Robotics and Autonomous Systems 107 (2018) 262–279

Contents lists available at ScienceDirect

Robotics and Autonomous Systems

journal homepage: www.elsevier.com/locate/robot

Visual–inertial estimation of velocity for multicopters based on vision motion constraint[☆]

Heng Deng[a,*], Usman Arif[a], Qiang Fu[a,b], Zhiyu Xi[a], Quan Quan[a], Kai-Yuan Cai[a]

[a] School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China
[b] School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China

HIGHLIGHTS

- A visual–inertial method to estimate velocity for multicopters.
- A Kalman filter which integrates monocular camera-based estimations with IMU data.
- No need for any prior knowledge of the environment or any external sensors.
- An efficient approach proposed based on MS algorithm to detect outliers.
- Observability analysis performed to verify the feasibility of the proposed method.

ARTICLE INFO

Article history:
Received 6 November 2017
Received in revised form 6 May 2018
Accepted 23 June 2018
Available online 4 July 2018

Keywords:
Visual–inertial
Velocity estimation
Multicopters
Observability analysis
Mean shift
Kalman filter

ABSTRACT

Velocity estimation is essential for multicopters to guarantee flight stability and maneuverability. For such a purpose, this paper proposes a new method for multicopter velocity estimation based on visual and inertial information in GPS-denied or confined environments. In this method, no map, artificial landmark of the environment is required, and only the off-the-shelf onboard sensors in a multicopter including a low-cost Inertial Measurement Unit (IMU), a downward-looking monocular camera and an ultrasonic range finder facing downwards are exploited to constitute the vision motion constraint. This constraint connects metric velocity with the point correspondences between successive images in which an efficient approach based on Mean Shift (MS) algorithm is developed to detect outliers and select optimal matching points. Then, it is theoretically verified that the estimation system is observable based on observability analysis. Furthermore, combined with the vision motion constraint and a multicopter dynamic model, the metric velocity is estimated using a standard Linear Kalman Filter (LKF). Finally, the proposed method is tested with a collection of synthetic data from simulation as well as flight experiments using real data from DJI Matrice 100 and Guidance. The simulation and experimental results indicate that the proposed method can accurately estimate the velocity of the multicopter in GPS-denied or confined environments.

1. Introduction

In recent years, multicopters, a kind of unmanned aerial vehicles (UAVs), have attracted increasing attention in the field of both academic research and industrial applications. Since multicopters are highly maneuverable and enable relatively safe and low-cost experimentation in navigation, mapping, and control strategies [1], they are widely used in a range of mission scenarios, such as search and rescue [2], load transportation [3], aerial manipulation [4], surveillance [5] and agricultural application [6]. However, there exist some scientific and technological challenges. Environment sensing and autonomous navigation, which is crucial to guarantee stability and safety for multicopters, remains an open and challenging issue. As stated in [7], accurate state estimation is always a fundamental necessity to implement fully autonomous manipulation in complex environments. Besides, accurate velocity estimation is required for multicopters since velocity feedback will increase damping to improve the stability and in return make multicopters more tractable.

As for velocity estimation of UAVs, a massive amount of research at the beginning has focused on indoor research using external motion capture systems such as Vicon [8,9] and outdoor applications using GPS signals [10–12]. However, these approaches rely mainly on external positioning systems, restricting UAVs to be used in a wide range of applications. While Vicon exhibits

---

[☆] This work is supported by the National Key Project of Research and Development Plan under Grant 2016YFC1402500.

[*] Corresponding author.
E-mail address: dengheng@buaa.edu.cn (H. Deng).

https://doi.org/10.1016/j.robot.2018.06.010
0921-8890/© 2018 Elsevier B.V. All rights reserved.

excellency with multiple high-resolution external cameras to track the pose of multicopters with submillimeter accuracy, it demands a complicated installation and calibration process, and it is infeasible for the vast outdoor environment. Besides, the operation range is limited within the field of view of the vision system. GPS signals may not be available or sufficiently reliable in some limited or confined areas without high-quality satellite signals, such as forests and buildings. Therefore, onboard sensing, especially the onboard vision has been a promising sensor modality for small autonomous multicopters since it does not require energy to interrogate the environment, and it can provide rich information and span wide field of view [13]. In general, visual–inertial fusion using onboard vision is a widely-used approach to provide more accurate velocity estimation for multicopters in limited environments [14,15]. Some attempts are using artificial landmarks or user-specified points of known position and appearance [16], but they can be reasonably accurate only if the target is detected successfully and quickly. Thus, they are limited to some scenarios with individual targets and relies pretty much on known features. As the work [17] shows, the automatic landing of a Micro Aerial Vehicle (MAV) on a moving vehicle was implemented with successful flight tests at up to 50 km/h. They fused together measurements from the MAV's onboard integrated navigation system, from multiple cameras tracking a visual fiducial marker and from the IMU and GPS data on the ground vehicle.

While the methods above require some knowledge or modifications to the environment, it is a better choice to develop a more general approach for the unknown scene. Simultaneous Localization And Mapping (SLAM) is a conventional technique for multicopter navigation in unknown environments. Weiss et al. in [18] enabled the Micro Aerial Vehicle (MAV) to determine its position autonomously and consequently stabilize itself. The work has been viewed as the first implementation for a MAV to navigate autonomously through an unknown environment, in which a monocular camera is used as the only exteroceptive sensor independent of any external aid like GPS or artificial landmarks. Following these results, Weiss has made more subsequent improvements to deal with practical issues such as scale drift, time delays, online estimation in [19,20]. Shen et al. in [21] proposed a method to estimate the velocity through a SLAM algorithm and an unscented Kalman filter which fused information from stereo cameras and inertial measurements. However, the data association and loop closure in SLAM required hundreds of points to be stored demanding more computational power and such process was quite complicated.

Compared to SLAM, optical flow is an alternative approach. Herissé et al. in [22] proposed a system to implement hovering and landing on a moving platform based on optical flow. However, it only provided the scaled linear velocity. Parrot AR.Drone [23] was the first commercial product to use an onboard downward-looking monocular camera and an ultrasonic range finder to measure metric velocity to stabilize itself based on optical flow , but the hardware design and software implementation were closed source. Similarly, PX4FLOW [24] also used a monocular camera to compute velocity with an ultrasonic sensor for scaling. However, it could only deal with a small resolution of 64 × 64 pixels limiting the measurement range and accuracy. Besides, the velocity measurements are only available when the multicopter flies over at most five meters above the ground and the ground is restricted to be relatively flat as discussed in the literature [25]. Grabe et al. in [26] proposed and experimentally verified an onboard velocity estimation and closed-loop control using the observed optical flow based on the continuous homography constraint. Homography, which contains rich information between two successive images, has been successfully applied to vision-based navigation missions. Zhao et al. in [27] designed a homography-based vision-aided inertial navigation system to provide drift-free velocity and attitude estimation for UAV stabilization. The work in [28] claimed that the attitude, velocity, and IMU measurement biases are observable during a time interval based on the assumption that multiple salient and repeatable feature points can be extracted and matched between two successive images according to their similarity. However, in practice, the matching points are usually contaminated by outliers. Therefore, it is necessary to detect and reject outliers to guarantee an accurate and robust estimation.

In contrast to existing visual estimation methods, we do not require any prior knowledge of the scene, nor do we need any external sensors like GPS or motion capture systems. It is assumed that the states of the multicopter flying in GPS-denied or confined spaces only come from the onboard sensors including an IMU, an ultrasonic range finder and a downward-looking monocular camera without any other exteroceptive sensor. For the monocular vision system, there is not any map or artificial landmark in the environment with the scene supposed as a flat plane. For the inertial system, unknown constant biases corrupt the measurements of the low-cost IMU so that they must be estimated and then compensated online.

To this end, this paper proposes a new visual–inertial estimation of metric velocity for multicopters based on vision motion constraint. The constraint is related to the corresponding features between two successive images and contains the velocity information directly. The mismatching of the features may indeed cause an error in the estimation. Therefore, an efficient approach based on Mean Shift (MS) algorithm is developed to detect outliers and select optimal matching points. It is proved that only one matching point is required to obtain the estimate based on observability analysis. More specially, combined with the vision motion constraint, the metric velocity is estimated using a standard Linear Kalman Filter (LKF) with a unique multicopter dynamic model. In contrast to existing studies, the major contributions of this paper are: (1) no need for any prior knowledge of the environment or any external sensors with only one feature correspondence required, (2) an efficient approach proposed based on MS algorithm to detect outliers and select optimal matching points between two successive images, and (3) an observability analysis performed to verify the feasibility of the proposed visual–inertial estimation.

The remainder of the paper is organized as follows. The problem formulation is given in Section 2. Section 3 presents the design of the proposed visual–inertial estimation system. In Section 4, vision motion constraint and optimal matching points selection based on MS algorithm are described and proved. In Section 5, the proposed method is theoretically verified through observability analysis, and then the procedure of discrete LKF is given. Section 6 shows the simulation and experimental results to validate the proposed estimation method and Section 7 gives the conclusions and future research plan.

2. Problem formulation

2.1. Preliminaries

2.1.1. Notations and definitions

Note that the notations and definitions in this paper are consistent with that in [29]. Let $\mathbb{R}^{m\times n}$ denote a real matrix with m rows and n columns while $\mathbb{R}^{n}$ an n-dimensional real column vector. Define $\mathbf{A}^{\mathrm{T}}$ and $\mathbf{A}^{-1}$ as transpose and inverse of the corresponding matrix $\mathbf{A}$, respectively. Let $\mathbf{I}_n$ denote an n-dimensional identity matrix and $\mathbf{0}_{m\times n}$ is a null matrix of dimension $m \times n$. The symbol $\mathbf{e}_3$ denotes a unit vector $[0 \quad 0 \quad 1]^{\mathrm{T}}$. For an arbitrary vector $\mathbf{a} = [a_1 \quad a_2 \quad a_3]^{\mathrm{T}} \in \mathbb{R}^3$, define the corresponding skew symmetric matrix

$$[\mathbf{a}]_\times = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \in \mathbb{R}^{3\times 3}$$