International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XXXVIII-1/C22, 2011
ISPRS Zurich 2011 Workshop, 14-16 September 2011, Zurich, Switzerland

MAV-BASED REAL-TIME LOCALIZATION OF TERRESTRIAL TARGETS WITH CM-LEVEL ACCURACY: FEASIBILITY STUDY

Yosef Akhtman, Abhinav Garg and Jan Skaloud

TOPO Lab, École polytechnique fédérale de Lausanne, Lausanne, Switzerland
{yosef.akhtman,abhinav.garg,jan.skaloud}@epfl.ch
http://topo.epfl.ch

KEY WORDS: UAVs, MAVs, Photogrammetry, Sensor Orientation, Direct Georeferencing

ABSTRACT:

We carry out a comprehensive feasibility study for a real-time cm-level localization of a predefined terrestrial target from a MAV-based autonomous platform. Specifically, we conduct an error propagation analysis which quantifies all potential error sources, and accounts for their respective contribution to the final result. Furthermore, we provide a description of a practical MAV system using the available technology and of-the-shelf components. We demonstrate that, indeed, the desired localization precision of a few centimeters may be realistically achieved under a set of necessary constraints.

1 INTRODUCTION

The field of micro-air-vehicles (MAVs) has been recently experiencing a surge in research and development, as well as rapid commercialization. The potential applications include aerial mapping, surveillance, remote sensing, support of emergency services, telecommunications, etc. The key appeal of the MAV concept is constituted by its mobility in conjunction with the small form-factor, license-free operation, and the convenience of a single operator handling.

Despite the considerable potential, the design of MAV platforms, which would be capable of matching the functional efficacy exhibited by some of the larger UAVs has proven to be a major challenge, largely because of the stringent constraints on the size and the weight of the on-board sensor and computing equipment (Pines and Bohorquez, 2006). A considerable progress in this context has been recently reported, for instance, in (Zufferey et al., 2010), where a high-resolution wide-area mapping and visualization have been achieved using a 500g Swinglet MAV platform, that employs of-the-shelf consumer-grade digital camera as the optical sensor. Subsequently, a relatively accurate georeferencing of the acquired visual data has been realized using a bundle adjustment method (Förstner and Wrobel, 2004). Nevertheless, a real-time georeferencing of terrestrial targets with a sub-decimeter precision, routinely required in some professional applications, from a MAV platform remains an open problem.

In this paper we discuss the design considerations for a MAV platform capable of a cm-level localization of terrestrial targets by the means of direct real-time localization of a predefined planar target using a combination of on-board sensors including a GPS, an IMU and a high-resolution camera. Specifically, in Section 2 we introduce an alternative formulation of a standard direct georeferencing problem, which utilizes the *a priori* knowledge of a terrestrial target's dimensions and internal orientation. In Section 3 we conduct a comprehensive analysis of the error propagation inherent to the system design and the algorithmic flow considered. In Section 4 we provide a description of a practical MAV georeferencing system in development and detail the numerical properties of all potential noise sources associated with the various system components. Finally, in Section 5 we provide a quantitative analysis of the achievable performance before drawing our conclusions in Section 6.

2 PROBLEM FORMULATION

We commence with the mathematical formulation of the standard georeferencing problem, namely that of determining the position of a physical point p in the Earth-centered Earth-fixed (ECEF) frame of reference from an airborne platform, as illustrated in Fig. 1 and described in (Schwarz et al., 1993)

$$\boldsymbol{x}_p^e(t) = \boldsymbol{x}_b^e(t) + \boldsymbol{R}_l^e(t)\boldsymbol{R}_b^l(t)\boldsymbol{R}_s^b(\omega_b, \phi_b, \kappa_b)\left(\boldsymbol{x}_s^b + \boldsymbol{x}_p^s(t)\right), \tag{1}$$

where $\boldsymbol{x}_b^e(t)$ is the navigation center of the onboard IMU, which represents the origin of the body (b) frame in the earth (e) frame of reference; $\boldsymbol{x}_s^b$ denotes the relative displacement between the optical sensor (s) and the body frames of reference origins, while the rotation matrix $\boldsymbol{R}_s^b$ describes the corresponding relative misalignment between the s and the b frames, and is defined by the three Euler angles ω_b, ϕ_b and κ_b. Finally, $\boldsymbol{R}_b^l$ and $\boldsymbol{R}_l^e$ denote the rotation matrices, which represent the body-to-local and the local-to-earth frame conversions, respectively. The notation (t) indicates the temporal variability of the preceding quantity.

In the context of this paper, however, we would like to substitute the singular point p with a planar target of predefined dimension, asymmetric marking and orientation, which introduces a corresponding right-handed frame of reference p having its z-axis pointing up perpendicular to the plain of the target as well as the x and y axes predefined in the plain of the target. Furthermore, we would like to introduce the notation $\boldsymbol{x}_{ij}^k$ to denote a vector from the origin of frame i to the origin of frame j and expressed using the k frame of reference.

Consequently, we may reformulate Eq. (1) based on the following analysis

$$\begin{aligned}\boldsymbol{x}_p^e(t) =& \boldsymbol{x}_b^e(t) + \boldsymbol{R}_l^e(t)\boldsymbol{x}_{bp}^l(t) = \boldsymbol{x}_b^e(t) - \boldsymbol{R}_l^e(t)\boldsymbol{x}_{pb}^l(t) \\ =& \boldsymbol{x}_b^e(t) - \boldsymbol{R}_l^e(t)\boldsymbol{R}_p^l(t)\boldsymbol{x}_b^p(t),\end{aligned} \tag{2}$$

where we have the target-to-local frame transformation

$$\boldsymbol{R}_p^l(t) = \boldsymbol{R}_b^l(t)\boldsymbol{R}_s^b\boldsymbol{R}_p^s(t), \tag{3}$$

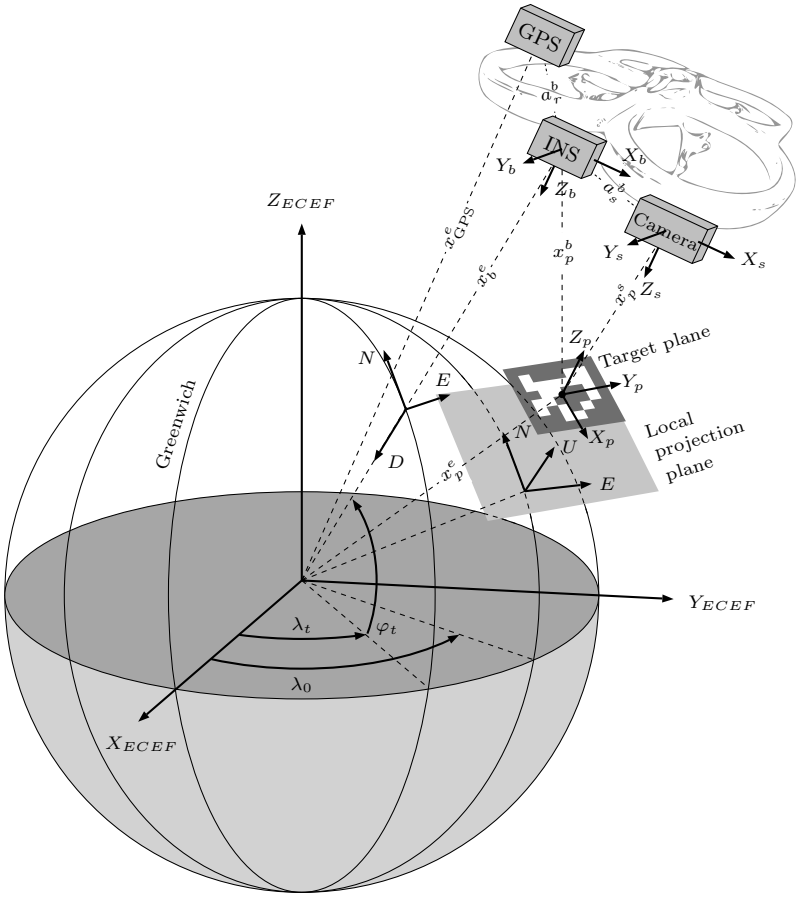

Figure 1: MAV-based direct georeferencing.

and the target-to-body vector

$$\boldsymbol{x}_b^p(t) = -\boldsymbol{x}_{bp}^p(t) = -\boldsymbol{R}_s^p(t)\left(\boldsymbol{x}_s^b + \boldsymbol{x}_p^s(t)\right). \tag{4}$$

Observe that substituting (4) and (3) into (2) makes the resultant expressions of Equations (2) and (1) identical. Importantly, however, the target's position estimation expression in (2) eliminates the direct dependency on the quantities $\boldsymbol{R}_b^l(t)$ and $\boldsymbol{x}_p^s(t)$ which are subject to rapid temporal variability due to the instantaneous changes to the MAV's orientation, in particular roll and pitch. Instead, in (2) we have the target-to-local frame transformation $\boldsymbol{R}_p^l(t)$, which only depends on the target's attitude that may be assumed time-invariant and will be therefore denoted as simply $\boldsymbol{R}_p^l$ from here on; as well as the target-to-body vector $\boldsymbol{x}_b^p(t)$, which has to account for the relatively slow translational movements of the MAV, but is independent of its orientation. Finally, we will drop the time-dependancy of the local-to-earth transformation $\boldsymbol{R}_l^e$, which in the scope of our analysis may be considered insignificant. We may thus further simplify the expression in (2) to read

$$\boldsymbol{x}_p^e(t) = \boldsymbol{x}_b^e(t) - \boldsymbol{R}_l^e\boldsymbol{R}_p^l\boldsymbol{x}_b^p(t). \tag{5}$$

It is the primary objective of this paper to analyze the error budget, and the corresponding attainable accuracy within the algorithmic flaw of calculating the result of Equation (5).

3 ERROR PROPAGATION ANALYSIS

Our analysis is based on the first-order approximation of the error exhibited by the composite observation y that is calculated as a function of a number of noisy observations $x_i = x_i^0 + \varepsilon_i,\ i = 0, 1, \ldots, n$, where for the sake of simplicity we will assume only Gaussian noise sources $\varepsilon_i \sim N(\mu_i, \sigma_i^2)$, and $y = f(x_0, x_1, \ldots, x_n)$. The first-order approximation of the resultant error may be expressed as

$$\varepsilon \approx \sum_{i=0}^{n} \frac{\partial f}{\partial x_i}\Big|_{\epsilon_j=0}\varepsilon_i. \tag{6}$$

Applying the principle of Eq. (6) to the evaluation of $\boldsymbol{x}_p^e(t)$ in (5) yields

$$\begin{aligned}\boldsymbol{\varepsilon}_p^e \approx \boldsymbol{\varepsilon}_b^e + \dot{\boldsymbol{x}}_b^e\tau - \boldsymbol{\Lambda}_l^e\boldsymbol{R}_p^l\boldsymbol{x}_b^p - \boldsymbol{R}_l^e\boldsymbol{\Lambda}_p^l\boldsymbol{x}_b^p \\ - \boldsymbol{R}_l^e\boldsymbol{R}_p^l\boldsymbol{\varepsilon}_b^p - \boldsymbol{R}_l^e\boldsymbol{R}_p^l\dot{\boldsymbol{x}}_b^p\tau,\end{aligned} \tag{7}$$

and furthermore from Equations (4) we have

$$\boldsymbol{\varepsilon}_b^p = -\left(\boldsymbol{\Lambda}_s^p + \tau\dot{\boldsymbol{R}}_s^p\right)\left(\boldsymbol{x}_s^b + \boldsymbol{x}_p^s\right) - \boldsymbol{R}_s^p\left(\boldsymbol{\varepsilon}_s^b + \boldsymbol{\varepsilon}_p^s\right) \tag{8}$$

and

$$\dot{\boldsymbol{x}}_b^p = -\dot{\boldsymbol{R}}_s^p\left(\boldsymbol{x}_s^b + \boldsymbol{x}_p^s\right) - \boldsymbol{R}_s^p\dot{\boldsymbol{x}}_p^s. \tag{9}$$

The objective of the analysis carried out in this section is to assess

the magnitude and the statistical properties of the expected error $\boldsymbol{\varepsilon}_p^e$ associated with the calculation of Equation (5) based on a instantaneous set of sensory observation at time t. In particular, the temporal properties of the encountered error are beyond the scope of this analysis. We have therefore omitted the time dependancies of the various quantities in Equations (7–8). We may furthermore simplify the expressions (7–8) using the following simple observations

1. We are utilizing GPS as the time reference and may therefore assume $\tau_{\mathrm{GNSS}} = 0$, which eliminates the corresponding error component in Eq. (7).
2. The residual misalignment between e and l frames denoted by the quantity $\boldsymbol{\Lambda}_l^e$ in (7) may be generally ignored relative to other error components due the accuracy in positioning.
3. The frame-to-frame conversion operators $\boldsymbol{R}_i^j$ in (7–8) are orthonormal matrices, which do not affect the magnitude of the resultant error components, and may therefore be substituted by a unity matrix $\boldsymbol{I}$ when focusing on the error's amplitude.
4. The quantity $\boldsymbol{\varepsilon}_s^b$ in Equation (8) describe the residual error after the calibration of the translation displacement (also called level-arm) between the body and the optical sensor's frames of reference. Although this quantity may potentially result in a some systematic bias, it can be usually mitigated by either improved calibration, or compensation by the sensor fusion filter. We may therefore assume this error components to be negligible relative to the other noise sources.
5. The length of the lever-arm vector $\boldsymbol{x}_s^b$ in (8), which is equal to a few centimetres in our case, may be considered to be negligible relative to the length of the vector $\boldsymbol{x}_p^s$, and may therefore be ignored. We may also assume $\boldsymbol{x}_p^b = \boldsymbol{x}_p^s$.
6. The residual misalignment $\boldsymbol{\Lambda}_p^l$ between the target and the local frame of reference may be mitigated to a desired level of precision by a sensor fusion and filtering. Here, we make an explicit assumption that the target's orientation is either known *a priori*, or exhibits a very slow temporal variability.

Applying observations 1–6 to Equations (7)–(9) and further substituting (8) and (9) into (7) yields

$$\boldsymbol{\varepsilon}_p^e \sim \boldsymbol{\varepsilon}_b^e + \boldsymbol{\varepsilon}_p^s + \boldsymbol{\Lambda}_s^p \boldsymbol{x}_p^s + \dot{\boldsymbol{R}}_s^p \boldsymbol{x}_p^s \tau + \dot{\boldsymbol{x}}_p^s \tau, \tag{10}$$

where the remaining time variable τ denotes the synchronization error between the timing of the optical sensor reading and the GPS fix time reference.

We may now examine the remaining error components of Eq. (10) in detail. In particular

- $\boldsymbol{\varepsilon}_b^e$ is a GNSS measurement error that may be assumed to to be unbiased and have a variance of σ_{GNSS}^2.
- $\boldsymbol{\varepsilon}_p^s$ is an error associated with the measurement of the vector between the optical sensor and the center of the target, which is calculated from the target's projection onto the image plain of the optical sensor. In this context we have to identify the following two distinctive error constituents, namely the error in the position of the target's projection center in the image plane, and the error in the calculation of the target's projection size in the image plane (scale error). Subsequently, the error ε_{pi} in the calculation of a point p_i of the target p relative to the image plane may be quantified as

$$\varepsilon_{pi} = \rho \varepsilon_{\theta i}, \tag{11}$$

where $\varepsilon_{\theta i}$ is the angular error associated with the observation of the point p_i in the image plane, while ρ is the distance between the camera and the target, or in other words the magnitude of the vector $\boldsymbol{x}_p^s$. The corresponding error variance for the calculation of a center of a target relative to the image plane based on the detection of N distinct points may be expressed as

$$\sigma_{pc}^2 = \frac{1}{N} \rho^2 \sigma_\theta^2, \tag{12}$$

where σ_θ^2 is the angular resolution of the optical sensor considered. Furthermore, the scale error will result in an error in the calculation of the magnitude ρ of the vector $\boldsymbol{x}_p^s$, having a variance of

$$\sigma_\rho^2 = \frac{\rho^2}{d^2} \sigma_{pc}^2 = \frac{1}{N} \frac{\rho^4}{d^2} \sigma_\theta^2, \tag{13}$$

where d denotes the size of the target. Importantly, the error $\boldsymbol{\varepsilon}_p^s$ is perpendicular to the image of the plane, while the error $\boldsymbol{\varepsilon}_{pc}$ lies in the image plane. We may therefor express the variance of the resultant horizontal translation error in the position of the target as

$$\sigma_{ph}^2 = \frac{1}{N} \frac{\rho^4 \sigma_\theta^2}{d^2} \sin^2 \phi + \frac{1}{N} \rho^2 \sigma_\theta^2 \cos^2 \phi, \tag{14}$$

where ϕ is the angle between the camera bore-sight and the local horizontal plane.

- $\boldsymbol{\Lambda}_s^p \boldsymbol{x}_p^s$ describes the error in the calculation of the position of the optical sensor in the target frame of reference. The corresponding orientation vector is calculated from the distortion of the target's projection in the image plane. We may therefore conclude that the error vector $\boldsymbol{\Lambda}_s^p \boldsymbol{x}_p^s$ lies in parallel to the image plane and has a magnitude variance of

$$\sigma_s^2 = \frac{\rho^2}{d^2} \sigma_{pc}^2 = \frac{1}{N} \frac{\rho^4}{d^2} \sigma_\theta^2, \tag{15}$$

while having a contribution to the horizontal error with a variance of

$$\sigma_{sh}^2 = \frac{1}{N} \frac{\rho^4 \sigma_\theta^2}{d^2} \cos^2 \phi, \tag{16}$$

- $\dot{\boldsymbol{R}}_s^p \boldsymbol{x}_p^s \tau$ describes the noise component introduced by the synchronization offset between the GNSS fix and the acquired data from the optical sensor. Specifically, the time derivative of the rotation matrix $\boldsymbol{R}_s^p$ may be expressed using the skew-symmetric matrix $\boldsymbol{\Omega}_{sp}^p$ as defined in (Skaloud and Cramer, 2011), such that we have

$$\begin{aligned} \left\| \dot{\boldsymbol{R}}_s^p \boldsymbol{x}_p^s \tau \right\| &\leq \left\| \begin{pmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\| \rho^2 \tau^2 \\ &= \left\| \begin{pmatrix} -\omega_z + \omega_y \\ \omega_z - \omega_x \\ -\omega_y + \omega_x \end{pmatrix} \right\| \rho^2 \tau^2 \\ &\approx 2 \left(\omega_x^2 + \omega_y^2 + \omega_z^2 \right) \rho^2 \tau^2. \end{aligned} \tag{17}$$

Consequently, the variance of the resultant error component $\dot{\boldsymbol{R}}_s^p \boldsymbol{x}_p^s \tau$ may be expressed as

$$\sigma_w^2 = 2\omega^2 \rho^2 \tau^2, \tag{18}$$

where ω is the instantaneous angular speed exhibited by the MAV.

- $\dot{\boldsymbol{x}}\tau$ is likewise the noise component introduced by the synchronization offset between the GNSS fix and the acquired data from the optical sensor, which has a variance of

$$\sigma_v^2 = v^2 \tau^2, \tag{19}$$

where v denotes the MAV horizontal speed.

To summarize we may substitute the results of Equations (14), (16), (18), (19) into (10), which yields the overall horizontal position variance of

$$\sigma_{ep}^2 = \sigma_{\text{GNSS}}^2 + \frac{1}{N}\frac{\rho^4 \sigma_\theta^2}{d^2} + \frac{1}{N}\rho^2 \sigma_\theta^2 \cos^2\phi + 2\omega^2 \rho^2 \tau^2 + v^2 \tau^2 \tag{20}$$

In Section 4 we detail the quantitative characteristics of all terms in Equation (20).

4 SYSTEM DESIGN

We employ a Pelican quadrotor MAV platform from Ascending Technologies, which was chosen primarily due to its best-in-class payload capacity and configurational flexibility (Ascending Technologies, 2010). The Asctec Pelican MAV platform comprises of a microcontroller-based autopilot board, as well as an Atom processor-equipped general purpose computing platform. The maximum available payload of the Pelican MAV is 500 g, which caters for considerable level of flexibility in the choice of a custom on-board sensor constellation.

The on-board computer has been installed with a customized Ubuntu Linux operating system. Furthermore, we have utilized a Robot Operating System (ROS, www.ros.org) running on top of Linux OS for the sake of performing high-level control, image processing, as well as data exchange, logging and monitoring tasks. In particular, the interface between the on-board computer and the AscTec autopilot board were implemented using the *asctec_drivers* software stack of ROS (Morris et al., 2011).

The planar target pose estimation has been developed by adopting the methodology commonly employed by the augmented reality research community. Specifically, we have utilized the open source *ARToolKitPlus* software library (Wagner and Schmalstieg, 2007), which implements the Robust Planar Pose (RPP) algorithm introduced in (Schweighofer and Pinz, 2006).

4.1 GNSS

We employ a Javad TR-G2 L1 receiver, with a support for SBAS corrections and RTK-enabled code and carrier position estimates at a maximum update rate of 100Hz. For the sake of this study we will assume an unbiased GNSS sensor output having a standard deviation, as detailed by the manufacturer's specifications document (Javad, 2010), of $\sigma_{\text{GNSS}} = 10$mm. Additionally, the TR-G2 sensor provides redundant velocity and speed measurements, which may be utilized to mitigate the corresponding noise components generated by other on-board sensors. The AscTec Pelican MAV platform also includes an redundant built-in Ublox LEA-5t GPS module, which is utilized for INS-GNSS sensor fusion and GNSS-stabilized flight control. The resultant additional position estimates may be fused with the primary GNSS data, but their accuracy appears to be inferior to the measurements of the Javad module, and therefore does not effect the assumed statistics of the position errors.

4.2 INS

The AscTec Pelican MAV platform includes an on-board IMU platform, which comprises a 3D accelerometer, three MEMS-based first-class gyros, 3D magnetometer, and a pressure-based altimeter. The serial port interface with the on-board sensors facilitates the pooling of both sensor fused and filtered, as well as raw sensory data at a maximum rate of 100Hz (Ascending Technologies, 2010).

The filtered attitude data after sensor fusion generated by this sensor exhibited a functional unbiased accuracy of about $0.1\,\text{deg} = 1.75 \times 10^{-3}$rad. Additional MEMS-IMU was considered as part of this study, which shall soon enter production. Based on a preliminary testing in static conditions the new sensor exhibited an accuracy of $0.01\,\text{deg} = 1.75 \times 10^{-4}$rad. As part of the orientation errors is coupled with the errors in accelerometers, geometric uncertainty in the sensor assembly and initialization errors, the accuracy stated above may be optimistic and is yet to be confirmed under dynamic conditions.

4.3 Optical sensor

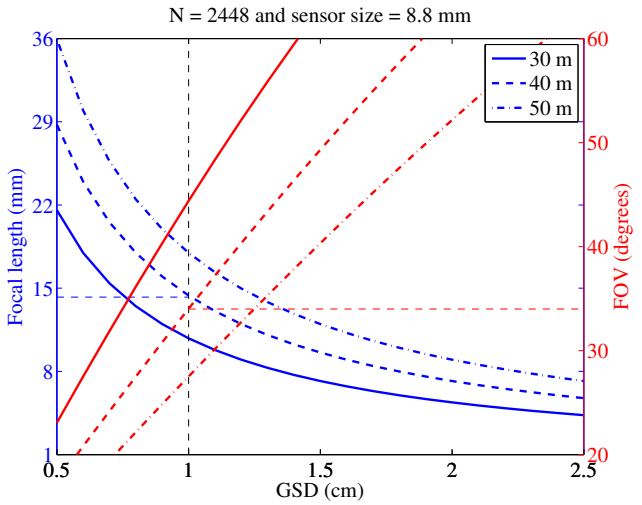

Figure 2: Focal length and Field of View (FoV) as a function of the Ground Sampling Distance (GSD) in relation to the flying height above terrain.

The optical sensor constitutes the critical component of the considered application. The properties which have to be taken into account include image quality, image sensor technology, connectivity, synchronization capability and accuracy, configuration flexibility, as well as mechanical coupling, dimensions and weight. These constraints exclude the possibility of employing a consumer-grade compact cameras that are typically used on MAV platforms (Zufferey et al., 2010). Furthermore, industrial-grade cameras with the necessary specifications has appeared only very recently. To the best of our knowledge only one such device satisfying all system requirement, including resolution, connectivity, and mechanical properties was commercially available at the time of submission

Camera parameters	
Interface	USB 2.0
Lens mount	C-Mount
Sensor technology	CCD
Resolution (h x v)	2448×2050
Resolution category	5 Megapixel
Color depth	8bit (14bit ADC)
Sensor size	2/3" (8.446×7.066 mm)
Shutter	Global
Dimensions	34.0×32.0×34.4 mm
Mass	74.0 g
Lens parameters	
Optical resolution rating	3 Megapixel
Focal length	12 mm
Back focal length	8.4 mm
Maximum aperture	1.4f
Angle of view @2/3" (h x v)	39.57°×30.14°
Optical distortion	≤0.7%
Dimensions	ø30.0×51.0 mm
Weight	68.2 g
Experimental optical resolution	10^{-4}rad ≈ 0.34arcmin

Table 1: UEye USB UI-2280SE camera and VS Technology B-VS1214MP lens specifications

of this paper. More specifically, we have opted for a 5MP C-Mount USB camera having a 2/3" CCD-type image sensor and connected to the on-board Atom board using an USB-2.0 cable.

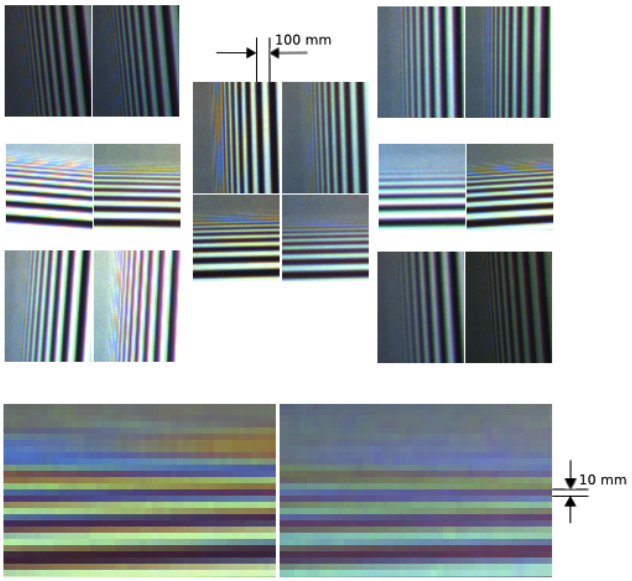

Figure 3: Resolution comparison between the 5M and the 3M lenses using the images of a resolution testing board taken from a distance of 100 m.

In order to guarantee a robust visual tracking of terrestrial targets in the presents of inevitable pitch/roll variations, as well as to maintain good geometry for mapping applications in post-processing mode (i.e. sufficient overlap and good intersections of rays), the on-board optical sensor is required to satisfy a tight trade-off between the attainable ground sample distance (GSD) and the sufficiently wide field-of-view (FoV). The resultant correspondence between the required characteristics is depicted in Fig. 2. In particular a lens having a focal length of 12 mm have been found to provide the optimum tradeoff between the attainable GSD and FoV.

Subsequently, a number of 12 mm lenses have been tested for attainable angular resolution and optical distortion. Fig. 3 portrays the resolution comparison between the 5MPixel and the 3MPixel lenses. The figure contains the full-zoom fragments from the images of a resolution testing board taken from a distance of 100 m. Each pair of images contains a 3MP lens-based image on the left and a 5MP lens-based image on the right taken at the respective edge of the image frame (for instance the top-right pair of images were taken by positioning the target board at the very top-right of the visible frame). Furthermore, at the bottom of the image there is a further 5x blow-up of the 3MP and 5MP images from the center of the frame, where a single-pixel resolution may be visible. The specifications of the camera and the 3MP lens being found to satisfy the necessary requirements are summarized in Table 1. In particular the camera-lens combination characterized in Table 1 exhibits an angular resolution of 10^{-4} radians. Consequently, for the sake of calculation of Eq. (20) we will assume the angular error standard deviation of $\sigma_\theta = 0.0001$.

4.3.1 Camera calibration procedure In order to minimize the contribution of the optical distortion errors to the overall localization accuracy, we have conducted a comprehensive calibration of the optical sensor. Specifically, a free-network based camera calibration procedure introduced in (Luhmann et al., 2006) was carried out to estimate the essential calibration parameters using bundle adjustment. A target network was constructed in outdoor conditions with targets scattered in three dimensions. This is important to achieve minimum correlation between the interior and exterior orientation parameters, which is crucial for achieving maximum positioning accuracy with the camera. Images were taken from multiple camera positions and at every location camera was rolled by 90° to ensure decorrelation of the camera perspective centre coordinates from other parameters. The targets were surveyed using a Leica TCR403 theodolite with 3mm positioning accuracy (as specified by the manufacturer).

4.4 Planar target design

(a) (b) (c) (d)

Figure 4: Planar marker systems utilized in computer vision (Köhler et al., 2010).

Several computer vision marker systems depicted in Fig. 4 and discussed in detail in (Köhler et al., 2010) have been considered for the design of the terrestrial planar target. We have selected the BCH ARTag marker set exemplified in Fig. 4 (c) due to its high marker library size, near-zero false positive identification rate, BCH code protected near zero inter-tag confusion rate, as well as good pose estimation properties (Fiala, 2010). Specifically, a $(1/20)$-pixel target acquisition accuracy has been reported in (Fiala, 2010). This result needs to be verified experimentally in dynamic conditions, but for the sake of this study, we will assume $N = 20$ in Equation (20).

5 QUANTITATIVE ANALYSIS

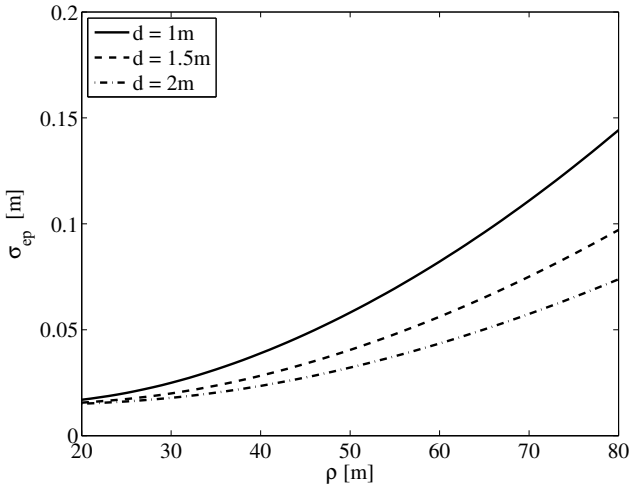

Figure 5: Terrestrial target horizontal position standard deviation versus MAV flight altitude ρ and target size d assuming the flight speed of $v = 5$m/s and synchronization error of $\tau = 1$ms.

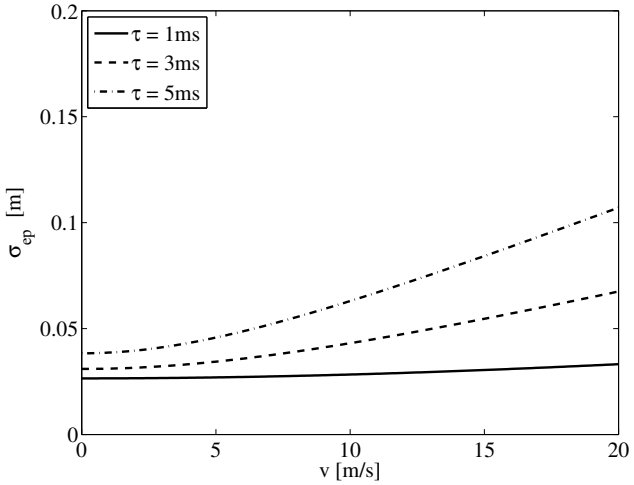

Figure 6: Target horizontal position standard deviation versus MAV horizontal speed v and sensor synchronization error τ assuming MAV flight altitude of $\rho = 40$m and target size of $d = 1.5$m.

Figures 5 and 6 characterize the expected standard deviation of the terrestrial target horizontal position, calculated from Eq. (20), as a function of the flight-above-terrain altitude ρ and the horizontal flight speed v, respectively. In particular, Fig. 5 suggests the feasibility of terrestrial target localization with a precision of about 3cm from an altitude of 40m and using a target of 1.5m in diameter. Observe that the quadratic relation between the estimated position standard deviation and the flight altitude constrains the altitude at which sub-decimetre georeferencing of terrestrial targets is attainable. Furthermore, Fig. 6 demonstrates the tolerance of the target localization accuracy to the MAV flight speed. Specifically, at the flight altitude of $\rho = 40$m a sub-3cm localization accuracy may be achieved with synchronization accuracy of $\tau = 1ms$ and flight speeds of up-to $v = 20m/s$.

6 CONCLUSIONS AND FUTURE WORK

We have carried out a simplified error propagation analysis when localizing a multi-point planar terrestrial target from a MAV platform in real-time. Our theoretical analysis is complemented by the quantitative characterization of the attainable performance based on specific hardware platform. Our feasibility study suggests that the cm-level (< 3cm) real-time localization is attainable using commercially available hardware and we intend to demonstrate such functionality as part of our future research.

REFERENCES

Ascending Technologies, 2010. AscTec Pelican with Autopilot User's Manual. URL: www.asctec.de/assets/Downloads/Manuals/AscTec-Autopilot-Manual-v10.pdf.

Fiala, M., 2010. Designing highly reliable fiducial markers. IEEE Transactions on Pattern Analysis and Machine Intelligence 32(7), pp. 1317–1324.

Förstner, W. and Wrobel, B., 2004. Mathematical concepts in photogrammetry. In: J. McGlone and E. Mikhail (eds), Manual of Photogrammetry, fifth edn, American Society of Photogrammetry and Remote Sensing, Bethesda, MA., pp. 15–180.

Javad, 2010. Javad TR-G2 Datasheet. URL: javad.com/downloads/javadgnss/sheets/TR-G2_Datasheet.pdf.

Köhler, J., Pagani, A. and Stricker, D., 2010. Detection and identification techniques for markers used in computer vision. In: M. Grabner and H. Grabner (eds), Visualization of Large and Unstructured Data Sets IRTG Workshop.

Luhmann, T., Robson, S., Kyle, S. and Harley, I., 2006. Close range photogrammetry: principles, techniques and applications. Whittles Publishing Dunbeath, Scotland.

Morris, W., Dryanovski, I., Bellens, S. and Bouffard, P., 2011. Robot Operating System (ROS) software stack: *asctec_drivers*. The City College of New York. www.ros.org/wiki/asctec_drivers.

Pines, D. J. and Bohorquez, F., 2006. Challenges facing future micro-air-vehicle development. Journal of Aircraft 43, pp. 290–305.

Schwarz, K. P., Chapman, M., Cannon, M. and Gong, P., 1993. An integrated ins/gps approach to the georeferencing of remotely sensed data. Photogrammetric Engineering and Remote Sensing 59(11), pp. 1167–1674.

Schweighofer, G. and Pinz, A., 2006. Robust pose estimation from a planar target. Pattern Analysis and Machine Intelligence, IEEE Transactions on 28(12), pp. 2024–2030.

Skaloud, J. and Cramer, M., 2011. GIS Handbook. Springer, chapter Data Acqusition.

Wagner, D. and Schmalstieg, D., 2007. ARToolKitPlus for Pose Tracking on Mobile Devices. In: M. Grabner and H. Grabner (eds), Computer Vision Winter Workshop.

Zufferey, J.-C., Beyeler, A. and Floreano, D., 2010. Autonomous flight at low altitude using light sensors and little computational power. International Journal of Micro Air Vehicles 2, pp. 107–117.