FUSION OF ABSOLUTE VISION-BASED LOCALIZATION AND VISUAL ODOMETRY FOR SPACECRAFT PINPOINT LANDING

Bach Van Pham1, Simon Lacroix1, Michel Devy1, Thomas Voirin2, Marc Drieux3, and Clement Bourdarias3

1 CNRS ; LAAS ; 7 avenue du colonel Roche, F-31077 Toulouse, France
1 Université de Toulouse ; UPS ; INSA ; INP ; ISAE ; LAAS ; F-31077 Toulouse, France
2 European Space Agency, ESTEC - Keplerlaan 1 - P.O. Box 299 - 2200 AG Noordwijk ZH, The Netherlands
3 EADS-ASTRIUM, 66 Route de Verneuil, Les Mureaux Cedex , 78133, France

ABSTRACT

In this paper, we present the vision-based absolute navigation (VIBAN) system that enables the global localization ability (also called "pinpoint landing") of the lander. The system integrates the absolute visual localisation (AVL) approach Landstel with visual odometry (VO), a relative visual motion estimation approach. The combination of the two vision-based sensors has several advantages. Firstly, a higher accuracy and a better robustness in localization ability is obtained, as in a GPS-INS fusion system. Secondly, the AVL system can exploit the points tracked by visual odometry in order to robustify the system itself. Finally, these tracked points are also used according to a fault detection approach to verify and check the correctness of the AVL algorithm. Extensive simulations illustrate the proposed approach.

Key words: Landstel; pinpoint landing; fault detection.

1. INTRODUCTION

Future space missions such as MoonNext require a precise landing position, far from obstacles on the planet surface and close to different scientifically interesting sites. However, the current Entry, Descent and Landing technologies are still far from this ability, and various solutions have been introduced to solve the “pinpoint landing” problem, i.e. precise spacecraft localization with respect to a known reference.

Approaches using a camera as the primary sensor include SIFT feature matching [11], crater detection and matching [1], optical flow [4], template matching [7] or the landmark constellation matching approach (“Landstel” [9]) – see an overview in [5]. The common principle of these vision based techniques is to extract surface landmarks on every descent image and to match them to landmarks extracted from the ortho-rectified image of the landing area previously acquired (e.g. by an orbiter). These matched points are then used to estimate the global position of the lander thanks to the Digital Elevation Map (DEM) associated to the ortho-image.

With respect to inertial navigation, vision-based approaches provide a significant improvement in solving the pinpoint landing problem. Nevertheless, most of the current work only presents results obtained with a standalone AVL function. A loose coupling mechanism between an AVL sensor and an inertial sensor (INS) is presented in [10]. The work of Mourikis et al [7] called VISI-NAV also describes a tight integration between the AVL method based on persistent features, and a visual odometry method using opportunistic features.

The goal of this article is to present a full integration of visual odometry and the AVL approach “Landstel”, previously depicted in [9], within the VIBAN system. In this system, positions estimated by the two functions are fused, and the feature points detected and tracked by visual odometry are also used to ease the AVL function. The first advantage is to provide a solution which is more precise, faster and more robust. The second advantage is that points tracked by visual odometry are also exploited to enhance the Landstel approach and to detect false estimates.

Outline: the next section presents the VIBAN system architecture in which the Landstel and the visual odometry methods are integrated, via the Consistency Checking module and the Global Navigation filter. The various steps involved in the Landstel algorithm are also briefly depicted in the same section. Section III introduces the approach applied by the Global Navigation filter, in order to fuse the absolute position estimated by Landstel with the relative motion estimated by the visual odometry (or by an inertial navigation sensor). Section IV firstly presents the vision-based fault detection principle. Then, it introduces how to exploit points tracked by visual odometry in order to robustify Landstel. Finally, validation results obtained with the VIBAN system in a Lunar landing scenario are presented in section V.
2. GLOBAL ARCHITECTURE

Figure 1 presents the VIBAN system architecture with two main components: the visual odometry and the Landstel functions. The outputs of visual odometry are firstly used to enhance the global position estimation provided by Landstel via the Global Navigation filter. Then, these outputs are also used to make the integrated system more robust, by directly feeding this information into Landstel. Finally, the visual odometry outputs are used to verify the output of the Landstel function in the Consistency Checking module.

2.1. Visual Odometry

The first visual odometry ever to be used in a spatial system is the Descent Image Motion Estimation System (DIMES) [2]. The sensor was implemented on the Mars Exploration Rovers (MER) in 2003 and tracked four corner points through three sequential images during the Mars descent to estimate the spacecraft’s ground-relative horizontal velocity.

Another visual odometry (the “NPAL” camera) is described in [3]. The system is based on an EKF filter in which the positions of features are included in the state vector. The output of the system is firstly the relative spacecraft movement in position and altitude and secondly, the list and the positions of tracked landmarks through a series of descent images.

Note however that the visual odometry used in this article is simulated using Sift points matches [6]. Given two consecutive images, Sift features of each image are extracted and compared with each other, which gives a list of matched points between these two images. The step is repeated from the first till the last image in the descent series to form a list of tracked landmarks.

2.2. Landstel

The Landstel module is composed of one off line and one on line functions. In the off line one, the DEM and the associated 2D ortho-image of the foreseen landing area are built from Orbiter images. Initial visual landmarks are then extracted in the ortho-image (further denoted as the “geo image”), using the Difference of Gaussian (DoG) feature points detector [6]. A signature is defined for each extracted feature point. The initial landmarks 2D positions, their signatures and their 3D absolute co-ordinates on the planet surface constitute a database stored in the lander memory before launch.

The on line function of the Landstel algorithm consists in 5 steps (figure 2). The first and second steps extract and transform the information in the descent image so that the similarity of the geometric repartition between the descent landmarks and the initial landmarks is maximized (using an homography computed from the altimetry and orientation information). Then, the third step allows to extract the signature of each descent landmark. The extracted signature of each descent landmark is compared with the initial landmarks signatures (step 4): a list of match candidates from the initial landmarks set is associated to each descent landmark. In the last step, a voting scheme is applied to access the correct matches: several affine transformations are extracted within the potential candidate list, and the best affine transformation (the one supported by the highest number of matches) is used to generate other matches between descent landmarks and initial ones.

Images on figure 3 show two examples of the final matches obtained by Landstel with different illumination conditions.
Figure 3. Two examples of matches provided by Landstel under different illumination conditions. The geo image (left) is acquired with 55° − 25° (azimuth-elevation) sun position, whereas the descent images (right) are acquired at 5710 m altitude with 145° − 10° Sun position (a) and at 3052 m altitude with 235° − 10° Sun position (b). Zoomed areas of the geo image (center) show the corresponding matched regions of the descent images in the geo one.

3. FUSION OF GLOBAL AND RELATIVE POSITION INFORMATION

The main objective of the fusion between the Landstel absolute position estimate and the visual odometry relative motion estimate, is to yield a more precise absolute estimation of the spacecraft’s position. This is achieved thanks to the complementary filter implemented in the Global Navigation filter (figure 4).

Moreover, the integration of the estimated motion provided by visual odometry, can allow Landstel to focus the match search within a specific region of the geo image instead of searching in the whole landing area. This focusing mechanism not only accelerates the algorithm by reducing the search area, but also improves the algorithm’s performance by limiting the false matches probability.

The Kalman filter implemented in the Global Navigation filter (figure 4) is defined as follows:

1. System state:

\[ x = [\delta \Psi^T \delta v^T \delta p^T] \]

where \( \delta \Psi \) is the system attitude’s error, \( \delta v \) the system speed’s error and \( \delta p \) the system position’s error. Each term is a 3-dimensional vector (3 × 1).

2. Transition matrix:

\[ \Phi_k = \begin{bmatrix} -\Omega_{ie}^e & 0_3 \nabla \times & -2\Omega_{ie}^e \\ 0_3 & I_3 & 0_3 \end{bmatrix} \]

where the \( 0_3 \) and \( I_3 \) symbols respectively denote the 3 × 3 null matrix and the 3 × 3 identity matrix, and \( \Omega_{ie}^e \) is a skew-symmetric matrix which represents the planet rotation given by the angular rate \( \omega_{ie}^e = [\omega_1, \omega_2, \omega_3] \) between the planet-centered inertial frame (i-frame) and the planet-centered planet-fixed frame (e-frame):

\[ \Omega_{ie}^e = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix} \]

and \( \Upsilon = -\Omega_{ie}^e \Omega_{ie}^e - \Gamma^e \), \( \Gamma^e \) being the short notation for the gravity gradient (the derivative of the gravity). The \( \nabla \times \) represents the misalignment of the transformation matrix between the i-frame and the e-frame.

3. Observation: the observation provided by Landstel is the spacecraft position \( p \) [9]:

\[ z_k = \begin{bmatrix} 0_{2 \times 3} \end{bmatrix} \]

and the observation matrix is:

\[ H_k = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \]

Figure 4. The Landstel-INS fusion principle.
The Kalman filter estimates the error in the current estimated position stored in the Global Navigation filter and the uncertainty of the estimated global position, provided by the covariance matrix $P_{k|k−1}$.

4. FUSION OF TRACKED POINTS

4.1. Vision-based fault detection

Figure 5. This figure illustrates the use of the camera projection function to predict the future positions of the Landstel matched points. With Landstel, the 3D surface points $D_{LS}$ are associated to each matched points $M$ in the descent image. Then, the lander position $P^t$ at time $t$ is estimated as $P^t_{LS}$. Using the motion estimation $δP^t_{INS}$ provided by visual odometry, the next position $P^t+1_{LS}$ is predicted. On the basis of $P^t_{LS}$ and $D_1$, $D_2$, $D_0$, the predicted positions of $M_1$, $M_2$, $M_3$ at time $t+1$ are calculated.

The role of the vision-based fault detection in the Consistency Checking module is to verify the correctness of the Landstel global position estimation. The fault detection function is based on the comparison between the observation set and the prediction set of the matched points set $M$.

- **Prediction Set:** using the estimated position returned by Landstel $P^t_{LS}$ and the motion estimation between $t$ and $t+1$ given by visual odometry, the position of the lander at $t+1$ is predicted as $P^t+1_{LS} = P^t_{LS} + δP^t+1_{INS}$. As illustrated in figure 5, each point $M_k$ in the set $M^t$, is associated to a 3D point $D_k$ on the surface. The set of 3D points matched by Landstel at time $t$ is named $D_{LS}$. Given the predicted position $P^t+1_{LS}$ and the group of 3D points $D_{LS}$, the positions of the $M^t$ points in the next image are predicted using the back projection function. The image (b) in figure 6 illustrates the predicted positions of the matched points at $t$ in the image acquired at $t+1$ (in red). The prediction set is calculated with:

$$\text{Pre}(M^t) = \text{BackProj}(M^t, P^t+1_{LS}, D_{LS})$$

$$= \text{BackProj}(M^t, P^t_{LS} + δP^t+1_{INS}, D_{LS})$$  \hspace{1cm} (7)

- **Observation Set:** given a set of the matched points $M^t$ returned by Landstel at time $t$, the set point is tracked to the next image at time $t+1$ with visual odometry. The left image in figure 6 shows a subset of the matched points returned by the Landstel algorithm (in blue). The right image shows the tracked points of the matched points in the next image (in green). The observation set is thus obtained via:

$$\text{Obs}(M^t) = \text{Track}(M^t, im^t, im^{t+1})$$  \hspace{1cm} (8)

Using the observation function in equation 8 and the prediction function in equation 7 (illustrated by the green and red points in figure 6), the Landstel output is considered as being incorrect if the prediction set is inconsistent with the observation set. Indeed, the observation set depends only on the nature of the two consecutive images or more precisely on the functionality of the visual odometry which is considered as highly precise and reliable at short term. In contrast, the prediction function does not exploit the content of the two images. In fact, the prediction set depends only on the output of the Landstel algorithm, i.e. the estimated position and the associated 3D points. The motion estimation provided by the visual odometry is also considered as being precise in comparison with the Landstel error during a short period of one second (Landstel is operating at 1 Hz).

Let $ξ$ the difference or the innovation between the observation set and the prediction set:

$$ξ(M^t) = \text{Obs}(M^t) - \text{Pre}(M^t)$$  \hspace{1cm} (9)

The normalized value $α$ of the innovation vector $ξ(M^t)$ is calculated with:

$$α = \frac{ξ(M^t)^T * P(M^t)^{-1} * ξ(M^t)}{\text{trace}(P(M^t))}$$  \hspace{1cm} (10)

where $P(M^t)$ is the covariance of the matched points $M^t$. The covariance value here indicates the precision of the Landstel matched points location, which depends on the parameters used in the Landstel algorithm and is calculated off line. The value $α$ is used to compare with a predetermined threshold to classify the Landstel output. If the $α$ value is bigger than the threshold, the Landstel output will be reported as erroneous and will be discarded, and only the visual odometry information will be used to propagate the global position estimation — until Landstel retrieves an absolute position estimate.

Nevertheless, when the prediction set is consistent with the observation set, it does not mean that the Landstel output is correct. The Landstel output is however then considered as correct and is fed to the Global Navigation filter, hoping that subsequent steps will retrieve a potential inconsistency.

4.2. Landstel Robustification

The main purpose of the Landstel Robustification procedure is to increase the number of matched points returned
by Landstel, which consequently improves its precision. Given a set of matched points returned by Landstel at one instant (blue points in figure 6), the set is tracked to next image with the visual odometry (green points in figure 6). After having successfully tracked the matched points, the tracked points are injected into the affine candidates set of the Landstel algorithm (output of step 5). Then, the new candidate will be considered as the other normal candidates. Among these affine candidates, the one with the highest number of matches will be considered as the best candidate and will be used to calculate the position of the spacecraft.

5. RESULTS

5.1. Experiment setup

The purpose of these experiments is to analyse the VIBAN system sensitivity for Lunar landing application with the PANGU simulator [8]. In these experiments, the system sensitivity in regard to different parameters as well as the system performance analysis are carried on. Two simulated terrains are used for these experiments. The two simulated terrains used in this experiment. A cratered surface (named "NH") and a mountainous terrain (named "MB"), are shown in figure 7.

In these experiments, the system is employed during the Moon transfer orbit coast phase where the lander descends from 100 kilometres down to 15 kilometres altitude as shown in figure 8. The lander traverses 5000 kilometres during 1 hour. Due to the long flight time, the system is only tested at three altitudes which are at 100 kilometres (point A), 58 kilometres (point B) and 29 kilometres (point C). For each altitude, the system is employed for a time lapse of 43 seconds with 1 Hz frequency. Therefore, there are 43 images taken for each lapse (also called "trial"). The orbital image resolution used for the two points A and B is 160 m/pixel while that for the point C is 80 m/pixel.

In order to verify the robustness of the system with different levels of sensor noise, the experiments are set up with the following configuration (besides the nominal parameters described in section 6.6):

1. Camera inclination angle: in this scenario, the embedded camera angle is about 50 degrees with respect to the horizon.
2. Image: white noise $\mathcal{N}(0, 0.05)$.
3. Radar altimeter: there is no radar altimeter used. However the lander altitude is known at the beginning using earth-based localization with an accuracy of 10 percent. Altitude information is then propagated by pure integration of IMU measurement and corrected by the VIBAN system.
4. Attitude noise: the INS attitude estimation is considered as being precise due to the usage of the visual odometry and the star tracker (standard deviation below 1° – the Landstel has however shown to be able to cope with 5° attitude angle errors [9]).
5. Velocity error: the INS velocity error is $[1, 1, 1.5]$ metres ($3\sigma$) in long track, cross track and radial axis.

6. Initial position uncertainty: the initial position uncertainty is set to 5 km for both the long track and the cross track for every point A, B and C (worst case scenario).

5.2. Analysis

5.2.1. Components performance analysis

In order to show the role of each component in the whole system, the different elements of the VIBAN system are analysed and validated one by one. In these tests, the altitude error is within 10 percent for each point A, B and C, while the long track and cross track error is within a range of 5 kilometres. The sun’s elevation is kept at 1 degree while the difference in azimuth angle between the descent image and orbital map varies from $\pm 45, \pm 90$ degrees to 180 degrees. In general, there are 5160 images (120 trials) used for these experiments. The cratered surface NH is used in this test.

Figure 9. Estimation result with the number of "false" and "correct" estimations and the number of images where the algorithm can not find matches (5160 images in total). The left chart shows the results obtained with Landstel in a standalone mode, the middle one shows the result obtained with visual odometry-Landstel position fusion. The right chart shows the results obtained with the vision-based fault detection and with the Landstel robustification.

Figure 9 illustrates the number of image (among 5160 images) where the algorithm can provide a correct estimation, a false estimation or where the algorithm cannot deliver a position estimation. As shown in the middle chart of the figure, the combination of the relative position estimation computed by the INS (or by a visual odometry) and of the global position estimation provided by Landstel shows better performance than Landstel in a standalone mode. In this case, the number of “false estimation” is decreased from 26 percent down to 6 percent thanks to the reduction in the geo image search area. The number of “no estimation” reduces from 29 percent down to 26 percent. In fact, the “no estimation” case is less dangerous than the “false estimation” since the lander can still navigate by integrating IMU measurements.

By comparing the right chart with the center one, the number of "false estimation" is reduced from 6 percent down to 4 percent which means that 38.8 percent of faults are detected. In this case, the number of the "correct estimation" increases to 70 percent from 68 percent.

As shown by these experiments, a full integration of the Landstel sensors with visual odometry yields a better robustness and accuracy. Therefore, the full system is used for the following analyses.

5.2.2. Estimation errors

Figures 10 and 11 show the position estimation error of the VIBAN system for the three start points A, B and C. The performance of the system at points A and B is equivalent: this is due to the fact that the orbital image resolution is the same for both case (160 m/pixel). At lower altitude (point C), the estimation is more precise since the orbital resolution is doubled. The final estimation error at the end of the trajectory starting at C is approximately 300 metres.

Figure 10. VIBAN performance with NH surface at 100, 58 and 29 km altitude.

Figure 11. VIBAN performance with mountainous terrain MB at 100, 58 and 29 km altitude.

However, with respect to the mountainous surface MB,
the performance of the system at 100 km altitude is better than that at 58 km and is equal to that at 29 km. The reason of this phenomenon is due to firstly the ratio between the lander altitude and the terrain height variation. In fact, the Landstel algorithm assumes that the surface is flat, since a homography is applied to the image to before establishing the matches. At small altitudes, the mountainous terrain MB violates this assumption. Secondly, at high altitude the camera perceives more terrain, which facilitates the landmark matching process.

5.2.3. Influence of the sun position

In this part, the VIBAN sensitivity to the sun parameters is analysed. The first experiments study the impact of the illumination difference between the orbital image and the landing image. The second experiments study the impact of the incidence angle between the camera orientation and the sun’s azimuth.

5.2.4. Lander-Orbital Images Illumination

Two scenarios are evaluated here. The first one analyses the illumination condition at the Moon poles: the cratered surface NH is more appropriate and is used in this test. The sun direction for the orbital image is kept at 5 degrees elevation and 0 degree azimuth while the sun elevation for the landing image is kept at 1 degree and the azimuth angles are set at 45, 90, 180, 270 (or -90) and 315 (or -45) degrees. The test is made with the image acquired at all of the three points A, B, C (100, 58 and 29 km altitude). Figure 12 shows the average estimation error with respect to different landing illumination conditions. The estimation error is minimized if the landing image illumination is identical to that of the orbital image. However, different illumination conditions have little impact on the system performance.

5.2.5. Lander/Sun Orientation

Here we analyse the incidence angle between the lander camera orientation and the sun direction. In this analysis, the cratered surface NH is used. The estimation error is calculated with the different incidence angle. Figure 14 illustrates the average estimation error with the different incidence angles, and shows that the VIBAN system perturbed when the sun is directly in front of the spacecraft. In this situation, the embedded camera can only perceive the craters rim. Therefore, the landing image content is poorer than with other incidence angles. Figure 14 also shows that there is no difference in the system performance with incidence angle smaller than 90 degrees. The final estimation error for the worst case is about 1500 metres (at 100 kilometres altitude) while the best final estimation error is about 1000 metres.

5.2.6. Nominal Scenarios

Contrary to the previous analyses where there is no connection between the three points A, B and C, the nominal scenario tries to analyse the VIBAN system within a whole coasting trajectory. The estimation error at the one of the two points B (respectively C) inherits from the
Figure 14. VIBAN estimation error with cratered surface NH at 100 kilometres altitude, using different sun/lander incidence angle.

VIBAN estimation applied for the trajectory starting at point A (respectively B).

Two scenarios are defined for this analysis (figure 15). In the first scenario, the lander follows the transition of the Moon dark and bright sides from the North Pole down to the South Pole. The sun/lander incidence angle is set to 45 degrees for the points A and C (corresponding to North and South Pole). For the second scenario, as illustrated in figure 15, the lander faces the sun at point A (i.e. 180 degrees sun incidence angle). At point C, the sun is directly behind the lander. The difference in illumination between the landing image and the orbital image varies from 45 degrees to 345 degrees. The initial position error at point A is set to (1.6; 1.6;0.1) km for long track, cross track and altitude.

Figure 16 shows the average estimation error of five runs with the two scenarios. Since at each point A, B and C, the Landstel sensor is used with a time lapse of 43 seconds, there are in total 129 estimation results. The navigation period using the INS or visual odometry between the two points A and B or B and C isn’t shown due to the large difference in time scale (129 seconds versus 2500 seconds). The relative navigation error results in high jumps in the estimation error at the 43rd second and at the 86th second. As explained above, the second scenario yields larger estimation errors at the beginning of the coasting phase due to the front sun situation. However, the difference in the estimation error at the end of point A is corrected at the beginning of point B. The final average estimation errors in magnitude for the two scenarios 1 and 2 are respectively 245 metres and 293 metres ([151 149 38] and [186 164 28] in long track, cross track and altitude). Trials with lower altitudes showed that VIBAN can yield smaller estimation errors.

6. CONCLUSIONS

In this paper, we have demonstrated the benefits of the coupling mechanism between an absolute vision-based localization algorithm (Landstel) and a visual odometry method. Similarly to an INS-GPS fusion problem, the fusion between the global estimation computed by the Landstel module and the local estimation provided by visual odometry, not only improves the localization precision but also enhances the system’s overall performance in term of speed and robustness. Besides these advantages, the use of the points tracked by visual odometry also robustifies the Landstel performances.

Moreover, using visual odometry in the system instead of conventional INS permits the introduction of a vision-based fault detection approach, which can prevent the system from using the incorrect information provided by the absolute visual localization method. In fact, this vision-based fault detection system can not only be used with the Landstel function but can also be coupled with other vision-based global position estimation solutions, like the crater detection, the template matching or the optical correlation approaches.
REFERENCES


