

基于视觉的移动机器人同时定位与建图研究进展

孙凤池¹, 黄亚楼¹, 康叶伟²

(1. 南开大学 软件学院, 天津 300071; 2. 南开大学 信息技术科学学院, 天津 300071)

摘要: 同时定位与建图(SLAM)是实现移动机器人真正自治的必要前提, 视觉传感器由于能够提供丰富的环境信息而在SLAM研究中受到重视, 本文从视觉传感器配置方式、视觉特征提取方法、视觉SLAM实现机制、地图表示类型以及环境对视觉SLAM的影响五个方面综述基于视觉传感器的同时定位与建图研究的发展现状, 对已有的典型视觉SLAM方法进行分析 and 比较, 并展望了未来的发展趋势。

关键词: 移动机器人; 同时定位与建图; 视觉传感器; 特征提取

中图分类号: TP242 **文献标识码:** A

Review on the achievements in simultaneous localization and mapping for mobile robot based on vision sensor

SUN Feng-chi¹, HUANG Ya-lou¹, KANG Ye-wei²

(1. College of Software, Nankai University, Tianjin 300071, China;

2. College of Information Technical Science, Nankai University, Tianjin 300071, China)

Abstract: Simultaneous localization and mapping (SLAM) are necessary prerequisites to make mobile robot truly autonomous. The vision sensor attracts more and more attention in the SLAM research for its rich information of the environment. This paper summarizes the state of the art in the vision-based SLAM from the aspects such as the deployment of vision sensors, the extraction of visual features, the implementation mechanism of visual SLAM, the representation of the map and the effect of environment on visual SLAM. Some classical visual SLAM methods are compared and analyzed. The research prospects are described.

Key words: mobile robot; simultaneous localization and mapping; vision sensor; feature extraction

1 引言(Introduction)

智能移动机器人可以代替人到某些特殊场合进行工作, 在军事、航天工业、服务业等许多领域具有广泛的应用前景。自主导航是智能移动机器人的一项首要功能。真正的自主导航要求机器人能够在一个完全未知的环境中探索, 建立环境地图, 进行精准定位, 以及执行相应的规划任务。长期以来, 建图研究与定位研究独立进行, 研究定位时需要先验地图指导, 研究建图时则假定机器人位姿已知。1987年Smith提出了一种同时解决二者的概率方法^[1], 揭开了同时定位与建图(simultaneous localization and mapping, SLAM)研究的序幕, SLAM研究最近一二十年取得了长足的进展。一方面归功于好的SLAM算法, 另一方归功于好的传感器, 特别是激光传感器。激光传感器能精确地测量机器人相对于被观测对象的相对距离和角度, 并且测量数据的噪

声比较直观。但是, 激光传感器由于体积大、能耗高, 不适合小型机器人使用。此外, 激光传感器的感知分辨能力低, 很难区分不同路标以及将新观测的特征与已观测到的特征关联起来, 并且在高度混乱或者拥挤的环境很难提取角和直线特征。同时, 激光传感器还具有感知漂移问题。超声传感器也被应用于一些SLAM方法中, 其价格虽然便宜^[2], 但是检测频率低、观测范围短、对反射以及声波入射角偏差敏感, 会导致重复检测以及镜面反射问题, 不适合高精度建图场合。随着微电子技术的进步, 视觉传感器性价比迅速提升, 在SLAM领域得到了广泛应用。视觉传感器能够获取的信息丰富, 容易提取环境中的多种特征信息, 如角点、直线边^[3]、SIFT^[4]特征等, 而且适合三维空间应用。

文[5,6]针对一般SLAM问题的研究现状进行了综述, 本文则综述基于视觉传感器进行同时定位

收稿日期: 2008-07-26; 收修改稿日期: 2009-08-05。

基金项目: 国家自然科学基金资助项目(60605021, 60805031); 国家高技术研究发展计划资助项目(2006AA04Z223); 教育部高等学校博士点基金新教师资助项目(200800551015)。

与建图的研究进展,按照视觉SLAM所涉及的基础理论和关键技术,从视觉传感器配置方式、视觉特征提取方法、SLAM实现机制、地图表示类型以及环境对视觉SLAM的影响等方面系统总结了视觉SLAM的研究现状,对已有的典型视觉SLAM方法进行了分析和比较,并展望了视觉SLAM的未来发展趋势。

2 视觉传感器的配置方式(Deployment of vision sensors)

用于SLAM研究的视觉传感器有多种配置方式,包括:单目视觉、双目立体视觉、多目立体视以及全景视觉等,此外,视觉传感器与激光、超声等传感器组合使用也受到重视。

2.1 单目视觉(Monocular vision)

与多目视觉或者立体视觉相比,单目视觉^[7~9]在方向测量中不确定性较大,而且深度恢复比较困难,但是它具有计算负荷低、灵活、价格低等特点,因此有很多SLAM研究基于单目视觉开展。基于单目视觉的SLAM在一次观测中只能获得方向信息,无法提取距离信息,是一种仅有方向信息(bearing-only)的方法,因此存在特征初始化问题。特征初始化是体现不同Bearing-only SLAM方法的一个重要方面。目前有两类初始化方法,一类是延迟初始化,一类是非延迟初始化。

最简单的延迟初始化将传感器观测延迟 N 次,然后处理 N 次观测信息,在获得完全的特征观测信息后再进行初始化^[10~12]。基于粒子滤波器的方法也是一种常用的方法^[13,14],开始时将特征初始化为一条3D直线,在直线上一定范围内均匀分布一些粒子表示深度假设或者深度信息概率密度函数。在随后的步骤中,每一种特征深度假设以不确定性椭圆形式投影到图像上,以此来检验并估计该深度信息正确的可能性,之后粒子基于贝叶斯计算权值规则(bayesian reweighting)重新分布,当深度分布能够很好地被高斯分布表示时,该直线变为具有适当高斯不确定性的特征点。这种方法计算负荷大,粒子的初始分布需要覆盖整个路标的深度范围,如果检测的路标多了,则比较困难。

非延迟初始化方法有基于GSF滤波器、近似GSF和变形GSF等几种。GSF滤波器采用的每一种深度假设都对应一个地图,最终地图的概率密度函数是各个高斯地图的加权和,这种方法计算代价比较高。基于近似GSF的方法,将不同的深度假设对应不同的路标,所有的假设与同一个地图相关,计算复杂度大大降低,但是这种假设会导致地图不一致。

由于单目视觉提取深度信息的困难,室内环境的

视觉SLAM中,一些先验知识被应用进来简化算法的实现。文[15]通过将视觉指向室内环境的天花板,提取天花板上的视觉特征,由于地面平滑,视觉距离天花板的距离已知,从而简化了特征的初始化过程。

2.2 双目立体视觉(Binocular stereo vision)

双目视觉在SLAM中的应用更为广泛,在这种配置下,可以通过外极线几何约束匹配左右摄像机的特征,从而能够在既定帧速率的前提下提供场景的深度信息,提取完整的特征信息(机器人和特征之间的距离和方向角),方便特征初始化。但是,立体视觉由于基线固定,虽然能够避免单目视觉微小运动造成位姿估计误差大的问题,但是难以重建远距离的物体。

除了常规的立体视觉外,文[16]采用主动双目立体视觉进行SLAM研究,实际上是一种主动SLAM方法。但是,不同于一般的主动SLAM(需要有选择地改变机器人的路径来提高地图估计精度),该文假定机器人的运动轨迹已知或者由一些目标驱动的过程提供,从而通过控制主动立体摄像头运动,在当前可见的特征中选择那些最能有效利用当前资源的特征。文[10]虽然是双目立体视觉,但是滤波器更新时只使用来自一个视觉的方向信息,来自立体视觉的初始距离信息只是用来为估计的新特征位置提供初始假设。

2.3 多目立体视觉(Stereo vision with multi-camera)

三目视觉比双目视觉更能提高准确性和鲁棒性,在视觉SLAM中也被应用^[3,17],这种方式在实现方面与双目立体视觉没有本质区别,只是多了一次匹配,只有被三个摄像机同时检测到的特征点才有可能成为路标。

多目视觉被用于获取更大的视野,文[18]采用了环绕一个平面四周分布8个摄像头的配置,并将SLAM问题转换为一个似然估计问题,地图和轨迹更新用固定滞后平滑算法,与小视野的视觉相比,采用大视野视觉的SLAM方法的优点是适合大规模环境,并且能够提高环境重建质量。

2.4 全景视觉(Panoramic vision)

全景视觉能提供更大的视野、远距离跟踪特征,因此能够适应具有远距离特征的场合,这是立体视觉做不到的。远距离SLAM容易产生不一致的位置估计,因此具有与位置估计独立的闭循环(loop closing)检测机制很重要,全景视觉很适合这一应用,因为它容易用图像索引技术解决这一问题。基于全景视觉的SLAM中,文[19]是第一个有效探索全景视觉SLAM的方法,通过全局图像与特征数据库建立

索引可以不依赖机器人和路标的位置估计而有效检测闭循环. 文[20]提出一种基于全景视觉的移动机器人同步定位与地图创建方法, 该方法提取颜色区域作为视觉路标, 在分析全景视觉成像原理和视觉不确定性的基础上建立起系统的观测模型, 定位出路标位置, 进而通过卡尔曼滤波同步更新机器人位置和地图信息, 该方法在建立环境地图的同时可以有效地修正由里程计造成的累积定位误差.

2.5 视觉与其它传感器的组合应用(The application of vision sensor in combination with other sensors)

对于一些复杂环境或者为了提高传感器的鲁棒性, 需要多传感器融合使用. 在SLAM中视觉传感器经常与激光传感器^[21, 22]、超声传感器组合使用. 文[21]用激光传感器建立3D点云地图, 用视觉传感器检测闭循环. 文[23]用激光传感器和单目视觉分别提取水平环境特征和垂直物体边缘特征, 从而基于扩展卡尔曼滤波器架构完成平面环境下的室内SLAM. 文[24]中声纳返回的信息则被用来初始化SLAM地图.

3 视觉特征提取方式(Methods of extraction of visual feature)

机器人在未知环境运动, 同一个路标可能被多次观测, 因此路标的可靠性尤为重要, 可靠的路标能够简化SLAM算法中的数据关联问题, 不可靠的路标则会导致数据关联的错误, 最终可能导致建图不一致或者滤波器发散. 如何应用视觉传感器获得对光照及观测视角变化、几何及仿射变换不敏感、处理代价小的特征作为路标, 是成功进行视觉SLAM的前提. 此外, 好的视觉特征应该具备具有精确的图像位置、便于精确建图、在连续的图像中容易跟踪的特点.

点特征在图像上突出, 具有好的不变性, 容易提取和跟踪, 计算代价小, 从而成为视觉路标的首选^[19]. Harris^[25], SIFT^[4], KLT^[26]是视觉SLAM中最常用的点特征检测算法.

Harris角检测子(detector)^[25]是视觉SLAM方法中常用的点特征检测方法, 它基于Moravec^[27]的思想, 即感兴趣的点对应于图像窗口四个方向上(水平、垂直和两个对角线方向)最大的梯度强度. Harris和Stephens应用图像微分扩展了Moravec检测子^[27], 用Gaussian平滑算子确定图像位置, 检测算子具有旋转不变性. 但是, Harris角点对图像伸缩变化敏感. 尺度自适应的改进Harris算法在文[28]被提出来, 然而, 如果不能获得尺度变化信息, 则特征匹配相当耗时, 因为尺度变为一个附加的搜索维度. 目前

所有基于Harris检测子提取视觉特征的SLAM算法, 基本都是非尺度变化的Harris算法. Harris特征的匹配一般采用NSSD(normalized sum of squares difference)^[16]和SSD(sum of squares difference)^[18]相关性方法.

在具有尺度伸缩变化的场合, SIFT(scale invariant feature transformation)^[4]成为首选的特征提取手段. SIFT首先检测由高斯差分DoG(difference of Gaussians)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. SIFT特征具有对图像平移、缩放和旋转不变性以及部分地对照明变化和局部图像变形不变性的特点, 在许多局部特征描述中性能最突出, 因而在基于视觉的SLAM研究中得到了广泛应用.

SIFT采用不同尺度空间上的DoG塔形结构上的极值点作为特征位置. DoG产生类似斑点(Blob)的图像结构, 方便物体识别, 但对于定位和重构精度不高. 因此一些改进的方法被提出来. 文[11]采用Harris-Laplace算法决定尺度空间上塔形结构的极值点, 这样使得特征在空间上更加精确, 便于结构重建和定位应用, 而且用Harris-Laplace平均产生的特征点数为DoG方法的1/4, 有利于快速计算. 文[12]同样采用Harris-Laplace决定尺度空间上塔形结构的极值点, 除此之外, 文中的SIFT方法保留梯度直方图的原样(原始SIFT在平滑的梯度直方图的极值点上分配归一化的方向), 从而避免了具有较大旋转差异的角点特征描述子相似的问题. 文[29]基于不同尺度的匹配比率分析了从特定的尺度空间提取SIFT特征, 不仅提高了SIFT特征的提取速度, 也大大减少了错误匹配.

一般SIFT特征是一个128维的向量, 其计算代价相对较高, 在SLAM应用中Kd树常被用来组织SIFT特征, 加速特征的匹配与检索. 文[17]为了减小计算代价, 在不影响建图质量的情况下, 只采用16维的SIFT检测子.

KLT^[26]算法是在一个包含了特征纹理的窗口内根据图像间的灰度差的平方和来实现特征跟踪的算法. KLT假定 $t + \tau$ 时刻获得的图像可以由 t 时刻的图像中每一个像素都移动一定量而获得, 算法的目的就是计算这个移动量. 容易发现, KLT算法不适合大基线运动的图像特征匹配. 在文[14, 15, 30]的SLAM方法中均用KLT算法检测路标特征.

此外, 不同检测子和描述子对的配合被应用于检

测特征. 文[21]使用Harris affine检测子与SIFT描述子配合检测到的图像特征, 完成SLAM中的闭环检测.

由于视觉特征的丰富性, 除了点特征之外, 颜色特征也被应用到SLAM算法中. 文[13]用归一化的颜色信息作为特征, 并采用归一化的强度分布, 从而对图像运动具有不变性, 但易受尺度、观测角、遮挡的影响. 在结构化环境, 直线特征比较常见, 基于直线构建的地图能够提供环境结构的相关信息, 而且直线匹配比较可靠, 一些视觉SLAM方法利用了这一特征. 文[23]在室内环境的SLAM中采用直线边作为特征, 文[3]用Canny算法检测到的线段作为特征.

4 SLAM实现机制(Implementing mechanism of SLAM)

SLAM通常被作为一个后验概率估计问题来解决, 扩展卡尔曼滤波器和粒子滤波器是常被采用的两种实现方式. 此外, 由于视觉领域的从运动恢复结构(structure from motion, SFM)与机器人领域的SLAM具有相似之处, 随着SFM技术在速度、准确性、效率方面取得了巨大进步, SFM在基于视觉的SLAM算法中得到了应用.

4.1 扩展卡尔曼滤波器(Extended Kalman filter)

扩展卡尔曼滤波器是卡尔曼滤波器在非线性系统的推广, 基于卡尔曼滤波的SLAM算法有严格的限制条件: 首先, 系统模型必须是线性的; 其次, 运动模型和观测模型概率分布符合单峰正态分布; 此外, 地图必须是由可以区分的路标表示. 基于卡尔曼滤波器的方法具有简单、收敛性能够严格证明的优点, 但是系统模型的线性假设在很多情况下不成立.

而基于扩展卡尔曼滤波器(EKF)的SLAM算法虽然能够应用于非线性的运动和观测模型, 但是单峰正态分布的假设仍然是必要前提.

应用视觉的SLAM方法中, 基于EKF的有很多^[10, 11, 14~17], 但由于视觉特征一般比较多, 而基于EKF的SLAM算法的复杂度是特征数量的平方级, 所以很难适应大规模环境. 因此目前这类SLAM算法的应用环境规模都比较小.

4.2 粒子滤波器(Particle filter)

粒子滤波器是一种序列蒙特卡洛滤波方法, 使用一些采样(粒子)集合逼近SLAM的后验概率. 在机器人应用领域, 粒子滤波器能够表示任何可用马尔可夫链表示的概率机器人模型, 而且具有容易实现的优点. 但是, 高维空间的概率分布需要大量的粒子(随分布的维数指数级增长), RBPF(Rao-blackwellised particle filter)^[31]滤波器被用来解决这个问题. 它将状态空间分离为独立的部分, 并边缘

化该部分的一个或者多个组成成分, 从而提供了一种简化概率估计的机制. RBPF对具有大量路标的地图可伸缩性强, 能够更有效地解决SLAM中的数据关联错误问题. FastSLAM^[32]是RBPF滤波器最为成功的实例, 也是应用最为广泛的SLAM方法. 应用视觉传感器建图, 由于检测到的特征多, 这种对于每个特征都要和地图中的所有特征比较的方法, 数据关联比较困难, 这种情况下允许多样假设数据关联的FastSLAM会造成粒子滤波器的鲁棒性下降, 文[33]实现了基于视觉的FastSLAM以及快速的数据关联, 为了提高了对大量SIFT特征的存储和访问的效率, 作者利用Kd树来管理SIFT特征, 以对数时间复杂度实现了特征之间的最近邻匹配, 使得一帧信息的数百个特征可以与数据库中数千个特征快速关联.

4.3 SFM技术的应用(The application of SFM)

从运动中恢复结构(SFM)是一个同时恢复摄像机运动轨迹和场景结构(场景特征在空间的描述)的过程, 在视觉领域也是一个“鸡生蛋, 蛋生鸡”问题, 即如果已知获取图像时摄像机的位置, 那么场景重构就是一个三角测量的问题; 反过来, 如果已知场景结构, 那么应用三角测量技术就能够定位摄像机的位置. 通常, 在SFM问题中, 场景结构和摄像机位置先验知识都未知. SFM技术可以分为两类: 一类是迭代的或者基于卡尔曼滤波器的; 另一类是同时处理多幅图像, 利用图像间内在的多视角几何约束同时估计3D结构和摄像机位置, 与迭代方法相比, 本类方法不需要明确的摄像机运动模型, 只需要几幅不同角度的场景图像, 不需要视频流.

SFM与SLAM目标一致, 但是二者并不相同, 首先SFM通常离线运行, 批量处理顺序获取的所有图像, 而SLAM要求迭代计算方法以适合在线实时运行, 其次, SFM不需要在SLAM中经常使用的里程计的反馈信息.

文[30]采用基于EKF方法的SFM技术估计摄像机运动. SFM用于SLAM的主要问题之一是精度依赖基线距离, 文[34]将基线自适应选择的SFM技术应用于SLAM过程, 提高了算法的估计精度.

5 地图类型(The types of map)

SLAM问题中, 地图如何表示是一个重要问题, 在机器人研究中, 环境地图的表示目前主要有三种方法, 分别是尺度地图(metric map)、拓扑地图(topologic map)、以及尺度地图和拓扑地图的混合地图.

栅格地图和特征地图是最常用的尺度地图. 栅格地图将环境分割为若干小的固定大小的区域, 并用确定值或概率值表示障碍物对栅格的覆盖情况.

栅格地图能够满足高层导航任务的需要, 比如路径规划、避碰等. 但是, 对于大规模环境, 这种方式计算和存储代价高.

特征地图用参数特征的全局位置表示环境, 参数特征可以是点或者是直线等. 环境由少数参数描述的特征构成, 因此环境表示非常有效. 但是, 由于环境的稀疏表示, 一般基于特征的地图不利于路径规划^[19]. 此外, 这种方式的地图仅适合能被基本几何特征模型表示的环境.

拓扑地图由表示环境特定位置的点和不同位置之间的边(路径)表示, 通常用图结构描述. 这种方式只维护不同特征位置的连接关系, 并不维护它们之间的尺度信息. 拓扑地图对环境规模适应性强、表示紧凑、占用空间小, 适合快速的导航和建图算法, 主要不足是位置之间可靠的导航, 以及位置识别.

混合地图具有拓扑地图和尺度地图的优点, 本质上是拓扑架构, 只不过是位置定义或者路径定义包含尺度信息, 这意味着位置点不再限制在离散的区域, 同时它能够像尺度地图一样描述任意大小和形状的区域.

目前, 绝大部视觉SLAM方法建立由点特征构成的环境地图^[3, 19, 34], 少数方法生成其它类型的地图. 文[35]通过图像外观特征可以识别的3D自然路标学习地图, 栅格地图作为这一算法的附属物获得, 文[36]的SLAM方法中生成GVG(generalized voronoi graph)形式的拓扑地图.

6 环境对视觉SLAM的影响(The effect of environment on visual SLAM)

随着SLAM研究的深入, 应用环境逐步从小规模、结构化、静态环境向大规模、非结构化、动态环境发展, 这些更为复杂的外部环境对SLAM研究的算法和关键技术有了更高的要求.

6.1 大规模环境下的视觉SLAM(Visual SLAM in large scale environment)

计算量大是大规模环境SLAM面临的主要难题, 利用分层或分块处理的方法是当前常用的解决方案. 文[37]利用普通广角立体摄像机得到的环境信息, 把全局地图划分成由SIFT特征确认的子地图, 在子地图水平, 使用贝叶斯方法进行低层次的机器人定位和路标建图; 在子地图层次之上增加了一个高层SLAM, 基于子地图的SIFT标记实现关联, 这种分层机制在大规模环境SLAM实验中获得了良好的实时效果.

对于大规模环境, 单机器人由于受到自身处理能力以及电源的限制很难高效建立环境地图, 这种情况下多机器人共同协调探索环境是一种显而易见

的方式, 因为多个机器人通过共享信息可以使得定位更加高效, 此外, 多机器人可以获得环境的冗余信息, 比单个机器人的容错能力更强.

文[38]中每个机器人配置一个单目视觉, 多个机器人共同完成SLAM任务. 初始时假设机器人相对位姿未知, 路标未知, 每个机器人从一个未知位置开始SLAM, 当两个机器人发生重叠时, 它们之间的联合地图建立. 从而使得两个机器人可以在联合地图中定位自身. 只要每个机器人至少和其它机器人重叠一次, 就可以建立整个团队的联合地图. 文[39]本质上是一种基于立体视觉的SLAM方法, 只是两个摄像机分别位于两个机器人上, 从而使得摄像机运动灵活, 建图更快.

文[39]介绍了一种变种的多机器人SLAM方法, 与常规多机器人SLAM不同, 该方法中一些机器人被用来作为路标. 用来自多个机器人的里程计数据和观测数据建立未知环境的地图. 其中用作路标的机器人一旦位置确定则静止不动, 其它机器人自由运动, 实验表明静止机器人的数量增加, 可以提高SLAM的精度, 而自由机器人对精度贡献不大.

6.2 非结构化环境下的视觉SLAM(Visual SLAM in unstructured environment)

针对难以从中提取路标的非结构化大规模环境, 文[40]提出了基于路标动态配置的主动SLAM方法, 实时计算配置动态路标对定位与建图的影响, 并将其引入到优化目标中, 使机器人根据最优化目标函数计算出符合探索需要的控制输入, 进而通过主动的运动方式动态在线配置路标, 并利用配置的路标完成定位与建图.

非结构环境的不规则性会大大提高对环境采样的需求量, 并导致SLAM计算量的剧增. 文[41]针对未知非结构化环境, 基于摄像机和激光传感器, 提出了一个具有较小计算代价的SLAM方案. 通过一个智能的算法让机器人从收集到的传感器数据中选择具有最大信息量的观测用于估计过程, 从而能够应对从真实的非结构化环境中采样得到的大量数据.

6.3 动态环境下的视觉SLAM(Visual SLAM in dynamic environment)

目前的SLAM方法大多假定机器人处在一个静态环境中, 需要环境的部分先验信息, 并依靠数据关联建立路标地图, 如果在SLAM过程中环境发生动态变化, 则容易引起数据关联错误, 从而导致SLAM发散. 针对这类问题的研究成果目前较少见, 但是可以预知的是, 传感信息的合理、充分应用是解决这个问题的可能途径. 文[42]引入遗传算法的概念, 扩展了适用于静态环境SLAM中的RBPF粒

子滤波器, 利用原始的测量数据恢复传感器信息, 从而在粒子地图中高效率地确定运动障碍物, 可以在动态环境中可靠的运行. 文[22]使用视觉传感器搭配激光传感器实现对未知环境中动态物体的识别和定位, 使用SIFT特征匹配识别物体, 利用激光测距实现运动目标定位.

7 结论与展望(Conclusions and prospects)

目前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. 从当前的研究看, 视觉SLAM研究的室内应用环境一般不超出 $15\text{ m} \times 16\text{ m}$ 的环境^[10, 14, 16], 基于单目视觉的3D SLAM方法适应的环境规模更小^[18], 而基于激光传感器的SLAM方法已经在约 $250\text{ m} \times 250\text{ m}$ 的环境成功建图^[35]. 应用于室外环境的视觉SLAM研究较少^[3, 13, 21], 环境规模也不大. 已有SLAM算法大多针对静态环境, 高度动态环境下的SLAM仍是一个难题, 而实际的环境往往不是静态的. 可以预知, 动态环境下, 特别是非结构化大规模动态环境下的视觉SLAM, 是一个具有挑战性的重大问题.

特征地图是基于视觉SLAM方法的主要环境表示形式, 其特征主要是“点”特征, 然而现实环境中可用的信息丰富, 一方面需要进一步挖掘新的特征表示形式, 另一方面需要不断探索性能良好且提取快速的视觉特征作为路标.

在SLAM研究中, 机器人的运动路径对最终的建图质量有重大影响, 当机器人自主探索未知环境时, 如何把探索策略和SLAM算法结合(即主动SLAM)是一个重要问题, 基于信息驱动的探索策略是当前主动SLAM的主要方式, 与基于激光传感器的研究相比, 基于视觉的主动SLAM研究相对滞后, 需要更多的关注.

实时性是SLAM算法的永恒话题, 尽管当前许多基于视觉的SLAM算法都称为实时算法^[14, 15, 28, 30], 但是适应的环境规模小. 对于基于视觉的SLAM方法, 如何高效地检测、匹配和管理视觉特征是影响其实时性的重要因素, 离线算法^[34, 35]和接近实时的算法^[21]都不是SLAM追求的目标. 视觉领域提出许多特征检测算法, 机器人领域存在许多SLAM实现机制, 在以视觉为主导的应用中如何将两者结合起来以进一步提高实时性还需要研究. SLAM技术要真正达到实用地步, 适应大规模环境是必须解决的问题, 而实时性在其中扮演着重要的角色.

当前基于视觉的SLAM主要产生稀疏的三维点特征地图, 适合机器人定位, 但难以直接应用于机器人路径规划以及人观测. 充分利用视觉传感器信息丰富的特性, 研究适合高层导航任务进行路径规划

的环境地图生成方式, 解决当前视觉SLAM单一特征地图的不足, 具有重要的意义.

参考文献(References):

- [1] SMITH R, SELF M, CHEESEMAN P. A stochastic map for uncertain spatial relationships[C] // *Proceedings of the 4th International Symposium on Robotics Research*. Cambridge, MA: MIT Press, 1987: 467 – 474.
- [2] TARDOS J D, NEIRA J, NEWMAN P, et al. Robust mapping and localization in indoor environments using sonar data[J]. *International Journal of Robotics Research*, 2002, 21(4): 311 – 330.
- [3] DAILEY M N, PARNICKUN M. Landmark-based simultaneous localization and mapping with stereo vision[C] // *Proceedings of the 2005 Asian Conference on Industrial Automation and Robotics*. Bangkok, Thailand: Springer LNCS, 2005: 108 – 113.
- [4] LOWE D G. Object recognition from local scale-invariant features[C] // *Proceedings of the IEEE International Conference on Computer Vision*. Corfu, Greece: IEEE, 1999: 1150 – 1157.
- [5] 罗荣华, 洪炳镕. 移动机器人同时定位与地图创建研究进展[J]. 机器人, 2004, 26(2): 182 – 186.
(LUO Ronghua, HONG Bingrong. The progress of simultaneous localization and mapping for mobile robot[J]. *Robot*, 2004, 26(2): 182 – 186.)
- [6] 陈卫东, 张飞. 移动机器人的同步自定位与地图创建研究进展[J]. 控制理论与应用, 2005, 22(3): 455 – 460.
(CHEN Weidong, ZHANG Fei. Review on the achievements in simultaneous localization and map building for mobile robot[J]. *Control Theory & Applications*, 2005, 22(3): 455 – 460.)
- [7] LUPTON T, SUKKARIEH S. Removing scale biases and ambiguity from 6 DOF monocular salm using inertial[C] // *Proceedings of the IEEE International Conference on Robotics and Automation*. Pasadena, USA: IEEE, 2008: 3698 – 3703.
- [8] ANGELI A, DONCIEUX S, MEYER J A, et al. Incremental vision-based topological SLAM[C] // *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*. Nice, France: IEEE, 2008: 1031-1036.
- [9] CIVERA J, DAVISON A J, MONTIEL J M M. Interacting multiple model monocular SLAM[C] // *Proceedings of the IEEE International Conference on Robotics and Automation*. Pasadena, USA: IEEE, 2008: 3704 – 3709.
- [10] MIRO J V, DISSANAYAKE G, ZHOU W Z. Vision-based SLAM using natural features in indoor environments[C] // *Proceedings of the IEEE International Conference on Intelligent Networks, Sensor Networks and Information Processing*. Melbourne, Australia: IEEE, 2005: 151 – 156.
- [11] JENSEFELT P, KRAGIC D, FOLKESSON J, et al. A framework for vision based bearing only 3D SLAM[C] // *Proceedings of the IEEE International Conference on Robotics and Automation*. Orlando, USA: IEEE, 2006: 1944 – 1950.
- [12] JENSEFELT P, FOLKESSON J, KRAGIC D, et al. Exploiting distinguishable image features in robotic mapping and localization[C] // *Proceedings of European Robot Symposium*. Heidelberg, Germany: Springer-Verlag, 2006: 143 – 158.
- [13] FITZGIBBONS T, NEBOT E. Bearing-only SLAM using colour-based feature tracking[C] // *Proceedings of Australian Conference on Robotics and Automation*. Auckland. New Zealand: Springer-Verlag, 2002: 234 – 239.
- [14] DAVISON A J. Real-time simultaneous localization and mapping with a single camera[C] // *Proceedings of IEEE International Conference on Computer Vision*. Nice, France: IEEE, 2003: 1403 – 1410.

- [15] JEONG W, LEE K M. CV-SLAM: a new ceiling vision-based SLAM technique[C] // *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*. Edmonton, Canada: IEEE, 2005: 3070 – 3075.
- [16] DAVISON A J, MURRAY D. Simultaneous localization and map-building using active vision[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(7): 865 – 880.
- [17] SE S, LOWE D G, LITTLE J J. Vision-based global localization and mapping for mobile robots[J]. *IEEE Transactions on Robotics*, 2005, 21(3): 364 – 375.
- [18] KAESSE M, DELLAERT F. Visual SLAM with a multi-camera rig[R]. Technical report, Georgia Institute of Technology. 2006, GIT-GVU-06-06.
- [19] LEMAIRE T, LACROIX S. SLAM with panoramic vision[J]. *Journal of Field Robotics*, 2007, 24(1/2): 91 – 111.
- [20] 许俊勇, 王景川, 陈卫东. 基于全景视觉的移动机器人同步定位与地图创建研究[J]. *机器人*, 2008, 30(4): 289 – 297.
(XU Junyong, WANG Jingchuan, CHEN Weidong. Omni-vision-based simultaneous localization and mapping of mobile robots[J]. *Robot*, 2008, 30(4): 289 – 297.)
- [21] NEWMAN P, COLE D, HO K. Outdoor SLAM using visual appearance and laser ranging[C] // *Proceedings of the IEEE International Conference on Robotics and Automation*. Orlando, USA: IEEE, 2006: 1180 – 1187.
- [22] LOPEZ D G, SJO K, PAUL C, et al. Hybrid laser and vision based object search and localization[C] // *Proceedings of the IEEE International Conference on Robotics and Automation*. Pasadena, USA: IEEE, 2008: 2636 – 2643.
- [23] 庄严, 王伟, 王珂, 等. 移动机器人基于激光测距和单目视觉的室内同时定位和地图构建[J]. *自动化学报*, 2005, 31(6): 925 – 933.
(ZHUANG Yan, WANG Wei, WANG Ke, et al. Mobile robot indoor simultaneous localization and mapping using laser range finder and monocular vision[J]. *Acta Automatica Sinica*, 2005, 31(6): 925 – 933.)
- [24] WILLIAMS S, MAHON I. Simultaneous localisation and mapping on the great barrier reef[C] // *Proceedings of the IEEE International Conference on Robotics and Automation*. New Orleans, USA: IEEE, 2004: 1771 – 1776.
- [25] HARRIS C, STEPHENS M. A combined corner and edge detector[C] // *Proceedings of the 4th Alvey Vision Conference*. Manchester, UK: Springer-Verlag, 1988: 147 – 151.
- [26] SHI J, TOMASI C. Good feature to track[C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Berlin, Germany: IEEE, 1993: 593 – 600.
- [27] MORAVEC H P. Towards automatic visual obstacle avoidance[C] // *Proceedings of International Joint Conference on Artificial Intelligence*. Cambridge, MA: AAAI Press, 1977: 584 – 590.
- [28] DUFOURNAUD Y, SCHMID C, HORAUD R P. Matching images with different resolutions[C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. South Carolina, USA: IEEE, 2000: 612 – 618.
- [29] WANG X, ZHANG H. Good image features for bearing-only SLAM[C] // *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*. Beijing, China: IEEE, 2006: 2576 – 2581.
- [30] MELTZER J, GUPTA R, YANG M H, et al. Simultaneous localization and mapping using multiple view feature descriptors[C] // *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*. Sendai, Japan: IEEE, 2004: 1550 – 1555.
- [31] DOUCET A, FREITAS J D, MURPHY K, et al. Rao-blackwellized particle filtering for dynamic bayesian networks[C] // *Proceedings of the 8th Conf on Uncertainty in Artificial Intelligence*. San Francisco: Morgan Kaufman, 2000: 499 – 516.
- [32] MONTEMERLO M. *FastSLAM: A factored solution to the simultaneous localization and mapping problem with unknown data association*[D]. Pittsburgh, USA: Carnegie Mellon University, 2003.
- [33] SIM R, GRIFFIN M, SHYR A, et al. Scalable real-time vision-based SLAM for planetary rovers[C] // *Proceedings of the IEEE IROS Workshop on Robot Vision for Space Applications*. Edmonton, Canada: IEEE, 2005: 16 – 21.
- [34] TOMONO M. 3D localization and mapping using a single camera based on structure-from-motion with automatic baseline selection[C] // *Proceedings of the IEEE International Conference on Robotics and Automation*. Barcelona, Spain: IEEE, 2005: 3342 – 3347.
- [35] ELINAS P, SIM R, LITTLE J J. σ SLAM: stereo vision SLAM using the Rao-Blackwellised particle filter and a novel mixture proposal distribution[C] // *Proceedings of the IEEE International Conference on Robotics and Automation*. Orlando, USA: IEEE, 2006: 1564 – 1570.
- [36] CHOSET H, NAGATANI K. Topological simultaneous localization and mapping (SLAM): toward exact localization without explicit localization[J]. *IEEE Transactions on Robotics and Automation*, 2001, 17(2): 125 – 137.
- [37] SCHLEICHER D, BERGASA L M, BAREA R, et al. Real-Time wide-angle stereo visual SLAM on large environment using SIFT feature correction[C] // *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*. San Diego, USA: IEEE, 2007: 3878 – 3883.
- [38] HAJJIDJAB H, LAGANIERE R. Vision-based multi-robot simultaneous localization and mapping[C] // *Proceedings of the First Canadian Conference on Computer and Robot Vision*. Ontario, Canada: IEEE, 2004: 155 – 162.
- [39] DELLAERT F, ALEGRE F, MARTINSON E B. Intrinsic localization and mapping with 2 applications: diffusion mapping and marco polo localization[C] // *Proceedings of the IEEE International Conference on Robotics and Automation*. Taipei: IEEE, 2003: 2344 – 2349.
- [40] 苑晶. 未知环境下移动机器人主动同时定位与建图研究[D]. 天津: 南开大学, 2007.
(YUAN Jing. *Study on active simultaneous localization and mapping of mobile robot in unknown environment*[D]. Tianjin: Nankai University, 2007.)
- [41] ZHOU W, VALLS J. Information-efficient 3-D visual SLAM for unstructured domains[J]. *IEEE Transaction on Robotics*, 2008, 24(5): 1078 – 1087.
- [42] DONG J F, WIJESOMA S. Extended Rao-Blackwellised particle genetic algorithm filter SLAM in dynamic environment with raw sensor measurement[C] // *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*. San Diego, USA: IEEE, 2007: 1473 – 1478.
- [43] HOE K G, SUNG K J, SANG H K. Vision-based simultaneous localization and mapping with two cameras[C] // *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*. Edmonton, Canada: IEEE, 2005: 3401 – 3406.

作者简介:

孙凤池 (1973—), 男, 副教授, 研究领域为智能机器人、嵌入式系统, E-mail: fengchisun@nankai.edu.cn;

黄亚楼 (1964—), 男, 博士生导师, 研究领域为智能机器人系统、智能信息处理, E-mail: huangyl@nankai.edu.cn;

康叶伟 (1977—), 男, 博士生, 研究领域为智能机器人系统、计算机技术与应用, E-mail: chinakang@mail.nankai.edu.cn.