

# 基于视觉的多旋翼无人机定位与地图构建

郭峰, 王国胜, 吕强, 张洋

装甲兵工程学院控制工程系, 北京 100072

**摘要** 在缺乏GPS等定位方式的未知环境中,采用视觉的方法对多旋翼无人机定位导航与地图构建逐渐成为该领域研究的热点,本文针对无人机的视觉定位与地图构建技术方法进行综述。首先,分析了两种自主飞行控制方式的优缺点,指出设计的要求和难点;其次,研究无人机定位与地图构建的不同视觉方法,对几种方法特点进行分析并总结研究现状;最后,讨论多旋翼无人机视觉控制的发展趋势。

**关键词** 多旋翼无人机;飞行控制;视觉定位;地图构建

一直以来自主移动机器人导航技术是智能机器人研究领域的热点问题。在环境信息未知、通信不畅的情况下,通过无人平台搭载传感器对未知地域进行探测,同时实现自身定位和导航有着实际的需要。多旋翼无人机具有垂直起降、灵活机动和可靠性高等特点,可以携带更多的探测、处理等设备,在未来军事侦察、灾难救援、工业检查和环境保护中,旋翼型无人机将会扮演重要的角色<sup>[1]</sup>。

在自主多旋翼无人机执行探索和侦察等任务的过程中,需要利用各类传感器完成避障、导航、定位和路径规划等动作。目前的无人机广泛采用GPS和惯性导航系统定位和导航,然而在建筑物内等GPS信号缺失的环境中,需要可靠的定位系统对无人机进行导航定位和位姿估计,而传统的位姿估计与地图构建方法采用激光雷达和超声波传感器,具有一定的局限性。视觉传感器能够获取环境的真实信息,具有成本低、获取信息量大和灵活性高等优点,这是其他传感器无法比拟的。

基于视觉传感器的多旋翼无人机视觉飞行控制逐渐成为近年的研究热

点。视觉算法可以突破一些特殊环境的局限性,其次在算法的复杂程度和精度上逐渐得到了改善,可在一些特殊的飞行任务中得到应用,具有很高的研究价值和广阔的应用前景。但到目前为止,还没有多旋翼无人机的视觉定位导航和地图构建系统在实际中得到应用,大部分研究还停留在理论和实验阶段。基于视觉的自主多旋翼无人机视觉定位与地图构建系统需要从无人机负载、系统实时性、鲁棒性和精确性等方面进行分析,本文将对上述内容进行综述。

## 1 自主飞行控制系统

基于视觉的多旋翼无人机的飞行控制系统根据处理数据的方式和自主控制方法的不同可分为离板控制(off-board)和在板控制(onboard)。离板控制即通过无线的形式在地面站计算机上对所有数据和控制命令完成处理;在板控制则是无人机搭载计算处理单元来完成数据处理和实时控制。

1) 离板控制。由于微小型多旋翼无人机负载和在板处理单元处理能力上的限制,传统的视觉控制方法都是基

于地面站架构,视觉数据和控制命令需要通过Wi-Fi等无线模块进行传输。法国派诺公司的ARDrone四旋翼就是搭载640×480分辨率的摄像头,采用了嵌入式Linux实时操作系统完成图像采集、预处理和Wi-Fi通信。通过Wi-Fi与地面站主控计算机连接,将图像信息和飞行数据传输回地面站计算机,地面站计算机通过图像处理后将控制指令发回ARDrone四旋翼实现自主飞行<sup>[2]</sup>。同样的,德国慕尼黑工业大学应用视觉同时定位与地图构建(SLAM)算法实现四旋翼的自主定位和导航,但是其视觉SLAM算法是在地面站上运行<sup>[3]</sup>。美国加州大学河滨分校使用视觉SLAM算法实现了无人机的定位,但是同样也是运行在地面站上,没有进行机载在板控制<sup>[4]</sup>。这些方法需要依赖一个可靠的无线传输,导致飞行的实时性和鲁棒性不高,同时无线传输距离的限制使其不能适用于所有的室内或者室外任务。

2) 在板控制。近几年计算机硬件和微电子传感器正在往小型化转变,微小型计算机处理能力逐渐增强,计算机处理单元逐渐开始被搭载在无人机上,

收稿日期:2015-10-19;修回日期:2016-10-28

作者简介:郭峰,工程师,研究方向为计算机视觉、自主移动机器人,电子信箱:gfeng1005@163.com

引用格式:郭峰,王国胜,吕强,等.基于视觉的多旋翼无人机定位与地图构建[J].科技导报,2017,35(7):83-87;doi:10.3981/j.issn.1000-7857.2017.07.010

所有数据的处理和飞行控制都实现了在板的实时运行,使得地面站架构只需要对飞行状态进行监控,这也大大保证了多旋翼无人机飞行的自主性和稳定性。瑞士苏黎世联邦理工学院开发的PIXHAWK项目设计在板视觉处理的微型四旋翼,能够以毫秒级的处理速度

对控制信息进行处理<sup>[5]</sup>。德国布伦兹维克大学 Gabel 等设计的四旋翼系统装备了四代 Intel Core i5 处理器,满足定位、地图构建、障碍物检测和路径规划等高级任务的计算需求<sup>[6]</sup>。天津大学搭建了一个采用嵌入式控制系统架构的四旋翼平台,其在板的计算机搭载主频

为 1.8 GHz 的 Intel Core i3 双核处理器。所以随着多旋翼无人机进行跟踪、避障、识别和探索等任务的多样化,各种复杂算法被加入到无人机平台上,对飞行控制系统实时性的要求也在逐渐提高,在板控制的方法开始流行。自主控制方式的对比如表 1。

表 1 自主控制方式比较

Table 1 Comparison of automatic control mode

控制方式	特点	控制效果	应用情况
离板控制	由地面计算机处理图像数据,通过无线形式传输数据并完成飞行控制	能够处理较复杂的数据,依赖无线模块进行控制和图像数据传输,控制存在一定延时,易受到干扰	应用于使用复杂、数据量大的视觉算法平台
在板控制	所有的图像数据处理和飞行控制命令由无人机搭载的计算机处理单元完成	控制实时性、鲁棒性高,需要无人机搭载高性能、低功耗的计算机处理单元	在执行多样化任务的自主多旋翼无人机平台上应用较多

## 2 无人机视觉定位与地图构建方法

1986 年 Smith 和 Cheeseman 等提出了视觉同时定位与地图构建(simultaneous localization and mapping, SLAM)问题。SLAM 是指移动平台在未知环境中运动时利用视觉传感器增量地构建周围环境地图,同时运用此地图对移动平台自身位姿进行估计<sup>[7]</sup>。多旋翼无人机是一个具有 6 自由度的无人平台,需要可靠且高帧率的位姿估计完成导航飞行,在实时性、鲁棒性和精度等方面进行研究。根据获取的环境信息和位姿估计的方法不同可将无人机视觉定位与地图构建分为 3 类:稀疏、稠密和半直接方法。图 1 为部分基于视觉控制的多旋翼无人机平台。

### 2.1 稀疏的方法

稀疏的视觉定位与地图构建仅使用了图像中的部分可用信息进行位姿估计,例如图像中某些点、线和小斑块

等,在连续的图像间进行跟踪。早期的单目视觉并行跟踪和地图构建(PTAM)系统,在第一帧图像必须对提取的特征点赋予初始深度值,继而新的后续特征点深度值能够通过随后变换计算出来。稀疏的视觉算法首要任务是从当前环境图像中提取合适的特征以描述无人机所处未知环境的典型局部特征,提取图像帧局部二维特征后构建描述符,新的特征和前一帧的特征对应关系通过描述符的匹配来建立。最后对每一对匹配特征点进行最小投影误差计算寻找相邻两帧的变换关系。

麻省理工学院的 Huang 等<sup>[8]</sup>较早开始研究四旋翼视觉自主导航,使用 RGB-D 传感器结合惯性测量单元(IMU)给四旋翼飞行器提供可靠的位姿估计和环境感知,利用实时的稀疏视觉定位方法实现飞行器自主导航,其视觉定位包括图像预处理,特征提取,初始旋转估计,特征匹配,内点检测和

运动估计 6 个步骤。该系统在 0.1 m/s 的飞行速度下满足了实时控制。清华大学自动化系<sup>[9]</sup>同样利用一个拓展的卡尔曼滤波(EKF)融合 RGB-D 传感器的视觉运动估计和惯性测量单元(IMU)数据为微型无人机提供精确姿态估计,完成导航任务,其视觉运动估计算法主要包括 4 个步骤:特征检测、帧对帧的特征配准、鲁棒的运动估计和漂移修正的全局估计优化。SURF 二维特征被用来完成帧与帧之间的匹配,特征点之间对应关系找到后,再加入深度距离信息完成运动估计。而纽约城市大学 Valenti 等<sup>[10]</sup>的稀疏方法首次提出帧对帧的策略,相对于 Huang 等的其他帧对帧方法大大减少了累计漂移误差,其核心是对环境构建 3D 特征模型,所有的特征点则采用 ICP 算法与模型进行配准。

稀疏的方法主要以一个合适的特征提取算法为基础,所以特征提取环节

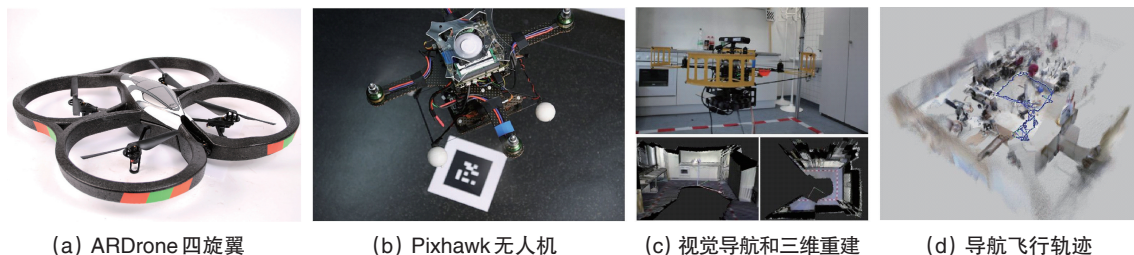


图 1 基于视觉控制的多旋翼无人机

Fig. 1 Multi rotor unmanned aerial vehicle based on vision control

的优劣影响着整个系统的性能。为保证视觉定位和无人机飞行控制系统的实时性和精确性,视觉图像特征算法逐渐成为研究重点。Engelhard等<sup>[11]</sup>提出实时的三维视觉SLAM方法,采用SIFT、SURF和ORB特征分别提取图像特征,进行对比分析,但从实验结果看,在精度和实时性上还不适用于无人机实时的位姿估计。Mouats等<sup>[12]</sup>在四旋翼的视觉里程计中,提出一个基于特征分组策略的梯度矢量流(GVF)方法,提高匹配的鲁棒性,对比实验表明其方法增加了四旋翼视觉运动估计的精度。Valenti等<sup>[10]</sup>的四旋翼稀疏视觉方法研究SURF、ORB和Shi-Tomasi 3种特征算法在模型配准中的性能,最终采用鲁棒性和实时性都较好的Shi-Tomasi特征构建模型。

由于精度与实时性的要求,稀疏的方法近几年在无人机上得到了较多运用,且随着算法的改进,处理速度和精度逐渐得到提升,但是由于这类方法需要从环境图像中提取合适的特征完成匹配,所以在特征不明显的飞行环境中鲁棒性不强,易导致特征的跟踪失败。

## 2.2 稠密的方法

稠密的视觉定位和地图构建方法与稀疏的方法的不同点在于,稠密方法利用了整个图像的信息完成帧与帧之间的匹配,省去了特征检测和匹配的步骤,比基于特征的稀疏方法在小纹理的场景中鲁棒性更强<sup>[13]</sup>。尤其在缺少特征的环境中,可以获得更多的几何信息,对机器人增强现实应用具有重要作用。

第一个稠密的视觉里程计是由Comport等<sup>[14]</sup>提出的基于立体图像的方法,之后德国慕尼黑科技大学<sup>[15]</sup>使用RGB-D数据进行相似的稠密跟踪,均是通过相邻两帧图像的最小光度误差来进行位姿估计。类似的使用最小光度误差的方法可以被3D曲面的最小几何误差代替,即称之为迭代最近点算法(ICP)<sup>[16]</sup>,但是这种方法也存在两方面的缺点,一个是需要了解3D曲面的结构,另一个是包含复杂的计算去搜索最近点来创建点对应关系。Henry等<sup>[17]</sup>的方法对二维彩色图像提取特征,加入深

度信息后形成三维特征点,再对特征点运用ICP计算出三维点的对应关系。一般来说,所有的视觉位姿估计在进行长距离的连续帧跟踪后会形成累计误差的漂移,KinectFusion和稠密跟踪与地图构建(DTAM)建立了一个稠密全局的环境地图模型,并行跟踪运动,优化跟踪线程的误差漂移,同时将跟踪线程中每个新的帧加入到地图模型中。KinectFusion运用一个ICP算法与RGB-D地图模型进行匹配<sup>[18]</sup>,而DTAM则利用了稠密的方法中最小光度误差方法进行跟踪匹配<sup>[19]</sup>。这些稠密的方法都运用了GPU进行加速优化。

德国慕尼黑科技大学提出一种直接的稠密单目视觉SLAM算法<sup>[20]</sup>,构建大规模的环境地图,算法基于稠密的图像配准得到精确位姿估计,同时构建环境的3D地图,主要有2点创新之处,一是在sim(3)控制进行关键帧的稠密配准,有效检测了尺度漂移;二是采用一种概率方法在跟踪过程中考虑深度噪声的影响,这种方法具有较高的精度。在四旋翼视觉导航和地图构建方面,计算机视觉实验室搭建了一个RGB-D传感器的稠密跟踪和地图构建系统<sup>[21]</sup>,获得室内环境高精度的3D模型来完成四旋翼位置控制的自主飞行,符号距离函数(SDF)被增加到系统中完成精确的位姿估计,但是其算法需要依赖一个额外的地面站GPU,没有实现实时的在板控制。

稠密的视觉算法能够获得更高的精度和鲁棒性,且在缺少特征的环境中能获取更多几何信息,但由于运算较复杂,处理速度不高,很多方法都需要依赖GPU,这些原因都限制了其在无人机等飞行平台上的应用。

## 2.3 半直接的方法

半直接的视觉定位和地图构建是近两年才提出的新方法,因为图像中稀疏的信息足够获取粗略的运动估计和找到特征对应关系,而一旦有了特征对应关系和初始运动估计,即可以继续采用稀疏点特征进行估计的优化。所以这种方法综合了稀疏和稠密特点,采用稀疏的点特征,然后用稠密的计算方法

对稀疏特征配准完成位姿估计,没有描述符的构建,在速度上有一定的优势。也就是说半直接方法在位姿估计时,虽使用的是图像的稀疏特征,但并未采用特征匹配来寻找对应关系,所以称之为半直接。

这类方法比较典型的代表是苏黎世联邦理工学院机器人与感知实验室提出的快速半直接单目视觉里程计<sup>[22]</sup>,使用2个并行的线程,一个估计摄像头的运动,另外一个对周围环境进行地图构建。运动估计线程中,对稀疏特征进行最小光度误差主要是进行初始化位姿估计,然后重投影的二维特征点对应关系对运动估计进行优化。地图构建线程采用一个深度滤波来初始化每一个二维特征,当完成新的关键帧选择时,新的深度滤波来初始化点特征,进而三维点加入到地图中并被投入到运动估计线程中。亚像素的特征对应也大大增加了其精度。该方法在飞行实验中跟踪了84 m距离后位置误差仍然控制在 $\pm 0.1$  m以内。在主频为1.6 GHz的Odroid-U2型嵌入式计算机上,在板实时运行速度为55帧/s,完成一次位姿估计的总耗时为18.17 ms。飞行的实时性和精度都达到了一个较高标准。

目前对半直接的视觉方法研究暂时还不是很多,主要还是纯粹的视觉里程计,且在大规模和长距离的飞行导航中精度和鲁棒性还有所欠缺。3类方法的比较见表2。

## 3 发展趋势

随着无人平台在各个领域的广泛运用和计算机视觉技术的飞速发展,无人飞行平台逐渐往智能化和多样化方向发展,不同的计算机视觉算法被应用在无人平台完成一些特殊任务。而利用无人机完成导航、侦察和搜救等任务也逐渐成为一个热门前沿的研究领域。从目前的研究现状看,基于视觉的无人机定位与地图构建技术还远未成熟,大多数研究还停留在理论阶段,且往往只考虑特定环境下的应用,未来对于无人机的视觉定位与地图构建将围绕如何提高控制的鲁棒性、实时性和精

表2 视觉方法比较  
Table 2 Comparison of visual methods

视觉方法	特点	实时性	精度和效果	应用情况
稀疏的方法	使用环境点、线等部分数据进行运动估计,需要依赖一个可靠的特征提取算法	对图像的特征提取耗时较多,一般采用多线程,基本能满足实时性	易受光照的影响,环境变化时鲁棒性不高	对环境光照和纹理比较敏感。在无人机平台上得到较多应用
稠密的方法	利用全部图像信息完成运动估计,计算复杂度高	大部分需要依赖 GPU,处理复杂,速度慢	计算精度较高,运动估计和地图构建效果好	三维重建中应用较多,需要 GPU,实时性欠缺,在无人机平台上应用受限
半直接的方法	利用环境点特征,没有构建描述符匹配,估计速度快	处理速度较快,实时性高	在长距离和大尺度环境中精度和鲁棒性不高	能满足无人机实时性和精度的要求,目前还只是一个纯粹的运动估计,无三维地图构建

确性展开,同时需要适应各类复杂任务,如三维地图构建、路径规划和自主搜索导航等。

1) 不同环境下控制的鲁棒性。在一些特殊的环境中视觉方法完成无人机的定位和导航存在局限性,多个不同的传感器进行数据的融合已成为一种新研究趋势,如结合 GPS、激光雷达和 RGB-D 传感器完成定位和导航,不仅能够提高定位和控制精度,而且增强了对不同环境的适应性和飞行控制的鲁棒性,采用多传感器的方式,提高了无人机平台承担更多元化任务的能力。

2) 三维重建与路径规划研究。无人机进入一个未知的环境,在实现自身定位的同时,需要对所处环境构建三维地图,实现导航和路径规划是在未知环境中进行自主探索、侦察、搜救和测量的基础。

3) 结合机器学习的视觉定位与地图构建。将机器学习以及模式识别中的成果引入视觉定位与地图构建中,将是无人机视觉定位与地图构建技术发展的一个新高度。目前大多数方法都是基于几何方法,将机器学习的新思想推广到无人机视觉导航中来,将为很多

目前方法存在的问题提供新的解决方案和途径。

#### 4 结论

本文讨论了基于视觉的多旋翼无人机平台优点和性能,其必将在未来的军事和民用领域发挥重要作用。但在目前多旋翼无人机的负载运算能力有限、飞行姿态不稳定和应用环境复杂等情况下,很多方法还不能适应这些要求。在设计多旋翼无人机的视觉定位与地图构建系统时,需要根据任务和需求的的不同对方法进行相应改进。

#### 参考文献 (References)

- [1] 岳基隆, 张庆杰, 朱华勇. 微小型四旋翼无人机研究进展及关键技术浅析[J]. 电光与控制, 2010, 17(10): 46-52.  
Yue Jilong, Zhang Qingjie, Zhu Huayong. Research progress and key technologies of Micro quadrotor UAV[J]. Electro Optic and Control, 2010, 17(10): 46-52.
- [2] Krajník T, Vonúšek V, Fišer D, et al. AR-drone as a platform for robotic research and education[M]//Research and Education in Robotics-EUROBOT. Heidelberg: Springer Berlin, 2011: 172-186.
- [3] Engel J, Sturm J, Cremers D. Camera-based navigation of a low-cost quadcopter[C]//Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on. New York: IEEE, 2012: 2815-2821.
- [4] Ghadiok V, Goldin J, Ren W. On the design and development of attitude stabilization, vision-based navigation, and aerial gripping for a low-cost quadrotor[J]. Autonomous Robots, 2012, 33(1-2): 41-68.
- [5] Meier L, Tanskanen P, Fraundorfer F, et al. Pixhawk: A system for autonomous flight using onboard computer vision[C]//Robotics and Automation (ICRA), 2011 IEEE international conference on. New York: IEEE, 2011: 2992-2997.
- [6] Gabel M, Krüger T, Bestmann U. Design of a quadrotor system for an autonomous indoor exploration[C]. IMAV 2014: International Micro Air Vehicle Conference and Competition 2014, Delft, The Netherlands, August 12-15, 2014.
- [7] Durrant-Whyte H, Bailey T. Simultaneous localization and mapping: Part I[J]. Robotics & Automation Magazine, IEEE, 2006, 13(2): 99-110.
- [8] Huang A, Bachrach A, Henry P, et al. Visual odometry and mapping for autonomous flight using an RGB-D camera[C]//International Symposium on Robotics Research (ISRR), USA, 2011: 1-16.
- [9] Li D, Li Q, Cheng N, et al. Combined RGBD-inertial based state estimation for MAV in GPS-denied indoor environments[C]// Control Conference (ASCC), 2013 9th Asian. IEEE, 2013: 1-8.
- [10] Valenti R G, Dryanovski I, Jaramillo C, et al. Autonomous quadrotor flight using onboard RGB-D visual odometry[C]//Robotics and Automation (ICRA), 2014 IEEE International Conference on. New York: IEEE, 2014: 5233-5238.
- [11] Engelhard N, Endres F, Hess J, et al. Real-time 3D visual SLAM with a hand-held RGB-D camera[C]//Proceeding of the RGB-D Workshop on 3D Perception in Robotics at the European Robotics Forum, Vasteras, Sweden. 2011: 180.

- [12] Mouats T, Aouf N. Robust feature matching for aerial visual odometry[C]//ELMAR (ELMAR), 2014 56th International Symposium. IEEE, 2014: 1-4.
- [13] Lovegrove S, Davison A J, Ibanez-Guzmán J. Accurate visual odometry from a rear parking camera[C]//Intelligent Vehicles Symposium (IV), 2011 IEEE. IEEE, 2011: 788-793.
- [14] Comport A I, Malis E, Rives P. Real-time quadrifocal visual odometry[J]. The International Journal of Robotics Research, 2010, 29(2-3): 245-266.
- [15] Steinbrücker F, Sturm J, Cremers D. Real-time visual odometry from dense RGB-D images[C]//Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on. IEEE, 2011: 719-722.
- [16] Besl J, McKay D. A method for registration of 3-D shapes[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1992, 14(2): 239-256.
- [17] Henry P, Krainin M, Herbst E, et al. RGB-D mapping using depth cameras for dense 3D modeling of indoor environments[C]. the 12th International Symposium on Experimental Robotics(ISER), 2010.
- [18] Newcombe R A, Izadi S, Hilliges O, et al. KinectFusion: Real-time dense surface mapping and tracking[C]//Mixed and Augmented Reality (ISMAR), 2011 10th IEEE International Symposium on. IEEE, 2011: 127-136.
- [19] Newcombe R A, Lovegrove S J, Davison A J. DTAM: Dense tracking and mapping in real-time[C]//Computer Vision (ICCV), 2011 IEEE International Conference on. IEEE, 2011: 2320-2327.
- [20] Engel J, Schöps T, Cremers D. LSD-SLAM: Large-scale direct monocular SLAM[M]//Computer Vision-ECCV. Heidelberg: Springer International Publishing, 2014: 834-849.
- [21] Sturm J, Bylow E, Kerl C, et al. Dense tracking and mapping with a quadcopter[C]//UAV-g 2013. 2013.
- [22] Forster C, Pizzoli M, Scaramuzza D. SVO: Fast semi-direct monocular visual odometry[C]//Robotics and Automation (ICRA), 2014 IEEE International Conference on. IEEE, 2014: 15-22.

## Vision based localization and mapping for multi-rotor UAV

GUO Feng, WANG Guosheng, LÜ Qiang, ZHANG Yang

Department of Control Engineering, Academy of Armored Forces Engineering, Beijing 100072, China

**Abstract** In GPS-denied unknown environment, the visual navigation and mapping for multi-rotor UAVs has become popular within the unmanned aircraft field. This paper gives an introduction to the technologies of UAV visual localization and mapping. First of all, the relative merits of two different autonomous flight control systems are analyzed and the difficulties and requirements of design are pointed out. Then, the different visual methods of UAV localization and mapping are investigated, the characteristics of some UAV visual methods are analyzed, and its present research development is summarized. Finally, the future research direction of UAV visual control is discussed.

**Keywords** multi-rotor UAV; flight control; visual localization; mapping

(责任编辑 王媛媛)