@126.com 收稿/录用/修回: 2016-12-05/2017-02-09/2017-02-18

SLAM) [8]、ORB-SLAM (oriented brief-SLAM) [9] 等, 在大型场景中获得了较好效果. 为了减小累积误差, 闭环检测 (loop closure detection) 已成为视觉 SLAM 研究的热点问题 [10]. 此外, 基于深度相机的视觉 SLAM 算法 [11] 也成为近年来 SLAM 研究的重点之一.

依据所采用特征的不同, 可对视觉 SLAM 算法重新进行分类. 点特征 [12] 在视觉 SLAM 中被广泛使用 [4-5,13]. 点特征匹配算法成熟, 适用面广. 基于点特征的 SLAM 算法仍然是目前的主流研究方向. 然而, 基于点特征的 SLAM 算法计算量较大, 以 3 维点云作为环境地图描述, 精度不高且存储地图所需空间较大. 此外, 基于点特征构建的环境地图不利于后续的场景理解及人机交互等应用, 而且点特征更容易受到光线变化及阴影等环境因素的影响. 在结构化的人造环境中, 线段特征丰富. 与点特征相比, 线段特征同样易于提取 [14] 但受环境影响更小, 线段特征不仅能够更好地刻画环境的结构信息, 建立更高层次的环境地图, 而且能够有效减少特征数量, 降低计算复杂度. 因此, 基于线特征的 SLAM 算法日益受到关注 [15-18]. 图像中线段端点的不确定性以及线段匹配问题往往会给基于线特征的 SLAM 算法带来挑战. 平面特征也被一些学者用于 SLAM 算法中 [19-20]. 上述研究均表明, 线段、平面等高层次特征在 SLAM 应用中越来越受到重视. 总的来说, 点、线段、平面等特征在视觉 SLAM 中各有特点, 点特征匹配算法稳定, 适用面广. 而在诸如室内或者城市等人造环境中, 线段和平面特征能以相对较小的时间和空间复杂度完成 SLAM 任务. 然而, 目前基于线特征和平面的 SLAM 算法大都将特征单独看待, 没有深入挖掘特征之间的关联和约束关系. 文 [21] 提出基于点线相合的 SLAM 算法, 使用线段特征描述环境信息, 将点线相合的增量式地图构建方法引入粒子滤波中. 文 [22] 将线段、直线、消失点等特征融合用于图像解析, 然而该方法还未被用于视觉 SLAM 工作中.

针对特征提取与描述子计算耗时等问题, 近年来出现了直接法 SLAM 方案. 随着 SVO (semi-direct monocular visual odometry) [23]、LSD-SLAM (large-scale direct SLAM) [24] 等直接法 SLAM 方案的流行, 直接法本身也受到越来越多的关注. 与基于特征的 SLAM 方法相比, 直接法可以省去计算特征和描述子的时间, 只要求有像素梯度即可, 无须特征点. 因此, 直接法可以在特征缺失的场合下使

用. 但因其完全依靠梯度搜索, 降低目标函数来计算相机位姿, 所以只适用于运动很小的应用中.

本文针对人造环境这类特定的应用场景, 提出了一种基于点、线段、平面特征融合的视觉 SLAM 算法——PLP-SLAM, 这是一种基于特征的 SLAM 方法. 本文的算法将点、线段、平面特征充分融合, 点特征用于对机器人当前位姿进行初始估计, 线段和平面特征用来描述环境信息, 同时平面特征可为线段特征提供共面的几何约束. 与基于点特征的 SLAM 方法相比, PLP-SLAM 计算量小, 机器人的定位精度较高, 环境信息更加丰富且存储所需空间更小.

## 2 SLAM 问题定义 (Definition of SLAM problem)

基于点、线、面特征融合的移动机器人单目视觉 SLAM 问题定义包括坐标系定义、状态向量、运动模型和观测模型.

### 2.1 坐标系定义

定义  $\{W\}$  为 3 维直角世界坐标系,  $\{C_k\}$  和  $\{I_k\}$  分别为  $k$  时刻的摄像机坐标系和图像坐标系,  $\{C_0\}$  与  $\{W\}$  重合.

### 2.2 状态向量

系统状态被定义为

$$\mathbf{X}(k) = \begin{bmatrix} \mathbf{X}_r(k) \\ \mathbf{M}(k) \end{bmatrix} \quad (1)$$

其中,  $\mathbf{X}_r(k) = [x_r(k), y_r(k), \theta_r(k)]^T$  为机器人在  $\{W\}$  中的位姿,  $\mathbf{M}(k) = [\boldsymbol{\pi}_i^w, \dots, \boldsymbol{l}_j^w, \dots]^T$ ,  $\boldsymbol{\pi}_i^w$  为  $\{W\}$  中的第  $i$  个平面特征,  $\boldsymbol{l}_j^w$  为  $\{W\}$  中位于平面  $\boldsymbol{\pi}_i^w$  上的第  $j$  个线特征.

### 2.3 运动状态模型

假设机器人在  $k$  时刻的控制输入为线速度  $v(k)$  和角速度  $\omega(k)$ ,  $\Delta T$  为相邻关键帧的时间间隔, 则机器人的运动状态方程为

$$\mathbf{X}_r(k+1) = \begin{bmatrix} x_r(k+1) \\ y_r(k+1) \\ \theta_r(k+1) \end{bmatrix} = \begin{bmatrix} x_r(k) + \Delta T v(k) \cos \theta_r(k) \\ y_r(k) + \Delta T v(k) \sin \theta_r(k) \\ \theta_r(k) + \Delta T \omega(k) \end{bmatrix} \quad (2)$$

假设环境路标均静止, 则环境路标的状态方程为

$$\mathbf{M}(k+1) = \mathbf{M}(k) \quad (3)$$

综上, 系统的状态模型为

$$\begin{aligned} X(k+1) &= f(X(k), v(k), \omega(k)) + q(k) \\ &= \begin{bmatrix} x_r(k) + \Delta T v(k) \cos \theta_r(k) \\ y_r(k) + \Delta T v(k) \sin \theta_r(k) \\ \theta_r(k) + \Delta T \omega(k) \\ M(k) \end{bmatrix} + q(k) \quad (4) \end{aligned}$$

其中,  $q(k)$  为均值为 0、协方差为  $Q(k)$  的高斯白噪声.

## 2.4 观测模型

### 2.4.1 空间点特征重建与机器人位姿估计

基本矩阵 (fundamental matrix) 包含了摄像机的运动信息, 可通过分解基本矩阵来获得机器人的位姿. 但视图之间匹配的图像线特征对基本矩阵完全没有约束, 即不能仅通过匹配的线特征计算得到基本矩阵. 因此, 本文采用基于点特征的方法估计基本矩阵, 进而获得机器人位姿.

基于匹配的点特征, 通过最小化式 (5) 中的重投影几何误差, 优化估计  $\{C_k\}$  中的空间点特征以及机器人在  $\{C_k\}$  中的位姿.

$$\begin{aligned} \arg \min_{\hat{F}_k, \hat{s}_i^k, \hat{s}_i^{k+1}} & \sum_i (d(s_i^k, \hat{s}_i^k) + d(s_i^{k+1}, \hat{s}_i^{k+1})) \quad (5) \\ \text{s.t. } & (\hat{s}_i^{k+1})^T \hat{F}_k \hat{s}_i^k = 0, \\ & \hat{s}_i^k = K[I|0] \hat{P}_i^k, \hat{s}_i^{k+1} = K[\hat{r}_k | \hat{t}_k] \hat{P}_i^k \end{aligned}$$

其中,  $d(\cdot)$  表示几何距离,  $s_i^k$  为  $\{I_k\}$  中一个特征点的齐次坐标,  $s_i^{k+1}$  为  $s_i^k$  在  $\{I_{k+1}\}$  中的匹配点,  $\hat{s}_i^k$  和  $\hat{s}_i^{k+1}$  分别为  $s_i^k$  和  $s_i^{k+1}$  的估计值,  $\hat{P}_i^k$  为在  $\{C_k\}$  中第  $i$  个特征点的估计值,  $K$  为摄像机内参数矩阵,  $I$  和  $0$  分别为单位矩阵和零向量,  $\hat{r}_k$  和  $\hat{t}_k$  分别表示  $k$  时刻摄像机旋转矩阵和平移向量的估计值,  $\hat{F}_k$  为  $k$  时刻基本矩阵的估计值, 且

$$\hat{F}_k = K^{-T} [\hat{t}_k]_{\times} \hat{r}_k K^{-1} \quad (6)$$

其中,  $[\cdot]_{\times}$  为叉乘的斜对称矩阵表示.

### 2.4.2 线特征 3 维重建

定义  $\mathbf{l}_j^k$  为  $\{I_k\}$  中第  $j$  条直线的齐次坐标,  $\mathbf{l}_j^k$  和  $\mathbf{l}_j^{k+1}$  为一对匹配的线特征, 则  $\{C_{k+1}\}$  中的线特征  $\mathbf{l}_j^{k+1}$  可表示为

$$\mathbf{l}_j^{k+1} \triangleq \begin{bmatrix} (\boldsymbol{\pi}_j^k)^T \\ (\boldsymbol{\pi}_j^{k+1})^T \end{bmatrix} \quad (7)$$

其中,  $\boldsymbol{\pi}_j^k$  和  $\boldsymbol{\pi}_j^{k+1}$  为线特征  $\mathbf{l}_j^k$  和  $\mathbf{l}_j^{k+1}$  的反向投影平面, 且

$$\boldsymbol{\pi}_j^k = (K[I|0])^T \mathbf{l}_j^k, \boldsymbol{\pi}_j^{k+1} = (K[r_k | t_k])^T \mathbf{l}_j^{k+1}$$

### 2.4.3 观测模型

观测模型将系统状态映射到环境路标在摄像机坐标系中的观测值.

定义摄像机坐标系  $\{C_k\}$  与世界坐标系  $\{W\}$  之间的齐次变换为  $\mathbf{t}_k^W$ , 且

$$\mathbf{t}_k^W = \begin{bmatrix} \mathbf{r}_k & -\mathbf{r}_k \tilde{\mathbf{C}}_k \\ \mathbf{0}^T & 1 \end{bmatrix}_{4 \times 4} \quad (8)$$

其中,  $\tilde{\mathbf{C}}_k$  表示  $k$  时刻摄像机中心在  $\{W\}$  中的非齐次坐标, 即  $\tilde{\mathbf{C}}_k = [x_r(k), y_r(k), 0]^T$ ,  $\mathbf{r}_k$  为一个  $3 \times 3$  的旋转矩阵, 表示  $\{C_k\}$  在  $\{W\}$  中的姿态.

定义点特征  $\mathbf{P}_i^W$  在  $\{C_k\}$  中的观测值为  $\mathbf{P}_i^k$ , 则  $\mathbf{P}_i^k$  的齐次坐标为

$$\mathbf{P}_i^k = \mathbf{t}_k^W \mathbf{P}_i^W \quad (9)$$

假设在  $\{W\}$  中,  $\mathbf{P}_a^W$  和  $\mathbf{P}_b^W$  是位于线特征  $\mathbf{l}_j^W$  上的 2 个点的齐次坐标, 则  $\mathbf{l}_j^W$  可表示为

$$\mathbf{l}_j^W = \begin{bmatrix} (\mathbf{P}_a^W)^T \\ (\mathbf{P}_b^W)^T \end{bmatrix}_{2 \times 4} \quad (10)$$

$\mathbf{l}_j^W$  在摄像机坐标系  $\{C_k\}$  下的观测值  $\mathbf{l}_j^k$  为

$$\begin{aligned} \mathbf{l}_j^k &= \begin{bmatrix} (\mathbf{P}_a^k)^T \\ (\mathbf{P}_b^k)^T \end{bmatrix} = \begin{bmatrix} (\mathbf{t}_k^W \mathbf{P}_a^W)^T \\ (\mathbf{t}_k^W \mathbf{P}_b^W)^T \end{bmatrix} \\ &= \begin{bmatrix} (\mathbf{P}_a^W)^T \\ (\mathbf{P}_b^W)^T \end{bmatrix} (\mathbf{t}_k^W)^T = \mathbf{l}_j^W (\mathbf{t}_k^W)^T \quad (11) \end{aligned}$$

定义平面特征  $\boldsymbol{\pi}_i^W$  在  $\{C_k\}$  中的观测值为  $\boldsymbol{\pi}_i^k$ , 则:

$$\boldsymbol{\pi}_i^k = (\mathbf{t}_k^W)^T \boldsymbol{\pi}_i^W \quad (12)$$

由于  $\mathbf{l}_j^W$  位于  $\boldsymbol{\pi}_i^W$  上, 且  $\{C_0\}$  与  $\{W\}$  重合, 则存在下列约束:

$$(K[r_k | t_k])^{-T} \mathbf{l}_j^k = \lambda (\mathbf{H}_i^k)^{-T} \mathbf{l}_j^0 \quad (13)$$

其中  $\mathbf{H}_i^k$  为由平面  $\boldsymbol{\pi}_i^k$  诱导的单应矩阵,  $\lambda$  为尺度系数.

综上, 系统的观测模型为

$$\mathbf{Z}(k) = \mathbf{H}(\mathbf{X}(k)) + \mathbf{r}(k) \quad (14)$$

其中,  $\mathbf{H}(\mathbf{X}(k)) = [\mathbf{l}_j^W (\mathbf{t}_k^W)^T, (\mathbf{t}_k^W)^T \boldsymbol{\pi}_i^W]^T$ ,  $\mathbf{r}(k)$  为均值为 0、协方差为  $\mathbf{r}(k)$  的高斯白噪声.

### 3 SLAM 流程 (Process of SLAM)

#### 3.1 特征提取

每帧图像中的点特征利用 SIFT (比例不变特征变换) 算法<sup>[4]</sup>提取, 线段特征使用 LSD (line segment detector) 算法<sup>[8]</sup>提取. 相邻关键帧中的点特征通过 SIFT 算法匹配, 线段特征通过文 [25] 中算法进行匹配. 平面特征只存在于 3 维空间, 通过相邻关键帧中共面的线特征估计得到. 由于位于同一平面上的线特征满足同一个单应矩阵模型, 所以可在 RANSAC (随机抽样一致性) 算法框架下根据不同平面的单应矩阵模型, 将所有线特征按共面关系分组, 然后再利用分组后共面的线特征对所处平面进行估计.

#### 3.2 初始化

(1) 在未知环境中, 以世界坐标系原点为初始位置, 即  $\{C_0\}$  与  $\{W\}$  重合.

(2) 当  $k=1$  时, 可用式 (5) 和 (7) 中的方法进行空间点、线、平面特征的 3 维重建以及机器人的位姿估计. 此时, 机器人第 1 步的平移量需要通过其他传感器测量获得, 以便得到后续位姿估计时的绝对平移量. 至此, 获得了  $X_r(1)$  及  $M(1)$ , 即获得了系统的初始状态  $X(1)$ .

(3) 以  $X(1)$  为初始状态, 进入预测和更新迭代环节, 新的环境特征被加入. 此时, 每一步获得的点、线、平面特征都是基于摄像机坐标系的, 需将其变换到世界坐标系下.

#### 3.3 数据关联

数据关联环节需要对观测特征和状态特征进行匹配. 定义  $m(k+1)$  为  $k+1$  时刻观测到的环境特征, 而  $M(k)$  是  $k$  时刻已存在的环境特征. 数据关联就是要计算  $M(k)$  和  $m(k+1)$  中共同的环境特征.

进行线特征数据关联的代价函数为

$$C(l_j^{k+1}, l_i^W) \\ l_j^{k+1} \sqsubset m(k+1), l_i^W \sqsubset M(k) \\ = \sum_{t=0,1} (d_{\perp}(E_t(T(l_j^{k+1})), l_i^W) + d_{\perp}(E_t(l_i^W), T(l_j^{k+1})))$$

其中, 符号 “ $\sqsubset$ ” 表示向量为矩阵的行向量,  $T(\cdot)$  表示摄像机坐标系到世界坐标系的变换函数,  $d_{\perp}(\cdot)$  表示点到直线的距离,  $E_t(\cdot)$  ( $t=0,1$ ) 表示 3 维空间中一条线段的 2 个端点.

如果满足

$$C(l_j^{k+1}, l_i^W) < \delta \quad (15)$$

且  $C(l_j^{k+1}, l_i^W), \forall j$  的值最小, 则认为 2 个线特征关

联. 值得注意的是, 以下 3 种情况都满足上述要求: 1) 2 条线段基本重合; 2) 2 条线段为来自同一空间特征的共线且有重叠部分的线段; 3) 2 条线段为来自不同空间特征的共线线段. 其中, 前 2 种情形为“正确关联”, 而 3) 为“错误关联”. 第 3) 种情形可以通过下面方法去掉: 将一条线段向另一条线段所在直线投影, 如果线段投影与另一条线段不存在重合部分, 则认为这 2 条线段不关联. 平面特征的关联可依据位于平面上的线特征的关联关系确定.

#### 3.4 预测

根据式 (4) 中的状态模型, 预测机器人下一时刻的位姿和环境地图特征:

$$\hat{X}(k+1|k) = f(X(k|k), v(k), \omega(k)) \quad (16)$$

系统状态的协方差矩阵的预测:

$$P(k+1|k) = F(k)P(k|k)F^T(k) + Q(k+1) \quad (17)$$

其中,  $F(k) = \frac{\partial f}{\partial X} \Big|_{\hat{X}(k|k), v(k), \omega(k)}$ .

#### 3.5 更新

首先计算卡尔曼增益:

$$K(k+1) = P(k+1|k)H^T(k+1)s^{-1}(k+1) \quad (18)$$

其中,

$$H(k+1) = \frac{\partial h}{\partial X} \Big|_{\hat{X}(k+1|k)}$$

$$s(k+1) = H(k+1)P(k+1|k)H^T(k+1) + r(k+1)$$

更新系统状态估计:

$$\hat{X}(k+1|k+1) = \hat{X}(k+1|k) + K(k+1) \cdot (Z(k+1) - H(\hat{X}(k+1|k))) \quad (19)$$

更新状态估计的协方差:

$$P(k+1|k+1) = (I - K(k+1)H(k+1))P(k+1|k) \quad (20)$$

其中,  $I$  表示单位矩阵.

为了去除虚假观测特征, 先将观测到的新状态特征放入临时存储区, 如果在后续 2 幅关键帧中都没有新的观测与其关联, 则认为该特征为假特征, 将其从临时存储区中删除. 否则将其转换到世界坐标系下作为新的状态特征加入状态向量, 同时对各个相关矩阵进行扩维、更新.

### 4 实验结果 (Experiment result)

本文算法已利用 C++ 和 OpenCV 实现, 图形用户界面采用 Qt 实现. 实验的硬件平台为: ThinkPad M8600 T 系列台式机, CPU 为 Intel i7 双

核 3.4 GHz, 4 GB 内存, 1TB 硬盘, AMD 独立显卡. 实验数据采用 HRBB4 数据集<sup>[26]</sup>. HRBB4 数据集包含通过车载单目相机采集的办公走廊环境的共计 12 000 幅图像, 图像原始分辨率为 1920×1080, 为了降低计算量, 本实验中将所有图像分辨率降至 640×360. 采集图像时, 机器人的运动轨迹长度约 70 m. 在 HRBB4 数据集中, 包含了通过人工路标计算出来的摄像机运动轨迹的真实值. 图 1 所示为 HRBB4 中一些典型的图像帧.

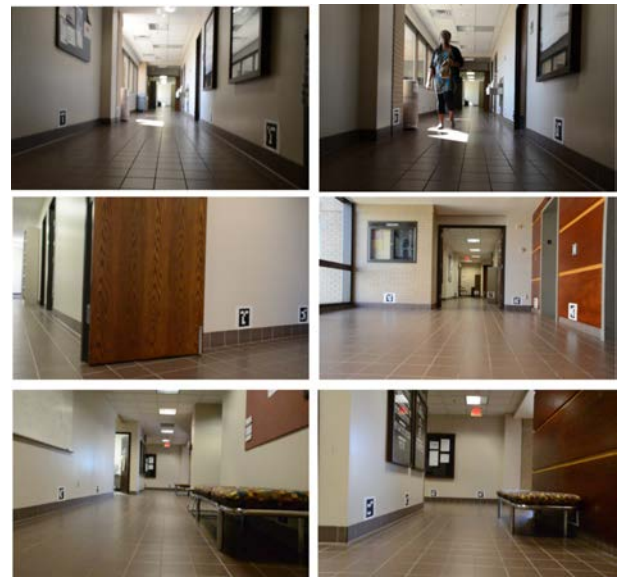


图 1 HRBB4 中的典型图像帧示例  
Fig.1 Some typical frames in HRBB4 dataset

实验中, 首先需要从数据集中提取关键帧, 同时基于相邻图像帧中特征点的数量变化检测抖动帧并去除, 即如果发现相邻帧中的 SIFT 点特征数量变化大于某一设定阈值, 则认为发生了抖动, 去除当前帧.

4.1 SLAM 结果

本文的算法成功完成了基于 HRBB4 数据集的 SLAM 任务. 算法生成的机器人轨迹以及 3 维环境地图信息如图 2 所示, 其中的虚线为估计的机器人运动轨迹, 环境地图中只显示了线段特征, 平面特征以对线特征提供共面约束形式体现, 即共面的线段以相同的颜色表示.

PLP-SLAM 算法提取的关键帧数、生成的环境地图中各种几何特征数量、算法所需时间等统计结果如表 1 所示. 可见, 在特征数量方面, PLP-SLAM 生成的环境地图中线段和平面特征与常见的 SLAM 算法生成的点云相比, 数量大幅减少. 在实时性方面, HRBB4 数据集一共包含 12 000 帧图像, 虽然只提取其中的 170 帧执行 SLAM 算法, 但在关

键帧提取、特征点提取方面需要对数据集中每一帧图像都进行操作, 因此平均每帧图像所需处理时间约为 398 ms. 与 ORB-SLAM 等方法相比, 本文耗时较长, 主要原因包括: 特征提取和匹配本身比较耗时、未采用多线程的处理方式、利用 Qt 进行 3 维环境地图实时显示较耗时等. 上述原因均是可提高 PLP-SLAM 算法实时性的努力方向.

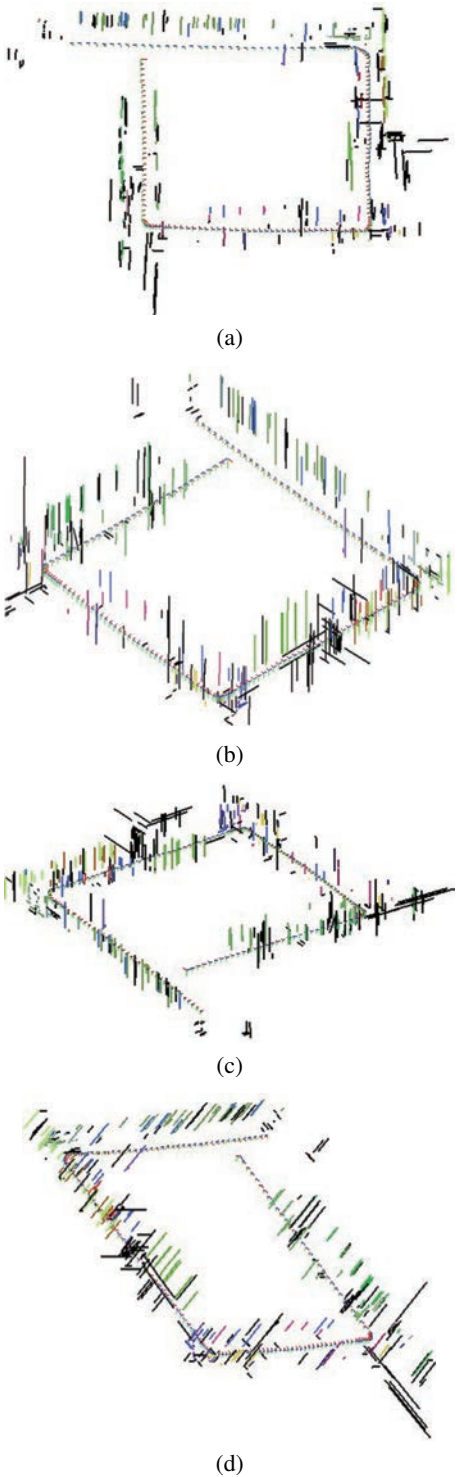


图 2 不同视角下的 SLAM 结果  
Fig.2 SLAM results from different views

表 1 SLAM 结果统计  
Tab.1 SLAM results

关键帧数	线特征数量	平面数量	路径 /m	算法所需时间 /s
170	2170	21	70.2	4773.7

4.2 定位精度对比实验结果

为了说明本文在特征选取及融合方面的贡献, 首先将本文 PLP-SLAM 算法与 1-point-EKF SLAM 算法<sup>[6]</sup>进行实验对比. 1-point-EKF 是 EKF 框架下利用点特征进行 SLAM 的经典算法. 1-point-EKF 的源代码可以从网上获得. 此外, 进一步将 PLP-SLAM 算法与目前较为流行的 ORB-SLAM 算法<sup>[9]</sup>进行了实验对比.

机器人运动轨迹估计结果如图 3 所示. 可以看出, 与 1-point-EKF 相比, PLP-SLAM 的定位精度更高. 由于 HRBB4 数据集中每幅图像所包含的点特征较少, 而且运动过程中有约 90° 的急转弯, 容易在 SLAM 过程中引发尺度漂移. 从图 3 中可见, 1-point-EKF 算法中, 估计的机器人轨迹在转弯处发生了更明显的漂移. 由于 PLP-SLAM 与 1-point-EKF 均采用 EKF 框架, 该实验结果可在一定程度上说明本文采用点—线段—平面多几何特征融合的优势. 与 ORB-SLAM 相比, 本文 PLP-SLAM 的估计精度稍差, 主要与本文未进行闭环检测和大规模优化等相关. 但本文算法在环境构图方面采用的是以线段和平面作为环境地图, 不同于 ORB-SLAM 的点云, 是一种环境构图的新思路, 除特征数量和所需存储空间更小外, 线段和平面所包含的几何信息也更加丰富, 更适合描述人造环境的 3 维几何地图特征.

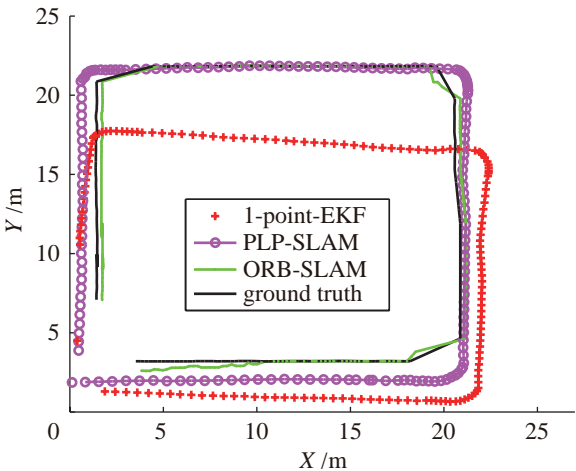


图 3 轨迹估计结果  
Fig.3 Trajectory estimation results

为了评估、比较算法定位的准确性, 本文使用

绝对轨迹误差 (absolute trajectory error, ATE) 作为评价指标.  $k$  时刻的 ATE 的定义为

$$\epsilon_k = \|e_k - g_k\| \tag{21}$$

其中,  $e_k$  和  $g_k$  分别表示机器人位置的估计值和真实值.

在不同的关键帧时刻, 机器人的定位误差如图 4 所示. 该定位误差也表明本文提出的 PLP-SLAM 的定位精度要优于 1-point-EKF 算法.

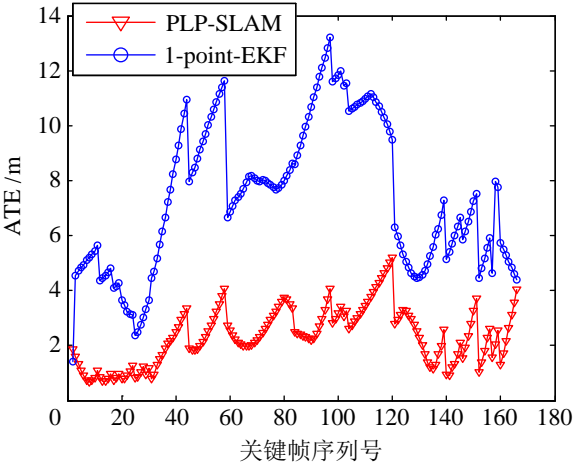


图 4 不同关键帧时刻机器人的定位误差  
Fig.4 Localization errors at different key frames

4.3 特征的贡献分析

由于缺少环境特征的真实值, 所以在 SLAM 结果中很难直接分析环境地图的构建精度. 本文通过比较采用不同特征时机器人的轨迹估计误差, 来说明所采用特征的贡献, 从而证明 PLP-SLAM 中采用点、线段、平面特征的必要性.

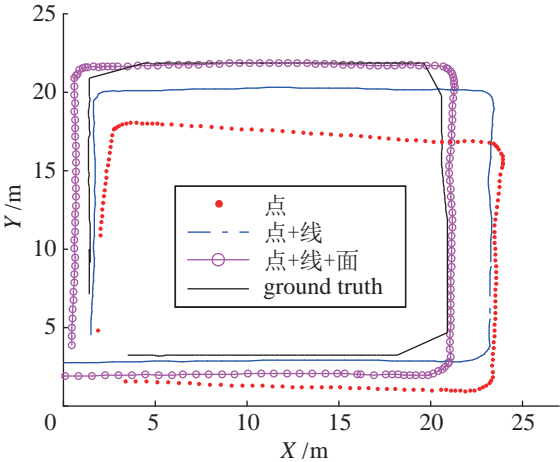


图 5 不同特征组合下运动轨迹估计结果对比  
Fig.5 Trajectory estimation results of SLAM with different feature combinations

不同特征组合下机器人轨迹估计结果如图 5 所示, 平均估计误差如表 2 所示. 特征组合为“点”,



表示在 SLAM 过程中只采用点特征, 特征组合为“点 + 线段 + 平面”表示采用 3 种特征融合, 即上述 PLP-SLAM 算法, 特征组合为“点 + 线段”是指在本文 PLP-SLAM 算法中去除对于特征共面约束的考虑, 即采用检测到的全部线特征完成 SLAM 任务. 从表 2 中的 MATE (mean absolute trajectory error) 对比可以看出, 定位误差由高到低分别是: 点、点 + 线段、点 + 线段 + 平面. 由此可以说明本文采用点、线、面特征组合的必要性.

表 2 采用不同特征组合的定位误差

Tab.2 ATE in SLAM with different feature combinations

特征组合	MATE /m
点	7.4
点 + 线段	4.5
点 + 线段 + 平面	2.3

## 5 结论 (Conclusion)

本文提出了一种基于点、线、面特征融合的视觉 SLAM 算法, 以共面的线段特征作为环境地图的描述, 通过在扩展卡尔曼滤波框架下推导带平面几何约束的点线特征观测模型、数据关联算法、状态更新模型等, 完成了机器人的位姿估计与环境地图构建. 利用公开数据集进行了实验, 结果表明, 本文提出的 PLP-SLAM 算法相比于 1-point-EKF SLAM 算法, 定位精度更高, 所需存储空间更小. 同时通过实验数据证明了特征融合的必要性的.

本文算法的实时性不高, 未来考虑多线程实现以及利用 GPU (图形处理器) 等对算法进行加速. 闭环检测也是急需完成的工作. 此外, 计划借助深度摄像机和惯性测量单元等传感器提高视觉 SLAM 的鲁棒性.

感谢美国得州农工大学 Netbot 实验室提供的实验数据集及部分源程序框架.

## 参考文献 (References)

- [1] Fuentes-Pacheco J, Ruiz-Ascencio J, Manuel Rendon-Mancha J. Visual simultaneous localization and mapping: A survey[J]. Artificial Intelligence Review, 2012, 43(1): 55-81.
- [2] 梁明杰, 闵华清, 罗荣华. 基于图优化的同时定位与地图创建综述 [J]. 机器人, 2013, 35(4): 500-512.  
Liang M J, Min H Q, Luo R H. Graph-based SLAM: A survey [J]. Robot, 2013, 35(4): 500-512.
- [3] 孙凤池, 黄亚楼, 康叶伟. 基于视觉的移动机器人同时定位与建图研究进展 [J]. 控制理论与应用, 2010, 27(4): 488-494.  
Sun F C, Huang Y L, Kang Y W. Review on the achievements in simultaneous localization and mapping for mobile robot based on vision sensor[J]. Control Theory and Applications, 2010, 27(4): 488-494.
- [4] Tardif J P, Pavlidis Y, Daniilidis K. Monocular visual odometry in urban environments using an omnidirectional camera[C]//IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway, USA: IEEE, 2008: 2531-2538.
- [5] Civera J, Grasa O G, Davison A J, et al. 1-point RANSAC for extended Kalman filtering: Application to real-time structure from motion and visual odometry[J]. Journal of Field Robotics, 2010, 27(5): 609-631.
- [6] Mullane J, Vo B N, Adams M D, et al. A random-finite-set approach to Bayesian SLAM[J]. IEEE Transactions on Robotics, 2011, 27(2): 268-282.
- [7] Adams M, Vo B N, Mahler R, et al. SLAM gets a PHD: New concepts in map estimation[J]. IEEE Robotics and Automation Magazine, 2014, 21(2): 26-37.
- [8] Klein G, Murray D. Parallel tracking and mapping for small AR workspaces[C]//6th IEEE and ACM International Symposium on Mixed and Augmented Reality. Piscataway, USA: IEEE, 2007: 250-259.
- [9] Mur-Artal R, Montiel J M M, Tardos J D. ORB-SLAM: A versatile and accurate monocular SLAM system[J]. IEEE Transactions on Robotics, 2015, 31(5): 1147-1163.
- [10] Zhang G X, Kang D H, Suh I H. Loop closure through vanishing points in a line-based monocular SLAM[C]//IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2012: 4565-4570.
- [11] Endres F, Hess J, Sturm J, et al. 3-D mapping with an RGB-D camera[J]. IEEE Transactions on Robotics, 2014, 30(1): 177-187.
- [12] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [13] 付梦印, 吕宪伟, 刘彤, 等. 基于 RGB-D 数据的实时 SLAM 算法 [J]. 机器人, 2015, 37(6): 683-692.  
Fu M Y, Lü X W, Liu T, et al. Real-time SLAM algorithm based on RGB-D data[J]. Robot, 2015, 37(6): 683-692.
- [14] Grompone von Gioi R, Jakubowicz J, Morel, J M, et al. LSD: A fast line segment detector with a false detection control[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(4): 722-732.
- [15] Lemaire T, Lacroix S. Monocular-vision based SLAM using line segments[C]//IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2007: 2791-2796.
- [16] Sola J, Vidal-Calleja T, Devy M. Undelayed initialization of line segments in monocular SLAM[C]//IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway, USA: IEEE, 2009: 1553-1558.
- [17] 武涛, 孙凤池, 苑晶, 等. 一种基于线段特征的室内环境主动 SLAM 方法 [J]. 机器人, 2009, 31(2): 166-170, 178.  
Wu T, Sun F C, Yuan J, et al. An active SLAM approach based on line segment feature in indoor environment [J]. Robot, 2009, 31(2): 166-170, 178.
- [18] Zhang G X, Lee J H, Lim J W, et al. Building a 3-D line-based map using stereo SLAM[J]. IEEE Transactions on Robotics, 2015, 31(6): 1364-1377.
- [19] Gee A P, Chekhlov D, Calway A, et al. Discovering higher level structure in visual SLAM[J]. IEEE Transactions on Robotics, 2008, 24(5): 980-990.

(下转第 229 页)

- [3] 席瑞, 李玉军, 侯孟书. 室内定位方法综述 [J]. 计算机科学, 2016, 43(4): 1-6, 32.  
Xi R, Li Y J, Hou M S. Survey on indoor localization[J]. Computer Science, 2016, 43(4): 1-6, 32.
- [4] Li W C, Wei P, Xiao X C. A robust TDOA-based location method and its performance analysis[J]. Science in China, Series F: Information Sciences, 2009, 52(5): 876-882.
- [5] Alismail H, Baker L D, Browning B. Continuous trajectory estimation for 3D SLAM from actuated Lidar[C]//IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2014: 6096-6101.
- [6] 李晓亮. CORS 系统的构建与应用 [D]. 北京: 中国地质大学, 2014.  
Li X L. Construction and application of CORS[D]. Beijing: China University of Geosciences, 2014.
- [7] 叶刚. 城市环境基于三维激光雷达的自动驾驶车辆多目标检测及跟踪算法研究 [D]. 北京: 北京理工大学, 2016.  
Ye G. Multi-target detection and tracking algorithm for autonomous driving car based on a 3D Lidar in urban traffic environment[D]. Beijing: Beijing Institute of Technology, 2016.
- [8] 杨丽, 曹志强, 张文文, 等. 一种基于局部感知的多机器人动态跟随方法 [J]. 自动化学报, 2010, 36(1): 101-106.  
Yang L, Cao Z Q, Zhang W W, et al. A multi-robot dynamic following approach based on local sensing[J]. Acta Automatica Sinica, 2010, 36(1): 101-106.
- [9] Maier D, Kleiner A. Improved GPS sensor model for mobile robots in urban terrain[C]//IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2010: 4385-4390.
- [10] Ryosuke T, Oscar B J, Kazunobu I, et al. Development of crawler-type robot tractor based on GPS and IMU[C]//3rd IFAC Conference in Modelling and Control in Agriculture, Horticulture and Post-Harvest Processing – Agricontrol. Amsterdam, Netherlands: Elsevier, 2010: 151-156.
- [11] Naranjo J E, Gonzalez C, Garcia R, et al. Lane-change fuzzy control in autonomous vehicles for the overtaking maneuver[J]. IEEE Transactions on Intelligent Transportation Systems, 2008, 9(3): 438-450.
- [12] Vincenty T. Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations[J]. Survey Review, 1975, 23(176): 88-93.
- [13] Karney C F F. Geodesics on an ellipsoid of revolution[C/OL]. Washington, USA: SRI International. (2015-05-14) [2016-12-28]. <http://geographiclib.sourceforge.net/geod.html>.
- [14] 周广宇, 茅旭初. 基于平方根卡尔曼滤波的微弱 GPS 信号跟踪方法 [J]. 上海交通大学学报, 2009, 43(7): 1149-1154.  
Zhou G Y, Mao X C. Square root filter-based tracking of weak GPS signals[J]. Journal of Shanghai Jiaotong University, 2009, 43(7): 1149-1154.
- [15] 徐林, 李世玲, 屈新芬. 一种稳健的自适应传递对准算法 [J]. 系统仿真学报, 2012, 24(11): 2324-2328.  
Xu L, Li S L, Qu X F. Robust adaptive algorithm for transfer alignment[J]. Journal of System Simulation, 2012, 24(11): 2324-2328.
- [16] Zhu Z, Liu S R, Zhang B T. An improved Sage-Husa adaptive filtering algorithm[C]//31st Chinese Control Conference. Piscataway, USA: IEEE, 2012: 5113-5117.

#### 作者简介:

谢德胜 (1993 -), 男, 硕士生. 研究领域: 智能车辆环境感知.

徐友春 (1972 -), 男, 博士, 教授, 博士生导师. 研究领域: 智能车辆技术.

(上接第 220 页)

- [20] Flint A, Mei C, Reid I, et al. Growing semantically meaningful models for visual SLAM[C]//IEEE Computer Society International Conference on Computer Vision and Pattern Recognition. Piscataway, USA: IEEE, 2010: 467-474.
- [21] 王文斐, 熊蓉, 褚健. 基于粒子滤波和点线相合的未知环境地图构建方法 [J]. 自动化学报, 2009, 35(9): 1185-1192.  
Wang W F, Xiong R, Chu J. A simultaneous localization and mapping approach by combining particle filter and dot-line congruence[J]. Acta Automatica Sinica, 2009, 35(9): 1185-1192.
- [22] Tretyak E, Barinova O, Kohli P, et al. Geometric image parsing in man-made environments[J]. International Journal of Computer Vision, 2012, 97(3): 305-321.
- [23] Forster C, Pizzoli M, Scaramuzza D. SVO: Fast semi-direct monocular visual odometry[C]//IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2014: 15-22.
- [24] Engel J, Schöps T, Cremers D. LSD-SLAM: Large-scale direct monocular SLAM[C]//13th European Conference on Computer Vision. Berlin, Germany: Springer, 2014: 834-849.
- [25] Fan B, Wu F C, Hu Z Y. Robust line matching through line-point invariants[J]. Pattern Recognition, 2012, 45(2): 794-805.
- [26] Lu Y, Song D Z. Visual navigation using heterogeneous landmarks and unsupervised geometric constraints[J]. IEEE Transactions on Robotics, 2015, 31(3): 736-749.

#### 作者简介:

李海丰 (1984 -), 男, 博士, 讲师, 硕士生导师. 研究领域: 计算机视觉, 机器人导航.

胡遵河 (1990 -), 男, 硕士生. 研究领域: 移动机器人视觉 SLAM.

陈新伟 (1984 -), 男, 博士, 讲师. 研究领域: 智能机器人控制.