http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2016.0965

无人机室内视觉/惯导组合导航方法

王亭亭[1], 蔡志浩[1,2,*], 王英勋[1,2]

(1. 北京航空航天大学 自动化科学与电气工程学院, 北京 100083;
2. 北京航空航天大学 飞行器控制一体化技术国防科技重点实验室, 北京 100083)

摘 要: 针对室内无卫星定位下的无人机自主导航问题,提出了一种融合惯导、光流和视觉里程计的组合导航方法。在速度估计上,采用基于 ORB 特征的光流法,该方法可以实时地估计出无人机的三轴线速度信息。方法采用基于特征点的稀疏光流,对金字塔 Lucas-Kanade 光流法进行了改进,采用前后双向追踪和随机采样一致的方法提高特征点追踪精度。在位置估计上,采用视觉/惯导融合的视觉里程计,以人工图标法为主,融合视觉光流信息和惯导数据实现无人机定位。通过与运动捕捉系统的定位信息、Guidance 和 PX4Flow 导航模块的测速信息进行对比,以及实际的飞行测试,验证本文方法的可行性。

关 键 词: 无人机; 视觉导航; 光流; ORB 特征; 多传感器融合
中图分类号: TP242.6; V249.32
文献标识码: A **文章编号**: 1001-5965(2018)01-0176-11

在传统的导航方法中,卫星导航是应用最广泛的导航方法。但是无人机在城市或者室内飞行时,由于楼宇、森林和墙壁等阻碍,卫星信号将不可用,加之普通的惯性测量元件存在精度低和严重漂移等问题,传统的导航方法将不能满足无人机的导航需求,因此基于视觉感知系统的无人机导航方法被广泛研究和应用。这些技术主要包括单目/双目视觉里程计(Visual Odometry, VO)[1-4]、单目/双目视觉同时定位与建图(Visual Simultaneous Localization and Mapping, V-SLAM)[5-6]、图像光流[7-10]等,且多采用组合导航方法。其中,文献[1]利用视觉、激光、惯导和 GPS 的组合导航实现多旋翼无人机室内/外自主飞行;文献[3-6]利用视觉/惯导组合导航实现无人机车/无人机位姿估计。部分消费级无人机,如大疆创新科技有限公司的精灵 4 利用 2 个双目视觉传感器和 1 个单目视觉传感器实现了自主避障、室内飞行和行人跟踪[9];零零无限科技有限公司的 Hover Camera 利用一个下视和前视视觉传感器实现室内指尖放飞、定点和跟拍飞行[10]。无人机是 1 个在三维空间运动,且只有 6 自由度的平台,对算法的实时性和鲁棒性要求较高,而一般的视觉导航方法普遍存在更新速率低和延迟大的缺点。

因此,本文提出了一种融合视觉传感器和惯性传感器的组合导航方法,采用基于 ORB(Oriented FAST and Rotated BRIEF)特征的光流法估计无人机速度信息,采用图标法里程计估计位置信息,并应用扩展卡尔曼滤波(Extended Kalman Filter, EKF)将惯导的角速度和加速度信息与视觉光流和里程计进行融合,利用单目/双目视觉系统和惯性传感器搭建导航平台,应用于小型无人机的室内导航。

收稿日期: 2016-12-23; **录用日期**: 2017-02-06; **网络出版时间**: 2017-03-23 16:44
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20170323.1644.003.html
基金项目: 航空科学基金 (20135851043)
* **通信作者**. E-mail: czh@buaa.edu.cn

引用格式: *王亭亭, 蔡志浩, 王英勋. 无人机室内视觉/惯导组合导航方法[J]. 北京航空航天大学学报, 2018, 44(1): 176-186.*
WANG T T, CAI Z H, WANG Y X. Integrated vision/inertial navigation method of UAVs in indoor environment[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(1): 176-186 (in Chinese).

1　基于 ORB 特征的光流法的速度估计

1.1　ORB 特征提取与描述

ORB[11]是一种实时性较高的特征提取与描述方法，它融合了 FAST[12]和 BRIEF（Binary Robust Independent Elementary Features）[13]2 种描述子。FAST 具有很强的实时性，但其不具备尺度不变性，且不包含方向因子。BRIEF 用简单的二进制形式描述特征，实时性比较高，但其无法保证旋转和尺度不变性，对噪声比较敏感。ORB 基于 FAST 和 BRIEF 方法，具体实现如下。

1）建立图像高斯金字塔，对每一层图像提取 FAST 特征点。

2）为特征点添加方向因子。以特征点为中心建立一个图像块，定义图像块各像素点的矩为

$$m_{\alpha\beta}=\sum_{x,y}x^{\alpha}y^{\beta}I(x,y) \tag{1}$$

式中：α、β 为矩的阶次；$I(x,y)$ 为图像灰度表达式。利用零阶矩和一阶矩来计算质心的坐标，即 $C_0=\left(\frac{m_{10}}{m_{00}},\frac{m_{01}}{m_{00}}\right)$。建立一个从图像块的中心到质心的向量作为方向因子，向量的方向可简化为 $\theta=\arctan(m_{01},m_{10})$，这里忽略此特征是亮点还是暗点。

3）引入旋转因子。利用 BRIEF 方法描述特征点，并在此基础上采用方向因子驱动法实现二进制字符串向量的旋转。

在 BRIEF 方法中，对于一个 $N\times N$ 大小的图像块 P，有如下定义：

$$\tau(I;x,y):=\begin{cases}1 & I(x)<I(y)\\ 0 & \text{其他}\end{cases}$$

选取 n 个特征点得到的 n 维向量如下：

$$f_n(p):=\sum_{1\leqslant i\leqslant n}2^{i-1}\tau(I;x_i,y_i) \tag{2}$$

定义一个 $2\times n$ 的矩阵：$\boldsymbol{S}=\begin{bmatrix}x_1 & x_2 & \cdots & x_n\\ y_1 & y_2 & \cdots & y_n\end{bmatrix}$，利用方向因子 θ 和基础旋转矩阵 $\boldsymbol{R}_\lambda$，定义 $\boldsymbol{S}_\lambda=\boldsymbol{R}_\lambda\boldsymbol{S}$。得到最终的特征点描述子如式（3）所示。只要 λ 确定了，便可快速获得描述子。

$$g_n(p,\lambda):=f_n(p)\,|\,(x_i,y_i)\in\boldsymbol{S}_\lambda \tag{3}$$

本文对 ORB、SIFT[14]和 SURF[15]的特征提取时间开销进行了试验分析，如表 1 所示。

在对不同场景进行特征点提取时，ORB 比 SIFT 和 SURF 的耗时小，相差一个数量级。这是因为：①FAST角点提取时间复杂度低。②SIFT 特征为 128 维描述子，占据 512 Bytes 空间；SURF 特征为 64 维描述子，占据 256 Bytes 空间。而用 BRIEF 描述每个特征点仅需一个长度为 256 的向量（二进制形式向量），存储空间为 32 Bytes，降低了后续特征匹配的时间复杂度。因此将 ORB 特征提取方法用于稀疏光流场特征提取，可以减小时间开销。

表 1　SIFT、SURF 和 ORB 特征提取时间开销对比

Table 1　Comparison of time consumption of feature extraction among SIFT, SURF and ORB

方　法	特征点数	时间/ms
SIFT[14]	171	18.89
	253	18.99
	254	19.66
SURF[15]	86	12.44
	254	16.21
	187	13.86
ORB	168	2.82
	299	4.11
	251	4.7

1.2　改进的 Lucas-Kanade 光流提取

基于特征点的 Lucas-Kanade 算法（简称 LK 算法）是经典的光流提取算法，该算法基于以下 3 种假设：目标图像亮度一致、图像空间连续和图像变化在时间上连续。$I(u,v,t)$ 表示 t 时刻像素点处图像灰度值，基于以上假设可以得到像素点的亮度守恒方程为

$$\boldsymbol{I}_u\dot{u}+\boldsymbol{I}_v\dot{v}+\boldsymbol{I}_t=\boldsymbol{0} \tag{4}$$

式中：$\dot{u}$ 和 $\dot{v}$ 分别为图像横纵向光流值；$\boldsymbol{I}_u$ 和 $\boldsymbol{I}_v$ 分别为像素点亮度在 x 和 y 方向的偏导数；$\boldsymbol{I}_t$ 为像素点亮度随时间的导数。这里待求的 $\dot{u}$ 和 $\dot{v}$ 将直接参与速度解算。

LK 算法的 3 个假设在实际场景中很难满足，传统的解决办法包括：①保证足够大的图像帧速率；②引入图像金子塔，通过对不同尺度的图像的光流提取保证连续性。为了提高前后帧特征点追踪准确性，本文在传统方法的基础上采用前后双向追踪策略和随机采样一致滤波（Random Sampling Consensus，RANSAC）提高光流场的提取精度。

1）前后双向追踪策略

如图 1 所示，针对前后 2 帧图像，首先提取第 1 帧图像的特征点，得到特征点集 A，然后通过金字塔 LK 算法在第 2 帧图像中得到对应特征点集 A 的特征点集 B，此为前向追踪。之后，针对特征点集 B，返回在第 1 帧追踪得到对应的特征点集，记为 C，此为后向追踪。

由于噪声影响，前后追踪的特征点并非完

http://bhxb.buaa.edu.cn

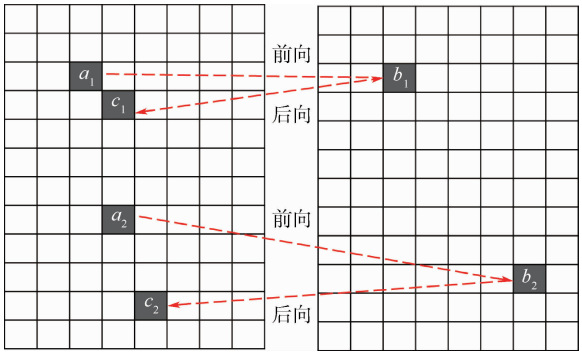


图 1 前后双向追踪示意图

Fig. 1 Schematic of forward-backward tracking

全一致。对此,选取误差和相似度 2 个标准对匹配点对进行滤波。其中,误差指 A、C 中对应点的距离,距离越小,误差越小。相似度指对 A、C 中对应的每一对点,分别选取 16×16 大小的邻域,对 2 个邻域进行模板匹配,选择归一化相关系数匹配法,将匹配的结果作为相似度。匹配值越大,表示越相似。最后选取相似度大且误差小的点对。

2)随机采样一致滤波

本文借助随机采样一致的思想,在前后双向追踪得到的特征点对中随机选取 8 对点对,算出基础矩阵,然后利用该基础矩阵对应的极性约束,对剩余的匹配进行测试。RANSAC 算法认为,支持集合越大,得到的矩阵正确的可能性越大。最终选取较大的支持集合的特征点对应的光流值来估计速度。

RANSAC[15] 算法的实现如下:

步骤 1 初始:假设 D 是 N 个特征对应的集合。

步骤 2 开启循环:

1)随机选取集合 D 中的 S 个采样点;

2)为该 S 个采样点拟合一个模型;

3)计算 D 集合中的其他点与该模型的距离;

4)所有距离小于阈值的点构成一个内联集;

5)存储内联集;

6)返回 1)继续,直到迭代到最大迭代次数。

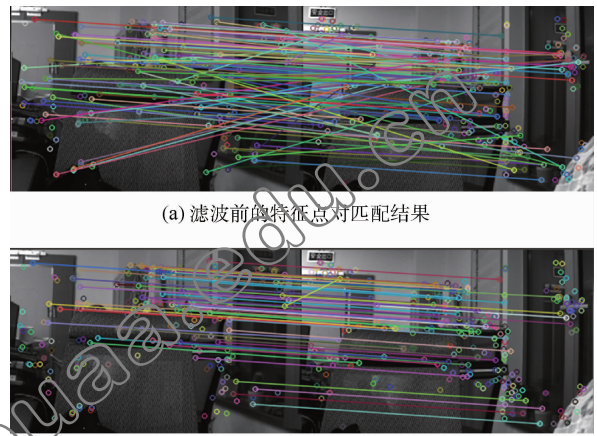
步骤 3 选择拥有最大内联集合的采样点 S^*,并用所有的内联集估计该模型。

图 2 显示了 RANSAC 算法前后的匹配效果。经过 RANSAC 算法,错误的匹配被有效滤除。

1.3 融合惯导角速度的速度解算

当相机没有水平移动而仅存在姿态变化时,依然会产生光流场,因此在实际由光流场解算速度时,需要考虑姿态变化的影响。

以下视安装的单目相机为例,如图 3 所示。



(a) 滤波前的特征点对匹配结果

(b) 滤波后的特征点对匹配结果

图 2 采用 RANSAC 算法对特征点对滤波

Fig. 2 Filtering features by RANSAC algorithm

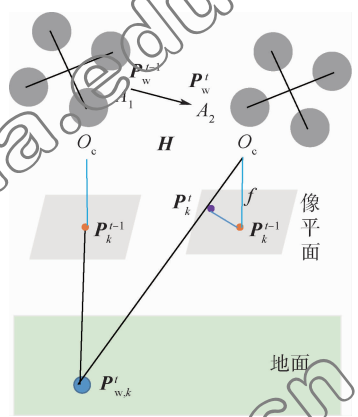


图 3 相机运动时空间点成像变化

Fig. 3 Point-mapping changing caused by moving camera

$t-1$ 到 t 时刻,相机从 A_1 点(坐标为 $\boldsymbol{P}_w^{t-1}$)运动到 A_2 点(坐标为 $\boldsymbol{P}_w^t$)。假设相机安装在飞机重心,像平面与机体 xy 平面平行,则相机到机体的旋转矩阵 $\boldsymbol{C}_c^b$ 可以求出,相机之间的移动可用单应矩阵 $\boldsymbol{H}$ 表示。令 t 时刻在图像中提取到 n 个特征点 $p_k^t, k = 1, 2, \cdots, n$,特征点在相机坐标系下的投影为 $\boldsymbol{P}_{c,k}^t$,对应于地面的点为 $\boldsymbol{P}_{w,k}^t$。已知相机内参矩阵为 $\boldsymbol{M}$,相机焦距为 f,f 和 $\boldsymbol{M}$ 可由标定相机得到,特征点在相机坐标系 O_c 轴上尺度为 Z_c。

鉴于图像更新频率很快(大于 80 Hz),认为连续两帧之间相机与地面的距离 h 未有明显的变化。则在 t 时刻有

$$\boldsymbol{P}_{c,k}^t = Z_c \boldsymbol{M}^{-1} p_k^t \tag{5}$$

$$\boldsymbol{P}_{w,k}^t = \boldsymbol{H} \boldsymbol{P}_{c,k}^t \tag{6}$$

$t-1$ 到 t 时刻 n 个特征点的稀疏光流为

$$\mathrm{OF}_k^t = (p_k^t - p_k^{t-1}) / \Delta t \tag{7}$$

单位时间相机在空间的移动(v_{bx}, v_{by})远远大于投影到像平面的像素的移动(Δu, Δv),相机距地面的高度 h 远远大于镜头的焦距 f。则光流与

http://bhxb.buaa.edu.cn

无人机线运动、角运动的关系可以描述为

$$\mathrm{OF}_k^t = (V_\mathrm{b} f \sin\eta)/h - w \tag{8}$$

则有

$$\begin{cases} v_{\mathrm{b}x} = \dfrac{\Delta u}{f\Delta t}h - qh \\ v_{\mathrm{b}y} = \dfrac{\Delta v}{f\Delta t}h - \gamma h \end{cases} \tag{9}$$

式中：V_b 为无人机线速度；η 为无人机 x 轴与无人机-目标之间连线的夹角；w 为无人机角速率；$v_{\mathrm{b}x}$、$v_{\mathrm{b}y}$ 分别为 x、y 方向的线速度；γ、q 分别为滚转和俯仰角速率，可通过精度较高的 IMU 测量得到。高度数据 h 可用超声传感器测量，对于前视双目相机的光流解算则需要获取每一对特征的深度 d，用 d 代替 h。如式(10)所示，(u_1, v_1)和(u_2, v_2)分别为左相机和右相机中的一对匹配点的图像层的坐标，以像素为单位。$\boldsymbol{M}_{\mathrm{iL}}$ 和 $\boldsymbol{M}_{\mathrm{iR}}$ 分别为左右相机的内参数，通过相机标定获得。(X, Y, Z)为该对匹配点对应的外部空间(即世界坐标系)点的实际坐标。在此，令世界坐标系与左相机的相机坐标系重合(见式(11))，求解出图像点对应的场景中的点的坐标(X, Y, Z)，即场景特征到相机的相对位置。工程应用中采用最小二乘法求解，以降低因数据噪声带来的解算误差。最后，采用卡尔曼滤波处理速度信息。

$$\begin{cases} Z_{c1}\begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11}^1 & m_{12}^1 & m_{13}^1 & m_{14}^1 \\ m_{21}^1 & m_{22}^1 & m_{23}^1 & m_{24}^1 \\ m_{31}^1 & m_{32}^1 & m_{33}^1 & m_{34}^1 \end{bmatrix}\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \boldsymbol{M}_{\mathrm{iL}}\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \\ Z_{c2}\begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11}^2 & m_{12}^2 & m_{13}^2 & m_{14}^2 \\ m_{21}^2 & m_{22}^2 & m_{23}^2 & m_{24}^2 \\ m_{31}^2 & m_{32}^2 & m_{33}^2 & m_{34}^2 \end{bmatrix}\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \boldsymbol{M}_{\mathrm{iR}}\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \end{cases} \tag{10}$$

$$\begin{cases} \boldsymbol{M}_{\mathrm{L}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \\ \boldsymbol{M}_{\mathrm{R}} = [\boldsymbol{R}_{\mathrm{r,l}} \quad \boldsymbol{T}_{\mathrm{r,l}}] \end{cases} \tag{11}$$

手持相机做平稳往复运动，图 4 显示了本文方法与 Guidance 和 PX4Flow[17] 等的测速性能对比。大的姿态角变化会引起较大的速度测量误差，使速度曲线产生大的“尖峰”，从图中可以看出(箭头指示)，因为引入了惯导角速率，本文方法能较好地降低姿态角变化带来的干扰。

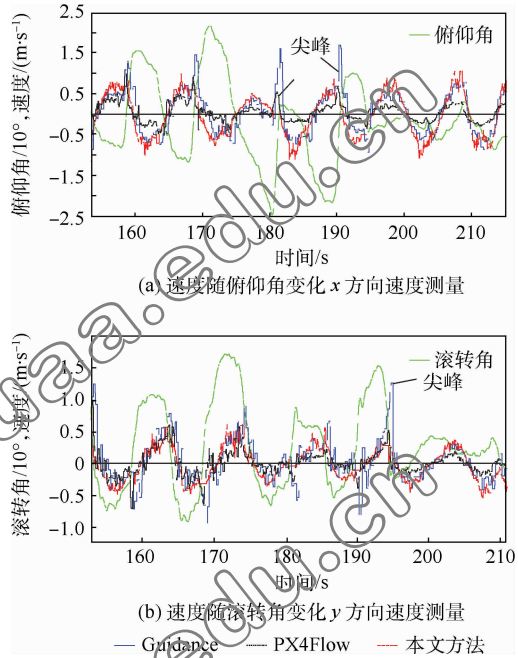


图 4　姿态变化时测速对比

Fig. 4　Velocity comparison with changing attitude

2　图标法里程计

针对室内地板格环境，本文采用图标法里程计，以地板格直线的交点为图标，通过追踪图标进行定位。算法基本流程如图 5 所示。

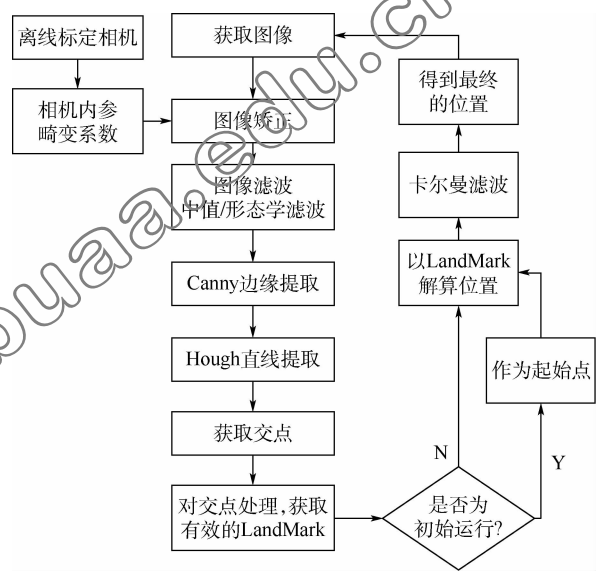
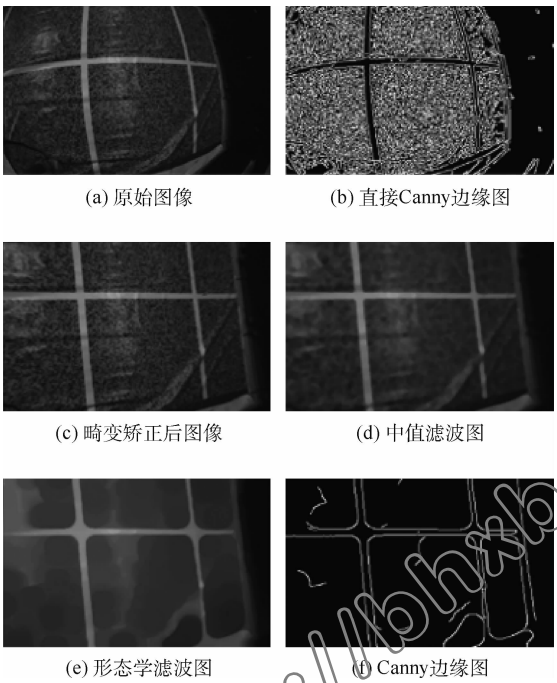


图 5　图标法里程计算法基本流程

Fig. 5　Algorithm flowchart of icon odometry

首先对摄像头获取的地板灰度图像进行滤波，利用 Canny 算子进行边缘提取(见图 6)。然后对获取的边缘图像利用 Hough 变换提取线段。根据以下原则对线段滤波：①合并“同线”的线段；②保留“水平”和“垂直”的线段；③滤除“较短”的线段。最后，求取线段所在直线的交点。

http://bhxb.buaa.edu.cn



(a) 原始图像 (b) 直接Canny边缘图

(c) 畸变矫正后图像 (d) 中值滤波图

(e) 形态学滤波图 (f) Canny边缘图

图 6 图像滤波与边缘提取

Fig. 6 Image filtering and edge extraction

地板直线的交点表现在图像中是一个图像块,取图像块的几何中心的像素值(u^*,v^*)用于二维图像平面到三维空间的映射[18]。

$$z_c\begin{bmatrix}u^*\\v^*\\1\end{bmatrix}=[\boldsymbol{A}\quad\boldsymbol{0}]\begin{bmatrix}\boldsymbol{R}&\boldsymbol{t}\\\boldsymbol{0}&1\end{bmatrix}\begin{bmatrix}X\\Y\\Z\\1\end{bmatrix}\tag{12}$$

式中:$\boldsymbol{A}$ 为相机的内参矩阵;$\boldsymbol{R}$ 和 $\boldsymbol{t}$ 分别为相机坐标系与导航坐标系的旋转矩阵和平移向量。鉴于飞行器飞行高度远远大于相机的焦距 $H\gg f$,因此可以令 $Z\approx H$,这样 Z 可以通过超声传感器获得。

相机与飞机固连且垂直向下安装。如图 7 所示,假设飞机在 1 个直线交点的附近起飞,其相机定位到的第 1 个交点作为原点。则定义交点 0 为原点,即导航坐标系的(0,0,0),飞机此时的位置为(0,0,H)。假设飞机沿着红色箭头的方向直线飞行到交点 3 的上空。在这个过程中,椭圆曲线范围内的点在不断变化,利用椭圆内点的定位平均值作为当前飞机的位置估计。之所以定义一个椭圆的范围,主要是因为考虑图像边缘畸变大,定位误差大,该椭圆的中心与图像的中心重合。而利用椭圆包围的点进行定位,误差小,只有当椭圆内搜索不到点时,才考虑椭圆外的点。已知地板格为边长为 1 m 的正方形,因此在飞机从 0 点到 3 点运动过程中,y 方向的值从 0 变为 1。

利用图标法里程计,只要保证视野里始终有交点,便可以实现对飞机的有效定位。

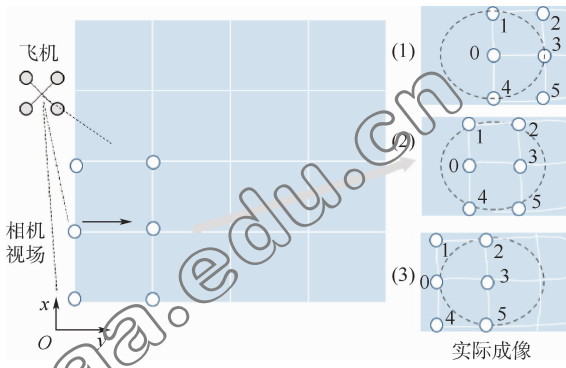


图 7 图标法里程计定位

Fig. 7 Positioning based on icon odometry

3 基于 EKF 的数据融合

第 1 节和第 2 节所述的光流法和图标法里程计都依赖于视觉传感器和外界特征(特征点和交点)。但在实际飞行中,由于种种原因,会出现视觉传感器失效的状况。为此,考虑视觉和惯导融合进行速度和位置的估计。

3.1 全状态下的视觉/惯导融合

假设图标法里程计和图像光流场比较稳定,将里程计、光流和惯导做数据融合,本文将这种情况定为全状态。

假设在极小的 Δt 时间内系统为非线性的匀变速过程。定义系统的状态量为导航坐标系下的飞机的三维全局位置,三维速度和三维加速度和飞机的俯仰角、滚转角、偏航角:

$$\boldsymbol{X}_k=[x,y,z,v_x,v_y,v_z,a_x,a_y,a_z,\theta,\phi,\psi]^{\mathrm{T}}\in\mathbf{R}^{12}$$

系统的观测量为视觉里程计测量的 x,y,z 方向的位置,光流法获得的速度,以及 MTI 的三轴线性加速度信息和姿态角:

$$\boldsymbol{Y}_k=[x_{\mathrm{vo}},y_{\mathrm{vo}},z_{\mathrm{vo}},v_{\mathrm{oF},x},v_{\mathrm{oF},y},v_{\mathrm{oF},z},a_{\mathrm{gyro},x},a_{\mathrm{gyro},y},a_{\mathrm{gyro},z},\theta_{\mathrm{gyro}},\phi_{\mathrm{gyro}},\psi_{\mathrm{gyro}}]^{\mathrm{T}}\in\mathbf{R}^{12}$$

定义过程模型为

$$\begin{cases}x_k=x_{k-1}+tv_{x,k-1}+\omega_{1,k-1}\\y_k=y_{k-1}+tv_{y,k-1}+\omega_{1,k-1}\\z_k=z_{k-1}+tv_{z,k-1}+\omega_{1,k-1}\\v_{x,k}=v_{x,k-1}+a_{x,k-1}t+\omega_{2,k-1}\\v_{y,k}=v_{y,k-1}+a_{y,k-1}t+\omega_{2,k-1}\\v_{z,k}=v_{z,k-1}-a_{z,k-1}t+\omega_{2,k-1}\\a_{x,k}=a_{x,k-1}+\omega_{3,k-1}\\a_{y,k}=a_{y,k-1}+\omega_{3,k-1}\\a_{z,k}=a_{z,k-1}+\omega_{3,k-1}\\\theta_k=\theta_{k-1}+\omega_{4,k-1}\\\phi_k=\phi_{k-1}+\omega_{4,k-1}\\\psi_k=\psi_{k-1}+\omega_{4,k-1}\end{cases}\tag{13}$$

http://bhxb.buaa.edu.cn

系统的观测模型为

$$\begin{cases} x_{\mathrm{vo},k} = x_k + \delta_{1,k} \\ y_{\mathrm{vo},k} = y_k + \delta_{1,k} \\ z_{\mathrm{vo},k} = z_k + \delta_{1,k} \\ v_{\mathrm{oF}x,k} = v_{x,k} + \delta_{2,k} \\ v_{\mathrm{oF}y,k} = v_{y,k} + \delta_{2,k} \\ v_{\mathrm{oF}z,k} = v_{z,k} + \delta_{2,k} \\ \begin{bmatrix} a_{\mathrm{gyrox},k} \\ a_{\mathrm{gyroy},k} \\ a_{\mathrm{gyroz},k} \end{bmatrix} = \boldsymbol{C}_{\mathrm{n}}^{\mathrm{b}} \begin{bmatrix} a_{x,k} \\ a_{y,k} \\ a_{z,k} \end{bmatrix} - g\boldsymbol{e} + \boldsymbol{C}_{\mathrm{n}}^{\mathrm{b}}\delta_{3,k} \\ \theta_{\mathrm{gyro},k} = \theta_k + \delta_{4,k} \\ \phi_{\mathrm{gyro},k} = \phi_k + \delta_{4,k} \\ \psi_{\mathrm{gyro},k} = \psi_k + \delta_{4,k} \end{cases} \tag{14}$$

式中：ω 为观测噪声；δ 为量测噪声；$\boldsymbol{C}_{\mathrm{n}}^{\mathrm{b}}$ 为导航-机体坐标系转换矩阵；g 为重力加速度；$\boldsymbol{e}$ 为单位向量。观测方程为非线性模型，根据 EKF 公式将 f_k 在最优估计点处展开，将观测方程线性化。

3.2　非全状态下的视觉/惯导融合

1）图标法里程计失效的情况

当地面没有可用的边线用于图标法里程计时，里程计的定位将失效。具体表现为视野内没有可用的交点。此时，置 VO_KEY 为 False，仅利用光流和惯导组合，估计定位信息。此时观测量为

$$\boldsymbol{Y}_k = [v_{\mathrm{oF},x}, v_{\mathrm{oF},y}, v_{\mathrm{oF},z}, a_{\mathrm{gyro},x}, a_{\mathrm{gyro},y}, a_{\mathrm{gyro},z}, \theta_{\mathrm{gyro}}, \phi_{\mathrm{gyro}}, \psi_{\mathrm{gyro}}]^{\mathrm{T}} \in \mathbf{R}^9$$

状态方程参考式（14），观测模型为

$$\begin{cases} v_{\mathrm{oF}x,k} = v_{x,k} + \delta_{2,k} \\ v_{\mathrm{oF}y,k} = v_{y,k} + \delta_{2,k} \\ v_{\mathrm{oF}z,k} = v_{z,k} + \delta_{2,k} \\ \begin{bmatrix} a_{\mathrm{gyrox},k} \\ a_{\mathrm{gyroy},k} \\ a_{\mathrm{gyroz},k} \end{bmatrix} = \boldsymbol{C}_{\mathrm{n}}^{\mathrm{b}} \begin{bmatrix} a_{x,k} \\ a_{y,k} \\ a_{z,k} \end{bmatrix} - g\boldsymbol{e} + \boldsymbol{C}_{\mathrm{n}}^{\mathrm{b}}\delta_{3,k} \\ \theta_{\mathrm{gyro},k} = \theta_k + \delta_{4,k} \\ \phi_{\mathrm{gyro},k} = \phi_k + \delta_{4,k} \\ \psi_{\mathrm{gyro},k} = \psi_k + \delta_{4,k} \end{cases} \tag{15}$$

2）视觉传感器失效的情况

如果视觉传感器失效，则利用惯导积分做短时的导航，直到光流速度或者里程计的测量恢复。这里判断光流速度失效的原则为连续图像光流点数量小于 4 个，或连续出现野值。

3.3　视觉、惯导的多速率问题

惯性传感器数据更新速率快，视觉传感器的数据更新速率慢，因此其中涉及了多速率融合的问题。对此的解决方法为：保证状态和噪声的定义与之前的定义相同，修改传感器的观测。系统的基本采样周期为 T_0，惯导和视觉传感器的采样周期分别为 T_1 和 T_2。这里保证 T_1 和 T_2 为 T_0 的整数倍，即 $T_1 = n_1 T_0$，$T_2 = n_2 T_0$。令 κT_0 为各个观测数据的采样周期的最小公倍数，则定义观测矩阵和观测噪声协方差如下：

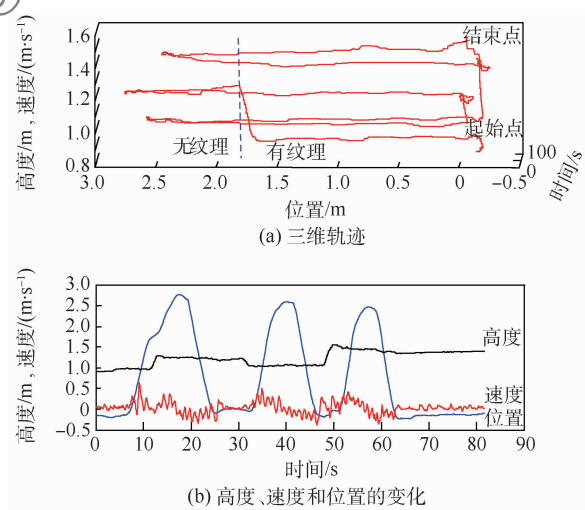
$$\boldsymbol{H}_k = \begin{cases} H_{ik} & \mathrm{mod}(k, n_i) = 0 \ \&\ \mathrm{mod}(k, \kappa) \neq 0 \\ \begin{bmatrix} H_{1k} \\ H_{2k} \end{bmatrix} & \mathrm{mod}(k, \kappa) = 0 \\ \boldsymbol{0} & \text{其他} \end{cases} \tag{16}$$

$$\boldsymbol{R}_k = \begin{cases} R_{ik} & \mathrm{mod}(k, n_i) = 0 \ \&\ \mathrm{mod}(k, \kappa) \neq 0 \\ \mathrm{diag}(R_{1k}, R_{2k}) & \mathrm{mod}(k, \kappa) = 0 \\ \boldsymbol{I} & \text{其他} \end{cases} \tag{17}$$

式中：mod(·)为求余函数。

已知试验所用的视觉传感器的数据更新速率约为 30 Hz，惯导的更新速率为 100 Hz，令 T_0 = 10 ms，则 T_1 = 20 ms，T_2 = 10 ms，$n_1 = 2$，$n_2 = 1$。在 ROS 操作系统中，多线程采集到的数据可以通过时间戳进行数据对准，并可以根据时间戳来判断是否采样得新数据，这样根据式（16）和式（17）便可以方便地编程实现多速率数据的融合。

图 8 显示了“地板格”从有（视觉里程计可用、下视光流可用）到无（纯白色地面、视觉里程计失效、下视相机失效）时的位置估计，包括三维轨迹及高度、速度和位置随时间的变化。手持相机运动到约 2 m 的位置时，开始移动出地板（朝着位置增加的方向）。在 3 次往返后，终点和起点近似重合，依然可以实现有效定位。



(a) 三维轨迹

(b) 高度、速度和位置的变化

图 8　定位测试

Fig. 8　Localization test

4 试验与分析

鉴于运动捕捉系统 OptiTrack 的数据延迟小(低于 5 ms),定位误差小(约 0.1 mm),可以 OptiTrack的定位数据作为基准,衡量本文方法的精度和延迟。本文采用美国 OptiTrack Prime 41 相机搭建运动捕捉系统进行试验。

图 9 ~ 图 11 分别为将本文方法与运动捕捉系统获得的位置和速度信息进行对比结果,可得本文方法延迟约 50 ms,位置误差约 0.15 m,速度误差约 0.05 m/s,基本满足室内环境下小型无人机的飞行需求。

本文搭建四轴八桨多旋翼无人机平台(见图 12),以手控的方式操纵飞机飞行(见图 13),实时记录导航数据,并将本文方法的测速数据与 Guidance 和 PX4Flow 导航模块的测速数据进行了对比。

图 14 显示了不同速度下本文方法与 Guidance 导航模块测速情况对比。无人机的飞行速度基本在 0 ~ 2.5 m/s 的范围。在此速度范围内本文方法可以有效实现对无人机的导航。当速度大于 2 m/s 时,有效特征点数下降,提到的光流场数量降低,当速度大于 2.5 m/s 时,有效光流场数量频繁出现小于 8 的状态,单纯的视觉信息将不再可靠。

图 15 和图 16 显示了某次飞行测试的导航数据。其中导航速度信息与 PX4Flow 的测速信息进行了对比。从图中的数据可以看出,本文方法在实际的无人机平台上比较稳定,可以尝试用于自动(闭环)飞行。

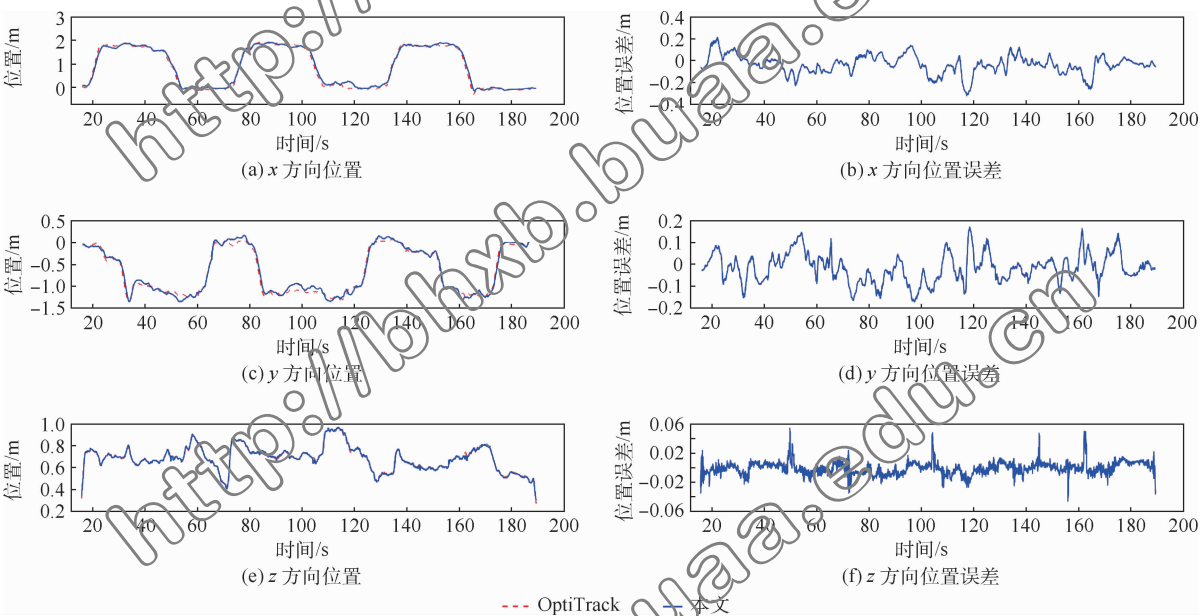


图 9 三维位置估计与位置误差

Fig. 9 Localization estimation and positioning error in 3D-directions

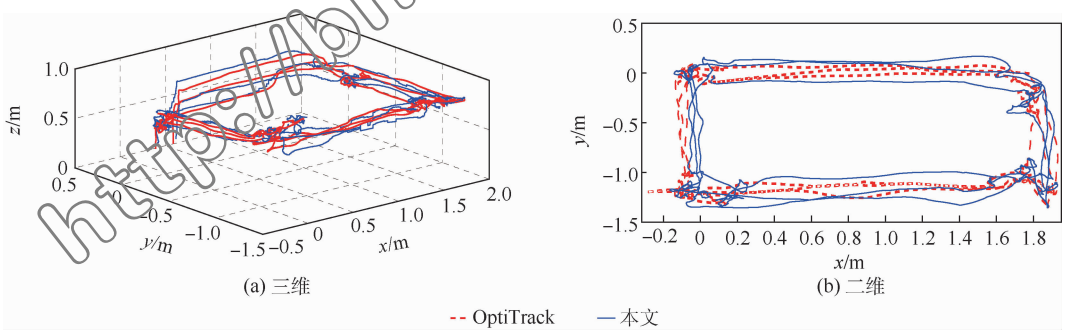


图 10 三维/二维运动轨迹

Fig. 10 Three/two-dimensional motion trajectory

http://bhxb.buaa.edu.cn

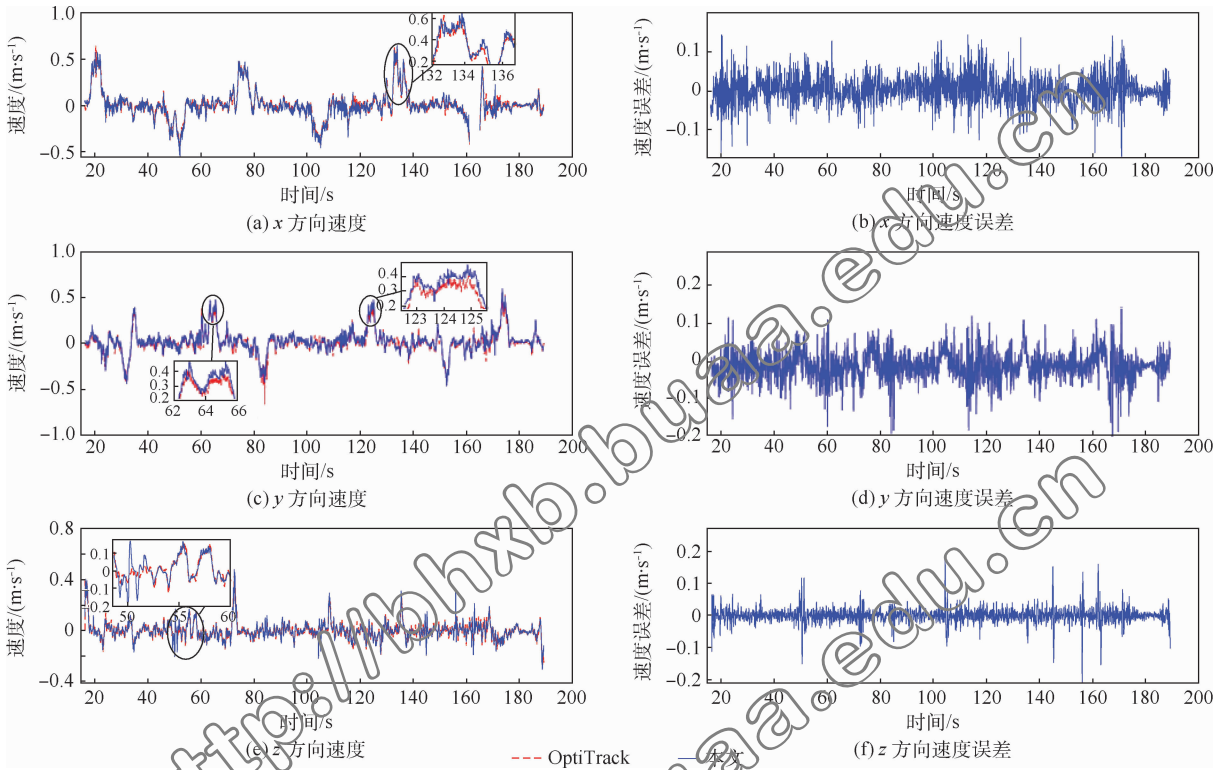


图 11　三维速度估计与速度误差

Fig. 11　Velocity estimation and velocity error in 3D-directions

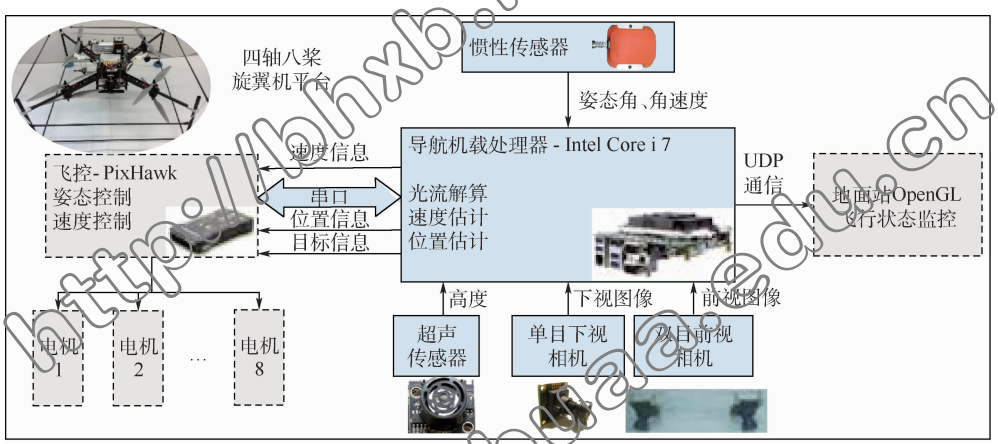


图 12　无人机平台及系统结构

Fig. 12　UAV platform and system structure



图 13　飞行视频截图

Fig. 13　Screenshots of flight video

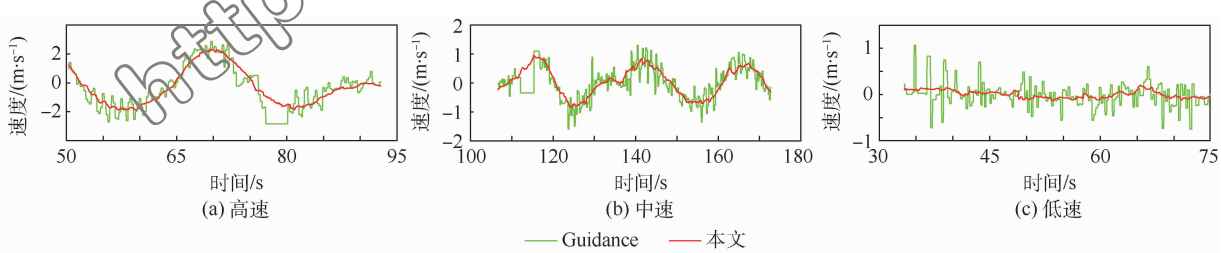


图 14　手控飞行速度测试

Fig. 14　Flight velocity test via manual control

http://bhxb.buaa.edu.cn

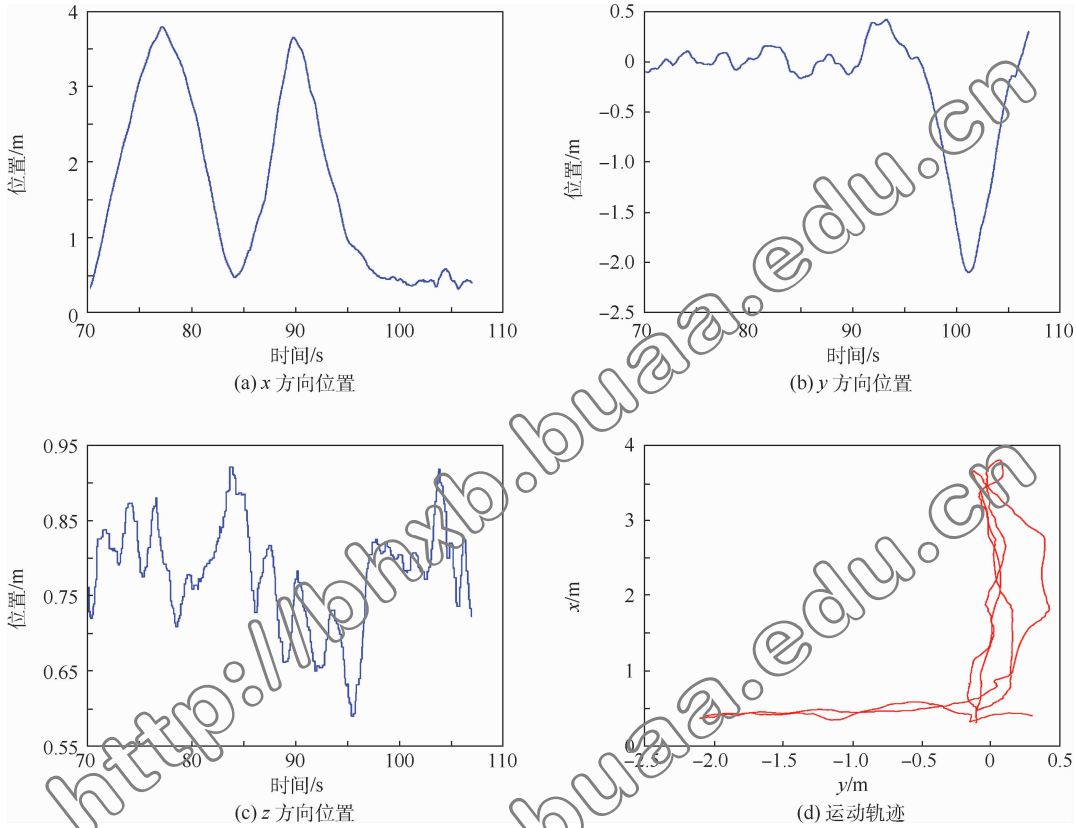


图 15 位置估计

Fig. 15 Position estimation

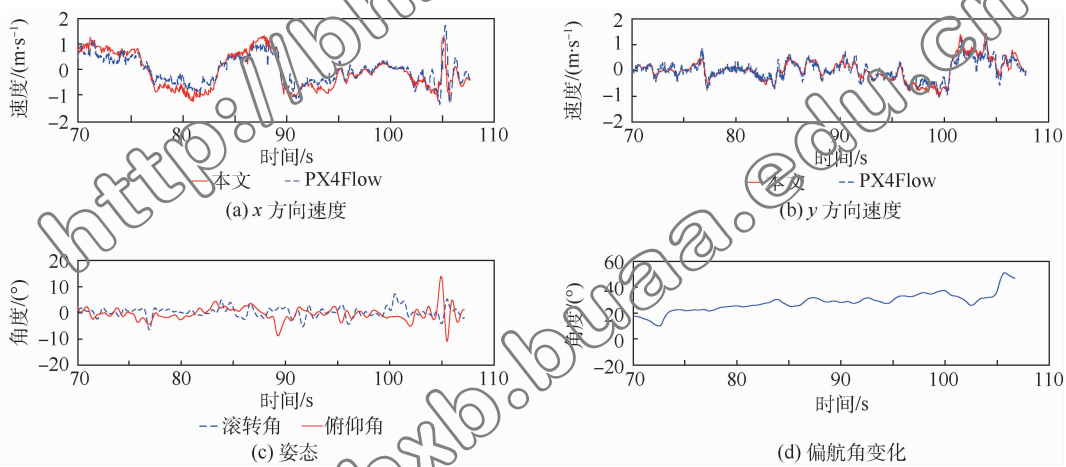


图 16 速度估计与飞机姿态

Fig. 16 Velocity estimation and aircraft attitude

5 结 论

本文提出了一种视觉和惯导融合用于估计小型无人机速度和位置的方法,可得到以下结论:

1) 基于 ORB 特征点的前后向追踪的光流提取算法,提高了光流场的提取精度,降低了算法的时间开销。

2) 基于图标法里程计的快速定位算法,可以有效应对带地板格的室内环境定位。该算法简单,实时性好。

3) 利用 EKF 将图标法里程计、光流速度和惯导信息进行融合,降低了方法对环境的依赖。

4) 以延迟小于 5 ms,精度小于 0.01 mm 的运动捕捉系统为参考基准,与本文提出的运动估计方法进行对比验证,得到本文方法延迟约 50 ms,定位误差约 0.15 m,速度误差约 0.05 m/s。同时搭建了多旋翼平台进行飞行验证。

http://bhxb.buaa.edu.cn

参考文献（References）

[1] SHEN S J. Autonomous navigation in complex indoor and outdoor environments with micro aerial vehicles[D]. Philadelphia: University of Pennsylvania,2014.

[2] 吴琦,蔡志浩,王英勋. 用于无人机室内导航的光流与地标融合方法[J]. 控制理论与应用,2015,32(11):1511-1517.
WU Q,CAI Z H,WANG Y X. Optical flow and landmark fusion method for UAV indoor navization[J]. Control Theory & Applications,2015,32(11):1511-1517(in Chinese).

[3] LI P,LAMBERT A. A monocular odometer for a quadrotor using a homogra-phy model and inertial cues[C]// IEEE Conference on Robotics and Biomimetics. Piscataway, NJ: IEEE Press, 2015:570-575.

[4] 叶长春. IARC 第 7 代任务中定位与目标跟踪方法研究[D]. 杭州:浙江大学,2016.
YE C C. Research on localization and object tracking for the IARC mission7[D]. Hangzhou, Zhejiang University, 2016 (in Chinese).

[5] MUR-ARTAL R, MONTIEL J M M, TARDOS J D. ORB-SLAM: A versatile and accurate monocular SLAM system[J]. IEEE Transactions on Robotics,2015,31(5):1147-1163.

[6] LEUTENEGGER S, FURGALE P, RABAUD V, et al. Keyframe-based visual-inertial SLAM using nonlinear optimization[C]// Robotics: Science and Systems,2013:789-795.

[7] CHAO H,GU Y,GROSS J,et al. A comparative study of optical flow and traditional sensors in UAV navigation[C]// American Control Conference(ACC). Piscataway, NJ: IEEE Press,2013: 3858-3863.

[8] MAMMARELLA M,CAMPA G,FRAVOLINI M L,et al. Comparing optical flow algorithms using 6-dof motion of real-world rigid objects[J]. IEEE Transactions on, Systems, Man, and Cybernetics, Part C: Applications and Reviews, 2012, 42(6): 1752-1762.

[9] DJI Innovations. PHANTOM 4 user's manual V1.2[EB/OL]. (2016-12-23) https://dl.djicdn.com/downloads/phantom_4/cn/Phantom_4_User_Manual_cn_v1.2_160328.pdf.

[10] Hover Camera 2016[EB/OL]. (2016-12-23) http://gethover.com.

[11] RUBLEE E, RABAUD V, KONOLIGE K, et al. ORB: An efficient alternative to SIFT or SURF[J]. Proceedings, 2011, 58(11):2564-2571.

[12] ROSTEN E, DRUMMOND T. Machine learning for high-speed corner detection[C]// European Conference on Computer Vision. Berlin: Springer-Verlag, 2006:430-443.

[13] CALONDER M, LEPETIT V, STRECHA C, et al. BRIEF: Binary robust independent elementary features[C]// European Conference on Computer Vision. Berlin: Springer-Verlag, 2010: 778-792.

[14] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2):91-110.

[15] BAY H, TUYTELAARS T, GOOL L V. SURF: Speeded up robust features[J]. Computer Vision & Image Understanding, 2006, 110(3):404-417.

[16] CHUM O, MATAS J, KITTLER J. Locally optimized RANSAC[J]. Lecture Notes in Computer Science, 2003, 2781:236-243.

[17] HONEGGER D, MEIER L, TANSKANEN P, et al. An open source and open hardware embedded metric optical flow cmos camera for indoor and outdoor applications[C]// International Conference on Robotics and Automation. Piscataway, NJ: IEEE Press, 2013:1736-1741.

[18] HARTLEY R I, ZISSERMAN A. Multi-view geometry in computer vision[M]. Cambrige: Cambridge University Press, 2004: 239-247.

作者简介:

王亭亭　女,硕士研究生。主要研究方向:机器视觉、无人机视觉导航。

蔡志浩　男,副教授,硕士生导师。主要研究方向:无人机自主控制与导航、多机协同与训练。

王英勋　男,研究员,博士生导师。主要研究方向:系统设计、自主控制与模拟训练。

http://bhxb.buaa.edu.cn

Integrated vision/inertial navigation method of UAVs in indoor environment

WANG Tingting[1], CAI Zhihao[1,2,*], WANG Yingxun[1,2]

(1. School of Automation Science and Electrical Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China;
2. Aircraft Control Integration National Defense Key Laboratory, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: A new integrated navigation method based on inertial sensor, optical flow and visual odometry is proposed for self-navigation indoor in GPS-denied environment. An ORB optical flow based method is also proposed for estimating real-time three-axis velocity of the UAV. The algorithm improves the traditional pyramid Lucas-Kanade method using sparse optical flow based on feature points. The tracking of feature points is made more accurate by applying forward-backward tracking and random sampling consensus strategies. For position estimation, a visual odometry method with integrated vision/inertial navigation is adopted, which uses the artificial icon method, visual optical flow information and inertial navigation data. Finally, the velocity and position estimations from the proposed method are validated via actual flight test and via comparison with velocity measurement information from a PX4Flow module and a Guidance module and with locating information from movement capture system.

Keywords: UAV; vision navigation; optical flow; ORB features; multi-sensor fusion

http://bhxb.buaa.edu.cn

Received: 2016-12-23; **Accepted**: 2017-02-06; **Published online**: 2017-03-23 16:44

URL: kns.cnki.net/kcms/detail/11.2625.V.20170323.1644.003.html

Foundation item: Aeronautical Science Foundation of China (20135851043)

* **Corresponding author**. E-mail: czh@buaa.edu.cn