Robotics and Autonomous Systems 62 (2014) 1646–1656

Contents lists available at ScienceDirect

Robotics and Autonomous Systems

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

# Scale-aware navigation of a low-cost quadrocopter with a monocular camera

Jakob Engel *, Jürgen Sturm, Daniel Cremers

*Department of Computer Science, Technical University of Munich, Germany*

## HIGHLIGHTS

- Complete visual navigation system for a low-cost quadrocopter.
- Uses an on-board, monocular camera as main sensor.
- Novel approach to estimate the scale of a monocular SLAM system.
- Extensive experimental evaluation.
- Code published as open-source in ROS.

## ARTICLE INFO

*Article history:*
Available online 1 April 2014

*Keywords:*
Quadrocopter
Visual navigation
Visual SLAM
Monocular SLAM
Scale estimation
AR.Drone

## ABSTRACT

We present a complete solution for the visual navigation of a small-scale, low-cost quadrocopter in unknown environments. Our approach relies solely on a monocular camera as the main sensor, and therefore does not need external tracking aids such as GPS or visual markers. Costly computations are carried out on an external laptop that communicates over wireless LAN with the quadrocopter. Our approach consists of three components: a monocular SLAM system, an extended Kalman filter for data fusion, and a PID controller. In this paper, we (1) propose a simple, yet effective method to compensate for large delays in the control loop using an accurate model of the quadrocopter's flight dynamics, and (2) present a novel, closed-form method to estimate the scale of a monocular SLAM system from additional metric sensors. We extensively evaluated our system in terms of pose estimation accuracy, flight accuracy, and flight agility using an external motion capture system. Furthermore, we compared the convergence and accuracy of our scale estimation method for an ultrasound altimeter and an air pressure sensor with filtering-based approaches. The complete system is available as open-source in ROS. This software can be used directly with a low-cost, off-the-shelf Parrot AR.Drone quadrocopter, and hence serves as an ideal basis for follow-up research projects.

## 1. Introduction

Research interest in autonomous micro-aerial vehicles (MAVs) has grown rapidly in the past years. Significant progress has been made, and recent examples include aggressive flight maneuvers [1], collaborative construction tasks [2], ball throwing and catching [3] or the coordination of large fleets of quadrocopters [4]. However, all of these systems require external motion capture systems. Flying in unknown, GPS-denied environments is still an open research problem. The key challenges here are to localize the robot purely from its own sensor data and to robustly navigate it even under temporary sensor outage. This requires both a solution to the so-called simultaneous localization and mapping (SLAM) problem as well as robust state estimation and control methods. These challenges are even more expressed on low-cost hardware with inaccurate actuators, noisy sensors, significant delays, and limited on-board computation resources.

For solving the SLAM problem on MAVs, different types of sensors such as laser range scanners [5], monocular cameras [6], stereo cameras [7], and RGB-D sensors [8,9] have been explored in the past. In our point of view, monocular cameras have two major advantages over other modalities: (1) they provide rich information at a low weight, power consumption, size, and cost and (2) in contrast to depth sensors, a monocular SLAM system is not intrinsically limited in its visual range, and therefore can operate both in small, confined and large, open spaces. The drawback however is, that the scale of the environment cannot be determined from monocular

---

* Corresponding author. Tel.: +49 17664082583.
*E-mail addresses:* engelj@in.tum.de (J. Engel), sturmju@in.tum.de (J. Sturm), cremers@tum.de (D. Cremers).

http://dx.doi.org/10.1016/j.robot.2014.03.012

**Fig. 1.** A low-cost quadrocopter navigates in unstructured environments using the front camera as its main sensor. The quadrocopter is able to hold a position, fly figures with absolute scale, and recover from temporary tracking loss. Picture taken at the TUM open day.

vision alone, such that additional sensors, such as an IMU or air pressure sensor, are required.

In this paper we follow up on our previous work in [10,11] where we presented our approach to use a monocular camera to navigate a low-cost quadrocopter in an unknown, unstructured environment. Computations are performed off-board on a ground-based laptop. For our experiments we used both the low-cost Parrot AR.Drone 1.0 and 2.0 which are available for $300 and, with a weight of only 420 g and a protective hull, safe to be used in public places (as illustrated in Fig. 1).

We extend upon our previous work in several ways: First, we provide a more in-depth explanation of the proposed scale estimation method, and show that it can accurately estimate the scale of a monocular SLAM system even from very noisy sensor data such as an air pressure sensor. We provide an extensive evaluation both on synthetic and on real-world data sequences, and perform a direct comparison with current state-of-the-art filtering based methods. Second, we provide additional experimental results, in particular we evaluate the flight and pose estimation accuracy using an external motion capture system as ground truth. Third, we provide the complete system as an open-source ROS package. Our software can be used directly with a low-cost, off-the-shelf Parrot AR.Drone 1.0 or 2.0, and no modifications to hardware or on-board software are required. It therefore serves as an ideal base for follow-up projects.

A video demonstrating the practical performance of our system, its ability to accurately fly to a given position and its robustness to loss of visual tracking is available online:

http://youtu.be/eznMokFQmpc

Further, we provide an open-source ROS implementation of the complete system:

http://ros.org/wiki/tum_ardrone

## 2. Related work

Previous work on autonomous flight with quadrocopters can be categorized into two main research areas. Several results have been published where the focus is on accurate and agile quadrocopter control [1,12]. These works however rely on advanced external tracking systems, restricting their use to a lab environment. A similar approach is to distribute artificial markers in the environment, simplifying pose estimation [13]. Other approaches learn a map offline from a previously recorded, manual flight and thereby enable a quadrocopter to reproduce the same trajectory [14]. For outdoor flights where accurate GPS-based pose estimation is possible, complete solutions are available as commercial products [15].

We are interested in autonomous flight without previous knowledge about the environment or GPS signals, while using only on-board sensors. Previous work on autonomous quadrocopter flight has explored lightweight laser scanners [5], RGB-D sensors [8,9] or stereo rigs [16] mounted on a quadrocopter as primary sensors. While these sensors provide absolute scale of the environment, their drawback is a limited range and large weight, size, and power consumption when compared to a monocular set-up.

In our work we therefore focus on a monocular camera for pose estimation. Stabilizing controllers based on optical flow from a monocular camera were presented e.g. in [17,18], and similar methods are integrated in commercially available hardware [19]. These systems however make strong assumptions about the environment such as a flat, horizontal ground plane. Additionally, they are subject to drift over time, and are therefore not suited for long-term autonomous navigation.

To eliminate drift, various monocular SLAM methods have been investigated on quadrocopters, both with off-board [5] and on-board processing [6,20,21]. A particular challenge for monocular SLAM is that the scale of the map needs to be estimated from additional metric sensors such as an air pressure sensor as it cannot be recovered from vision alone. This problem has been addressed in recent publications such as [22], where the scale is added to the extended Kalman filter as an additional state variable. In contrast to this, we propose in this paper a novel approach which directly computes the unknown scale factor from a set of observations: using a statistical formulation, we derive a closed-form, consistent estimator for the scale of the visual map. Our method yields accurate and robust results both in simulation and practice. As metric sensors we evaluated both an air pressure sensor as well as an ultrasound altimeter. The proposed method can be used with any monocular SLAM algorithm and sensors providing metric position or velocity measurements.

In contrast to previous work [6], we deliberately refrain from using expensive, customized hardware: the only hardware required is the AR.Drone, which comes at a cost of merely $300—a fraction of the cost of quadrocopters used in the previous work. Released in 2010 and marketed as a high-tech toy, it has been used and discussed in several research projects [23–25].

The remainder of this article is organized as follows: in Section 3, we briefly introduce the Parrot AR.Drone and its sensors. In Section 4 we derive the proposed maximum-likelihood estimator for the scale of a monocular SLAM system. We also describe in detail all necessary preprocessing steps as well as how the variances can be estimated from the data. In Section 5 we describe our approach as a whole, in particular we describe the EKF, how we estimate the required model parameters and how we compensate for time delays. In Section 6 we present an extensive evaluation of our scale estimation method and compare it to a state-of-the-art filtering based method. We also provide extensive experimental results on the flight agility, accuracy and robustness of our system using an external motion capture system. Finally, we conclude the paper in Section 7 with a summary and outlook to future work.

## 3. Hardware platform

For the experiments we use the Parrot AR.Drone 2.0, a commercially available quadrocopter as platform. Compared to other modern MAVs such as Ascending Technology's Pelican or Hummingbird quadrocopters, its main advantages are the low price, its robustness to crashes, and the fact that it can safely be used indoor and close to people. This however comes at the price of flexibility: neither the hardware itself nor the software running on-board can easily be modified, and communication with the quadrocopter is only possible over wireless LAN. With battery and hull, the AR.Drone measures 53 cm × 52 cm and weights 420 g.

### *3.1. Sensors*

The AR.Drone 2.0 is equipped with a 3-axis gyroscope and accelerometer, an ultrasound altimeter and two cameras. Furthermore it features an air pressure sensor and a magnetic compass.