HAL
open science

A 3 Points Vision Based Approach for MAV Localization in GPS Denied Environments

Chiara Troiani, Stefano Al Zanati, Agostino Martinelli

To cite this version:

Chiara Troiani, Stefano Al Zanati, Agostino Martinelli. A 3 Points Vision Based Approach for MAV Localization in GPS Denied Environments. 6th European Conference on Mobile Robots, Sep 2013, Barcelona, Spain. hal-00909613

HAL Id: hal-00909613

https://inria.hal.science/hal-00909613v1

Submitted on 26 Nov 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A 3 Points Vision Based Approach for MAV Localization in GPS Denied Environments

Chiara Troiani, Stefano Al Zanati and Agostino Martinelli

***Abstract*— This paper introduces a new method to localize a micro aerial vehicle (MAV) in GPS denied environments and without the usage of any known pattern. The method exploits the planar ground assumption and only uses the data provided by a monocular camera and an inertial measurement unit. It is based on a closed solution which provides the vehicle pose from a single camera image, once the roll and the pitch angles are obtained by the inertial measurements. Specifically, the vehicle position and attitude can uniquely be determined by having two point features. However, the precision is significantly improved by using three point features. The closed form solution makes the method very simple in terms of computational cost and therefore very suitable for real time implementation. Additionally, because of this closed solution, the method does not need any initialization. Results of experimentation show the effectiveness of the proposed approach.**

I. Introduction

In recent years, flying robotics has received significant attention from the robotics community. The ability to fly allows easily avoiding obstacles and quickly having an excellent birds eye view. These navigation facilities make flying robots the ideal platform to solve many tasks like exploration, mapping, reconnaissance for search and rescue, environment monitoring, security surveillance, inspection etc. In the framework of flying robotics, micro aerial vehicles (MAV) have a further advantage. Due to the small size they can also be used in narrow out- and indoor environment and they represent only a limited risk for the environment and people living in it. However, for such operations today's systems navigating on GPS information only are not sufficient any more. Fully autonomous operation in cities or other dense environments requires the MAV to fly at low altitude or indoors where GPS signals are often shadowed.

A crucial problem on an airborne vehicle is the stabilization and control in attitude and position, i.e. in six degrees of freedom. The most popular approaches to solve this problem are mainly based on the fusion of the data provided by an IMU and a GPS [1], [12]. However, these approaches require a reliable GPS signal. Laser range finders have also been adopted in this framework [2], [5], [6]. However, range finders sensors have two drawbacks: they have a restricted perception area (limited distance and field of view) and they are still too heavy for MAVs. A viable option for GPS denied environments is obtained by fusing visual and inertial data. This option has become very popular for Micro Aerial Vehicle (MAV) navigation due the low cost, power consumption and weight.

C. Troiani, S Al Zanati and A. Martinelli are with INRIA Rhone Alpes, Montbonnot, France e-mail: agostino.martinelli@ieee.org, chiara.troiani@inria.fr, stefano.alzanati@inria.fr

Fig. 1. AscTec Pelican quadcopter equipped with a monocular camera.

In this paper we propose a new approach to perform MAV localization by only using the data provided by an Inertial Measurement Unit (IMU) and a monocular camera. The proposed method does not use any known pattern but only relies on three natural point features belonging to the same horizontal plane. It is based on a closed solution which provides the vehicle pose from a single camera image, once the roll and the pitch angles are obtained by the inertial measurements. The first step of the approach provides a first estimate of the roll and pitch (through the IMU data) and then the vehicle heading by only using two of the three point features and a single camera image. In particular, the heading is defined as the angle between the MAV and the segment made by the two considered point features. Then, the same procedure is repeated two additional times, i.e., by using the other two pairs of the three point features. In this way, three different heading angles are evaluated. On the other hand, these heading angles must satisfy two geometrical constraints, which are fixed by the angles of the triangle made by the three point features. These angles of the triangle are estimated in parallel by an independent Kalman Filter. The information contained in the geometrical constraints is then exploited by minimizing a suitable cost function. This minimization provides a new and very precise estimate of the roll and pitch and consequently of the yaw and the robot position.

The rest of this paper is organized as follows. The next section provides the state of the art on visual-inertial sensor fusion and its application to MAV localization. Section III provides a description of the system. The approach is described in section IV. The performance of the approach is evaluated by using both synthetic (V-A) and real data (V-B).

Finally, conclusions are provided in VI.

II. Related Works

The problem of fusing vision and inertial data has been extensively investigated in the past. A special issue of the *International Journal of Robotics Research* has recently been devoted to this important topic [8]. In [7], a tutorial introduction to the vision and inertial sensing is presented. This work provides a biological point of view and it illustrates how vision and inertial sensors have useful complementarities allowing them to cover the respective limitations and deficiencies. Specifically, as it has been derived very recently in [11], the fusion of these sensors allows us to obtain the speed and the scale factor in closed form, allowing real time applications and robustness with respect to kidnapping.

Visual UAV pose estimation in GPS-denied environments is still challenging. Many implementations rely on visual markers, such as patterns or blobs, located in known positions [16], [18], [20]. Those approaches have the drawback that can work only in structured environment. In [19] Visual-Inertial Attitude Estimation is performed using image line segments for the correction of accumulated errors in integrated gyro rates when an unmanned aerial vehicle operates in urban areas. The approach will not work in environments that do not present a strong regularity in structure.

In [21], [23] the authors developed a very robust Vision Based Navigation System for micro helicopters. Their pose estimator is based on a monocular VSLAM framework (PTAM, Parallel Tracking and Mapping [22]). This software was originally developed for augmented reality and improved with respect to robustness and computational complexity. The resulting algorithm can be used in order to make a monocular camera a real-time onboard sensor for pose estimates. This allowed the first aerial vehicle that uses onboard monocular vision as a main sensor to navigate through an unknown GPS-denied environment and independently of any external artificial aids [24], [23].

The goal of our investigation is to find a new pose estimator which minimizes the computational complexity. We focus our attention on the problem of relative localization, which makes possible the accomplishment of many important tasks (e.g. hovering, autonomous take off and landing). In this sense, we minimize the number of point features which are necessary to perform localization. While 2 point features is the minimum number which provides full observability, by adding an additional feature, the precision is significantly improved, provided that the so-called planar ground assumption is honoured. This assumption has recently been exploited on visual odometry with a bundle adjustment based method [9].

III. The Considered System

Let us consider an aerial vehicle equipped with a monocular camera and IMU sensors. The IMU consists of three orthogonal accelerometers and three orthogonal gyroscopes. We assume that the transformation among the camera frame and the IMU frame is known (we can assume that the vehicle frame coincides with the camera frame).

We assume that three reliable point-features are detected on the ground (i.e. they belong to the same horizontal plane). As we will see, two is the minimum number of features necessary to perform localization. Figure 2 displays our global frame, which is defined by only using two points, P_1 and P_2. First, we define P_1 as the origin of the frame. The z-axis coincides with the gravity axis but with opposite direction. Finally, P_2 defines the x-axis [1].

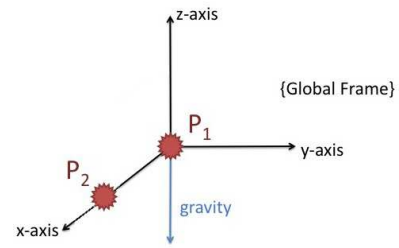


Fig. 2. Global frame. Two is the minimum number of point features which allows us to uniquely define a global reference frame. P_1 is the origin, the vertical axis is defined by the gravity and P_2 is on the x-axis

Then, by applying the method in [11], the distance between these point features can be roughly determined by only using visual and inertial data (specifically, at least three consecutive images containing these points must be acquired).

IV. The Method

The first step of the method consists in estimating the roll and the pitch angles. This is done by an EKF which estimates the gravity in the local frame by only using inertial data. Once the three components of the gravity in the local frame are estimated the roll and the pitch angles are obtained. The second step returns the yaw angle and the position of the vehicle starting from the roll and the pitch angles and a single camera image. This is obtained by running the *3p-algorithm* (sect. IV-B). This algorithm starts by running three times the *2p-algorithm*, which is described in the next subsection.

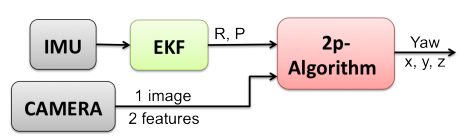


Fig. 3. The 2p-algorithm

A. 2p-Algorithm

This algorithm only uses two point features. Figure 3 shows the algorithm's inputs and outputs.

For each point-feature, the camera provides its position in the local frame up to a scale. The knowledge of the

[1]Note that the planar assumption is not necessary to define a global frame. It is sufficient that P_1 and P_2 do not lie on the same vertical axis (defined by the gravity). The x-axis can be defined by requiring that P_2 belongs to the xz-plane. In other words, P_2 has zero y coordinate.

absolute roll and pitch, allows us to express the position of the point-feature in a new local frame, where the z-axis has the same orientation as the z-axis of the global frame. Figure 4 displays all the reference frames: the global (represented by xyz), the local (represented by $x_l y_l z_l$) and the new local (represented by $x_n y_n z_n$). In the figure, the n-axis is the projection of the z_n-axis on the $y_l z_l$-plane. The pitch angle is the angle between the z_n-axis and the n-axis and the roll angle is the angle between the n-axis and the z_l-axis. Our goal is to determine the coordinates of the origin of the local frame in the global frame (x, y, z) and the orientation of the x_n-axis with respect to the x-axis (yaw angle).

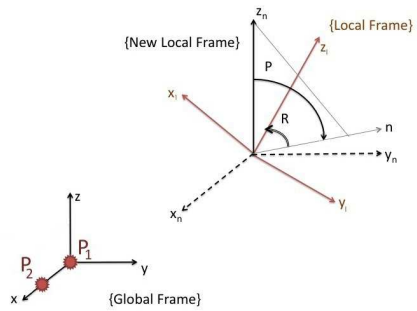


Fig. 4. The three reference frames adopted in our derivation.

Let us denote with $[x_1, y_1, z_1]^T$ and $[x_2, y_2, z_2]^T$ the coordinates of P_1 and P_2 in the new local frame. The camera provides $\mu_1 = \frac{x_1}{z_1}$, $\nu_1 = \frac{y_1}{z_1}$, $\mu_2 = \frac{x_2}{z_2}$ and $\nu_2 = \frac{y_2}{z_2}$. Additionally, the camera also provides the sign of z_1 and z_2[2].

Since the z_n-axis has the same orientation as the z-axis, $z_1 = z_2 = -z$, where z is the position of the origin of the local frame in the global frame, we obtain:

$$P_1 = -z \begin{bmatrix} \mu_1 \\ \nu_1 \\ 1 \end{bmatrix} \qquad P_2 = -z \begin{bmatrix} \mu_2 \\ \nu_2 \\ 1 \end{bmatrix} \tag{1}$$

Let us denote by D the distance between P_1 and P_2. We have:

$$z = \pm \frac{D}{\sqrt{\Delta\mu_{12}^2 + \Delta\nu_{12}^2}} \tag{2}$$

with $\Delta\mu_{12} \equiv \mu_2 - \mu_1$ and $\Delta\nu_{12} \equiv \nu_2 - \nu_1$. In other words, z can be easily obtained in terms of D. The previous equation provides z up to a sign. This ambiguity is solved considering that the camera provides the sign of z_1 and z_2. Then, we obtain $x_1 = -z\mu_1$, $y_1 = -z\nu_1$, $x_2 = -z\mu_2$ and $y_2 = -z\nu_2$. It is therefore easy to obtain $\alpha = \arctan 2(\Delta\nu_{12},\ \Delta\mu_{12})$ (Figure 5). Hence,

$$Yaw = -\alpha = -\text{atan}(\Delta\nu_{12}/\ \Delta\mu_{12}) \tag{3}$$

[2]For a camera with a field of view smaller than $180deg$ the $z-$component is always positive in the original $x_l y_l z_l$ local frame.

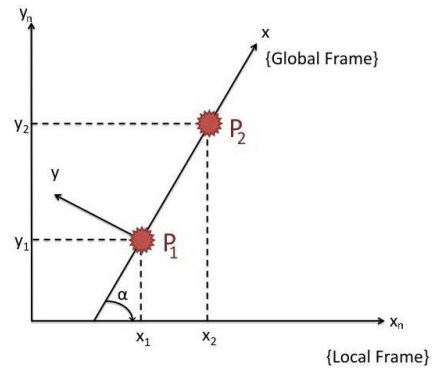


Fig. 5. The yaw angle $(-\alpha)$ is the orientation of the x_n-axis in the global frame.

Finally we obtain the coordinates of the origin of the local frame in the global frame,

$$\begin{aligned} x &= -\cos(\alpha)\ x_1 - \sin(\alpha)\ y_1 \\ y &= \ \ \sin(\alpha)\ x_1 - \cos(\alpha)\ y_1 \end{aligned} \tag{4}$$

Note that the position x, y, z is obtained in terms of the distance D. Specifically, the position scales linearly with D. As previously said, a rough knowledge of this distance is provided by using the method in [11]. We remark that a precise knowledge of this distance is not required to accomplish several important tasks (e.g. to hover on a stable position).

B. 3p-Algorithm

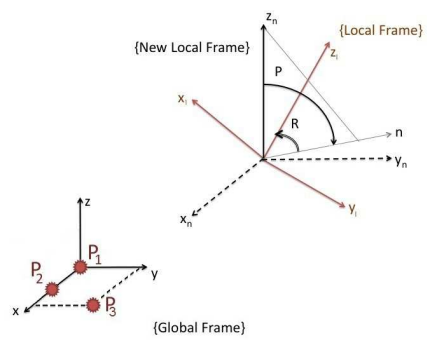


Fig. 6. The three reference frames adopted with 3 point features.

The three point features form a triangle in the (x, y)-plane. For the sake of clarity, we start our analysis supposing that we know the angles γ_1 and γ_2 shown in Figure 7 (i.e. we know the internal angles of the triangle). Then, we will show how we estimate on line these angles (IV-C).

We use the $2p-$algorithm three times, respectively with the sets of features (P_1, P_2), (P_1, P_3) and (P_2, P_3). We obtain three different values for the yaw. Yaw_{12} is the yaw given in (3). The other expressions are:

$$Yaw_{13} = -\text{atan}(\Delta\nu_{13}/\Delta\mu_{13})$$

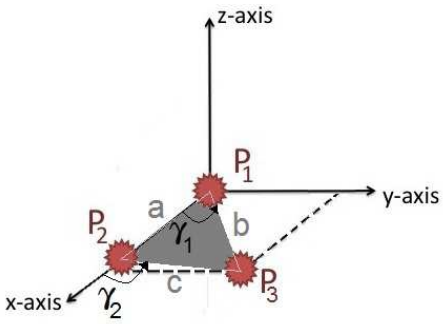


Fig. 7. The triangle made by the 3 point features.

$$Yaw_{23} = -\text{atan}(\Delta\nu_{23}/\Delta\mu_{23})$$

Looking at the Figure 7, we can see that the three Yaw angles previously calculated must satisfy the following constraints:

$$\begin{aligned} \gamma_1 &= Yaw_{13} - Yaw_{12} \\ \gamma_2 &= Yaw_{23} - Yaw_{12} \end{aligned} \tag{5}$$

Let us denote the known values of these angles with γ_1^0 and γ_2^0. We correct the estimation of the roll and pitch angles by exploiting these constraints.
To reach this goal, we solved a nonlinear least-squares problem. Specifically, we minimize the following cost function:

$$c(\zeta) = \tag{6}$$
$$= (Yaw_{13} - Yaw_{12} - \gamma_1^0)^2 + (Yaw_{23} - Yaw_{12} - \gamma_2^0)^2$$

in which the variables Yaw_{ij} are nonlinear functions of $\zeta = [Roll, Pitch]^\top$.

Once the least-squares algorithm finds the roll and pitch angles that minimizes the cost function previously described, we can estimate the Yaw angle and the coordinates x, y and z as described in $2p-$algorithm (Figure 8).

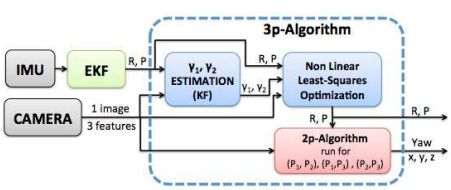


Fig. 8. Flow chart of the proposed pose estimator

C. Estimation of γ_1 and γ_2

In order to estimate the angles γ_1 and γ_2 of Figure 7 we run a Kalman filter.
The state that we want to estimate is $\Gamma = [\gamma_1, \gamma_2]^\top$.
During the prediction step the filter does not update the state Γ and its covariance matrix because the angles are constant in time. For the observation step we need the estimated roll and pitch and the camera images. Knowing those two angles, we can rotate the local frame in the new local frame shown in Figure 6. Being $[x_i, y_i, z]^T = z[\mu_i, \nu_i, 1]^T$ for $i = 1, 2, 3$ the observations of the three features in the current camera image, we can easily compute the angles of the triangle as follows. The three sides of the triangles are: $a = z\sqrt{\Delta\mu_{12}^2 + \Delta\nu_{12}^2}$, $b = z\sqrt{\Delta\mu_{13}^2 + \Delta\nu_{13}^2}$, $c = z\sqrt{\Delta\mu_{23}^2 + \Delta\nu_{23}^2}$.

Applying the law of cosine we can easily compute the two required angles:

$$\gamma_1 = \text{acos}\left(\frac{a^2+b^2-c^2}{2ab}\right)$$
$$\gamma_2 = \pi - \text{acos}\left(\frac{a^2+c^2-b^2}{2ac}\right)$$

Note that these angles are independent on z. We can use γ_1 and γ_2 as a direct observation of the state Γ of the Kalman Filter in order to correct the previous Γ.

V. Performance Evaluation

A. Simulations

In order to evaluate the performance of the presented method, we simulated different $3D$ trajectories and scenarios.

1) Simulated Scenario: The considered scenarios to test the 2p-Algorithm is shown in Figure 2. The features are $P1 = [0, 0, 0]$, $P2 = D * [1, 0, 0]$, where $D = 0.1m$. To compare the 2p-Algorithm with the 3p-Algorithm, we added a third feature $P3 = D*[0.5, \sqrt{3}/2, 0]$ (Figure 7). The angles γ_1 and γ_2 are respectively $60deg$ and $120deg$.

2) Simulated Trajectories: The trajectories are generated with a quadrotor simulator that, given the initial conditions, the desired position and desired Yaw, performs a hovering task [13]. The initial vehicle position is $x = y = z = 0\ m$, the initial vehicle speed is $v_x = v_y = v_z = 0\ ms^{-1}$ in the global frame.

3) Simulated Sensors: Starting from the performed trajectory, the true angular speed and the linear acceleration are computed each $0.01s$ We denote with $\boldsymbol{\Omega_i^{true}}$ and $\boldsymbol{A_{v\ i}^{true}}$ the true value of the body rates and linear accelerations at time stamp i. The IMU readings are generated as following: $\boldsymbol{\Omega_i} = N\left(\boldsymbol{\Omega_i^{true}} - \boldsymbol{\Omega_{bias}}, P_{\Omega_i}\right)$ and $\boldsymbol{A_i} = N\left(\boldsymbol{A_{v\ i}^{true}} - \boldsymbol{A_g} - \boldsymbol{A_{bias}}, P_{A_i}\right)$ where:

- N indicates the Normal distribution whose first entry is the mean value and the second one is the covariance matrix;
- P_{Ω_i} and P_{A_i} are the covariance matrices characterizing the accuracy of the IMU;
- $\boldsymbol{A_g}$ is the gravitational acceleration in the local frame and $\boldsymbol{A_{bias}}$ is the bias affecting the accelerometer's data;
- $\boldsymbol{\Omega_{bias}}$ is the bias affecting the gyroscope's data.

In all the simulations we set both the matrices P_{Ω_i} and P_{A_i} diagonal and in particular: $P_{\Omega_i} = \sigma_{gyro}^2 I_3$ and $P_{A_i} = \sigma_{acc}^2 I_3$, where I_3 is the identity 3×3 matrix. We considered several values for σ_{gyro} and σ_{acc}, in particular: $\sigma_{gyro} = 1\ deg\ s^{-1}$ and $\sigma_{acc} = 0.01\ ms^{-2}$.

The camera is simulated as follows. Knowing the true trajectory of the vehicle, and the position of the features in the global frame, the true bearing angles of the features in the camera frame are computed each $0.3s$. Then, the camera readings are generated by adding zero-mean Gaussian errors (whose variance is set to $(1\ deg)^2$) to the true values.

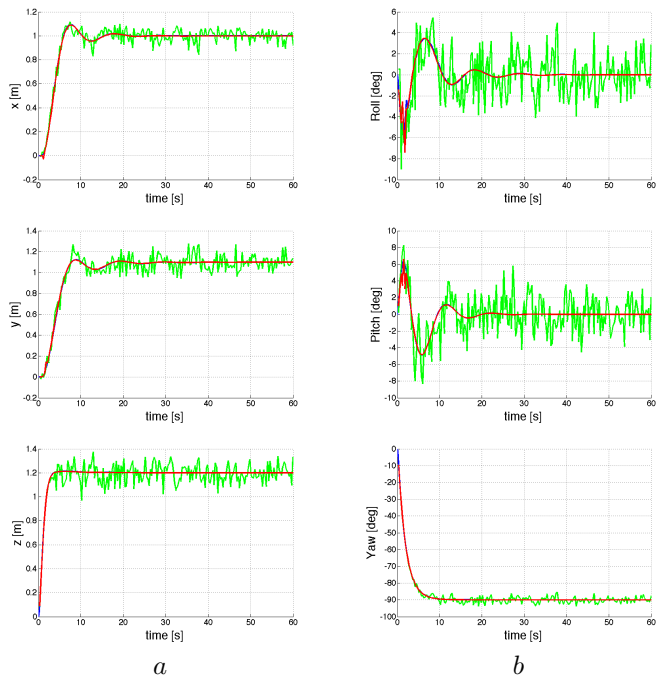


Fig. 9. Estimated x, y, z (a), and $Roll$, $Pitch$, Yaw (b). The blue line indicate the ground truth, the green one the estimation with the 2p-Algorithm and the red one the estimation with the 3p-Algorithm

Figures 9.a show the results regarding the estimated x, y and z. Figures 9.b show the results regarding the estimated $Roll$, $Pitch$ and Yaw. In each figure we represent the ground truth values in blue, the values estimated with the 2p-Algorithm in green and the values estimated with the 3p-algorithm in red.

Table I summarizes these results by providing the mean error on the estimated position and attitude.

	x	y	z	Roll	Pitch	Yaw
3pAlgo	0.26%	0.24%	0.08%	0.07 deg	0.04 deg	0.01 deg
2pAlgo	4.08%	5.41%	5.23%	1.63 deg	1.72 deg	1.36 deg

TABLE I

MEAN ERROR ON THE ESTIMATED STATES IN OUR SIMULATIONS. FOR THE POSITION THE ERROR IS GIVEN IN %

B. Experimental Results

This section describes our experimental results. The robot platform is a *Pelican* from *Ascending Technologies* [15] equipped with an Intel Atom processor board (*1.6 GHz, 1 GB RAM*) (Figure 1). Our sensor suite consists of an Inertial Measurement Unit (*3-Axis Gyro, 3-Axis Accelerometer*) belonging to the Flight Control Unit (FCU) AscTec Autopilot , and a monocular camera (*Matrix Vision mvBlueFOX*, FOV: 130 deg). The camera is calibrated using the Camera Calibration Toolbox for Matlab [14]. The calibration between the IMU and the camera has been performed using the Inertial Measurement Unit and Camera Calibration Toolbox in [10]. The IMU provides measurements update at a rate of $100Hz$, while the camera framerate is $10Hz$.

The Low Level Processor (LLP) of our Pelican is flashed with the *2012 LLP Firmware* [15] and performs attitude data fusion and attitude control. We flashed the High Level Processor (HLP) with the asctec_hl_firmware [17]. The onboard computer runs linux 10.04 and ROS (Robot Operating System). We implemented our method using ROS as a middleware for communication and monitoring . The HLP communicates with the onboard computer through a FCU-ROS node. The communication between the camera and the onboard computer is achieved by a ROS node as well. The presented algorithms are running online and onboard at $10Hz$.

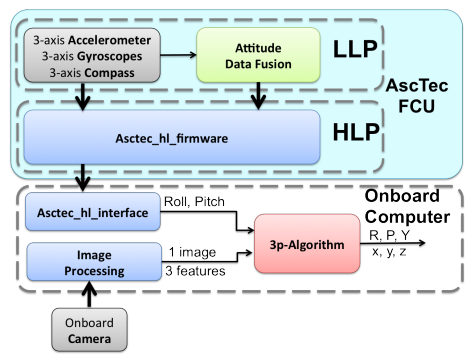


Fig. 10. Our Pelican quadcopter: a system overview

The scenario setup is shown in Figure 11. Since our lab is not yet equipped with a Motion Capture System, we used an ARToolKit Marker with the only aim of having a ground truth to evaluate the performance of our approach. The estimation of the camera pose provided by the marker is not used to perform the estimation. The marker is positioned such that it's reference frame is coincident with the configuration shown in Figure 7. The three features considered are the center of the three little balls in Figure 11. The use of three blob markers instead of natural features is only related to the need to get a ground truth. The information related to the pattern composed by the 3 features ($D = 0.25m$, $\gamma_1 = 60deg$, $\gamma 2 = 120deg$) is only used to evaluate the performance of our approach. The algorithm does not require any information about the features configuration.

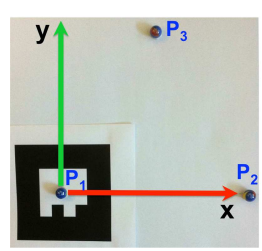
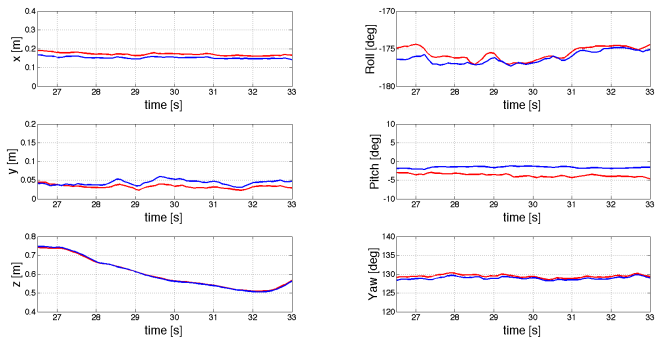


Fig. 11. Scenario: The AR Marker and the 3 balls are used only with the aim to get a rough ground truth. The AR Marker provides the camera $6DOF$ pose in a global reference frame according to our conventions.

Figures 12.a and 12.b show respectively the position and the attitude estimated by using the proposed approach and compared with the ground truth obtained with the ARToolkit marker. From Figure 12.a we see that the difference between our estimates and the ground truth values is of the order of $2cm$ for x and y and less than $0.5cm$ for z. From Figure 12.b we see that the difference between our estimates and the ground truth values is of the order of $2deg$ for $Pitch$ and less than $0.5deg$ for $Roll$ and Yaw.



a b

Fig. 12. Estimated position (a), respectively x, y, z and estimated attitude (b), respectively $Roll$, $Pitch$, Yaw. The red lines represent the estimated values with the 3p-Algorithm, the blue ones represent a rough ground truth (from ARToolkit Marker).

We believe that the main source of error is due to the distortion of the lens, which is not fully compensated by the calibration. Note that this distortion also affects our ground truth. We plan to test our approach in an environment equipped with a Motion Capture System.

VI. Conclusion

In this paper we proposed a new approach to perform MAV localization by only using the data provided by an Inertial Measurement Unit and a monocular camera. The approach exploits the so-called planar ground assumption and only needs three natural point features. It is based on a closed solution which provides the vehicle pose from a single camera image, once the roll and the pitch angles are obtained by the inertial measurements. This makes the approach very simple in terms of computational complexity and robust since the closed form makes unnecessary any initialization. We evaluated the performance of the proposed approach by using both synthetic and real data. We also described the results obtained by implementing the approach on our quadrotor in real-time and onboard.

The very low computational cost of the proposed approach makes it suitable for pose control in tasks like hovering, autonomous take off and landing.

References

[1] N. Abdelkrim, N. Aouf, A. Tsourdos, and B. White. Robust nonlinear filtering for ins/gps uav localization. In Mediterranean Conference on Control and Automation, pages 695702, June 2008.

[2] M. Achtelik, A. Bachrach, R. He, S. Prentice, and N. Roy. Stereo Vision and Laser Odometry for Autonomous Helicopters in GPS-denied Indoor Environments. In Proceedings of the SPIE Unmanned Systems Technology XI, 2009.

[3] Ahrens, S.; Levine, D.; Andrews, G.; How, J.P., Vision-based guidance and control of a hovering vehicle in unknown, gps-denied environments, IEEE International Conference on Robotics and Automation (ICRA 2009), Kobe, Japan, May, 2009.

[4] L. Armesto, J. Tornero, and M. Vincze Fast Ego-motion Estimation with Multi-rate Fusion of Inertial and Vision, The International Journal of Robotics Research 2007 26: 577-589

[5] A. Bachrach, R. He, and N. Roy. Autonomous flight in unstructed and unknown indoor environments. In European Conference on Micro Aerial Vechicles (EMAV), 2009

[6] A. Bachrach, R. He, and N. Roy. Autonomous flight in unknown indoor environments. International Journal of Micro Air Vehicles, 2009

[7] P. Corke, J. Lobo, and J. Dias, An Introduction to Inertial and Visual Sensing, International Journal of Robotics Research 2007 26: 519-535

[8] J. Dias, M. Vinzce, P. Corke, and J. Lobo, Editorial: Special Issue: 2nd Workshop on Integration of Vision and Inertial Sensors, The International Journal of Robotics Research, June 2007; vol. 26, 6: pp. 515-517.

[9] B. M. Kitt, J. Rehder, A. D. Chambers, M. Schonbein, H. Lategahn, and S. Singh. Monocular visual odometry using a planar road model to solve scale ambiguity. In Proceedings of the European Conference on Mobile Robots, Orebro, Sweden, 2011.

[10] J. Lobo and J. Dias, Relative pose calibration between visual and inertial sensors, International Journal of Robotics Research, 26(6):2007, 561-575.

[11] A. Martinelli, Vision and IMU Data Fusion: Closed-Form Solutions for Attitude, Speed, Absolute Scale and Bias Determination, Transaction on Robotics, Vol 28, pp. 44–60, February 2012

[12] B. Yun, K. Peng, and B. Chen. Enhancement of gps signals for automatic control of a uav helicopter system. In International Conference on Robotics and Automation, pages 11851189, June 2007.

[13] P. Castillo, R. Lozano and A. E. Dzul, Modelling and Control of Mini-Flying Machines, Springer 2005

[14] J.Y. Bouguet, Camera Calibration toolbox for Matlab. www.vision.caltech.edu/bouguetj.

[15] Ascending Technologies GmbH, http://www.asctec.de

[16] T. Zhang, Y. Kang, M. Achtelik, K. Kiihnlenz and M. Buss Autonomous Hovering of a Vision/IMU Guided Quadrotor, International Conference on Mechatronics and Automation, August 9 - 12, Changchun , China

[17] M. W. Achtelik, M. Achtelik, S. Weiss, Roland Siegwart Onboard IMU and Monocular Vision Based Control for MAVs in Unknown In- and Outdoor Environments In International Conference on Robotics and Automation (ICRA), 2011

[18] D. Eberli, D. Scaramuzza, S. Weiss, and R. Siegwart, Vision based position control for mavs using one single artificial landmark, International Conference & Exhibition on Unmanned Aerial Vehicles, (UAV), 2010.

[19] M. Hwangbo and T. Kanade Visual-Inertial UAV Attitude Estimation Using Urban Scene Regularities International Conference on Robotics and Automation, May 9-13, 2011, Shanghai, China

[20] T. Cheviron, T. Hamel, R. Mahony and G. Baldwin Robust Nonlinear Fusion of Inertial and Visual Data for position, velocity and attitude estimation of UAV International Conference on Robotics and Automation Roma, Italy, 10-14 April 2007

[21] S. Weiss., Vision Based Navigation for Micro Helicopters PhD Thesis, 2012

[22] G. Klein and D. Murray, Parallel tracking and mapping for small ar workspaces, in International Symposium on Mixed and Augmented Reality, Nov. 2007, pp. 225234.

[23] S. Weiss, M. W. Achtelik, M. Chli and R. Siegwart. Real-time Onboard Visual-Inertial State Estimation and Self-Calibration of MAVs in Unknown Environments. in IEEE International Conference on Robotics and Automation (ICRA), 2012

[24] S. Weiss, D. Scaramuzza and R. Siegwart, Monocular-SLAMbased navigation for autonomous micro helicopters in GPS-denied environments, Journal of Field Robotics (JFR), Vol. 28, No. 6, 2011, 854-874.