A vision-based approach for high accuracy assessment of satellite attitude

Alessandro Bevilacqua, Alessandro Gherardi and Ludovico Carozza

ARCES - Advanced Research Centre on Electronic Systems - University of Bologna

via Toffano 2/2 - 40125 Bologna, Italy

`{abevilacqua,agherardi,lcarozza}@arces.unibo.it`

Abstract

In this work we present the achievements of the approach we have conceived to estimate in real time the attitude of a satellite endowed with an integral camera looking to the Earth. We have implemented the dynamical model that describes how to reconstruct the spatial orientation of the camera using couples of subsequent views, passing through six reference frames. An in-depth analysis of the influences on the accuracy of the different sources of error arising when implementing the dynamical model has been carried out, from the theoretical framework to the laboratory prototype. Through these studies we have been able to characterize, and often overcome, the traps hidden at the development stage to propose the first vision based approach to determine with very a high accuracy the attitude of a satellite in a stand-alone system. Although the achievements refer to a classical mission critical application, the methodological approach, the choices accomplished and the results presented are of general interest.

1. Introduction

The automatic guidance and control of remote sensing devices in different fields (marine, aerospace, biomedical, etc.) represents a complex and challenging task, often involving the intelligent integration of a huge variety of information. Depending on the applications, the accuracy required, the power consumption, the computational burden and the economic budget are important factors that must be taken into account, trying to reach an acceptable trade-off among these aspects. Applications of computer vision to these fields are nowadays standing out for the advantages they can carry. The employment of a CCD camera together with a processing unit can replace ensembles of different sensors requiring a subsequent data integration. Navigation systems based on *visual odometry* have been conceived to achieve measurements of the motion parameters of the controlled object. This problem has been addressed using egomotion estimation, where the sensor's pose in an inertial reference frame is recovered by processing images of the surrounding environment acquired from one or more cameras mounted on-board.

The problem of recovering the attitude by using vision based approaches has been studied since the late Nineties, for Unmanned Aerial Vehicles (UAVs). The typical needs in UAVs are autonomous safe landing or target hitting. In the first case, often the attitude is recovered by tracking targets placed on the landing strip [4], while in the second case the system looks for known natural terrestrial landmarks to match (*e.g.* roads, rivers, etc.) [23]. Nevertheless, when vision-based only systems are used, often the accuracy needed (and provided) is of the order of degrees or few decimal of degrees [1]. Usually, in this application field there is no need to go deep into the sources of error of the *software* system in order to attain a higher accuracy, simply because what achieved is suitable for the application itself.

On the contrary, in stand alone mission critical systems for satellite applications is crucial to obtain very accurate information in order to recover attitude parameters. An experimental laboratory setup has to be built to provide direct and accurate measurements for ground truth data that permit to identify the most critical sources of error and face them to achieve the highest accuracy as possible. In this paper we propose the implementation of a dynamical model deriving from the application of registration techniques to satellite terrestrial images in order to estimate with a *high accuracy* the attitude of the satellite. The attitude parameters are recovered from the geometrical transformation that maps views of the Earth acquired at different epochs along the orbit. Since attitude estimation represents a mission critical step in satellite applications, it must present high robustness and precision, even working in real time requiring, at the same time, low computational resources. For these reasons, the model of the system has been analyzed in depth and specific solutions have been adopted to cope with the different sources of error. Furthermore, in order to assess the accuracy of our method with the highest precision, we have built an experimental test bed using a Numerical Control Unit (NCU) having $0.06°$ of accuracy. This

type of experimental assessment is necessary in these contexts, being more consistent than the other approaches usually employed, based on the well known pixel *re-projection error* [20]. Experiments on real world images proves the excellent results achieved by our approach.

It is worth remarking that our approach is the first attempt to exploit the Earth as a native target to be tracked in a satellite navigation system, using an on-board monocular camera to recover the satellite attitude. Also, our method only relies on "corner" points naturally present in the image and does not require any prior knowledge of specific targets to be tracked (like in [1]).

This work is organized as follows. Section 2 deals with the other approach utilized to face the problem. The geometric model relating different views and our registration algorithm are described in Section 3, as well as the reference frames adopted. In Section 4 some results concerning experiments with synthetic and real world images are shown, together with some important remarks. Finally, Section 5 draws conclusions and considerations about future developments.

2. Previous work

The core issue of designing autonomous systems for the automatic guidance of remote unmanned devices is represented by a pose estimation problem. The most recent works report basically two different approaches. The first one requires the integration of information provided by an ensemble of sensors, mounted on board or external (*e.g.* [17], right for satellite applications). We follow the second approach referring to visual ego-motion estimation [18], that in turns subdivides into the main fields of Structure From Motion (SFM) and Simultaneous Localization and Mapping (SLAM).

In [11], ego-pose estimation for unmanned ground vehicles is performed using an on-board stereo rig. The authors use EM over motion hypothesis to provide high robustness, but the accuracy of ground-truth data (provided by DGPS) is not satisfactory, besides being quite time consuming for many real time applications. In [6], an efficient stereo algorithm is proposed for the estimation of position and attitude of Mars Exploration Rovers. Terrain stereo pairs acquired on field are processed and results are compared with the ground-truth data obtained by integrating gyro data gathered during vehicle motion. This yields an indirect measure of the attitude accuracy, going not below than sub-degree (decimal) order. In [10], sequences of monocular aerial images acquired in real time are compared with a geo-referenced Earth image database to estimate position and velocity of the aircraft. This also helps to reduce drift errors in position estimation. However, although the author states to recover also the attitude, no results are reported. In addition, loading and processing in real time from a stor-

age device large images from a database could be too computing intensive with respect to the resources available on board of a satellite. Authors in [15] use an approach based on homography estimation of piecewise planar landscapes in UAVs navigation, starting from a geometrical framework similar to the one employed in this paper. Although few details are given regarding feature points extraction and robust estimation of homography between the views, test performed with manually placed synthetic planar patterns show an accuracy in attitude estimation of some degrees.

SLAM approaches were born to be used in robotics and indoor environments to cope with dead reckoning effects for long looping image sequences [14]. As a matter of fact, the application of SLAM to large outdoor environments is challenging, also due to the large need of computational power. In addition, although been effective to reduce error accumulation in looping paths, it suffers from a lack in the accuracy of the attitude estimation. In [16], a SLAM approach is successfully used to reduce drift errors in a looping path in an aerial flight. As for the attitude results, although they are promising, the maximum error during the flight can be of the order of several degrees. A joint approach is presented in [5], where again SLAM is employed just to reduce error drift in positioning, while both attitude and position estimations are provided using a SFM algorithm. On-field experiments are carried out by using inertial sensors to provide ground truth data with accuracy of $0.5°$, resulting in a measurement error for the attitude of the order of degrees.

As far as the satellite application is concerned, there have been some attempts to use a CCD camera to monitor position and satellite attitude. In [8], the relative pose of a reference (the "mother") satellite is recovered by tracking visible markers affixed at known positions on a "slave" satellite. The attitude accuracy has been measured in the laboratory prototype using a robotized camera providing ground truth data and it is of the order of some degrees.

3. The method

The control of the attitude is a crucial task to be performed in satellite missions. A three-axis orientation feedback has to be provided continuously to a closed loop control system in order to guarantee the correct orientation of the spacecraft with a very high accuracy (up to tens of arcsec for the state-of-the-art technology). In order to reach such a high accuracy, we can exploit the knowledge of the orbital position - it can be externally provided by satellite position sensors with an accuracy compliant with our purposes - since it can enforce constraints on the attitude estimation [20].

In the following we will refer to a nadir pointing (always directed toward the Earth) satellite moving along circular orbits around the Earth. It is supposed to be equipped with a CCD camera integral with the satellite and nadir point-

ing as well. A roll-pitch-yaw displacement with respect to a nominal attitude have to be estimated in real time by processing the terrestrial images acquired by the CCD sensor along the orbit.

The dynamical model adopted is discrete-time, that is position and attitude changes at each *epoch* along the orbit. The image at the generic epoch is registered with the view at the previous epoch. An image registration algorithm matches corresponding points on each couple of different views of the scene - Frame to Frame (F2F) registration - and find the geometrical transformation that links them. Attitude parameters are then extracted from the estimated transformation according to the dynamical model of the image formation due to the movements of the satellite along the orbit.

In the following subsections the reference frames involved, the geometrical model adopted and the image registration algorithm devised are described in details.

3.1. Reference frames

In our system six reference frames are adopted to describe the problem:

- the Earth-Centered Inertial reference frame (ECI);

- the Earth-Centered Earth-Fixed reference frame (ECEF);

- the Local Orbital reference frame (LORF);

- the Satellite (body) reference frame (SRF);

- the Camera reference frame (CRF).

- the Image reference frame (IRF).

The first four reference frames are commonly used in the aerospace field [12]. ECI reference frame is a quasi-inertial geo-centered system having the x-axis pointing towards the vernal equinox, the z-axis along the rotation axis of the Earth and the y-axis to form a right-handed triad. ECEF and ECI reference frames are equivalent up to the Earth rotation, and we refer to them as the *world* reference frame afterwards. LORF has its origin in the centre of mass of the satellite, the z-axis along the straight line joining the Earth centre and the satellite centre pointing toward the satellite, the y-axis perpendicular to the satellite orbit plane in the direction that sees the satellite rotates in the counter-clockwise way and the x-axis such that the system forms a right-handed triad – in case of circular orbit the x-axis points toward the satellite velocity vector.
SRF is integral with the satellite, with its origin in the satellite centre of mass and axes nominally aligned with LORF. The displacement in orientation from LORF (the nominal attitude) is measured as roll, pitch and yaw angles.

CRF is integral with SRF, with the optical centre of the camera (modeled as a full projective pinhole camera) coinciding with the satellite centre of mass (or being at a known displacement, without loss of generality).
Finally, IRF is related to CRF through the well-known matrix of the camera intrinsic parameters.

3.2. The dynamical geometrical model

In general, it can be proved [21] that two views of the same scene under general camera rigid motions are linked through the well known Eq. 1:

$$\begin{bmatrix} x' \\ y' \\ 1' \end{bmatrix} = KR'R^{-1}K^{-1} \left(\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} - \frac{KR(C'-C)}{Z_{cam}} \right) \quad (1)$$

where $[x \ y \ 1]^T$ and $[x' \ y' \ 1]^T$ represent the homogeneous coordinates of the corresponding points in the previous epoch t and in the current epoch t', respectively. K is the matrix of the intrinsic parameters of the camera. The rotation matrices of CRF with respect to the *world* (scene) reference frame, that is the *extrinsic* camera parameters, are represented by R' and R in the two epochs, respectively. The symbols C and C' stand for the three dimensional coordinates of the camera optical centre expressed in the *world* reference frame in the first and in the second epoch, while Z_{cam} is the component along the z-axis of CRF of each of the projection rays of the image points in the first epoch. The ratio in Eq. 1 forms the *parallax shift* term which must be subtracted from the coordinates of the first image reference frame. The difference in camera optical centre due to the translation of the satellite is responsible for the *parallax effect*, whereas the term Z_{cam} takes into account the three-dimensional structure of the scene. Accordingly, after the *parallax shift* term compensation, corresponding points in the two image reference frames are related through a plane homography H retaining the attitude variation $R'R^{-1}$ (Eq. 2):

$$H = \lambda \, KR'R^{-1}K^{-1}, \ \lambda \neq 0 \quad (2)$$

The Earth is modeled as a sphere of known radius ($R_E = 6371 \ km$), so that given satellite position and orientation it is possible to compute Z_{cam} for every pixel of the camera by intersecting the visual rays with the terrestrial sphere.

The parameter λ can be retrieved from Eq. 2 once H has been estimated, yielding Eq. 3:

$$det(H) = \lambda^3 det(KR'R^TK^{-1}) = \lambda^3 det(K)det(K^{-1})$$
$$= \lambda^3$$

$$(3)$$

The rigid-body motion group is denoted as usual as $SE(3)$ and the group of three dimensional rotation $SO(3)$

will be described using unit quaternions [13]. We have expressly chosen this representation because of the specific advantages it carries. In fact, unit quaternions have a group structure (S^3 from now onward), they are closely related to the intuitive axis-angle representation of rotations and the composition of rotations is performed by 4-components vector products, resulting in a more computational advantageous and numerically stable solution. Attitude estimation from noisy data usually yields non-orthonormal rotation matrices, which must be somehow re-orthogonalized, while quaternions need only to be normalized to the unit vector. These properties face the numerical effects introduced by the model, thus representing the best choice [12].

Unit quaternions q are related to the rotation matrix through an homomorphism $R(q) : S^3 \rightarrow SO(3)$. Usually, In satellite dynamics the rotation matrix associated with the satellite attitude quaternion satisfies the relation (Eq. 4):

$$\left(\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} - \begin{bmatrix} X_C \\ Y_C \\ Z_C \end{bmatrix} \right)_{ECI} = R(q) \begin{bmatrix} X_{Sat} \\ Y_{Sat} \\ Z_{Sat} \end{bmatrix} \qquad (4)$$

thus referring the satellite orientation to ECI.

By using the Direction Cosine Matrix $DCM(q) = R(q)^T$ and taking into account the Earth rotation velocity ω, after some passages Eq. 2 can be rewritten as Eq. 5:

$$H = \lambda \, K DCM(q(t')) R_3(\omega(t'-t))^T DCM(q(t))^T K^{-1} \qquad (5)$$

Therefore, if in Eq. 6 we set :

$$R(\Delta q)(= DCM(\Delta q)^T) \dot{=} \frac{1}{\lambda}(K^{-1}HK)^T \qquad (6)$$

then the *relative orientation quaternion* Δq can be computed by the estimated homography H [24]. Accordingly, starting from the estimation of the previous attitude quaternion $q(t)$, the current *absolute* attitude quaternion can be estimated using the relation expressed in Eq. 7:

$$q(t') = q_\omega(\omega(t'-t)) \circ q(t) \circ \Delta q \qquad (7)$$

where ($\circ$) represents the canonical product between quaternions.

3.3. The registration algorithm

In order to recover the matrix H, *sparse* motion field measurements have been preferred to *dense optical flows* methods [19], essentially for computational reasons. The dynamical model adopted is discrete-time and consequently *image displacements* have to be measured. To perform such measurements with a sub pixel accuracy, the registration algorithm we conceived works by following a *coarse-to-fine* strategy. For each couple of satellite images, Shi and

Tomasi features descriptors [22] are extracted from the image $I(t)$ acquired at the first epoch. The global 2-D translational components of the image motion field ($\Delta x, \Delta y$) are mostly due to the movement of the satellite along the orbit and they are estimated with pixel accuracy using the *phase correlation* algorithm [9], over decimated images to speed up the process. This estimation is used as a *bootstrapping* phase to feed the Lukas-Kanade tracker in his pyramidal implementation [3], in order to measure the residual *local* motion field vector at a sub-pixel accuracy for each of the features. The two corresponding sets of points achieved in this way are processed. After the correction with the parallax term (see Eq. 2), using the DLT algorithm [21] in conjunction with the RANSAC [7] outlier rejection method, a robust estimation of the homography H is achieved.

It is worth remarking that this registration algorithm is quite general and reliable. In fact, it does not need a model of terrestrial landmarks to be learned in advance and tracked down along the image sequence (*e.g.* in [1]), yet conversely it can cope with general scene patterns, provided that a sufficient number of corners can be extracted. Furthermore, it is robust to minimal partial changes in the scene that can cause false matching (due, for example, to the quite uniform appearance of clouds), thanks to the statistical removal of outliers. We have chosen this implementation rather than other registration techniques, like for example SURF [2] in conjunction with other statistical estimators since anyhow it can provide the required accuracy with a lower computational cost, typical in satellite applications where embedded hardware platforms are often one-generation older.

3.4. Sources of error

Eq. 7 in Section 3.2 describes the dynamics of the propagation of the attitude estimates to next epochs. Different sources of error have been faced in our approach, some of them inherent to the model (4) or its implementation (3, to the testing process (1). In addition, having the ground truth for the attitude in our experiments has permitted to perform experiments nullifying the effects of error 2. The following is the list of the errors:

1. the *registration* error, due to the registration algorithm chosen to measure the image motion field, *i.e.* to obtain the set of corresponding points in two partially overlapping images;

2. the previous estimate R over the *parallax shift* term, as it results from Eq. 1, affects the computation of the homography H;

3. *numerical* errors, on the whole addressed with the choice of unit quaternions, as previously discussed in Section 3.2;

4. a *pure propagation* error, due to the term $q(t)$ in Eq. 7, *i.e.* the current quaternion is computed *starting* from the previous "wrong" attitude estimation;

The first three kinds of errors affect the estimation of the relative attitude term (Δq) in Eq. 7. The last source of error is intrinsic in the F2F strategy, thus it cannot be addressed without changing the model. Moreover, it is responsible for the accumulation of the overall system error, mainly when the (Δq) estimation error is biased, resulting in drift effects. For these reasons, an in-depth analysis has been carried out on the errors introduced by the image registration algorithm in order to cope with these issues, yet preserving accuracy and computational requirements.

The largest part of the *registration error* is correlated to the difference between the expected (according to ground-truth data) and the measured (using images' corner tracking) motion field. Actually, as we show in Section 4, feature tracking is affected by the process of image generation, in particular by the interpolation techniques. Accordingly, real scene images should not be affected by the bias in the estimation error.

The *parallax shift* term could yield drift, although its contribution is small for the magnitudes of the parameters involved. It has already been discussed in Section 3.2. Of course, this term is absent if the view position does not change over the two epochs.

4. Experimental results

Both synthetic and real world images experiments have been performed in order to test the accuracy and the robustness of our method and system, using *ground truth* data.

Simulated image sequences are extracted from a real satellite image database to have a realistic evaluation of the performances of our algorithm even in presence of different environmental patterns (clouds, rivers, etc.). Tests on real world images have been carried out in laboratory conditions, by using photographic views of an airborne scene. In this setup, the attitude of the camera can be controlled using a NCU.

The attitude retrieved at each epoch by our system is compared to ground-truth data provided by an orbital simulator, for the synthetic data test, and by the NCU, for the lab tests. The *attitude error quaternion* is defined in Eq. 8 as:

$$err_q(t') = (q_{est}(t'))^{-1} \circ q_{est}(t'); \qquad (8)$$

and converted in roll-pitch-yaw offsets to compute displacements from ground-truth data.

4.1. Simulations with synthetic images

Simulations on synthetic data have been devised following two steps.

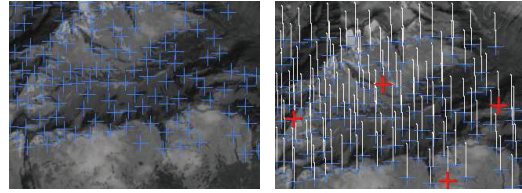| Sensor Dimensions	320x240 pixel
Pixel Size	$8\ \mu m$
Focal Length	$336.7\ mm$

Table 1. Synthetic parameters of the sensor

Figure 1. Couple of consecutive frames. *Left*: Previous frame with the extracted key-points. *Right*: Current frame with the tracked key-points. The big four crosses represent the 4-points subset selected by RANSAC.

Attitudes and positions of the satellite over epochs (that is, its *state vector*) are generated using a realistic orbital simulator developed in our University, for which the initial state is provided by the user.

A sequence of geo-referenced images are extracted from the WMS Landsat-7 ETM+ database, that meets the quality requirements of our project, by intersecting the camera's Field Of View (FOV) at the given state with the spherical model of the Earth. It is supposed to rotate with an angular velocity $\omega = 7.272 \cdot 10^{-5}\ rad/s$. In this way, the FOV can be geo-referenced and the correspondent view extracted from the database using image interpolation techniques.

For the adopted database, the ground resolution is $15\ m/pix$, correspondent to an instantaneous field of view of $4\ arcsec$ in geographic coordinates. The optical parameters of the (pinhole) camera are reported in Table 1. These values have been chosen so as to provide the ground resolution of the image database, besides being compatible with the devices commonly employed in the field of satellite imaging. In this simulation, lens distortion was not included. The first set of simulations (SET1Synth) covers a portion of a slightly perturbed low Earth *near polar* circular orbit, with a radius equal to $650\ km$. The satellite ground velocity is about $6.9\ km/s$, spanning on the area within the geographical coordinates $[44N, 8E] - [48N, 12E]$. The working frequency of the orbital simulator and our algorithm is $10Hz\ (10fps)$. In Figure 1, a couple of frames tracked along this orbit is shown.

Results for the frame to frame attitude error of the first 250 frames of this simulation are shown in Figure 2 (pitch results are not shown, since comparable to roll ones). It worths noticing that the accuracy is of the order of decimals of arcsec for roll (and pitch) angles while it is of tens of arcsec for the yaw angle, because of the far lower sensitivity of the image motion field to yaw (tilt) perturbations.
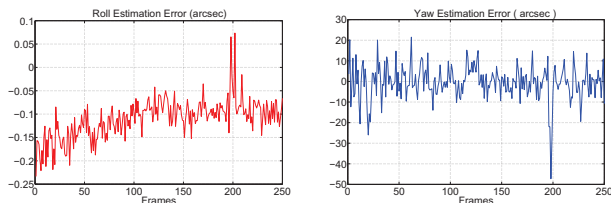
Figure 2. Trends of the F2F errors expressed in arcsec for SET1Synth orbit.
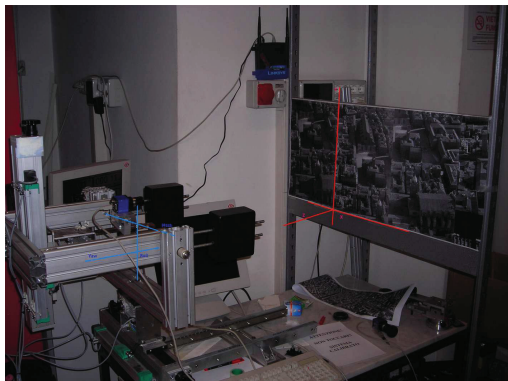

Figure 3. A picture of the NCU system together with the reference test image.

However, we realized that the estimation errors are not zero-mean (this is more appreciable for the roll error, due to the lower scale) and this bias can yield drift when the absolute orientation is computed and propagated to the next iteration. In particular, we have realized that for small attitude errors the contribution of the parallax shift term is low if compared with the main contribution due to the biasing of the *registration error*, for the values of the parameters involved in this simulation. After some investigations, we found out that this is due to a biased sub-pixel offset stemming from the interpolation necessary to generate the synthetic frames. We tested sequences generated using bilinear and Natural Cubic Spline (NCS) interpolation, the latter performing better. As expected, simulations on real world images, where interpolation is absent, are not affected by this kind of error.

4.2. Simulations with real world images

A CCD commercial camera has been mounted on a NCU, capable of performing three-axis rotations with an accuracy of $\sigma = 0.06°$, watching a printed airborne scene. The camera lens distortion and intrinsic parameters have been obtained once and for all after an off-line calibration process. Images acquired at a resolution of 1024×768 pixel are processed with our algorithm on a consumer PC (AMD2000+, 1.66GHz, 1GB RAM). Figure 3 shows the experimental test bed used. The camera optical axis is look-
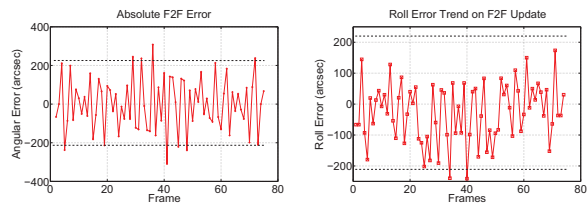

Figure 4. Absolute (left) and with update (right) F2F roll error measured between NCU ground-truth data and the image estimated attitude. Errors are in arcsec.

ing at the planar image of an airborne scene stitched onto a flat panel. The NCU allows to measure *ground-truth* attitude angles with respect to the NCU reference frame.

In this setting, a series of tests have been performed by varying angles along one axis. Thus, we can measure roll, pitch and yaw perturbations one at a time, leaving fixed the other two components. In order to compare the ground-truth data from the NCU with the measures expressed in terms of camera reference frame, it is necessary to align the two reference frames. An axis-angle notation (rotation vector) has been adopted accordingly to overcome the misalignment between the two reference frames. The first set of perturbations we present are roll (pan) perturbations. To nullify any parallax effect in the experimental setup, the camera has been arranged into the NCU so that its principal point belongs to the rotation axis of the NCU reference frame.

The first test sequence we present here consists of 75 frames acquired by varying roll perturbations angles within a range of $[-0.98°... + 13.15°]$ in a looping path with an angular step of about 30 hundredths degree. In Figure 4, the experimental absolute (left) and with update (right) errors measured between NCU ground-truth data and the estimated image attitude angles for the whole test sequence is shown together with the NCU uncertainty range (dotted lines). The F2F registration angular error almost always lays inside the NCU uncertainty range and it spreads across the whole range of accuracy provided by the NCU system. In particular, now the absolute F2F error has standard deviation $\sigma = 136.52$ and mean close to zero ($\mu = 0.4$ arcsec), thus being free of drift errors, as it can be seen in Figure 4, right, where the range of the cumulative errors keeps the same as in the left graph.

Results concerning a simulation of 68 frames for yaw perturbations are shown in Figure 5. The distances involved in this simulation does not affect dramatically the sensitivity of the image motion field, as conversely it happens for the simulation with synthetic images (see Figure 2, right). The camera rotates around the NCU yaw axis with a non-zero radius. Since the equipment built does not provide directly the position of the camera centre, it has been mechanically estimated with a rough accuracy. This is the reason for the biasing in the F2F absolute error in Figure 5, with a
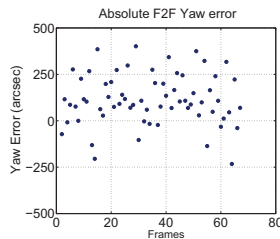
Figure 5. Absolute F2F yaw error measured between NCU ground-truth data and the image estimated attitude. Errors are in arcsec.

mean value of $\mu = 113.4$ arcsec and a standard deviation of $\sigma = 136.87$ arcsec, this being comparable with the value obtained for roll and pitch.

5. Conclusion and future works

This work describes the vision-based only approach we have implemented to achieve a real time high accuracy estimate of the attitude of an autonomous system, here a satellite. The camera, integral with the satellite, captures images of the Earth at different epochs along the orbit. A frame-to-frame registration approach is implemented exploiting natural corner points to track in couple of images. An accurate experimental stage on both synthetic and real world images has permitted to discover, analyze and possibly face different sources of errors that could even yield drift. The experimental instrumentation we utilized to built the laboratory prototype and to generate ground truth data for the orientation parameters has allowed us to achieve very reliable measures of the attitude as well as to deal with the different sources of errors. The high accuracy achieved constitutes a prove of concept for a vision-based stand alone control system when a high accuracy is required meanwhile using low power devices.

The improvement of the accuracy of the experimental instrumentation (at least, one order of magnitude) will permit to go deeper and deeper in the error analysis to improve the results even for the yaw perturbation. Including global registration using real time mosaicing is being considered, compatibly with the computational resources available on board, to increase the robustness of the system to error drift.

Acknowledgment

References

[1] A. A.Proctor, E. N.Johnson, and T. B.Apker. Vision-only control and guidance for aircraft. *Journal of Field Robotics*, 23(10):863–890, 2006.

[2] H. Bay, A. Ess, T.Tuytelaars, and L. Van Gool. Speeded-up robust features (SURF). *Computer Vision and Image Understanding archive*, 110(3):346–359, 2008.

[3] J. Y. Bouguet. Pyramidal implementation of the Lukas Kanade feature tracker: Description of the algorithm. *In Intel Research Laboratory, Technical Report*, 1999.

[4] O. Bourquardez and F. Chaumette. Visual servoing of an airplane for auto-landing. *Proceedings of the 2007 IEEE/RSI ICIRS, San Diego, CA, USA, Oct 29 - Nov 2, 2007*, pages 1314–1319, 2007.

[5] F. Caballero, L. Merino, J. Ferruz, and A. Ollero. Vision-based odometry and SLAM for medium and high altitude flying UAVs. *Journal of Intelligent and Robotic Systems*, 54:137–161, 2009.

[6] A. E.Johnson, S. B.Goldberg, Y.Cheng, and L. H.Matthies. Robust and efficient stereo feature tracking for visual odometry. *IEEE International Conference on Robotics and Automation(ICRA'08)*, pages 39–46, 2008.

[7] M. A. Fischler and R. C. Bolles. Random sample and consensus: A paradigm for model fitting with application to image analysis and automated cartography. *Comm. of the ACM*, 24(6):381–395, 1981.

[8] D. F.Malan and W. H.Steyn. Remote satellite position and pose estimation using monocular vision. *Small Satellite for Earth observation 5th International Symposium for IAA*, 2005.

[9] H. Foroosh, J. B.Zerubia, , and M. Berthod. Extension of phase correlation to subpixel registration. *IEEE Transactions on Image Processing*, 14(1):12–22, 2002.

[10] G.Conte and P.Doherty. An integrated uav navigation system based on aerial image matching. *Proceedings of the IEEE Aerospace Conference*, pages 1–10, 2008.

[11] G.Dubbelman, W. van der Mark, and F. C. A.Groen. Accurate and robust ego-motion estimation using expectation maximization. *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3914–3920, 2008.

[12] C. D. Hall. *Spacecraft Attitude Dynamics and Control*. February 2003.

[13] B. K. P. Horn. Closed form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America A*, 4(4):629–642, 1987.

[14] A. J.Davison, I. D.Reid, N. D.Molton, and O.Stasse. MonoSLAM: Real-time single camera SLAM. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(6):1052–1067, 2007.

[15] K. Kaiser, N. Gans, and W. Dixon. Localization and control of an aerial vehicle through chained, vision- based pose reconstruction. *American Control Conference*, pages 5934–5939, 2007.

[16] J. Kim and S. Sukkarieh. Real-time implementation of airborne inertial-slam. *Robotics and Autonomous Systems*, 55(1):62–71, 2007.

[17] X. Ning and J. Fang. An autonomous celestial navigation method for LEO satellite based on unscented Kalman filter and information fusion. *Aerospace Science and Technology*, 11:222–228, 2007.

[18] D. Nister, O. Naroditsky, and J. Bergen. Visual odometry for ground vehicle applications. *Journal of Field Robotics*, 23(1):3–20, 2006.

[19] R.Ben-Ari and N. Sochen. A general framework and new alignment criterion for dense optical flow. *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, 1:529–536, 2006.

[20] R.Carceroni, A.Kumar, and K.Daniilidis. Structure from motion with known camera positions. *IEEE International Conference on Computer Vision and Pattern Recognition(CVPR)*, 1:477–484, 2006.

[21] R.Hartley and A.Zisserman. *Multiple view geometry in computer vision*. Cambridge Academic Press, 2003.

[22] J. Shi and C.Tomasi. Good features to track. *IEEE International Conference on Computer Vision and Pattern Recognition*, pages 593–600, June 1994.

[23] S.Rathinam, P.Almeida, K.ZuWhan, S.Jackson, A.Tinka, W.Grossman, and R.Sengupta. Autonomous searching and tracking of a river using an uav. *American Control Conference 2007(ACC'07)*, pages 359–364, 2007.

[24] I. Y.Bar-Itzhack. New method for extracting the quaternion from a rotation matrix. *Journal of guidance, control, and dynamics*, 23(6):1085–1087, 2000.