

基于四旋翼无人机的单目建模系统

庄晓明,李 涛,付 龙

(中科院合肥物质科学研究院先进制造技术研究所,江苏 常州 213164)

摘 要:为了实现四旋翼无人机的自主飞行、环境单目稠密重建与即时导航功能,文中设计了该单目建模导航系统。飞行控制器采用 MPU6050 六轴传感器采集六轴信号,通过算法拟合出飞行器姿态,然后通过 STM32F103RBT6 芯片分别控制 4 个电机驱动模块,即时通过 PID 算法输出 PWM 值进行调速来实现飞行姿态的稳定和改变。建模系统通过控制 GPS 模块,可以获得飞行器的实时位置;飞行器通过图像采集模块,可以获得飞行器的环境信息。设计了上位机软件,能够接收图像并且进行单目稠密重建。稠密重建主要由从运动到结构、基础表面构建、约束场景流稠密重建、稠密对齐和局部模型集成共五方面组成,然后通过无线串口与无人机通讯,实现无人机的智能自主飞行。实验结果表明,该系统能够实现单目建模的功能。

关键词:四旋翼无人机;智能导航;三维重建;姿态控制

中图分类号:TP302

文献标识码:A

文章编号:1673-629X(2016)07-0104-05

doi:10.3969/j.issn.1673-629X.2016.07.022

Monocular Modeling System Based on Four Rotors UAV

ZHUANG Xiao-ming, LI Tao, FU Long

(Institute of Advanced Manufacturing Technology, Hefei Institute of Physical Science,
Chinese Academy of Science, Changzhou 213164, China)

Abstract: In order to realize autonomous flight, monocular dense reconstruction and auto navigation of four rotors UAV, the monocular reconstruction and modeling system is designed. The flight controller uses six axis sensors, MPU-6050, to detect the attitude of the four rotors UAV; the speed of the 4 motor drive module is controlled by the STM32F103RBT6 chip to change the attitude of the UAV with PID algorithm. The monocular reconstruction system acquires real-time position and environment information of the UAV by controlling the GPS module and image acquisition module. The software on PC can receive the images and make monocular dense reconstruction. Dense reconstruction contains five main part including structure from motion, basic surface construction, constrained scene flow dense reconstruction and local model integration. The intelligent autonomous flight of UAV is realized through communications between the PC and the UAV with wireless serial port. Experiments show that the system can realize monocular dense reconstruction.

Key words: four-rotor micro UAV; smart navigation; 3D reconstruction; attitude control

0 引 言

四旋翼无人机具有垂直起降、自由悬停的特点,并且体积小,操控方便,携带方便,可以进入人不易进入的复杂环境。目前,四旋翼无人机在航拍、娱乐、勘测、救灾等恶劣环境中发挥着极其重要的作用^[1]。

对于复杂不易进入的环境,更多时候可以使用无人机进入。无人机要具有自主飞行的功能,并且能对未知环境进行建模以方便外部人员观察。传统的四旋翼飞机需要使用遥控器进行位置和姿态控制,但无法感知环境的图像细节和深度细节。传统的三维重建方

法主要有结构光方法和立体视觉方法,代表产品如 Kinect 体传感器和 Bumblebee 相机等。

文中设计了一种基于从运动信息中恢复三维场景结构(structure from motion)的三维重建算法,通过使用单目相机采集图像,传输至远程电脑进行实时三维重建,四旋翼无人机根据三维环境结合 IMU 信息进行姿态控制,实现自主智能导航。

1 飞行器的基本工作原理

四旋翼飞行器是一种有四个旋翼并且旋翼呈十字

收稿日期:2015-06-16

修回日期:2015-09-23

网络出版时间:2016-06-22

基金项目:国家青年科学基金项目(61401437);中国科学院合肥物质科学研究院院长基金青年火花项目(YZJJ201323)

作者简介:庄晓明(1987-),男,硕士研究生,研究方向为控制工程与机器视觉。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20160621.1701.008.html>

形交叉的飞行器^[2]。飞行器的机械结构简单,旋翼对称分布在机体的前、后、左、右四个方向,四个旋翼处于同一高度平面,且四个旋翼的结构和半径都相同。四个电机对称地安装在飞行器的支架端,支架中间安放飞行控制计算机和外部设备。常见结构如图1所示。

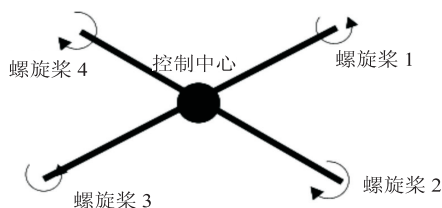


图1 四旋翼飞行器的结构

四旋翼飞行器通过调节四个电机的转速来改变旋翼转速^[3],从而实现升力的变化,控制四旋翼飞行器的位姿。四旋翼飞行器是一种六自由度的垂直升降机,但只有四个输入力,同时却有六个状态输出,所以是一种典型的欠驱动系统。其主要飞行姿态如图2所示。

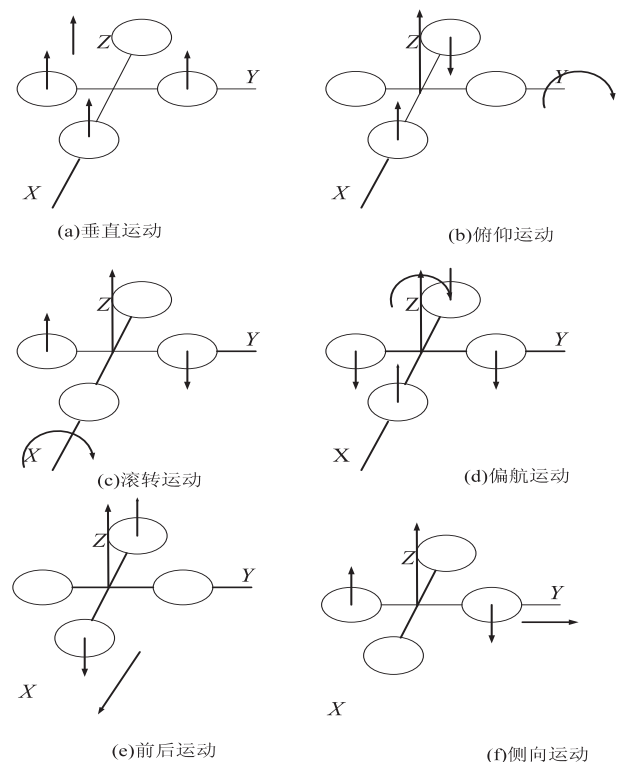


图2 四旋翼飞行器的飞行姿态

(1)垂直运动:同时增加四个电机的旋转速度使总拉力增大,当拉力克服整机的重量时,四旋翼飞行器就可以垂直上升;反之,同时减小四个电机的旋转速度,四旋翼飞行器就垂直下降,直至平衡落地,实现了沿 z 轴的垂直运动。当外界扰动量为零时,当旋翼产生的升力等于飞行器的重力时,飞行器就能保持悬停。

(2)俯仰运动:使电机1的转速上升,电机3的转速下降,电机2和4的转速保持不变。由于旋翼1的升力增大,旋翼3的升力减小,产生的不平衡力矩使机

身绕 y 轴旋转,实现飞行器的俯仰运动。

(3)滚转运动:改变电机2和4的转速,保持电机1和3的转速不变,就可以使机身绕 x 轴旋转,实现飞行器的滚转运动。

(4)偏航运动:当电机1和3的转速上升,电机2和4的转速下降时,机身就在反扭矩的作用下绕 z 轴转动,实现飞行器的偏航运动。

(5)前后运动:增加电机3的转速,并相应减小电机1的转速,同时保持其他两个电机转速不变,以保持反扭矩不变。飞行器就先发生一定程度的倾斜,使旋翼拉力产生水平分量,从而实现飞行器的前后运动。

(6)倾向运动:由于结构对称,因此倾向飞行的工作原理与前后运动的原理一致。

2 系统整体设计

四旋翼飞行器的系统结构如图3所示。

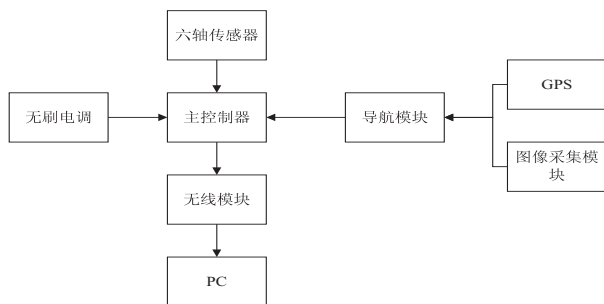


图3 系统结构图

GPS模块和六轴传感组成姿态测量系统,主控制器将姿态测量获得的数据进行卡尔曼滤波得到当前姿态,再与从无线模块发送过来的自主导航指令一起作为姿态控制器的输入,姿态控制器通过PID算法输出对应的PWM值^[4],驱动四个电调,进而改变4个电机的转速实现姿态控制。

同时系统搭建了基于单目建模的SLAM系统,可以实时单目稠密重建并在PC端进行渲染和导航。

3 飞行控制系统设计

四旋翼飞行器能够沿着规划的轨迹飞行,并且能够在飞行过程中同步进行SLAM,实时修正轨迹,并且能够返回初始位置^[5]。导航系统主要是由GPS模块、六轴传感模块和图像采集模块组成。该系统具有精确修正系统误差的功能,从而提高导航精度^[6]。

主控制芯片要进行数据的采集和滤波、图像的采集和传输等大量复杂运算,故选用了STM32F103-RBT6作为主控芯片,STM32F103XX基于高性能32位RISC的ARM Cortex-M3核,工作频率为72MHz。片上集成了高速存储器,通过APB总线连接了丰富、增强的外设和I/O。所有的设备都提供标准的通信

接口^[7]。

3.1 GPS 模块

GPS 模块型号为 LEA-5T, 作为导航系统的主要模块, GPS 对导航精度及灵敏度起着十分重要的作用, 所以必须使用高性能、高可靠性的 GPS 定位模块。该系统采用的 GPS 模块具有以下特点:

(1) 拥有一个 -160 dBm 捕获和跟踪灵敏度的 Super Sense;

(2) 具备 Kick Start 功能, 信号微弱时可实现加速启动;

(3) 热启动和辅助启动首次定位时间小于 1 s;

(4) 支持 AssistNow Online 和 AssistNow Offline 等 A-GPS 服务;

(5) GPS、GALILEO、SBAS(WAAS、EGNOS、MSAS、GAGAN) 混合引擎。

3.2 六轴传感模块

六轴传感模块使用的是 MPU6050, MPU6050 的角速度全格感测范围为 ± 250 、 ± 500 、 $\pm 1\ 000$ 与 $\pm 2\ 000$ $^{\circ}/\text{sec}$ (dps), 可准确追踪快速与慢速动作, 并且用户可程式控制的加速器全格感测范围为 $\pm 2\text{ g}$ 、 $\pm 4\text{ g}$ 、 $\pm 8\text{ g}$ 与 $\pm 16\text{ g}$ 。产品传输可透过最高至 400 kHz 的 IC。

3.3 图像采集模块

图像采集模块使用的是 OV2640。OV2640 像素输出最大支持 200 万像素, 支持 QCIF (176 * 144)、QVGA (320 * 240)、VGA (640 * 480)、1 027 * 768、1 600 * 1 200 等像素输出。具备两种输出格式:

(1) 原始数据如 RGB565, RGB RAW, YUV422;

(2) JPEG 压缩图像格式 (可极大减少传输带宽, 例如 640 * 480 分辨率的原图片大小在 300 kB 左右, JPEG 编码输出后大小仅约为 16 kB)。

在该系统中, 由于使用原始数据格式, 图像文件过大 (以 RGB RAW 为例, 一帧 640 * 480 的图像文件大小为 $640 * 480 * 3 = 900\text{ Kbytes}$), 影响数据传输速率和更新图像的频率。TM32FI03RBT6 处理能力有限, 无法做复杂的图像压缩算法, 所以选择 OV2640 内部 DSP 压缩后的 JPEGs 图像格式是最好的选择。

MCU 与 OV2640 通信采用串行与并行结合, OV2640 带有 SCCB (Serial Camera Control Bus) 双线串行接口, MCU 通过 SCCB 接口配置和读取 OV2640 的信息; MCU 通过并行总线的方式接收 OV2640 图像数据。

4 系统软件分析

文中提出了一种可以使用单个实时相机进行场景的快速稠密重建的方法。使用基于点的实时从运动中恢复结构 (SFM) 作为起始点, 生成精确的 3D 相机姿态估计和稀疏点云。主要创新在于使用了一种源于

SFM 的基础网格生成方法, 预测在一束自动选择的跨越场景的参考帧附近的姿态中的视野, 然后使用一种视野预测的光流和有约束的场景流更新的方法将基础网格变形为高度精确的深度图^[8]。由于产生深度图的高质量, 全局场景模型可以简单将它们并列排放, 然后去除重叠区域。笔者展示了一个现实环境可以通过一个实时相机在数秒内重建, 其处理使用的仅是普通桌面级别的硬件^[9]。

4.1 从运动到结构

与所有稠密重建系统类似, 都需要在全局帧中获得高质量的相机位姿估计。使用 SFM 来提供实时相机位姿的估计和作为稠密场景重建的起始点的稀疏点云。Davison 介绍了一种简单的相机追踪和映射系统, 其使用了连续概率滤波器来构建一致的场景图和实施相机位姿估计, 但是绝对精确对受限于每帧所追踪到的数量较少的特征点。Klein 和 Murray 的平行追踪和映射系统 (PTAM) 转而依赖基于每帧几百个特征点的测量的帧率位姿估计, 并迭代使用最优化平行方法的重复束调整。它能够实现几千个场景点位置和历史摄像帧序列的位姿的高精度估计, 因此能够满足需求。

只收集从 PTAM 获得的高质量点, 并将其送入稠密重建流水线, 期间丢弃那些异常值大于 10 的点。另一种从 PTAM 获得的有用的值是在关键帧集中场景点的联合可见性。通过对关键帧中的每个点的光轴方向取平均的方法来进行每个点粗略的初始表面法线估计。在上述的关键帧中, 这些点都是可见的。

4.2 基础表面构建

已有从 SFM 中得到的起始 3D 点云, 现在想要估计起始连续场景表面来获得稠密细化的基础。从有向点采样进行稠密 3D 表面重建的问题已经在计算机图形学社区获得了足够的关注度^[10], 并且这个问题是模型曲面补洞、优化表现模式和模型美化的基础, 隐式曲面重建技术更是获得了足够的关注。通过这些方法, 表面可以通过将函数拟合成点数据的方法估计得到, 其中, $R^3 \rightarrow R$, $f(x) = 0$ 。可以通过多边形化函数的零点集合 (zero level set) 的方法来提取重建网格^[11]。在获得即时的基础模型中, 速度是至关重要的。而全局优化的无参数的表面拟合技术需要极大的计算代价用于解很大的稠密系统的方程组。近年来, 重建的计算消耗有了大幅度的减少, 主要是将全局优化算法替换为分层和由粗到精的方法^[12]。尤其是有限支持径向基函数可以使有约束的稠密系统变得稀疏。

文中使用了一种最先进的多尺度紧支撑径向基函数的方法用于 3D 离散数据内插^[13]。该方法结合了全局和局部估计方法的优点, 适合于稀疏的、方向估计的从 SFM 过程中获得的点云, 特别是保留了在 I_{Ref} 低稠

密度的大区域插值的能力。

零点集合的多边形化,使用了 Bloomenthal 的方法,也由 MSCSRBF 支持,因为当估计隐式曲面时计算的大部分是基于上一层的函数层次。实际上,可以在每次生成一个新的关键帧时都进行基础表面重建,以便维护即时的基础模型。

4.3 约束场景流稠密重建

在该系统中,使用了简单迭代稠密深度图重建算法。该算法依赖于高质量的相机追踪与在上一章节中获得的基础模型。

每个参考帧都有一个灰度图 I_{Ref} 、投影矩阵 $\mathbf{P}_{\text{ref}} = \mathbf{K}[\mathbf{R}_{\text{cw}}^{\text{Ref}} | -\mathbf{R}_{\text{cw}}^{\text{Ref}} \mathbf{t}_{\text{cw}}^{\text{Ref}}]$ 和一组近邻帧 $I_{i \in (1 \dots n)}^C$ 。其中 $n \geq 1$ 。 \mathbf{K} 是相机特征矩阵。一系列帧组成了一个相机束。

假设有 n 个静止的摄像机观察一个可变形的表面。目标是运用稠密的跨摄像机束的图像测量来使基础表面估计变形到一个更加新而准确的形状上去。

一个可见的表面上的点: $\mathbf{x}_i = [x_i, y_i, z_i]^T$ 提供了一个在摄像机束里第 i 个视角的投影矩阵 $[u_j^i, v_j^i] = \mathbf{u}_j^i = \mathbf{P}^i(\mathbf{x}_i)$ 。其中 $\mathbf{P}^i(\mathbf{x}_i)$ 表示透视投影矩阵,基于投影矩阵 \mathbf{P}^i 表面点的一个小的三维运动 $\Delta \mathbf{x}_j$ 导致一个新的点位置 \mathbf{x}_j' 足够小,然后图像位移 $\Delta u_j^i = u_j^i - u_j^i$ 可以用在 \mathbf{x}_j 附近对 \mathbf{P}^i 进行一阶线性化而近似得到。在这种情况下, $\Delta \mathbf{x}_j$ 被称作场景流。

$$\Delta u_j^i = \mathbf{J}_\gamma^i \Delta \mathbf{x}_j \quad (1)$$

其中, \mathbf{J}_γ^i 是对带有投影参数 \mathbf{P}^i 的 \mathbf{x}_j 点处估计的投影函数求一阶偏导得到的雅克比矩阵。

稠密的基础网面提供了一个关于在摄像机束中每帧可见的表面元素的位置和法向量的预测。希望使用图像对应信息来获取端点更新 $\Delta \mathbf{x}_j$ 以用来将当前基础模型的每个端点 \mathbf{x} 变形到一个新的准确体现真实场景的估计中。

假设可以测量图像位移向量 $\Delta \mathbf{u}_j$, 它与 \mathbf{x}_j 和 \mathbf{x}_j' 相关联。然后对于 $n \geq 2$ 个视角, 给定一个 \mathbf{x}_j , 就可以为一个静止的摄像机计算一个雅克比矩阵。这样, 可以对一个过定的线性系统求解得到 $\Delta \mathbf{x}_j$, 进而给出每一个表面点的新位置 $\mathbf{x}_j' = \mathbf{x}_j + \Delta \mathbf{x}_j$ 。对于参考视角中的每个像素 u^{ref} , 将光线投影回去, 并且和基础模型相交来决定一个三维端点位置 \mathbf{x}_j 。然后, 在被投影到摄像机束中的每个参考帧, 以获取一个可预测的图像对应 u_j 。根据定义, 任何可见的基础模型的变形都没有改变在参考视角本身的预测图像对应。在参考视角中, 一个表面点的投影不管在变形前还是变形后都满足 $u^{\text{ref}} = u^{\text{ref}}$ 。这将构成在参考视角和每个比较视角的对极约束。场景流向量受到限制而位于光束中。这个光束是从参考摄像机处射出, 并且与基础模型在 \mathbf{x}_j 处相

交, 这个表面点为:

$$\Delta \mathbf{x}_i = r_i \lambda_i \quad (2)$$

其中, 光束具有单位向量方向 r_i 。每一个图像位移向量 $\Delta u_j^i = u_j^i - u_j^i$, 可以通过被预测的图像投影和 u_j^i 计算得到。通过组合两个线性方程来建立过定的线性系统。这些线性方程来自于每个位移向量的两个成分, 而每个位移向量与观察模型端点 \mathbf{x}_j 到列向量的每个摄像机姿态 $i = 1, 2, \dots, n$ 有关:

$$\begin{pmatrix} u_j^1 \\ v_j^1 \\ \vdots \\ u_j^{2n} \\ v_j^{2n} \end{pmatrix} = \Delta \mathbf{u}_j \quad \begin{pmatrix} K[u_j^1] \\ K[v_j^1] \\ \vdots \\ K[u_j^{2n}] \\ K[v_j^{2n}] \end{pmatrix} = \mathbf{K}_j \quad (3)$$

最小二乘解通过下列标准方程求得:

$$\lambda_i = \mathbf{K}_i^T \mathbf{K}_i^{-1} \Delta u \quad (4)$$

形成稠密的逐端点的表面更新:

$$\mathbf{x}_i' = \mathbf{x}_i + r_i \lambda_i \quad (5)$$

这个计算只包括了两个规模为 $2n$ 的内积运算, 以及一个标量除法, 可以在 GPU 硬件上进行并行化运算。在实践中, 通过使用之前的式(1), 获取最后的端点更新。在初始基础模型的端点已经无解的情况下, 优化过程继续最多 3 次迭代, 或者直至平均重投影误差小于 $e = 1e^{-4}$ 个归一化像素。

4.4 稠密对齐

稠密网格提供了相对于给定帧的位置和法线的预测, 然后将参考帧图像反向投影至基础模型表面上, 并且将纹理模型投影至比较相机帧上, 用以同步在此相机内的图像预测。这种方法可以使用 GPU 进行快速高效的同步运算。然后使用变分光流法来获得 Δu_j^i 。

4.5 局部模型集成

已经有大量算法可以用于多个深度图的融合, 这些深度图传统上是从结构光和范围扫描器中获得的^[14]。这些方法可大致分为两类。第一类方法使用计算深度图、密集点样本或者体素符号距离函数来拟合全局连续函数, 此函数可用于多边形化。但是很难用于大规模的重建, 因为受到普通 PC 内存和 CPU 的限制。第二类方法直接在局部网格上进行重建。链式多边形网格, 作为最原始的方法, 通过去除网格之间的重叠区域的方法来获得连续拓扑网格。文献[9]贡献了有噪声的深度图实时融合的方法。该方法可以获取一系列的深度图, 同时可以降低这些深度图的误差并且去除空间碰撞, 然后将其多边形化。

因为前面的工作保证了深度图的连续高质量, 只要使用一种非常简单的方法就可以满足需求。

假定已知一个从参考帧 \mathbf{P}^{ref} 获得的三角化网格,

将刚集成的稠密重建深度图渲染至 P^{ref} 并且在误差范围内去除邻近顶点。

5 仿真与调试

为了验证四旋翼无人机的飞行与单目导航功能建模,首先制作了样机,在地面调试 PID 参数,最后在室内进行实际飞行测试,并且在上位机上进行同步建模。飞行效果图如图 4~6 所示。



图 4 四旋翼飞行器样机调试

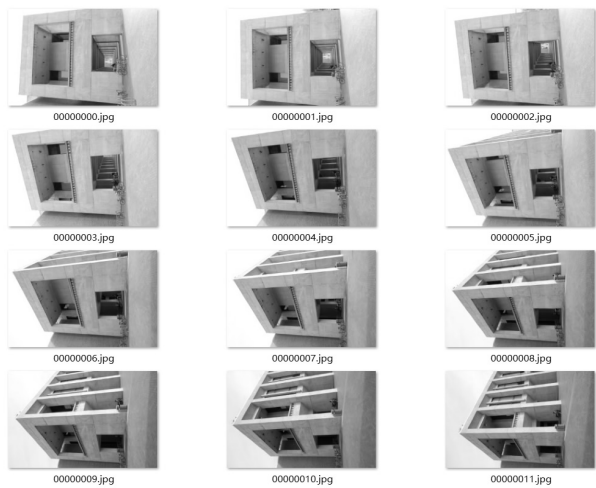


图 5 四旋翼飞行器传输图片

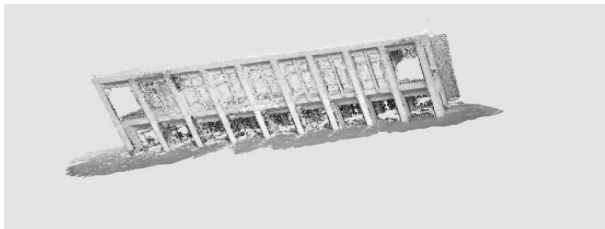


图 6 单目稠密重建效果图

6 结束语

文中设计了一种基于四旋翼无人机的单目导航

建模系统。该系统能够通过摄像头对环境进行单目稠密重建,在 PC 端显示并发送给四旋翼无人机,结合无人机的 GPS 和惯导系统,能够使无人机对极端环境进行探测。设计并制作了四旋翼无人机样机,并进行了实际飞行测试。试飞结果表明:该系统可以用于四旋翼无人机系统进入复杂环境并且进行建模。

参考文献:

- [1] 詹 镭,贺人庆,谢 阳,等.基于微型四旋翼无人机的智能导航系统[J].电子测量技术,2011,34(6):1-3.
- [2] 王晓明.电动机的单片机控制[M].第2版.北京:北京航空航天大学出版社,2007.
- [3] 杨明志,王 敏.四旋翼微型飞行器控制系统设计[J].计算机测量与控制,2008,16(4):485-487.
- [4] 刘金琨.先进PID控制及其MATLAB仿真[M].北京:电子工业出版社,2004.
- [5] 聂博文,马宏绪,王 剑,等.微小型四旋翼飞行器的研究现状与关键技术[J].电光与控制,2007,14(6):113-117.
- [6] 刘慧英,范宝山.基于STM32的多步进电机控制系统研究[J].测控技术,2010,29(6):54-57.
- [7] 符 冰,方宗德,侯 宇.一种新型微旋翼飞行器的设计与控制[J].航空制造技术,2006(5):91-94.
- [8] Wang L, Mordohai P, Frahm J M, et al. Real-time visibility-based fusion of depth maps[C]//Proceedings of the international conference on computer vision. [s. l.]:[s. n.],2007.
- [9] Newcombe R A, Strasdat H, Kelly P H J, et al. SLAM++: simultaneous localisation and mapping at the level of objects [C]//Proc of 2013 IEEE conference on computer vision and pattern recognition. [s. l.]:IEEE,2013:1352-1359.
- [10] Wu Hao, Bu Jiajun, Chen Chun, et al. Locally discriminative topic modeling [J]. Pattern Recognition, 2012, 45(1):617-625.
- [11] Klein G, Murray D W. Parallel tracking and mapping for small AR workspaces[C]//Proceedings of the international symposium on mixed and augmented reality. [s. l.]:[s. n.],2007.
- [12] Yu F, Gallup D. 3d reconstruction from accidental motion [C]//Proceedings of 27th IEEE conference on computer vision and pattern recognition. [s. l.]:IEEE,2014.
- [13] Wagman R S, Barmpoutis A, Bozia E. A novel framework for 3d reconstruction and analysis of ancient inscriptions [J]. Journal of Machine Vision and Applications, 2010, 21(6):989-998.
- [14] Zhang Lijun, Chen Chun, Bu Jiajun, et al. Active learning based on locally linear reconstruction [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(10):2026-2038.