

DOI:10.3969/j.issn.1673-4785.201503003

网络出版地址: <http://www.cnki.net/kcms/detail/23.1538.TP.20150630.1625.007.html>

# 单目视觉同步定位与地图创建方法综述

顾照鹏<sup>1</sup>, 刘宏<sup>1,2</sup>

(1. 北京大学深圳研究生院 物联网智能感知工程实验室, 广东 深圳 518055; 2. 北京大学 机器感知与智能教育部重点实验室, 北京 100871)

**摘要:**随着计算机视觉技术的发展,基于单目视觉的同步定位与地图创建(monocular SLAM)逐渐成为计算机视觉领域的热点问题之一。介绍了单目视觉 SLAM 方法的分类,从视觉特征检测与匹配、数据关联的优化、特征点深度的获取、地图的尺度控制几个方面阐述了单目视觉 SLAM 研究的发展现状。最后,介绍了常见的单目视觉与其他传感器结合的 SLAM 方法,并探讨了单目视觉 SLAM 未来的研究方向。

**关键词:**单目视觉;同步定位与地图创建;扩展卡尔曼滤波器;计算机视觉;特征检测与匹配

**中图分类号:** TP391 **文献标志码:** A **文章编号:** 1673-4785(2015)04-0499-09

中文引用格式:顾照鹏,刘宏.单目视觉同步定位与地图创建方法综述[J].智能系统学报,2015,10(4):499-507.

英文引用格式:GU Zhaopeng, LIU Hong. A survey of monocular simultaneous localization and mapping[J]. CAAI Transactions on Intelligent Systems, 2015, 10(4): 499-507.

## A survey of monocular simultaneous localization and mapping

GU Zhaopeng<sup>1</sup>, LIU Hong<sup>1,2</sup>

(1. Engineering Lab on Intelligent Perception for Internet of Things (ELIP), Shenzhen Graduate School, Peking University, Shenzhen 518055, China; 2. Key Laboratory of Machine Perception, Peking University, Beijing 100871, China)

**Abstract:** With the development of computer vision technology, monocular simultaneous localization and mapping (monocular SLAM) has gradually become one of the hot issues in the field of computer vision. This paper introduces the monocular vision SLAM classification that relates to the present status of research in monocular SLAM methods from several aspects, including visual feature detection and matching, optimization of data association, depth acquisition of feature points, and map scale control. Monocular SLAM methods combining with other sensors are reviewed and significant issues needing further study are discussed.

**Keywords:** monocular vision; simultaneous localization and mapping; extended Kalman filter; computer vision; feature detection and matching

随着计算机技术和人工智能的发展,机器人的研究得到越来越多的关注和投入,智能自主移动机器人成为机器人领域的一个重要研究方向和研究热点。同步定位与地图创建(simultaneous

localization and mapping, SLAM) 问题是指移动机器人在未知环境中,利用自身的传感器创建一个与环境相一致的地图,并同时确定自身在地图中的位置。SLAM 问题最先是由 Smith Self 和 Cheeseman<sup>[1]</sup>于 1987 年提出,被认为是实现移动机器人真正自主的关键。近年来随着计算机视觉技术的发展以及计算机运算能力的增强,出现了大量利用摄像机作为传感器的视觉 SLAM 方法。视觉传感器具有信息量大、灵活性高、成本低等优点,这是其他传感器所无法比拟的。单目视觉

收稿日期:2015-03-03. 网络出版日期:2015-06-30.

基金项目:国家 863 计划资助项目(2006AA04Z247);国家自然科学基金资助项目(60875050,60675025,61340046);深圳市科技计划项目及基础研究计划资助项目(201005280682A, JCYJ20120614152234873);教育部高等学校博士学科点专项科研基金资助项目(20130001110011).

通信作者:顾照鹏. E-mail: guzp@pku.sz.edu.cn.

SLAM 仅使用一个摄像头作为传感器完成同步定位与地图创建操作,具有系统结构简单、成本低且易实现等优点。因此,基于单目视觉 SLAM 技术逐渐成为近年来的研究热点,并且在许多领域中体现出巨大的应用价值。

### 1 单目视觉 SLAM 方法分类

根据算法是否采用概率框架,文献中的单目视觉的 SLAM 方法主要可以分为两类,即基于概率框架的单目视觉 SLAM 方法和非概率框架的单目视觉 SLAM 方法。

#### 1.1 基于概率框架的单目视觉 SLAM 方法

由于 SLAM 问题涉及未知的环境描述和传感器噪声,因此大部分基于单目视觉的 SLAM 方法采用概率框架来描述该问题。其总体思路是在

给定了从初始时刻到当前时刻的控制输入以及观测数据的条件下,构建一个联合后验概率密度函数来描述摄像机姿态和地图特征的空间位置。通过递归的贝叶斯滤波方法对此概率密度函数加以估计,从而实现摄像机同步定位和地图创建。在该类方法中,利用扩展卡尔曼滤波器(extended Kalman filter)<sup>[2]</sup>实现同步定位与地图创建是一种常用的解决方案<sup>[3-10]</sup>。例如, Davison 等<sup>[3-8]</sup>实现了一种基于扩展卡尔曼滤波器的实时单目视觉 SLAM 系统(MonoSLAM), Lemaire 等<sup>[9]</sup>利用 MATLAB 实现了一个基于扩展卡尔曼滤波器的单目视觉 SLAM 仿真实验系统, Roussillon 等<sup>[10]</sup>利用高速摄像机实现了基于 EKF 的实时单目视觉 SLAM 系统 RTSLAM。图 1 给出了典型基于 EKF 的单目视觉 SLAM 的流程图。

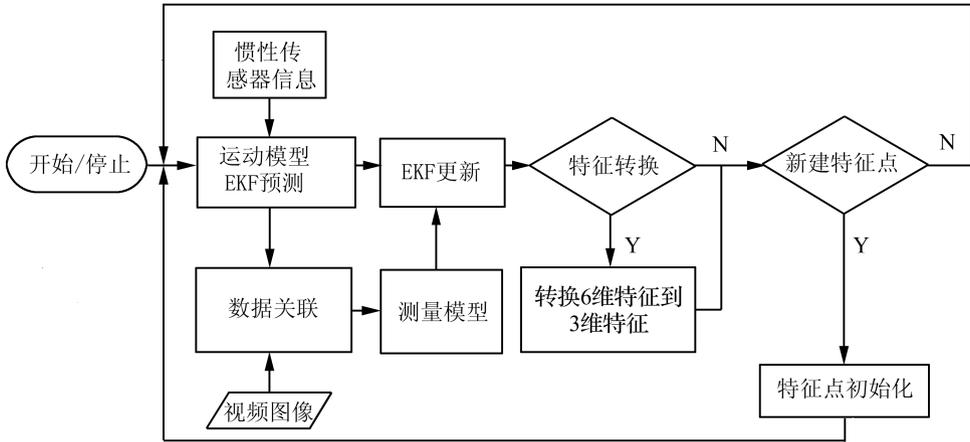


图 1 基于 EKF 的单目视觉 SLAM 流程图

Fig. 1 Flowchart of EKF based Monocular SLAM

此外,一些研究人员利用粒子滤波等概率滤波器实现了单目视觉 SLAM 方法<sup>[11-16]</sup>。例如, Sim 等<sup>[11]</sup>提出了一种基于粒子滤波的 SLAM 方法,其基本原理是将运动路径与地图的联合后验分布估计问题分解为用粒子滤波器对运动路径的估计问题和在路径已知的条件下的路标估计问题。然而,为了保证定位精度,基于粒子滤波的 SLAM 方法在复杂场景和运动的情况下需要较多的粒子,这大大增加了算法时间复杂度,并且其重采样过程中容易导致样本贫化等问题,因此这种方法没有得到大范围的应用。另外为了改进系统非线性给 EKF 带来的影响,无迹卡尔曼滤波器也被引入到单目视觉 SLAM 算法中<sup>[17-18]</sup>,但是先比 EKF 增加了计算复杂度。

#### 1.2 基于非概率框架的单目视觉 SLAM 方法

非概率框架的单目视觉 SLAM 系统目前主要是基于关键帧和稀疏捆绑调整等优化算法实现。例如, Mouragnon 等<sup>[19]</sup>利用摄像机最近采集的固定数目的图像作为关键帧,进行局部捆绑调整优化,进而实现单目摄像机的定位与地图创建。Klein 等<sup>[20-21]</sup>提出了一种基于关键帧的单目视觉 SLAM 系统 PTAM。该系统利用一个线程对摄像机姿态进行跟踪,同时利用另外一个线程对关键帧数据以及所有特征点的空间位置进行捆绑调整,进而实现了实时的同步定位与地图创建。PTAM 的双线程系统框图如图 2 所示。另外,基于图优化的 SLAM 方法<sup>[22-23]</sup>也逐渐被引入到单目视觉 SLAM 中。

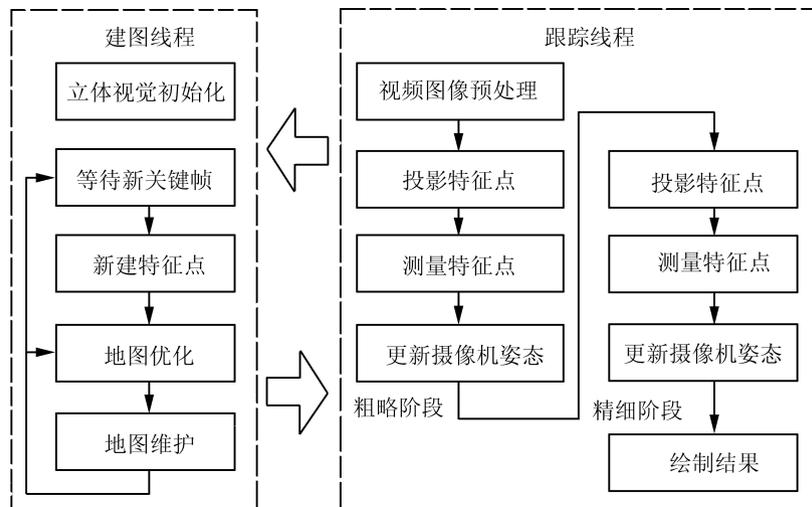


图2 PTAM 双线程系统框图

Fig. 2 Dual-thread system diagram of PTAM

## 2 单目视觉 SLAM 中的几个关键问题

### 2.1 特征的检测与匹配

在典型的 SLAM 系统中,地图建模以及地图匹配都是以环境中的路标点为基础的。因此路标的可靠性尤为重要,可靠的路标能够简化 SLAM 中的数据关联问题,不可靠的路标容易导致错误的关联,最终导致整个 SLAM 系统发散。不同于传统 SLAM 系统,单目视觉系统无法从单帧图像恢复深度。因此,如何利用图像特征的检测与匹配获得高质量的路标,关系到单目视觉 SLAM 算法的鲁棒性。

一般图像特征可分为点特征、直线特征、以及边缘、轮廓等特征。相比其他特征,点特征具有对遮挡相对鲁棒,识别性好,提取速度快等特点。目前绝大多数视觉 SLAM 系统的数据关联都是利用点特征的检测与匹配来实现。点特征的检测与匹配是 SLAM 系统中的关键研究问题。

#### 2.1.1 角点特征

角点在保留图像图形重要特征的同时,可以有效地减少信息的数据量,使其信息的含量很高,有效地提高了计算的速度,有利于图像的可靠匹配,使得实时处理成为可能。因此角点是 SLAM 中应用最广泛的一类特征。

Harris 角点是 Harris 等<sup>[24]</sup>在 1988 年提出的一种基于信号的点特征提取算子。其前身是 Moravec<sup>[25]</sup>角点。Harris 角点检测方法使用自相关矩阵改进了 Moravec 的方法。这种方法避免了使用离散的方向和偏移,它在窗口内使用高斯函数加权导数,取代了简单的求和。如果自相关矩阵有 2 个较大的特征值就表示该点为检测到的特征点。Har-

ris 角点对亮度和对比度的变化不敏感,具有旋转不变性,但不具备尺度不变性。文献[26-29]中单目视觉 SLAM 算法均采用了 Harris 角点作为特征检测方法。

Shi 等<sup>[30]</sup>于 1994 年采用与 Harris 完全不同的角点响应函数来提取图像角点,其方法是如果 2 个特征值中较小的一个大于最小阈值,则判定为角点。Shi 等的角点判定方法比较充分,并且在很多情况下可以得到比使用 Harris 算法更好的结果。文献[3-8,31]等均采用了 Shi-Tomasi 角点提取方法。

FAST(features from accelerated segment test)角点是 Rosten 等<sup>[32]</sup>在 2006 年提出的一种简单快速的角点检测算法,该算法检测的角点定义为在像素点的周围邻域内有足够多的像素点与该点处于不同的区域。应用到灰度图像中,即有足够多的像素点的灰度值大于该点的灰度值或者小于该点的灰度值。由于 FAST 算法能在短时间内生成大量角点,因此被广泛用于对实时性要求较高的 SLAM 系统中<sup>[20-21,33-35]</sup>。其中文献[20-21,35]为了减少大量非角点的干扰,对每个检测到的 FAST 角点分别进行了 Shi-Tomasi 阈值检测,以确保获得更加鲁棒的特征角点。

#### 2.1.2 SIFT 与 SURF 特征

SIFT(scale invariant feature transform)是 Lowe 等<sup>[36-37]</sup>在 1999 年提出的一种特征点提取方法。SIFT 特征对图像的尺度变化、图像缩放、旋转甚至仿射变换具有不变性,而且对光照的变化和图像变形具有较强的适应性。因此 SIFT 被广泛应用物体辨识、机器人定位知与导航、图像拼接、3D 重建、手势辨识、视频跟踪等领域。

SIFT 特征检测方法首先检测由高斯差分 (difference of Gaussians, DoG) 构建的尺度空间上的局部极值点, 并将其作为候选特征点, 经过筛选, 排除不太稳定的以及低对比度的候选特征点, 从而确定关键点, 接下来为关键点分配方向, 并通过采样关键点周围图像块图像梯度的幅值和方向来建立 SIFT 特征描述子。

一些研究人员曾尝试直接将 SIFT 特征引入单目视觉 SLAM 系统<sup>[38-40]</sup>。由于 SIFT 特征的提取和匹配存在大量的浮点运算, 时间复杂度较高, 不能保证实时性, 因此直接在单目视觉 SLAM 中应用的并不多。而利用 GPU 加速浮点运算在一定程度上克服了这一缺点, 因此一些单目视觉 SLAM 系统<sup>[41-42]</sup> 采用了 GPU-SIFT 作为特征检测与匹配方法。文献 [43] 利用 PCA 主成分分析的方法将 SIFT 特征向量降低至 36 维, 以加速特征匹配速度。

SURF (speeded up robust features)<sup>[44]</sup> 特征是对 SIFT 算法的改进, 其基本结构、步骤与 SIFT 相近, 但具体实现的过程有所不同。SURF 算法的优点是速度远快于 SIFT 且稳定性好。SURF 在单目视觉 SLAM 系统中也获得比较广泛的应用<sup>[45-46]</sup>。

### 2.1.3 BRIEF 与 ORB 特征

BRIEF (binary robust independent elementary features) 特征是 Calonder 等<sup>[47]</sup> 在 2010 年提出的一种可以快速计算且表达方式为二进制编码的描述子。主要思路是在特征点附近随机选取若干点对, 将这些点对的灰度值的大小, 组合成一个长为 256 的二进制串, 并将这个二进制串作为该特征点的特征描述子。由于其描述子利用二进制编码, 因此在特征匹配时只需计算 2 个特征点描述子的 Hamming 距离。经过大量实验数据测试, 不匹配特征点的描述子的 Hamming 距离在 128 左右, 匹配点对描述子的 Hamming 距离则远小于 128。由于 BRIEF 的匹配速度远高于 SURF 和 SIFT, 因此被很多系统采用。由于 BRIEF 特征不具备旋转不变性, 因此很难用于旋转比较剧烈的环境。为了解决这一问题, Ethan Rublee 等<sup>[48]</sup> 在 2011 年提出了 ORB (oriented brief) 特征, 较好地解决了尺度不变性。在尺度变化不大的环境, ORB 特征在 SLAM 系统中获得了大量应用<sup>[49-50]</sup>。

除了点特征之外, 直线特征、颜色特征、超像素特征等也被应用到单目视觉 SLAM 算法中。文献 [51] 利用直线段作为特征进行匹配, 能够有效解决点特征在运动模糊情况下跟踪失败的问题。文献 [52] 利用反深度参数化的直线段特征实现了单目

视觉 SLAM 算法。文献 [53] 利用了超像素特征解决了点特征匹配在弱纹理环境定位失败的情况。

## 2.2 数据关联的优化

如上节所述, 基于图像的特征检测与匹配不可避免地带有误差。如果直接使用图像特征匹配的结果作为数据关联的结果, 往往会带来很大误差, 以至于定位系统发散。如何保证选择正确的数据关联, 对于 SLAM 定位系统至关重要, 因此需要对图像特征匹配的结果进行优化来获得准确的数据关联结果。

### 2.2.1 固定区域匹配

这种方法假设特征点在前后两帧图像像素距离在一个范围内, 超过此范围即认为匹配错误。PTAM<sup>[20-21]</sup> 以及一些改进方法<sup>[54-55]</sup> 主要利用这种方法来进行特征匹配的优化, 降低特征匹配计算量。这种方法实现简单, 对于深度变化不大、特征点距离摄像机稍远的场景比较合适。但对于摄像机快速运动或者运动平行于图像平面的情况并不合适。

### 2.2.2 Active Matching

Davison 等<sup>[56-58]</sup> 提出 Active Matching。这种方法主要用在基于 EKF 滤波的视觉 SLAM 系统<sup>[3-10]</sup> 中, 利用 EKF 系统运动模型获得系统状态预测, 进而估计环境特征点在摄像机估计姿态中的投影图像位置, 获得特征点的图像分布区域。这种方法能够适应摄像机的绝大部分运动情况, 但对于摄像机姿态估计协方差较大的情况, 容易产生较大的特征匹配的估计区域, 仍然面临着错误匹配的可能。

### 2.2.3 1-Point RANAC

为了更好地去除 Active Matching 中的错误匹配, Civera 等<sup>[59-60]</sup> 提出了基于 EKF-SLAM 系统的 1-Point RANSAC 算法。不同于传统的 RANSAC 方法, 1-Point RANSAC 主要面向基于 EKF 滤波的 SLAM 方法。该方法首先随机选取一个匹配点, 用该点的匹配信息更新摄像机姿态。然后判断其他匹配点与估计图像位置的距离是否小于一定阈值来剔除外点。最后利用获得的内点集来更新整体滤波器状态。对于可能被剔除的卡尔曼增益较大的正确匹配点, 重新加入内点集并再次更新系统状态。该方法由于要多次对滤波器进行更新, 因此会损耗一定的运算时间。

### 2.2.4 几何约束

另外, 对于单目视觉 SLAM 系统, 可以利用视觉几何的一些约束条件, 例如对极几何<sup>[61]</sup>、PNP<sup>[62-63]</sup> (perspective N points) 等剔除误匹配点。这种方法从几何求解出发, 相对于 1-Point RANSAC 来说, 不

需要频繁更新系统状态,因此能获得较好的系统运行速度。但是这些方法对于不同的情况需要利用不同的几何约束条件,因此增加了系统的复杂性。

### 2.3 特征点深度的获取

由于单目摄像机相当于 Bearing Only 类传感器,不能立刻感知深度信息,因此如何获取环境特征点深度直接关系到单目视觉 SLAM 算法的成败。特征点深度获取方法主要有以下 3 种方法:几何三角化、反深度(inverse depth)以及粒子滤波(particle filter)方法。

#### 2.3.1 几何三角化

目前采用关键帧与稀疏捆绑调整(sparse bundle adjustment, SBA)框架的单目视觉 SLAM 系统<sup>[20-21, 54-55]</sup>,普遍采用类似运动恢复结构(structure from motion, SFM)中的多关键帧匹配点三角化,重建特征点的三维位置,如图 3 所示。对于非关键帧则不进行特征点的初始化和更新,这类方法,在定位精度要求不高的情况下能够取得比较好的运行效率。

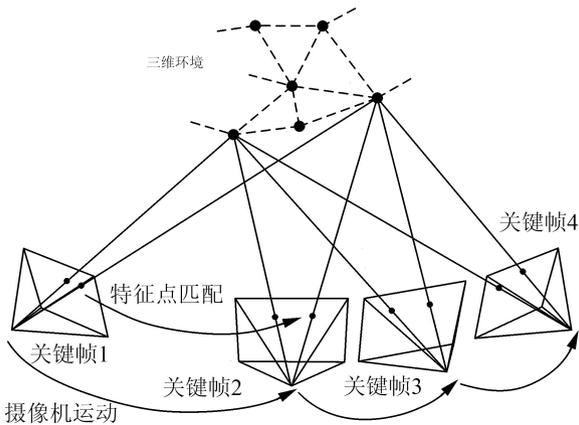


图 3 多关键帧三角化

Fig. 3 Triangulation of multiple keyframes

#### 2.3.2 Inverse Depth

为了减小深度分布非高斯性的影响, Civera 等<sup>[64]</sup>提出了 Inverse Depth 方法,将深度的倒数在 EKF 系统状态里进行更新,这种方法获得了比较好的线性效果。Inverse Depth 方法利用一个 6 维向量来表示一个特征点的三维位置,包括摄像机的位置以及特征点方向和深度信息。随着滤波器不断地进行更新,当特征估计协方差小于某一阈值的情况下,将特征点的 6 维表示转换为 3 维的欧式坐标(如图 4)。

Inverse Depth 方法相比 3 维欧式坐标增加了系统状态向量的长度,亦增加了系统时间复杂度。为了解决这一问题, Pietzsch 等<sup>[65-66]</sup>提出了各自的解决方案,利用在同一帧图像建立多个特征点,将属于同一帧特征点的表示形式缩减为一个摄像机姿态加多个深度的形式,有效地降低了系统状态的长度。

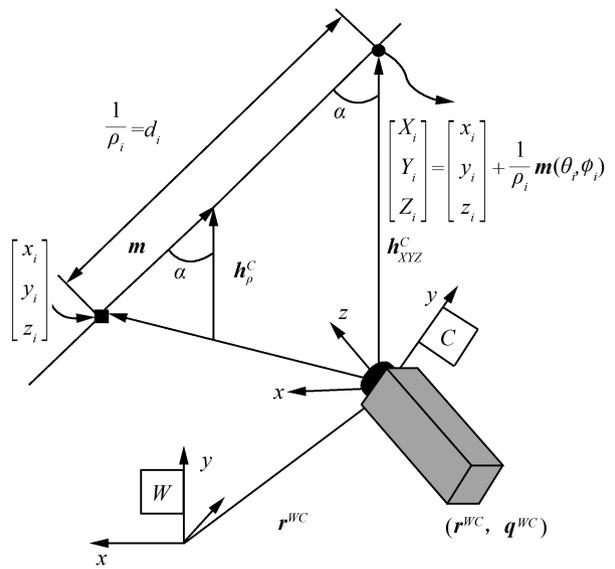


图 4 Inverse Depth 特征参数化

Fig. 4 Inverse Depth feature parameterization

#### 2.3.3 Particle Filter

ParticleFilter 方法最先被 Davison 等<sup>[3]</sup>采用,当特征初始化以后,在特征所在深度方向生成多个粒子。通过粒子的匹配与更新,恢复特征点深度的概率分布,其深度估计的收敛过程如图 5 所示。Forster 等<sup>[67]</sup>实现的单目测距系统, SVO 系统中的 Depth Filter 也采用了类似的思想。这种方法将特征点的深度从 EKF 滤波器中分离出来,单独进行估计,虽然降低了滤波器整体的更新时间,但是容易增加系统的不一致性,最终导致概率估计的发散。

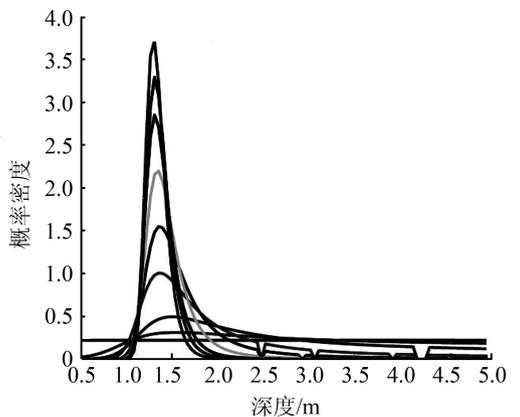


图 5 基于粒子滤波的深度估计收敛过程

Fig. 5 Convergence process of depth estimation based on particle filter

### 2.4 地图的尺度控制

单目视觉 SLAM 中的环境深度是通过多帧图像的匹配信息来恢复的,因此恢复深度的过程中不可避免地会出现尺度漂移。对于单目视觉中的尺度控制主要有以下几种方法。

### 2.4.1 7DOF SLAM

7DOF SLAM 的方法由 Strasdat<sup>[68]</sup> 首次提出, 主要解决基于 EKF 滤波器的单目视觉 SLAM 的尺度漂移问题。该方法将地图尺度作为一个估计量加入到系统状态中, 使三维摄像机姿态由 6 自由度变为 7 自由度, 尺度的估计量随滤波器一起更新, 有效地控制了尺度漂移。

### 2.4.2 Global SBA

Global SBA 方法主要应用于基于关键帧的单目视觉 SLAM 方法<sup>[20-21]</sup> 中, 利用全局 SBA 优化以统一整个地图的尺度。

### 2.4.3 利用 IMU 等其他传感器

单目视觉无法获取确切的尺度信息, 而 IMU (inertial measurement unit) 等惯性传感器能够实时地输出真实的加速度信息。因此, 可以通过分析加速度信息与摄像机运动之间的相关性, 以获得地图真实的尺度。文献[69]中利用 2 次加速度计输出峰值之间的加速度信息, 通过运动学方程估计出相对运动距离, 进而估计出摄像机运动与 IMU 获得的运动之间的尺度。最后, 通过多次测量, 利用非线性优化估计出环境地图的真实尺度。

## 3 单目视觉 SLAM 与其他传感器组合

对于一些复杂环境, 为了弥补单目视觉的不足, 提高单目视觉 SLAM 系统的鲁棒性, 往往需要配合其他传感器组合使用。

### 3.1 单目视觉+惯性传感器

惯性传感器 IMU 一般由三轴加速度计、三轴陀螺仪以及三轴磁场计组成, 可以实时获取自身加速度、角速度以及所处环境磁场强度信息。Azuma 等<sup>[70]</sup> 利用扩展卡尔曼滤波器融合摄像机和 IMU 的数据进行摄像机姿态跟踪, 实验结果表明, 单目视觉结合 IMU 的定位方法相比一般单目视觉定位方法能够更有效地降低系统误差。Nützi 等<sup>[71]</sup> 提出了一种融合 IMU 和单目视觉, 能够获得绝对尺度信息的 SLAM 方法。由于 IMU 采用惯性坐标系, 因此该类 IMU-摄像机组合系统在使用前需要处理 IMU 与摄像机之间的标定问题。例如, Johnson 等<sup>[72]</sup> 利用激光测距仪配合标定板对 IMU-摄像机组合定位系统进行了标定。Lang 等<sup>[73]</sup> 提出了一种基于手眼标定方程的 IMU-摄像机组合定位系统标定方法, 该方法假设摄像机和 IMU 之间的位移非常近, 没有对二者之间的平移进行标定, 因此不适用于摄像机与传感器相对位置较大的情况。另外, 针对低成本 IMU 航向角 Yaw 输出误差较大的问题, 文献[74-75] 提出

了在 IMU 部分信息不可靠情况下的标定方法。

### 3.2 单目视觉+激光测距仪

激光测距仪是一种基于飞行时间 (time of flight, TOF) 原理的高精度、高解析度外部传感器, 与摄像机相比具有不受光线影响、处理方法简单、数据精度高和采样周期短等特点。激光测距仪是目前移动机器人高精度测距的主要手段之一。文献[76] 给出了一种利用单目视觉和激光测距仪实现的 SLAM 系统。针对激光测距仪与摄像机的标定问题, 文献[77] 给出了一种直接的标定方法。

## 4 结论与展望

尽管单目视觉近些年已经取得了很大的研究进展, 但是仍然有许多问题和难点需要解决:

1) 目前单目视觉 SLAM 的应用环境非常广泛, 但是多数研究仍以室内结构化静态环境为主。另外一些应用主要以某些特定运动, 例如室内单目天花板视觉<sup>[78]</sup>、室外无人机视觉导航<sup>[67]</sup> 以及车载视觉导航<sup>[79]</sup> 等。由于地图规模的限制, 基于单目视觉 SLAM 适应环境规模较小。应用于室外单目视觉 SLAM 的研究相对较少, 地图规模也不大。目前的研究主要面向静态环境, 而实际环境往往是动态的。因此对于动态环境的单目视觉 SLAM 将是一个重要研究方向。

2) 单目视觉中数据关联主要由视觉特征的检测与匹配完成, 目前的视觉特征匹配还存在着匹配精度低、速度慢以及无法适应重复纹理等缺点。因此需要进一步研究发掘新的视觉特征检测与匹配方法, 以满足单目视觉 SLAM 在实际环境中的定位需求。

3) 实时性是单目视觉 SLAM 系统方法永恒的话题, 尽管当前许多算法都称之为实时算法, 但是仍然受限于不同的硬件设备以及环境规模的大小。如何提高单目视觉 SLAM 算法各个组成部分的计算效率, 降低计算复杂度, 使之能够适应低成本、低功耗的计算平台, 是单目视觉 SLAM 发展的一个重要方向。

## 参考文献:

- [1] SMITH R, CHEESEMAN P. On the representation and estimation of spatial uncertainty[J]. The International Journal of Robotics Research, 1987, 5(4): 56-68.
- [2] KALMAN R E. A new approach to linear filtering and prediction problems[J]. Journal of Basic Engineering, 1960, 82(1): 35-45.
- [3] DAVISON A J. SLAM with a single camera[C]// Proceed-

- ings of Workshop on Concurrent Mapping and Localization for Autonomous Mobile Robots in Conjunction with ICRA. 2002.
- [4] DAVISON A J, MURRAY D W. Simultaneous localization and map-building using active vision[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(7): 865-880.
- [5] DAVISON A J, MAYOL W W, MURRAY D W. Real-time localisation and mapping with wearable active vision[C]// *Proceedings of 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality*. Washington, DC: IEEE, 2003: 315-316.
- [6] DAVISON A J. Real-time simultaneous localisation and mapping with a single camera[C]// *Proceedings of the Ninth IEEE International Conference on Computer Vision*. Washington, DC: IEEE, 2003: 1403-1410.
- [7] DAVISON A J, CID Y G, KITA N. Real-time 3D slam with wide-angle vision[C]// *Proceedings of IFAC Symposium on Intelligent Autonomous Vehicles*. Lisbon, 2004.
- [8] DAVISON A J, REID I D, MOLTON N D, et al. Monocular SLAM: Real-time single camera SLAM[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, 29(6): 1052-1067.
- [9] LEMAIRE T, LACROIX S, SOLÀ J. A practical 3D bearing-only SLAM algorithm[C]// *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Edmonton: IEEE, 2005: 2449-2454.
- [10] ROUSSILLON C, GONZALEZ A, SOLÀ J, et al. RT-SLAM: a generic and real-time visual SLAM implementation[M]// CROWLEY J L, DRAPER B A, THONNAT M, et al. *Computer Vision Systems*. Berlin: Springer, 2011: 31-40.
- [11] SIM R, ELINAS P, GRIFFIN M, et al. Vision-based SLAM using the Rao-Blackwellised particle filter[C]// *IJ-CAI Workshop on Reasoning with Uncertainty in Robotics*. Edinburgh, Scotland, 2005: 9-16.
- [12] PUPILLI M, CALWAY A. Real-time camera tracking using a particle filter[C]// *Proceedings of British Machine Vision Conference*. 2005: 519-528.
- [13] PUPILLI M, CALWAY A. Real-time camera tracking using known 3d models and a particle filter[C]// *IEEE International Conference on Pattern Recognition*. Washington, DC: IEEE, 2006, 1: 199-203.
- [14] PUPILLI M, CALWAY A. Particle filtering for robust single camera localisation[C]// *International Workshop on Mobile Vision*. 2006: 1-14.
- [15] LI Maohai, HONG Bingrong, CAI Zesu, et al. Novel Rao-Blackwellized particle filter for mobile robot SLAM using monocular vision[J]. *International Journal of Intelligent Technology*, 2006, 1(1): 63-69.
- [16] YAP T, LI Mingyang, MOURIKIS A I, et al. A particle filter for monocular vision-aided odometry[C]// *2011 IEEE International Conference on Robotics and Automation (ICRA)*. Shanghai: IEEE, 2011: 5663-5669.
- [17] ANDRADE-CETTO J, VIDAL-CALLEJA T, SANFELIU A. Unscented transformation of vehicle states in SLAM[C]// *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*. Barcelona: IEEE, 2005: 323-328.
- [18] MARTINEZ-CANTIN R, CASTELLANOS J A. Unscented SLAM for large-scale outdoor environments[C]// *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Edmonton: IEEE, 2005: 3427-3432.
- [19] MOURAGNON E, LHUILLIER M, DHOME M, et al. Monocular vision based SLAM for mobile robots[C]// *18th International Conference on Pattern Recognition*. Hong Kong: IEEE, 2006: 1027-1031.
- [20] KLEIN G, MURRAY D. Parallel tracking and mapping for small AR workspaces[C]// *6th IEEE and ACM International Symposium on Mixed and Augmented Reality*. Nara: IEEE, 2007: 225-234.
- [21] KLEIN G, MURRAY D. Improving the agility of keyframe-based SLAM[M]// FORSYTH D, TORR P, ZISSERMAN A, et al. *Computer Vision-ECCV 2008*. Berlin: Springer, 2008: 802-815.
- [22] EADE E, DRUMMOND T. Monocular SLAM as a graph of coalesced observations[C]// *IEEE 11th International Conference on Computer Vision*. Rio de Janeiro: IEEE, 2007: 1-8.
- [23] STRASDAT H, MONTIEL J M M, DAVISON A J. Scale drift-aware large scale monocular SLAM[J]. *Robotics: Science and Systems*, 2010, 2(3): 5.
- [24] HARRIS C, STEPHENS M. A combined corner and edge detector[C]// *Proceedings of the 4th Alvey Vision Conference*. 1988, 15: 147-151.
- [25] MORAVEC H P. *Obstacle avoidance and navigation in the real world by a seeing robot rover*[R]. Stanford: Stanford University, 1980.
- [26] MUNGUÍA R, GRAU A. Monocular SLAM for visual odometry[C]// *IEEE International Symposium on Intelligent Signal Processing*. Alcalá de Henares: IEEE, 2007: 1-6.
- [27] LEMAIRE T, BERGER C, JUNG I K, et al. Vision-based slam: Stereo and monocular approaches[J]. *International Journal of Computer Vision*, 2007, 74(3): 343-364.
- [28] GIL A, REINOSO O, MOZOS O M, et al. Improving data association in vision-based SLAM[C]// *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006: 2076-2081.
- [29] AVANZINI P, ROYER E, THUILOT B, et al. Using monocular visual SLAM to manually convoy a fleet of automatic urban vehicles[C]// *2013 IEEE International Conference on Robotics and Automation (ICRA)*. Karlsruhe: IEEE, 2013: 3219-3224.
- [30] SHI J, TOMASI C. Good features to track[C]// *1994 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 1994: 593-600.
- [31] SHEN S, MULGAONKAR Y, MICHAEL N, et al. Vision-

- based state estimation for autonomous rotorcraft MAVs in complex environments[C]//2013 IEEE International Conference on Robotics and Automation (ICRA). Karlsruhe: IEEE, 2013: 1758-1764.
- [32] ROSTEN E, DRUMMOND T. Machine learning for high-speed corner detection[M]// LEONARDIS A, BISCHOF H, PINZ A, et al. Computer Vision-ECCV 2006. Berlin: Springer, 2006: 430-443.
- [33] WILLIAMS B, KLEIN G, REID I. Real-time SLAM relocalisation[C]//IEEE 11th International Conference on Computer Vision. Rio de Janeiro: IEEE, 2007: 1-8.
- [34] WILLIAMS B, KLEIN G, REID I. Automatic relocalization and loop closing for real-time monocular SLAM[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(9): 1699-1712.
- [35] SONG S Y, CHANDRAKER M, GUEST C C. Parallel, real-time monocular visual odometry[C]//2013 IEEE International Conference on Robotics and Automation (ICRA). Karlsruhe: IEEE, 2013: 4698-4705.
- [36] LOWE D G. Object recognition from local scale-invariant features[C]// The Proceedings of the Seventh IEEE International Conference on Computer Vision, 1999. Kerkyra: IEEE, 1999, 2: 1150-1157.
- [37] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [38] CHEN C H, CHAN Y P. SIFT-based monocular SLAM with inverse depth parameterization for robot localization[C]// IEEE Workshop on Advanced Robotics and Its Social Impacts. Hsinchu: IEEE, 2007: 1-6.
- [39] FOLKESSON J, CHRISTENSEN H. Sift based graphical slam on a packbot[M]//Roland C L, ed. Field and Service Robotics: Results of the 6th International Conference. Berlin: Springer, 2008, 42: 317-328.
- [40] ZHAO L, HUANG S D, YAN L, et al. Large-scale monocular SLAM by local bundle adjustment and map joining[C]// 2010 11th IEEE International Conference on Control Automation Robotics & Vision (ICARCV). Singapore: IEEE, 2010: 431-436.
- [41] SÁNCHEZ J R, ALVAREZ H, BORRO D. Towards real time 3D tracking and reconstruction on a GPU using Monte Carlo simulations[C]//2010 9th IEEE International Symposium on Mixed and Augmented Reality (ISMAR). Seoul: IEEE, 2010: 185-192.
- [42] WARREN M, MCKINNON D, HE H, et al. Large scale monocular vision-only mapping from a fixed-wing sUAS[M]// YOSHIDA K, TADOKORO S, et al. Field and Service Robotics. Berlin: Springer, 2014, 92: 495-509.
- [43] YAN K, SUKTHANKAR R. PCA-SIFT: A more distinctive representation for local image descriptors[C]// Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE, 2004, 2: II-506-II-513.
- [44] BAY H, TUYTELAARS T, VAN GOOL L. SURF: Speeded up robust features[M]// LEONARDIS A, BISCHOF H, PINZ A, et al. Computer Vision-ECCV 2006. Berlin: Springer, 2006: 404-417.
- [45] ZHANG Z Y, HUANG Y L, LI C, et al. Monocular vision simultaneous localization and mapping using SURF[C]// 7th World Congress on Intelligent Control and Automation, 2008. Chongqing: IEEE, 2008: 1651-1656.
- [46] CIVERA J, GÁLVEZ-LÓPEZ D, RIAZUELO L, et al. Towards semantic SLAM using a monocular camera[C]// 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). San Francisco: IEEE, 2011: 1277-1284.
- [47] CALONDER M, LEPETIT V, OZUYSAL M, et al. BRIEF: Computing a local binary descriptor very fast[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(7): 1281-1298.
- [48] RUBLEE E, RABAUD V, KONOLIGE K, et al. ORB: an efficient alternative to SIFT or SURF[C]//2011 IEEE International Conference on Computer Vision (ICCV). Barcelona: IEEE, 2011: 2564-2571.
- [49] JUN L, PAN T S, TSENG K K, et al. Design of a monocular simultaneous localisation and mapping system with ORB feature[C]//Multimedia and Expo (ICME), 2013 IEEE International Conference on IEEE, 2013: 1-4.
- [50] LI J, PAN T S, TSENG K K, et al. Design of a monocular simultaneous localisation and mapping system with ORB feature[C]//2013 IEEE International Conference on Multimedia and Expo (ICME). San Jose: IEEE, 2013: 1-4.
- [51] LEMAIRE T, LACROIX S. Monocular-vision based SLAM using line segments[C]//2007 IEEE International Conference on Robotics and Automation. Roma: IEEE, 2007: 2791-2796.
- [52] SOLÀ J, VIDAL-CALLEJA T, DEVY M. Undelayed initialization of line segments in monocular SLAM[C]// IEEE/RSJ International Conference on Intelligent Robots and Systems. St. Louis: IEEE, 2009: 1553-1558.
- [53] CONCHA A, CIVERA J. Using superpixels in monocular SLAM[C]//2014 IEEE International Conference on Robotics and Automation (ICRA). Hong Kong: IEEE, 2014: 365-372.
- [54] WEISS S, SIEGWART R. Real-time metric state estimation for modular vision-inertial systems[C]//2011 IEEE International Conference on Robotics and Automation (ICRA). Shanghai: IEEE, 2011: 4531-4537.
- [55] JAMA M, SCHINSTOCK D. Parallel tracking and mapping for controlling VTOL airframe[J]. Journal of Control Science and Engineering, 2011: 413074.
- [56] DAVISON A J. Active search for real-time vision[C]// Tenth IEEE International Conference on Computer Vision. Beijing: IEEE, 2005, 1: 66-73.
- [57] CHLI M, DAVISON A J. Active matching[M]// FORSYTH D, TORR P, ZISSERMAN A, et al. Computer Vision-ECCV 2008. Berlin: Springer, 2008: 72-85.
- [58] CHLI M, DAVISON A J. Active matching for visual track-

- ing[J]. *Robotics and Autonomous Systems*, 2009, 57(12): 1173-1187.
- [59] CIVERA J, GRASA O G, DAVISON A J, et al. 1-point RANSAC for EKF-based structure from motion [C]// *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2009. St. Louis: IEEE, 2009: 3498-3504.
- [60] 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.
- [61] EADE E, DRUMMOND T. Scalable monocular SLAM [C]// *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. New York: IEEE, 2006, 1: 469-476.
- [62] GUERRA E, MUNGUIA R, BOLEA Y, et al. Validation of data association for monocular SLAM [J]. *Mathematical Problems in Engineering*, 2013: 1-11.
- [63] MUNGUÍA R, GRAU A. Monocular SLAM for visual odometry: A full approach to the delayed inverse-depth feature initialization method [J]. *Mathematical Problems in Engineering*, 2012: 676385.
- [64] CIVERA J, DAVISON A J, MONTIEL J M M. Inverse depth to depth conversion for monocular SLAM [C]// *2007 IEEE International Conference on Robotics and Automation*. Roma: IEEE, 2007: 2778-2783.
- [65] PIETZSCH T. Efficient feature parameterisation for visual SLAM using inverse depth bundles [C]// *EVERINGHAM M, NEEDHAM C, et al. Proceedings of the British Machine Vision Conference 2008*. BMVC Press, 2008: 1-10.
- [66] IMRE E, BERGER M O, NOURY N. Improved inverse-depth parameterization for monocular simultaneous localization and mapping [C]// *IEEE International Conference on Robotics and Automation*. Kobe: IEEE, 2009: 381-386.
- [67] FORSTER C, PIZZOLI M, SCARAMUZZA D. SVO: fast semi-direct monocular visual odometry [C]// *2014 IEEE International Conference on Robotics and Automation*. Hong Kong: IEEE, 2014.
- [68] STRASDAT H, MONTIEL J M M, DAVISON A J. Scale drift-aware large scale monocular SLAM [C]// *Proceedings of Robotics Robotics: Science and Systems*. 2010.
- [69] TANSKANEN P, KOLEV K, MEIER L, et al. Live metric 3d reconstruction on mobile phones [C]// *2013 IEEE International Conference on Computer Vision (ICCV)*. Sydney: IEEE, 2013: 65-72.
- [70] AZUMA R, BISHOP G. Improving static and dynamic registration in an optical see-through HMD [C]// *Proceedings of the 21st ACM Conference on Computer Graphics and Interactive Techniques*. New York: ACM, 1994: 197-204.
- [71] NÜTZI G, WEISS S, SCARAMUZZA D, et al. Fusion of IMU and vision for absolute scale estimation in monocular SLAM [J]. *Journal of Intelligent & Robotic Systems*, 2011, 61(1-4): 287-299.
- [72] JOHNSON A, WILLSON R, GOGUEN J, et al. Field testing of the mars exploration rovers descent image motion estimation system [C]// *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*. Barcelona: IEEE, 2005: 4463-4469.
- [73] LANG P, PINZ A. Calibration of hybrid vision/inertial tracking systems [C]// *Proceedings of the 2nd InerVis: Workshop on Integration of Vision and Inertial Sensors*. Barcelona, Spain, 2005.
- [74] 顾照鹏, 董秋雷. 基于部分惯性传感器信息的单目视觉同步定位与地图创建方法 [J]. *计算机辅助设计与图形学学报*, 2012, 24(2): 155-160.
- GU Zhaopeng, DONG Qiulei. Monocular SLAM based on partial inertial measurement unit information [J]. *Journal of Computer-Aided Design & Computer Graphics*, 2012, 24(2): 155-160.
- [75] LIU H, ZHOU Y L, GU Z P. Inertial measurement unit-camera calibration based on incomplete inertial sensor information [J]. *Journal of Zhejiang University: Science C*, 2014, 15(11): 999-1008.
- [76] FU S, LIU H F, GAO L F, et al. SLAM for mobile robots using laser range finder and monocular vision [C]// *14th International Conference on Mechatronics and Machine Vision in Practice*. Xiamen: IEEE, 2007: 91-96.
- [77] ZHANG Q, PLESS R. Extrinsic calibration of a camera and laser range finder (improves camera calibration) [C]// *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Sendai: IEEE, 2004, 3: 2301-2306.
- [78] JEONG W, LEE K M. CV-SLAM: A new ceiling vision-based SLAM technique [C]// *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Edmonton: IEEE, 2005: 3195-3200.
- [79] GEIGER A, ZIEGLER J, STÍLLER C. StereoScan: Dense 3d reconstruction in real-time [C]// *2011 IEEE Intelligent Vehicles Symposium (IV)*. Baden-Baden: IEEE, 2011: 963-968.

#### 作者简介:



顾照鹏,男,1981年生,博士后,IEEE会员,主要研究方向为单目视觉SLAM、移动机器人视觉导航。发表学术论文10余篇,被EI、SCI检索10篇。



刘宏,男,1967年生,教授,博士生导师,IEEE会员,国际机器人足球联盟(FI-RA)竞赛委员会委员,中国人工智能学会常务理事、副秘书长、青年工作委员会主任,中国自动化学会智能自动化专业委员会委员。主要研究方向为模式识别、智能信息处理与计算机视觉。发表学术论文100余篇,其中被EI、SCI检索50余篇。