

基于多模态融合的自动驾驶定位研究

冯逸潇

(北京化工大学 北京 100029)

摘要: 为了提高视觉 SLAM 系统的精度和鲁棒性, 将 IMU 传感器引入到可视化 SLAM 系统中。在图像处理过程中, 如果图像错误将会引起 IMU 的快速漂移。由于缺乏有效的加速度激励, 导致了四个方向不能被观察到, 且计算的标度系数也不准确。本文采用轮式里程计, 可以有效地解决视觉惯性里程计在移动机器人上的低可视化问题, 并在此基础上提出了基于视觉信息的非可视化处理方法。对卡尔曼滤波算法中的标度复原进行了改进, 采用半直接法对视觉惯性里程进行单眼观察, 并用 EuRoC 数据对改进后的算法进行了精度检验。利用树莓派完成视觉、IMU、轮式里程计的紧密耦合, 有效地解决了视觉辅助惯性导航 (VINS) 的方向不可视和规模较小的问题。

关键词: 多模态融合; 自动驾驶; 视觉 SLAM; 扩展卡尔曼滤波

0 引言

自动驾驶技术的发展速度很快, 但是想要真正地实现自动驾驶却是一件非常困难的事情。受各种复杂的环境和突发事件的影响, 自动驾驶汽车必须要对周围的环境进行精确的感知, 并根据周围的环境来做出正确的判断。此外, 自动驾驶汽车还必须具有在未知环境中定位的能力, 这是能够完成更加高级任务的基础。

在固定的局部线性加速度下, 自动驾驶汽车不能分辨出实际的加速度和加速度传感器的偏差; 在没有转动的情况下, 自动驾驶汽车无法分辨出重力加速度的方向与加速度传感器的位置。针对目前存在的问题, 提出了一种基于轮式里程计的多传感器融合方法, 以提高移动车辆的运动精度。

采用不同传感器进行多模态融合定位的自动驾驶汽车, 其定位精度受到传感器分辨率、制作工艺以及周边环境等因素的影响。为了解决自动驾驶汽车定位时, 单眼视觉难以适应光照变化、里程计累计误差等问题, 采用多个边缘传感器进行边缘融合定位。在单眼视觉中, 采用稀疏直接法 (半直接法) 进行单目相机的位姿实时估计, 利用惯性传感器进行姿态恢复, 再利用卡尔曼滤波算法进行融合, 获得 IMU 的加速度和偏航角, 实现更准确的定位。

边缘侧多个传感器的信息融合使得自动驾驶汽车可以在各种复杂的环境中准确、实时地定位。视觉 SLAM 技术能够根据系统的位置快速生成图像, 从而实现对场景的初步构造^[1]。利用物理算法对 ORB-SLAM2 融合技术进行了试验研究, 证明了该方法在弱纹理、光照变化等条件下的卡尔曼滤波效果良好。

本文重点介绍了将 SVO 半直接法与 ORB-SLAM2 相结合的方法, 在此基础上利用卡尔曼滤波器进行了自动驾驶汽车定位的试验, 对扩展卡尔曼滤波多传感器的融合定位精度进行了检验。

1 基于扩展卡尔曼滤波的多传感器融合定位

1.1 ORB-SLAM2 融合半直接法视觉里程计

传统的 SLAM 前端可视里程计 (Visual Odometry, VO)^[2] 是利用相邻的两个帧来计算所获取的图像的移动位姿。经过长期的计算, 前方里程计的相位姿态也会产生累积的误差, 所以对摄像机的位姿进行了大量的非线性优化。利用全局姿态信息, 对摄像机的姿态进行更新, 从而获得了具有全局一致性的轨迹。在 ORB-SLAM2 中, 回环检测作为一个单独的组件, 以判定当前检测完成的位置是不是先前所经过的位置。假定回环检测线程在检测到单眼摄像机所获取的数据为回环时, 向后端发送当前的数据, 并利用词袋模型进行回环融合, 以提高定位准确度。利用扩展卡尔曼滤波技术实现了 ORB-SLAM2 多传感器融合定位的构建。在视觉 SLAM 中, 主要通过摄像机来读取传感器的数据, 而在多个传感器的数据融合过程中, 与搭载的传感器进行同步读取。

SVO 算法是一种仅对当前时刻摄像机姿态进行估计的可视里程计算法, 并不具有后端优化和闭环检测功能。为了使半直接法具有较强的适应性, 本文在 ORB-SLAM2 中融合了 SVO 半直接追踪法, 对跟踪部分进行了重点改进, 并解决在嵌入式平台上不能迅速进行特征点定位的缺陷。融合半直接法的初始化位姿流程, 如图所示。

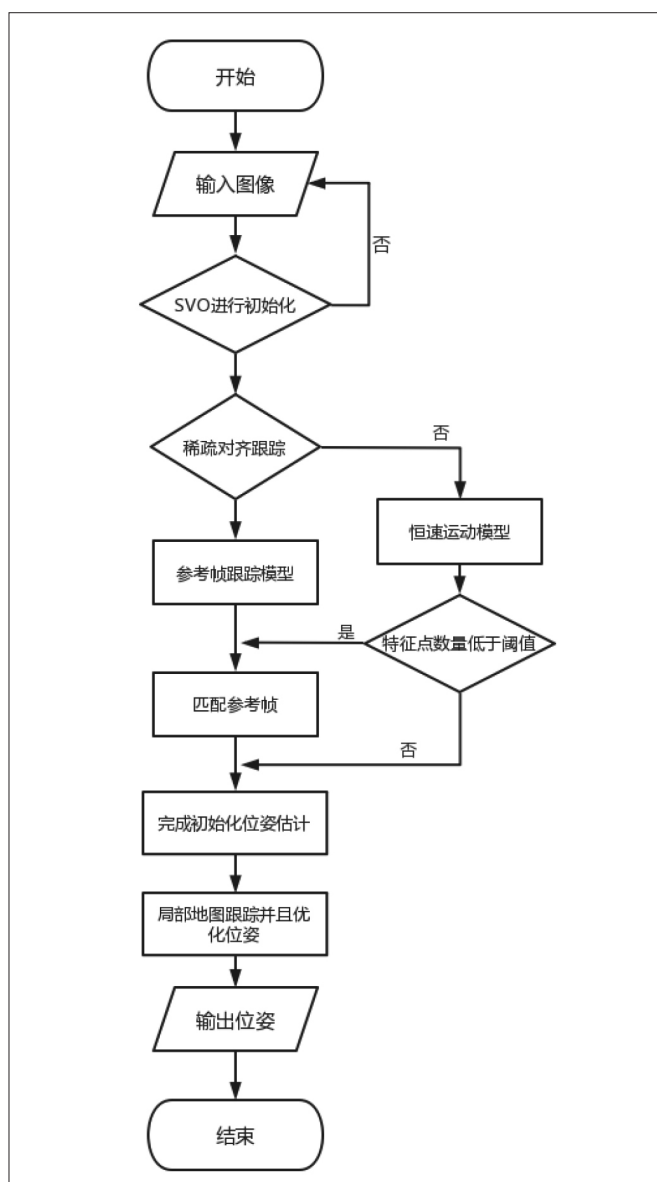


图 Tracking 线程流程图

在两个相邻帧之间的位姿变换矩阵 $T_{k|k-1}$ 是已知的, 半直接方法是将前一个邻近时间的位置假定为单元矩阵。在多帧间, 通过检测特征点和深度估算, 来确定在 $k-1$ 帧中特征点的位置和深度。所获得的 I_{k-1} 中的特征点在图像平面上的位置为 (u, v) , 其深度为 d , 并投影到定义在 I_{k-1} 相机坐标系的三维空间 P_{k-1} 。

然后, 将所述特征点投影至当前的帧 I_k , 并利用所述位姿变换矩阵 $T_{k|k-1}$ 获得所述当前帧坐标系中的三维坐标 P_k 。利用单目摄像机参数将特征点变换为目前的 I_k , 并在图像平面上进行再投射。由于其特征点的重投射存在着一定的误差, 因此需要对现有的重投影位进行优化以减少残差。最小残余损耗函数是:

$$T_{k|k-1} = \arg \min_{T_{k|k-1}} \frac{1}{2} \sum_{i \in R} \left\| \sigma I(T_{k|k-1}, u_i) \right\|^2 \quad (1)$$

为了解决此处产生的漂移, 利用特征点对齐来确定新帧中的实际位置。在光度不变假设的基础上, 通过重构残差, 对特征点的预测位置进行了优化:

$$u'_i = \arg \min_{u'_i} \frac{1}{2} \left\| I_k(u'_i) - A_i I_r(u_i) \right\|^2 \quad (2)$$

与公式 (1) 相比, 公式 (2) 所预测的摄像机的姿态位置更加精确。然而, 在公式 (2) 的最优解和公式 (1) 所预测的特征点的位置上也会有错误。为此, 必须构建新的优化目标函数。通过公式 (3) 进行了最优的计算, 得到了更加精确的摄像机姿态:

$$T_{k|w} = \arg \min_{T_{k|w}} \frac{1}{2} \sum_i \left\| u_i - \pi(T_{k|w} - wp_i) \right\|^2 \quad (3)$$

1.2 基于卡尔曼滤波器的多传感器融合定位算法的研究

多传感器的融合模式有松耦合和紧耦合两种。松耦合是通过对各个传感器的数据进行处理, 然后用融合算法对它们进行融合, 从而得到相应的位姿。紧耦合是利用视觉与惯性传感器的数据进行耦合, 并综合考虑各传感器的耦合特性, 但需要计算载体的运算能力。对于单一的传感器故障, 松耦合可以实现准确的定位。

滤波器的实质是将预测与观察相结合, 从而获得较为准确的后验数值^[3]。在这两种方法中, 预测和观察都来自于传感器, 所以滤光片的功能就是与传感器相结合, 从而获得最佳的效果。预测来自于 IMU、编码器及其他的传感器。观测来自于 GPS、雷达、照相机等传感器。后验为定位组件的输出 (即融合后的结果)。利用卡尔曼滤波器实现了多个传感器的位置融合, 从而检验了在实际的环境下, 当遇到紧急状况时, 该系统能否实现融合的精确位姿估计。为此, 本节应用了一种基于扩展卡尔曼滤波的松耦合融合算法。

1.2.1 自动驾驶系统状态初始化

当检测到第一个数据时, 自动驾驶的多传感器融合系统会对系统的状态和状态协方差进行初始化, 从而得到初始状态和初始状态的协方差。

1.2.2 扩展卡尔曼滤波预测过程

在扩展卡尔曼滤波中, 移动机器人的预测是利用状态变换矩阵对当前的状态变量 X_k 进行预测:

$$X_k = f(X_{k-1}) \quad (4)$$

式中: X_{k-1} — 上一时刻的系统状态变量。

利用状态转换函数的一阶泰勒展开式, 利用雅各比矩阵 F 与过程噪声协方差 Q 对目前的系统状态协方差 P_k 进行了预测:

$$P_k = F_{k-1} P_{k-1} F_{k-1}^T + Q \quad (5)$$

1.2.3 扩展卡尔曼滤波更新过程

自动驾驶汽车在进行扩展卡尔曼滤波计算过程中,

其系统的增益矩阵 K 是由观测矩阵 H 、观测噪声协方差 R 和预测系统状态协方差 P_k 来计算更新的。通过增益矩阵 K 更新扩展卡尔曼滤波系统状态量 X_k 以及状态协方差 P_k 。通过公式 (6) ~ (8) 计算所需状态量数值:

$$K_k = P_k H_k^T (H_k^T P_k H_k^T + R_k)^{-1} \quad (6)$$

$$X_k = X_k + K_k (Z_k - H_k X_k) \quad (7)$$

$$P_k = (1 - K_k H_k) P_k (1 - K_k H_k)^T + K_k R_k K_k^T \quad (8)$$

上述传感器的输入数据 R_k 受到过程噪声协方差 Q 和协方差 P_{k-1} 所影响。

1.2.4 单目视觉尺度因子的计算

在智能汽车上安装的单目照相机和车轮里程表是不会改变的。用 c 系表示单眼摄像机坐标系, 用标定方法得到摄像机和里程计坐标系的转换矩阵。选择了系统起始时间的载体坐标系作为导航坐标系, 用 w 系表示; 将单目视觉里程计初始化成功的瞬间的影像姿态, 用 cw 系表示; SLAM 系统输出的摄像机的姿态转换矩阵是 T_c^{cw} , 利用左乘法进行单眼视觉修正, 实现对导航坐标系的修正。

在完成了可视 SLAM 的初始化之后, T_s^w 代表了摄像机在从摄像机到导航坐标系中获取的第 2 帧图像时的转换矩阵。利用卡尔曼滤波方法, 将智能车的运动机器人与 IMU 的惯性传感器进行了扩展。在这种情况下, T_b^w 代表摄像机在导航坐标系中的姿态, 并将其赋值 0, 从而获得 Q_c^b 。

利用运动机器人轮型编码器与 IMU 惯性传感器的数据进行卡尔曼滤波融合, 利用融合后的数据计算运动距离 x_c ; 同时, 利用单眼视觉 SLAM 法, 在一定的时间间隔内, 得到了运动距离 x_c 。利用这两种方法的比例, 实现一种视觉尺度的复原, 并实现了对尺度因子 $\lambda = x_b/x_c$ 的实时估算。在完成单一目标度系数的估算后, 利用该

方法对目前单目标值系数进行实时运算。

2 结语

本文利用卡尔曼滤波松耦合方法, 将 IMU 惯性传感器、轮式里程计与单目相机的位姿进行了试验研究。通过对 ORB-SLAM2 融合半直接法进行试验验证, 证明了在缺乏特征点和光照条件下, 视觉里程计仍能正常工作。试验结果表明, 利用卡尔曼滤波器进行多个传感器的融合可以有效地改善自动驾驶车辆的定位精度。由于扩展卡尔曼滤波 (EKF) 的马尔可夫特性, 仅可对后验概率进行一次线性化处理。由于当前帧的数据与上一帧的数据相关, 并且当前的位置和上一次的位置有一定的距离, 所以一阶泰勒展开式并不能完全逼近该函数。在视觉里程计中, 扩展卡尔曼滤波器存在松耦合区恢复比例信息不精确的缺点。接着, 利用摄像机与 IMU 的视觉惯性紧密耦合方法, 实现了对图像的标度信息的恢复。

参考文献:

- [1] 刘文杰, 刘小娇, 付猛, 等. 基于视觉 SLAM 的动态图像处理研究方法 [J]. 数字通信世界, 2022 (06): 14-16.
- [2] 张燕咏, 张莎, 张昱, 等. 基于多模态融合的自动驾驶感知及计算 [J]. 计算机研究与发展, 2020, 57 (9): 1781-1799.
- [3] Hu F, Wu G. Distributed Error Correction of EKF Algorithm in Multi-Sensor Fusion Localization Model [J]. IEEE Access, 2020, PP (99): 93211-93218.

作者简介: 冯逸潇 (2000.02-), 男, 汉族, 河南焦作人, 本科在读, 研究方向: 多模态融合感知与视觉 SLAM。