

文章编号:1001-9081(2017)S2-0127-07

## 基于单目视觉与惯导融合的无人机位姿估计

熊敏君, 卢惠民\*, 熊 丹, 肖军浩, 吕 鸣

(国防科技大学 机电工程与自动化学院, 长沙 410073)

(\*通信作者电子邮箱 lhmnew@nudt.edu.cn)

**摘 要:**针对无卫星信号条件下仅利用惯性导航进行无人机(UAV)位姿估计精度低的问题,提出了一种基于扩展卡尔曼滤波(EKF)的单目视觉/惯导组合位姿估计算法。首先,借助超声波传感器,基于最小二乘法获取单目视觉的绝对尺度信息,从而计算得到视觉位姿估计结果;其次根据惯性测量单元(IMU)的运动学模型建立系统方程,进行惯导解算,并推导误差状态方程对系统误差状态建模;最后,将视觉位姿估计结果作为观测量,基于扩展卡尔曼滤波得到系统误差状态的最优估计值,从而对惯导解算的结果进行修正。使用该算法在室外150 m范围场景开展实验研究,位置估计的均方根误差(RMSE)低于0.995 m,航向角估计的RMSE为2.235°,水平姿态角估计的RMSE低于1.915°。

**关键词:**单目视觉;惯性导航;位姿估计;尺度估计;扩展卡尔曼滤波

**中图分类号:** TP242.6 **文献标志码:** A

### Pose estimation of UAV based on monocular vision and inertial navigation

XIONG Minjun, LU Huimin\*, XIONG Dan, XIAO Junhao, LYU Ming

(College of Mechatronic Engineering and Automation, National University of Defense Technology, Changsha Hunan 410073, China)

**Abstract:** To deal with the low accuracy of pose estimation for Unmanned Aerial Vehicles (UAVs) when only using inertial navigation without available Global Navigation Satellite System (GNSS), this paper proposed an integrated pose estimation algorithm combining monocular vision and Inertial Measurement Unit (IMU) based on Extended Kalman Filter (EKF). Firstly, the absolute scale in monocular vision system was obtained by using ultrasonic sensor and least square method, so the pose estimation results could be obtained using monocular vision; secondly, the system equation was built according to the kinematic model of the IMU to realize inertial navigation calculation, and the error state equation was derived to model the system's error states; finally, the optimal estimation of the system's error states was completed based on EKF using the pose estimation results of monocular vision as the measurement information, so the pose estimation results obtained by inertial navigation were corrected. Experiments were conducted in large scale environments with the range of 150 meters. The Root Mean Square Error (RMSE) of position estimation is less than 0.995 meters, the RMSE of heading estimation is 2.235 degrees, and the RMSE of horizontal attitude is less than 1.915 degrees.

**Key words:** monocular vision; inertial navigation; pose estimation; scale estimation; Extended Kalman Filter (EKF)

## 0 引言

无人机的位姿估计是其执行其他复杂任务的基本前提,针对无人机的位姿估计问题,卫星导航是比较传统且常用的方法。然而在很多环境下卫星信号微弱甚至缺失,以致卫星导航系统不能正常工作。为了解决这一问题,可采用其他多种自主导航方式,包括惯性导航、视觉导航、雷达定位导航等。其中单目视觉传感器体积小、重量轻、功耗低,特别适合在小型多旋翼无人机上使用。其获得的图像信息,不仅可以用于自定位,还可以用于目标检测、跟踪、障碍物识别等任务。近几年,视觉同步定位与建图(Visual Simultaneous Localization And Mapping, V-SLAM)技术<sup>[1]</sup>的发展非常迅速,为研究基于单目视觉的无人机位姿估计奠定了基础。单目视觉定位输出频率低,为了满足小型无人机高动态特性的位姿估计需求,本

文采用输出频率高的惯性测量单元(Inertial Measurement Unit, IMU)与视觉进行组合,从而获得输出频率与精度较好的位姿估计结果。

## 1 相关研究

V-SLAM是指通过相机获得的图像序列得到自身的位姿估计,同时构建周围的环境地图。目前主要存在两大类V-SLAM方法:一种是基于特征提取的方法,它通过前后两帧图像的特征点匹配来计算当前相机的位姿,典型的算法有PTAM算法<sup>[2]</sup>和ORB\_SLAM算法<sup>[3]</sup>。另外一种是直接法,它通过对图像中某些像素或者全部像素的像素值进行帧间优化求取相机的位姿变化,典型的算法有SVO算法<sup>[4]</sup>以及LSD-SLAM算法<sup>[5]</sup>。直接法比基于特征提取的方法计算量更大,为了满足无人机位姿估计的实时性要求,本文采用基于特征

收稿日期:2017-05-02;修回日期:2017-07-01。

基金项目:国家自然科学基金资助项目(61403409,61503401);中国博士后科学基金资助项目(2014M562648)。

作者简介:熊敏君(1993—),女,湖南常德人,硕士研究生,主要研究方向:机器人视觉、多传感器信息融合;卢惠民(1980—),男,福建南平人,副教授,博士,主要研究方向:机器人视觉、自主移动机器人;熊丹(1986—),男,湖南岳阳人,博士研究生,主要研究方向:机器人视觉;肖军浩(1984—),男,河北保定人,讲师,博士,主要研究方向:SLAM、自主移动机器人;吕鸣(1957—),男,吉林吉林人,高级工程师,主要研究方向:导航、制导与控制。

提取的 ORB\_SLAM 算法作为视觉位姿估计算法。

在小范围、短时间且无人机机动不高的情况下,单纯依靠视觉实现位姿估计是可行的,对于更加复杂的情况则需要融合其他传感器的信息才能获得更加准确的位姿估计结果。造成定位失败的原因主要有两个:一方面,当环境中特征较稀疏或者像素值变化不明显时,仅仅依靠视觉无法进行定位;另一方面,视觉里程计需要依靠前后帧的差别进行位姿估计,因此对旋转运动和高速运动特别敏感,在这些情况下位姿估计的效果不是很理想。IMU 能很好地解决视觉定位中存在的这些问题,因此,采用视觉和 IMU 融合的方式进行无人机的位姿估计近年来受到了普遍关注。

视觉传感器和 IMU 融合的方法通常可分为松耦合<sup>[6-7]</sup>和紧耦合<sup>[8-9]</sup>两种。其中,松耦合是指将视觉传感器和 IMU 作为两个单独的模块,分别计算得到各自的位姿信息,然后通过扩展卡尔曼滤波<sup>[10-11]</sup>或无迹卡尔曼滤波(Unscented Kalman Filter, UKF)<sup>[12]</sup>进行融合。紧耦合则是指将视觉和 IMU 得到的中间数据通过一个优化滤波器进行处理,最终得到位姿信息的过程。与紧耦合算法相比,松耦合算法相对简单,且扩展性很强,可以加入多种传感器信息进行融合,但融合得到的位姿估计精度相对较低。文献[13]与文献[14]描述的系统属于松耦合,文献[13]使用单目相机与 IMU 融合进行位姿估计,并采用气压计对 PTAM 算法的尺度进行估计,该算法初始化过程较为复杂,在高动态的情况下应用稳定性较差;文献[14]采用 EKF 对 ORB\_SLAM 算法与惯导进行信息融合,并进行了车载实验,通过 GNSS 辅助与闭环检测估计尺度值,尺度估计过程相对繁琐。文献[15]与文献[16]描述的系统属于紧耦合,文献[15]提出了一种多状态约束的卡尔曼滤波算法(Multi-State Constraint Kalman Filter, MSCKF)来实现视觉和 IMU 的信息融合,位姿精度较高,但算法实现复杂,计算量很大;文献[16]采用 MSCKF 算法并在手机上实现,但该系统只能在室内使用,应用范围比较局限。

本文采用松耦合的方式对单目视觉与 IMU 的信息进行融合,从而实现位姿估计,主要基于以下两点考虑:首先,由于小型多旋翼无人机具有高机动性特点,对位姿估计输出频率要求较高,而松耦合的计算复杂度更低,可以满足无人机位姿估计高实时性的需求;其次,无人机上传感器相对较多,如气压传感器、全球导航卫星系统(Global Navigation Satellite System, GNSS)等,松耦合方式的可扩展性较强,后续研究可加上气压传感器进行融合,或在可获得 GNSS 信号的情况下进一步融合 GNSS 信息,从而进一步提高定位精度。本文主要工作如下:

1) 针对单目 ORB\_SLAM 算法无法获得绝对尺度的问题,采用无人机上常用的超声波传感器进行尺度估计,基于最小二乘算法对尺度估计值进行优化。

2) 将单目视觉得到的位姿估计值作为观测量,采用扩展卡尔曼滤波算法对 IMU 位姿估计值进行修正,从而解决大场景环境下卫星信号缺失时无人机位姿估计精度低的问题。

## 2 系统组成及初始化处理

### 2.1 系统组成

本文用于位姿估计的传感器包括单目相机、超声波传感器和 IMU 传感器。如图 1 所示,单目相机和超声波传感器组成一个带尺度的视觉里程计,频率为 25 Hz;IMU 传感器通过惯导解算得到六自由度的位姿信息,频率为 200 Hz。上述两

个通道的位姿估计结果通过扩展卡尔曼滤波算法进行信息融合,最终得到频率为 200 Hz 的位姿估计结果。

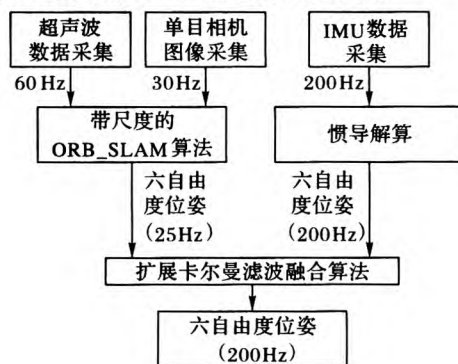


图1 系统整体框图

### 2.2 坐标系定义

在对系统进行建模之前,首先需要定义相关的坐标系,本文中相关坐标系定义如图 2 所示。其中, $o_i x_i y_i z_i$  表示 IMU 的体坐标系; $o_c x_c y_c z_c$  表示单目相机的体坐标系; $o_w x_w y_w z_w$  表示世界坐标系,世界坐标系与 IMU 初始时刻的体坐标系重合。 $p_c^w, q_c^w$  分别表示 IMU 在世界坐标系下的位置和姿态; $p_i^c, q_i^c$  分别表示相机的位置和姿态在 IMU 坐标系下的表示,为固定值,可通过 IMU 与相机位姿的标定得到; $p_c^w, q_c^w$  表示相机的位置和姿态(通过 ORB\_SLAM 算法获得)在世界坐标系下的表示。由于 ORB\_SLAM 算法将其初始化成功后第一帧的相机坐标系作为视觉输出的基准坐标系  $w_c$ , 而初始时刻 IMU 与相机坐标系的相对关系可通过标定获得,这样即可得到 ORB\_SLAM 算法的位姿输出在  $w$  系下的表示。

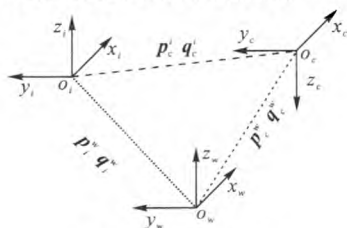


图2 系统坐标系定义

### 2.3 视觉和 IMU 内外参标定

本文将 IMU 初始时刻的坐标系作为世界坐标系,而单目相机和 IMU 安装的空间位置是固联的,如图 3 所示,它们的坐标系定义均符合右手定则。从图中可以看出,相机和 IMU 坐标系之间存在一个固定的旋转平移关系,可通过初始的外参标定获得。



图3 固联的视觉惯导系统

在对相机与 IMU 进行初始标定前,首先需要单独对二者进行标定获得相机的内参以及 IMU 的零漂等参数。相机的内参标定采用棋盘格标定板进行,可借助机器人操作系统

(Robot Operating System, ROS) 提供的软件包或 Matlab 标定工具箱进行;IMU 的零漂标定需借助于三轴转台进行。

相机和 IMU 之间的初始位姿采用开源的 ROS 包进行标定,使用棋盘格标定板或二维码标定板均可进行标定,本文实验中发现二维码标定板标定的效果更好。标定过程与相机内参标定类似,主要区别是在标定相机内参时标定板的姿态需要变化,而相机和 IMU 位姿标定时需要标定板固定,固联的相机和 IMU 沿着不同的方位加速运动。

## 2.4 IMU 的初始对准

惯性器件在使用前需要进行初始对准,初始对准的目的是为了估计 IMU 的初始姿态,从而得到其初始姿态在导航坐标系下的表示,具体到本系统中则是为了得到重力矢量在世界坐标系下的表示。

对于高精度惯导而言,通常是在静止情况下利用地球重力以及地球自转角速度通过解析对准的方法来获得初始姿态。然而对于本文使用的低精度 IMU 来说,其噪声较大,甚至大于地球自转角速度,无法使用解析对准进行计算。由于系统采用的 IMU 自带磁强计,可输出地磁信息,因此可利用磁强计数据来计算偏航角,利用静止状态下当地重力矢量在三个轴向上的分量以及 IMU 测得的加速度来计算俯仰角和滚转角如式(1)~(2):

$$\theta_x = \arctan(g_y/g_z) \quad (1)$$

$$\theta_y = -\arcsin(g_x/g) \quad (2)$$

其中:  $\theta_x, \theta_y$  分别代表滚转角和俯仰角,  $g$  代表当地的重力值,  $g_x, g_y, g_z$  分别代表 IMU 在三个轴上的加速度输出,通过滚转角、俯仰角以及磁强计的信息由式(3)~(5)即可得到偏航角  $\theta_z$ :

$$\theta_z = \arctan(M_{H_y}/M_{H_x}) \quad (3)$$

$$M_{H_x} = \cos \theta_y M_x + \cos \theta_z \sin \theta_y M_y - \cos \theta_x \sin \theta_y M_z \quad (4)$$

$$M_{H_y} = \cos \theta_x M_y + \sin \theta_x M_z \quad (5)$$

其中,  $M_x, M_y, M_z$  分别代表磁强计在三个轴向上的输出。通过上述方法可得到一个粗对准的结果,为使对准结果更加精确,本文还通过卡尔曼滤波进行精对准。

## 3 基于最小二乘的单目视觉尺度估计

单目 V-SLAM 算法在位姿估计的过程中无法得到绝对尺度信息,需要依靠其他信息辅助实现尺度估计,才能在无人机上实际应用。下面将详细介绍基于最小二乘法的单目视觉 ORB\_SLAM 的绝对尺度估计。

### 3.1 ORB\_SLAM 概述

ORB\_SLAM 是目前效果最好的单目 V-SLAM 算法之一,按照前述分类属于间接法 V-SLAM,通过提取环境图像中的 ORB(Oriented fast and Rotated BRIEF)特征进行特征匹配来实现定位,所创建的地图属于稀疏地图。其框架与 PTAM(Parallel Tracking and Mapping)算法框架相似,均采用了多线程的思想,定位跟踪与建图分别在两个不同的线程进行,提高了程序的运行效率。ORB\_SLAM 算法在此基础上加入了闭环检测,整个算法框架分为三个主线程,分别是 Tracking、LocalMapping 和 LoopClosing。跟踪线程(Tracking)主要负责定位,获得两帧之间的相机位姿变化;局部建图线程(LocalMapping)主要负责局部优化与建图,包括关键帧与地图点的插入、存储及优化,同时进行局部 BA(Bundle Adjustment)优化;闭环检测线程(LoopClosing)主要负责闭环检测,从而修正定位与建图结果的漂移。由于本文主要研究大场景下的小型无人机位姿估计,对于算法的实时性要求很高,因此本文对 ORB\_SLAM 算法进行了修改,关闭闭环检测线程并对关键帧与地图点的数量进行限制,从而大大提高了算法的实时性。

### 3.2 ORB\_SLAM 的尺度估计

视觉测量无法得到深度/尺度信息是单目相机的共性问题,ORB\_SLAM 算法使用单目相机作为传感器时同样存在这个问题。初始化过程中,ORB\_SLAM 通过对所有特征点的深

度取中值得到一个深度的初始值,因而并不能得到绝对尺度信息。本文提出利用超声波传感器获得飞行高度信息,并基于最小二乘法对单目 ORB\_SLAM 算法的绝对尺度进行估计。

在小型无人机的起飞阶段,单目 ORB\_SLAM 可得到固定时间间隔内无人机上升的高度差估计值  $z_i$ ,同时超声波传感器也可测得相同时间间隔内无人机上升的高度差  $h_i$ ,由于超声波传感器测得的高度为相对地面的绝对高度,因此尺度  $\lambda$  近似满足  $\lambda \approx h/z$ 。

超声波传感器的测量噪声近似服从高斯分布,通过多组数据样本可减少噪声带来的估计误差,在得到一系列样本数据  $\{(z_1, h_1), (z_2, h_2), \dots, (z_n, h_n)\}$  ( $n$  取值为 8~10)后即可对尺度进行参数估计,最简单的方法是对得到的样本数据求平均,其形式如式(6):

$$\lambda = \left( \frac{h_1}{z_1} + \frac{h_2}{z_2} + \dots + \frac{h_n}{z_n} \right) / n \quad (6)$$

上述方法估计的尺度误差较大,尤其是在前后两次尺度估计值相差较大的情况下无法得到最优解。因此,本文采用最小二乘估计来得到误差平方和最小的参数估计,具体过程如下。

根据现有测得的  $n$  组独立观测值以及尺度与高度的关系式可以得到式(7):

$$\begin{cases} h_i = \lambda z_i + \varepsilon_i \\ E(\varepsilon_i) = 0, D(\varepsilon_i) = \sigma^2; i = 1, 2, \dots, n \end{cases} \quad (7)$$

其中:  $\lambda$  为需要估计的参数,  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$  为误差值,且相互独立,最小二乘法的目的是估计出一个最优  $\hat{\lambda}$ , 使得其误差的平方和最小:

$$\hat{\lambda} = \arg\min_{\lambda} \sum_{i=1}^n (h_i - \lambda z_i)^2 \quad (8)$$

实际应用中,单目视觉 ORB\_SLAM 尺度存在漂移,需要对不同时间段测量到的数据施加不同的权重,给就近的测量值较大的权重,适当减小过去时刻的测量值权重,采用加权最小二乘法实现尺度估计,即在残差平方和时加上权重  $w$ ,得到最终的尺度估计为:

$$\hat{\lambda} = \sum_{i=1}^n w_i z_i h_i / \sum_{i=1}^n w_i z_i^2 \quad (9)$$

通过多次实验发现,该方法得到的尺度估计计算简单、精度高、稳定性好,能够实时进行尺度估计。

## 4 基于 EKF 的视觉惯导融合定位

V-SLAM 与惯导解算均可获得无人机的六自由度位姿信息。二者的区别为: V-SLAM 输出频率慢,姿态误差较大;惯导解算输出频率快,随着时间的推移位置累积误差较大。因此,考虑基于 EKF 实现两者的信息融合,通过惯导解算过程进行状态预测,用 V-SLAM 的输出作为观测值,对惯导解算得到的位姿进行修正。本章将详细介绍基于 EKF 的视觉惯导信息融合算法。

### 4.1 惯导解算过程

惯导解算是指通过加速度计测得的载体加速度和陀螺测得的载体相对于惯性坐标系的角速度来进行载体的位姿估计。使用 IMU 进行惯导解算时必须考虑到噪声的影响,其实际测量模型如式(10):

$$\begin{cases} \dot{\mathbf{a}}_m = \mathbf{a} - \mathbf{g} + \mathbf{n}_a + \mathbf{b}_a \\ \dot{\boldsymbol{\omega}}_m = \boldsymbol{\omega} + \mathbf{n}_g + \mathbf{b}_g \end{cases} \quad (10)$$

其中:  $\mathbf{n}_a, \mathbf{n}_g$  为零均值的高斯白噪声,服从正态分布;  $\mathbf{b}_a, \mathbf{b}_g$  分

别代表加速度计和陀螺的零漂,可通过三轴转台进行标定,零漂可以看作由高斯白噪声导致的一个随机游走过程。

系统的状态量包括 IMU 的位置  $p_i^w$ 、姿态  $q_i^w$ 、速度  $v_i^w$  以及加速度计和陀螺仪的零漂  $b_a$ 、 $b_g$ , 因此状态量可表示为 16 维的列向量如式(11):

$$x = [p_i^w \ q_i^w \ v_i^w \ b_a \ b_g]^T \quad (11)$$

根据定义的状态量可以得到 IMU 的运动方程如式(12):

$$\begin{cases} \dot{p}_i^w = v_i^w \\ \dot{q}_i^w = \frac{1}{2}\Omega(\omega_m - b_g - n_g) \otimes q_i^w \\ \dot{v}_i^w = C_i^w(a_m - b_a - n_a) - g \\ \dot{b}_a = n_{b_a} \\ \dot{b}_g = n_{b_g} \end{cases} \quad (12)$$

$C_i^w$  代表由当前 IMU 体坐标系到世界坐标系的旋转矩阵, 根据前三个运动方程可以对惯导进行解算, 其实质就是求解三个微分方程, 首先用式(13)对姿态进行解算:

$$q_i^w(t_k) = \left[ \cos \frac{\sigma}{2}; \frac{\sin(\sigma/2)}{\sigma} \sigma \right] \otimes q_i^w(t_{k-1}) \quad (13)$$

其中:  $q_i^w(t_k)$  代表当前时刻 IMU 相对于世界坐标系的姿态,  $q_i^w(t_{k-1})$  代表上一时刻的姿态;  $\sigma$  代表角度旋转矢量, 由陀螺仪输出的角速度积分得到,  $\sigma$  为  $\sigma$  的模, 此处的旋转矢量会由于圆锥运动引进一个常值漂移项, 采用圆锥算法可以在一定程度进行消除, 具体参考文献[17]。

速度解算过程如式(14), 其中,  $v_i^w(t_k)$ 、 $v_i^w(t_{k-1})$  分别为当前时刻和上一时刻 IMU 的速度;  $\Delta v_{gk}^w$  为加速度积分增量;  $\Delta v_g^w$  为重力加速度积分增量。

$$v_i^w(t_k) = v_i^w(t_{k-1}) + \Delta v_{gk}^w + \Delta v_g^w \quad (14)$$

位置解算直接由速度积分得到, 如式(15):

$$p_i^w(t_k) = p_i^w(t_{k-1}) + T[v_i^w(t_k) + v_i^w(t_{k-1})]/2 \quad (15)$$

#### 4.2 EKF 预测过程

惯导系统的误差状态方程是进行卡尔曼滤波的基础。根据前面得到的惯导系统方程可以推导出误差状态方程如式(16), 定义误差状态为  $\tilde{x} = x - \hat{x}$ , 其中,  $\tilde{x}$  为误差状态向量,  $\hat{x}$  为状态向量期望值。

$$\begin{aligned} \hat{x} &= [\hat{p}_i^w \ \hat{q}_i^w \ \hat{v}_i^w \ \hat{b}_a \ \hat{b}_g]^T \\ \tilde{x} &= [\delta p_i^w \ \delta q_i^w \ \delta v_i^w \ \delta b_a \ \delta b_g]^T \\ \begin{cases} \dot{\hat{p}}_i^w = \hat{v}_i^w \\ \dot{\hat{q}}_i^w = \frac{1}{2}\Omega(\omega_m - \hat{b}_g) \otimes \hat{q}_i^w \\ \dot{\hat{v}}_i^w = \hat{C}_i^w(a_m - \hat{b}_a) - g \\ \dot{\hat{b}}_a = 0 \\ \dot{\hat{b}}_g = 0 \end{cases} \end{aligned} \quad (16)$$

误差状态向量中姿态不再使用四元数进行表示, 而是采用欧拉角表示, 因此要得到四元数误差与欧拉角误差之间的关系。在旋转角度很小时, 本文采用小角度近似的方法来表述四元数误差  $\delta q$  和角度误差  $\delta \theta$  之间的关系如式(17):

$$\delta q = \begin{bmatrix} \delta q_1 \\ \delta q_2 \\ \delta q_3 \\ \delta q_4 \end{bmatrix} = \begin{bmatrix} k \sin(\delta \theta / 2) \\ \cos(\delta \theta / 2) \end{bmatrix} \approx \begin{bmatrix} \delta \theta / 2 \\ 1 \end{bmatrix} \quad (17)$$

采用同样的方法由式(18)~(19)可以得到方向余弦矩阵与角度误差之间的关系:

$$C(\delta q) \approx I_3 - [\delta \theta \times] \quad (18)$$

$$C_i^w = C(\delta q) \hat{C}_i^w \quad (19)$$

本文中  $[\delta \theta \times]$  表示方向误差向量  $\delta \theta$  的反对称阵, 其表达形式定义如式(20):

$$[\delta \theta \times] = \begin{bmatrix} 0 & -\delta \theta_x & \delta \theta_y \\ \delta \theta_x & 0 & -\delta \theta_z \\ -\delta \theta_y & \delta \theta_z & 0 \end{bmatrix} \quad (20)$$

通过误差状态的定义  $\tilde{x} = x - \hat{x}$ , 惯导系统方程经过一系列的线性化处理后可以得到误差状态方程如式(21):

$$\begin{cases} \delta \dot{p}_i^w = \delta v_i^w \\ \delta \dot{q}_i^w = -[\omega \times] \delta q_i^w - \delta b_g - n_g \\ \delta \dot{v}_i^w = -C_i^w[a \times] \delta q_i^w - C_i^w \delta b_a - C_i^w n_a \\ \delta \dot{b}_a = n_{b_a} \\ \delta \dot{b}_g = n_{b_g} \end{cases} \quad (21)$$

误差状态方程表示形式如式(22):

$$\dot{\tilde{x}} = F\tilde{x} + G\tilde{n} \quad (22)$$

噪声状态  $\tilde{n}$  表示为  $[n_a \ n_{b_a} \ n_g \ n_{b_g}]$ , 从而可以得到误差状态方程  $F$ 、 $G$  矩阵的表达式如式(23)~(24):

$$F = \begin{bmatrix} 0_3 & 0_3 & I_3 & 0_3 & 0_3 \\ 0_3 & -[\hat{\omega} \times] & 0_3 & 0_3 & -I_3 \\ 0_3 & -\hat{C}_i^w[\hat{a} \times] & 0_3 & -\hat{C}_i^w & 0_3 \\ 0_3 & 0_3 & 0_3 & 0_3 & 0_3 \\ 0_3 & 0_3 & 0_3 & 0_3 & 0_3 \end{bmatrix} \quad (23)$$

$$G = \begin{bmatrix} 0_3 & 0_3 & 0_3 & 0_3 \\ 0_3 & 0_3 & -I_3 & 0_3 \\ -\hat{C}_i^w & 0_3 & 0_3 & 0_3 \\ 0_3 & I_3 & 0_3 & 0_3 \\ 0_3 & 0_3 & 0_3 & I_3 \end{bmatrix} \quad (24)$$

为了求解状态转移矩阵, 在实现过程中还需要对系统矩阵进行离散化处理, 本文采用泰勒展开的方式如式(25), 并只取到二阶项, 具体过程参考文献[18]:

$$\varphi(t + \Delta t, t) = \exp(F\Delta t) = I_{15} + F\Delta t + \frac{1}{2}F^2\Delta t^2 + \dots \quad (25)$$

根据系统噪声矩阵  $G$  和状态转移矩阵  $\varphi$ , 可以求得系统噪声协方差矩阵  $Q_d$  的离散化表示如式(26):

$$Q_d = \int_t^{t+\tau} \varphi(\tau + \Delta t, \tau) G Q_c G^T \varphi^T(\tau + \Delta t, \tau) d\tau \quad (26)$$

其中:  $Q_c$  为系统噪声协方差矩阵的连续表达形式, 表示为  $Q_c = \text{diag}(\delta n_a^2 \ \delta n_{b_a}^2 \ \delta n_g^2 \ \delta n_{b_g}^2)$ 。

通过求得的系统状态转移矩阵和误差协方差矩阵可以进行卡尔曼滤波的预测过程如式(27):

$$\begin{cases} \hat{X}_{k+1|k} = \varphi \hat{X}_{k|k} \\ \hat{P}_{k+1|k} = \varphi \hat{P}_{k|k} \varphi^T + Q_d \end{cases} \quad (27)$$

其中:  $\hat{X}_{k+1|k}$  表示误差状态的预测值,  $\hat{P}_{k+1|k}$  表示系统状态协方差矩阵的预测值,  $\hat{X}_{k|k}$ 、 $\hat{P}_{k|k}$  则分别代表上一时刻的最优估计值。

#### 4.3 EKF 更新过程

采用 IMU 进行位姿估计虽然输出频率高, 但随着时间的增长累计误差严重, 因此采用低频输出的单目视觉位姿估计结果来校正惯导解算的输出, 对位姿估计进行更新, 视觉位姿估计值作为滤波器的观测值  $z$ , 即:

$$z = [p_c^w \ q_c^w]$$

单目相机和 IMU 之间的空间位置关系是固定的, 通过标



定可以得到它们之间的旋转平移矩阵  $q_i^c, p_i^c$ , 这样可以得到系统观测方程如式(28):

$$\begin{cases} z_p = p_i^w = p_i^w + C_i^w p_i^c \\ z_q = q_i^w = q_i^w \otimes q_i^c \end{cases} \tag{28}$$

由误差状态的定义,  $\tilde{z} = z - \hat{z}$  可以得到观测方程的误差状态表达式(29):

$$\begin{cases} \tilde{z}_p = \delta p_i^w - \hat{C}_i^w p_i^c \times \delta \theta_i^w \\ \tilde{z}_q = \delta q_i^w \approx [\delta \theta_i^w / 2 \quad 1] \end{cases} \tag{29}$$

由  $\tilde{z} = H\tilde{x}$ , 可以得到观测方程  $H$  的表示如式(30):

$$H = \begin{bmatrix} I_3 & -\hat{C}_i^w [p_i^c \times] & 0_3 & 0_3 & 0_3 \\ 0_3 & I_3 & 0_3 & 0_3 & 0_3 \end{bmatrix} \tag{30}$$

观测噪声协方差矩阵  $R$  可根据观测噪声来给定, 一般作为滤波的可调参数。由观测矩阵和观测噪声协方差矩阵就可以得到卡尔曼滤波的更新过程, 步骤如式(31)~(33):

首先计算系统的增益矩阵  $K$ :

$$K = P_{k+1|k} H^T (H P_{k+1|k} H^T + R)^{-1} \tag{31}$$

然后对系统状态向量进行更新:

$$X_{k+1|k+1} = X_{k+1|k} + K\tilde{z} \tag{32}$$

最后对系统状态协方差矩阵进行更新:

$$P_{k+1|k+1} = (I_{15} - KH)P_{k+1|k}(I_{15} - KH)^T + KRK^T \tag{33}$$

需要注意的是, 姿态在误差状态中使用欧拉角  $\delta\theta_i^w$  表示, 而实际状态向量中姿态采用四元数表示, 因此在状态更新时不能直接使用叠加的方式进行更新, 需要使用 4.2 节中的小角度近似方法进行转换。姿态的更新过程如下: 首先求得反对称阵  $[\delta\theta_i^w \times]$ ; 然后通过式(18)~(19)求得方向余弦矩阵; 最后将方向余弦矩阵形式转化为四元数表示。

5 实验结果分析

本文研究背景为大场景下的小型多旋翼无人机的自定位, 根据需求, 设计了如图 3 所示的多传感器系统, 其中超声波传感器位于相机正下方, 三个传感器固联, 实验在室内及室外开阔的场景下进行。

5.1 尺度估计结果分析

尺度估计实验对无人机的运动有一定的要求, 由于超声波传感器垂直向下安装, 无人机在初始化的过程中需要垂直运动从而获得高度差数据。为了保证尺度估计的精度, 需要在初始化成功后的一段时间内继续进行尺度估计优化, 通过多个数据样本得到更加稳定的尺度估计结果。本文分别进行了短距离(包含闭环检测)条件下和长距离(没有闭环检测)条件下的 ORB\_SLAM 尺度估计实验, 实验中无人机位姿的真值分别由 VICON 运动捕捉系统(小场景)和高精度组合导航系统(大场景)提供。

实验结果如图 4~5 所示, 图中圆点为起始位置, 三角形为终止位置, 星形为闭环检测成功的位置。图中三条轨迹分别代表无人机位置的真实值、不进行尺度估计时单纯 ORB\_SLAM 的位置估计结果以及进行尺度估计后得到的位置轨迹, 详见图例。图 4 为短距离(室内)尺度估计结果, 图 5 为长距离(室外)尺度估计结果。

表 1 尺度估计值

实验	距离/m	尺度估计值
室内实验	约 15	1.625
室外实验	约 150	18.265

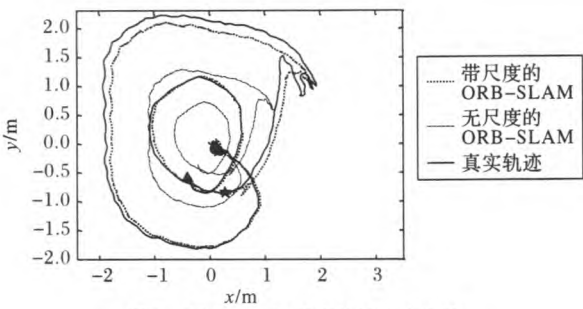


图 4 ORB\_SLAM 尺度估计室内实验

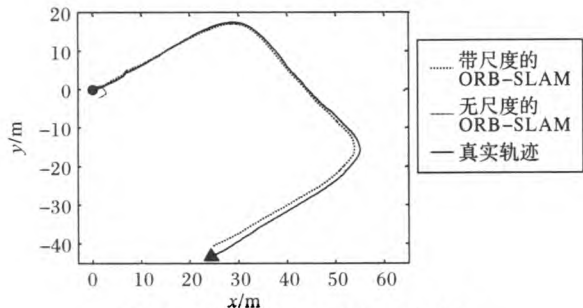


图 5 ORB\_SLAM 尺度估计室外实验

从结果可以看出, 尺度估计效果很明显, 刚开始一段时间 ORB\_SLAM 尺度估计后得到的轨迹与真实轨迹基本重合, 尺度估计准确。随着时间的推移, 尺度会有一些的漂移, 因此 ORB\_SLAM 得到的轨迹和真实轨迹产生了一定误差, 这时如果无人机产生高度变化使得尺度估计重启则可以在一定程度上修正尺度的漂移。产生这种现象的另外一个原因是, 随着时间的推移, ORB\_SLAM 本身也会产生累积误差, 在闭环检测成功之后, 尺度估计的精度又有明显的提高, 因为通过闭环检测对前面的累积漂移误差进行了修正。

5.2 位姿估计结果分析

前面一节已经介绍了单目 ORB\_SLAM 算法进行尺度估计后得到的位姿估计结果, 接下来开展纯惯导解算实验从而获得惯导输出的位姿信息, 最后进行融合实验得到最终的位姿估计结果。

5.2.1 惯导解算结果分析

本文采用的 IMU 型号为 MTI300, 其输出频率为 200 Hz。首先进行静止实验, 获得 IMU 的初始姿态及零漂; 然后进行 IMU 的动态实验, 将 IMU 在  $xy$  平面绕一个半径大约为 2 m 的圆形轨迹。实验结果如图 6 所示, 从结果可以看出, IMU 姿态解算结果准确度较高, 基本上与真值一致, 而位置解算结果漂移很大, 尤其是当 IMU 的加速度运动不明显时会出现非常大的误差漂移。一般小型多旋翼无人机上使用的是低精度 IMU, 位置解算误差漂移非常大, 因此需要将 IMU 和其他传感器进行信息融合从而对其位置进行修正, 得到更加准确的位姿估计结果。

5.2.2 EKF 融合结果分析

为了提高位姿估计的精度, 本文采用 ORB\_SLAM 的位姿估计结果作为观测值对惯导解算值进行修正。融合过程中需要注意数据的对齐, 由于本系统采用的相机无法使用硬件触发, 本文使用软件的方式来进行数据对齐, 通过 ROS 发布节点的时间戳来对齐视觉与惯导数据。

本文开展了室外大场景下视觉和 IMU 融合的位姿估计实验, 实验轨迹如图 5 所示, 实验条件及误差分析见表 2, 实验结果如图 7~8 所示。

表 2 融合实验误差分析值(距离约 150 m)

算法种类	误差种类	位置误差/m		姿态误差/(°)		
		x	y	横滚	俯仰	偏航
本文算法	Max	0.941	1.871	5.157	3.839	5.500
	Std	0.312	0.401	0.014	0.021	0.025
	RMSE	0.495	0.995	1.915	1.878	2.235
ORB-SLAM	RMSE	0.473	1.167	2.943	3.799	3.384
惯导	RMSE	947.4	411.6	2.454	1.823	3.349

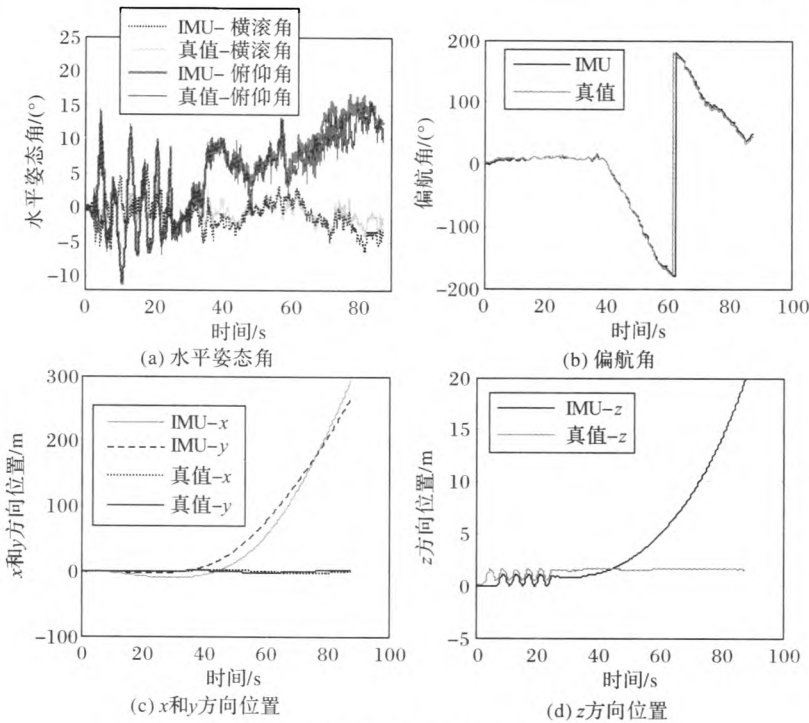


图 6 IMU 动态实验解算结果

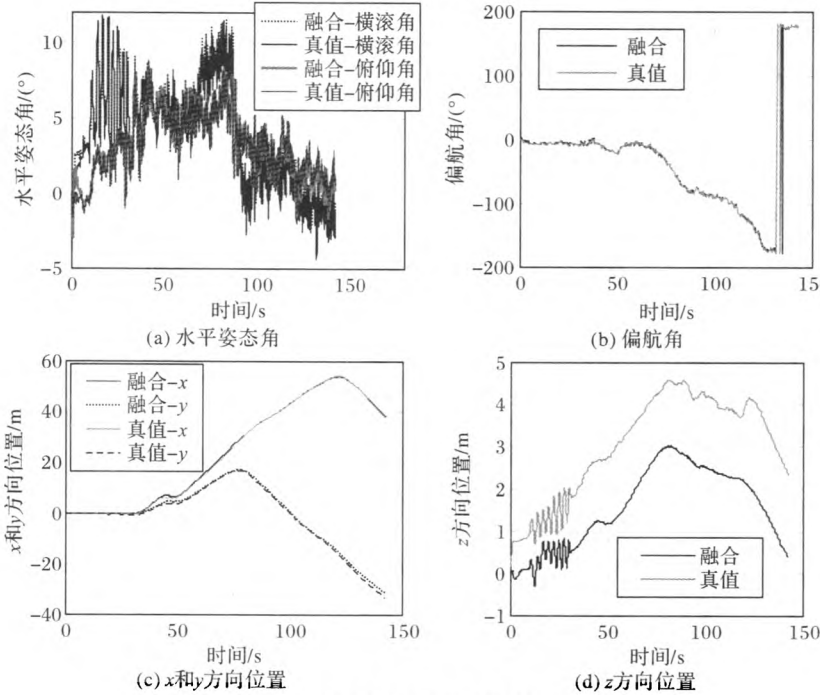


图 7 EKF 融合位姿真值对比图

实验中的真值由高精度组合导航系统提供,由于真值的 z 轴方向收敛较慢,本身误差就较大,因此实验中误差分析未考虑 z 轴。由图 7 的实验结果分析可知,与 IMU 动态实验结果(图 6)相比,融合后的位置曲线与真值曲线基本重合,精度

得到提高,位置不再随时间漂移;由图 8 的实验结果可以看出,与 ORB\_SLAM 的位姿估计结果相比,融合后位置误差有一定的降低,姿态误差明显降低。

表 2 为误差对比结果,其中 Max 为最大误差,Std 为误差的标准差,RMSE 为均方根误差。在 150 m 范围实验中,融合后位置估计均方根误差低于 0.995 m,水平姿态角估计均方根误差低于 1.915°,航向角估计均方根误差为 2.235°,相比于纯惯导解算与单目 ORB\_SLAM 算法单独的位姿估计结果,

融合后精度均有一定的提高。实验结果表明通过视觉与惯导融合,解决了卫星信号缺失时无人机定位困难的问题,一方面,融合后位姿估计精度得到了提高;另一方面,解决了视觉 SLAM 输出频率较低的问题,融合后的位姿输出频率为 200 Hz,满足小型无人机高动态特性位姿估计的需求。

6 结语

本文以小型无人机在无卫星条件下的位姿估计为背景,设计了一个多传感器组合定位系统,其中包括单目视觉传感器、IMU 以及超声波传感器,实现了单目视觉的绝对尺度估计、视觉与惯导的融合定位。

提出了基于最小二乘法实现单目视觉尺度的最优估计;基于扩展卡尔曼滤波实现视觉惯导融合定位,使得位姿估计的输出频率、精度与鲁棒性均有较大的提高,通过理论推导与实际实验验证了本系统及相关算法的有效性。

本文设计实现的系统还可进一步完善:一方面,扩展卡尔曼滤波选取状态向量时,并未把尺度作为一个滤波状态,这样就无法解决长时间工作情况下尺度漂移的问题;另一方面,本系统的状态观测值由单目视觉提供,视觉在长时间工作的情况下会产生一定的累积误差,从而导致融合结果精度下降,由于无人机的工作环境为室外,一般情况下不会出现卫星信号一直缺失的情况,因此可以考虑在有卫星信号的情况下采用卫星导航数据来校正位姿估计结果,从而提高小型无人机长时间位姿估计的精度。

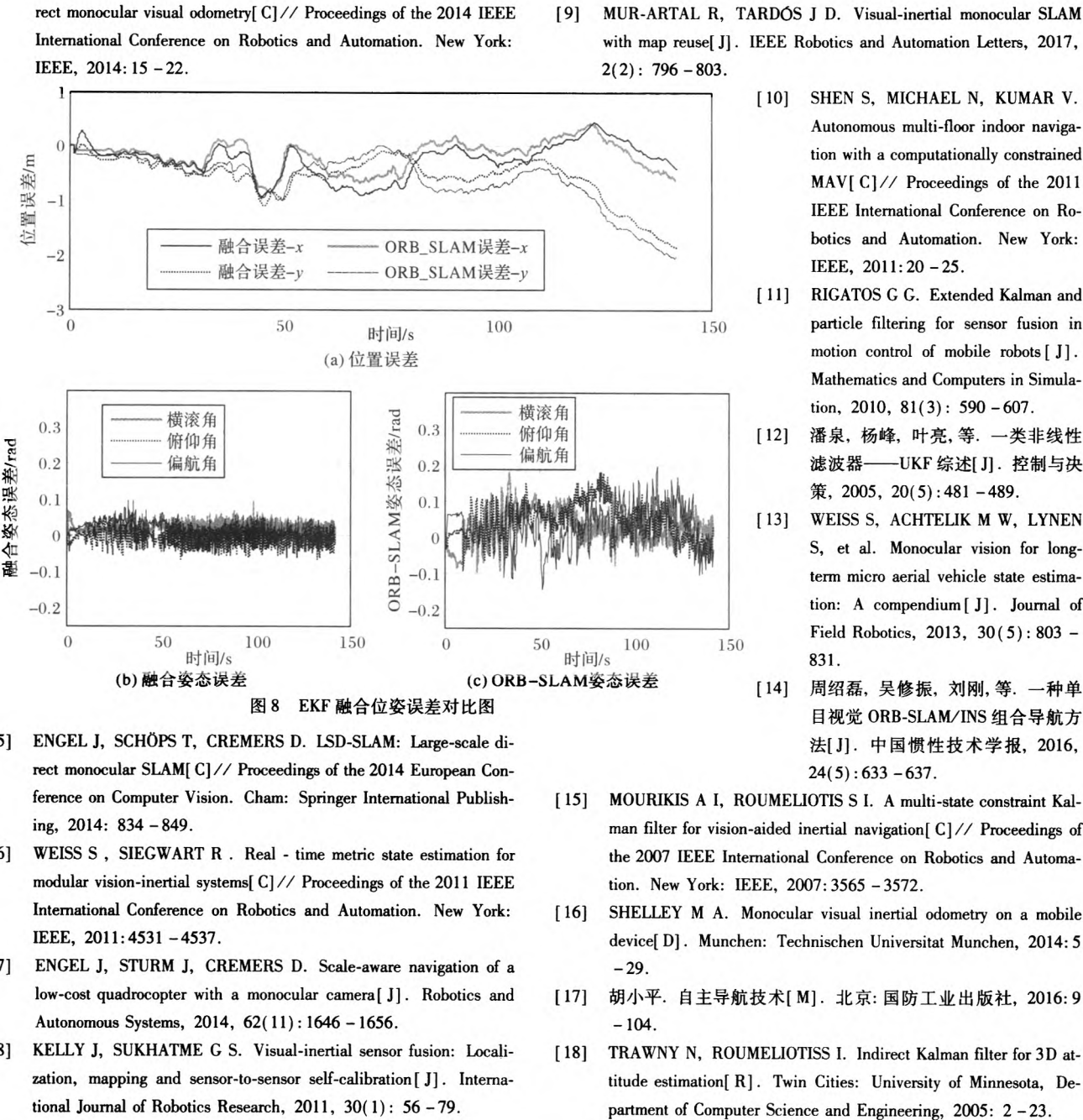
参考文献:

[1] DURRANT-WHYTE H, BAILEY T. Simultaneous localisation and mapping: part I[J]. IEEE Robotics & Automation Magazine, 2006, 13 (2): 99-110.

[2] KLEIN G, MURRAY D. Parallel tracking and mapping for small AR workspaces[C]// Proceedings of the 2007 IEEE and ACM International Symposium on Mixed and Augmented Reality. Washington, DC: IEEE Computer Society, 2007: 1-10.

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

[4] FORSTER C, PIZZOLI M, SCARAMUZZA D. SVO: Fast semi-di-



(上接第 94 页)

[4] 张大春, 王国民. 正射影像图技术在铁路勘测中的应用与发展[ J]. 铁道工程学报, 2007, 24(增刊): 20 – 26.

[5] 李芳芳. 一种新型轨道检测技术的基础理论探讨[ D]. 成都: 西南交通大学, 2010: 26 – 30.

[6] 牛发发, 陈莉, 张永新, 等. 基于鲁棒主成分分析的 Canny 边缘检测算法[ J]. 计算机应用, 2014, 34(6): 1727 – 1730.

[7] CANNY J. A computational approach to edge detection[ J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, 8(6): 679 – 698.

[8] 王小俊, 刘旭敏, 关永. 基于改进 Canny 算子的图像边缘检测算法[ J]. 计算机工程, 2012, 38(14): 196 – 198.

[9] 李牧, 闫继红, 李戈, 等. 自适应 Canny 算子边缘检测技术[ J]. 哈尔滨工程大学学报, 2007, 28(9): 1002 – 1007.

[10] 张帆, 彭中伟, 蒙水金. 基于自适应阈值的改进 Canny 边缘检测方法[ J]. 计算机应用, 2012, 32(8): 2296 – 2298.

[11] 刘新亮, 李星野. 基于 Radon 变换的多尺度虹膜特征提取算法[ J]. 光电子·激光, 2008, 19(4): 532 – 536.

[12] 吴丽丽, 余春燕. 基于 Sobel 算子和 Radon 变换的车牌倾斜校正方法[ J]. 计算机应用, 2013, 33(S1): 220 – 222.

[13] CARLSON B D, EVANS E D, WILSON S L. Search radar detection and track with the Hough transform Part I: system concept[ J]. IEEE Transactions on Aerospace and Electronic Systems, 1994, 30(1): 102 – 108.

[14] 侯和平, 郭凯铭, 刘凯, 等. 基于 Radon 变换与灰度投影积分值方法的矩形检测[ J]. 西安理工大学学报, 2014, 30(2): 133 – 138.

[15] 代勤, 王延杰, 韩广良. 基于改进 Hough 变换和透视变换的透视图像矫正[ J]. 液晶与显示, 2012, 27(4): 552 – 556.

[16] 何援军. 透视和透视投影变换——论图形变换和投影的若干问题之三[ J]. 计算机辅助设计与图形学学报, 2005, 17(4): 734 – 739.