ALPER: VISION BASED ABSOLUTE LOCALIZATION FOR PLANETARY EXPLORATION ROVERS

Loïc Le Cabec, Thierry Germa, Vincent Delort, Philémon Fieschi, and Emma Villanueva Rourera

Magellium, 1 rue Ariane, 31520 Ramonville St Agne, FRANCE (firstname.surname@magellium.fr)

ABSTRACT

This paper presents the outcomes of the ALPER project executed by Magellium under the ESA's GSTP program. Throughout this project, an essential component for absolute localization in planetary exploration rovers has been conceptualized, developed, and evaluated in a representative Martian environment. In the context of extended traverse missions, where the sole reliance on relative localization methods inevitably leads to error accumulation, absolute localization algorithms hold paramount importance. To address this, three distinct algorithms were designed, each characterized by an increasing degree of autonomy – ranging from an operator-assisted approach to a fully autonomous one. These algorithms are based on three complementary strategies for matching orbital and rover data: 1) Operator-guided visual tie-point tracking; 2) Constellation matching through rock detection on rover DEM; 3) Dense image co-registration involving local ortho-mosaics and orbital orthoimages. The primary focus of this paper centers on the statistical assessment of the operational capabilities of these algorithms. Furthermore, their practical application was demonstrated during field trials conducted in the Bardenas Reales in July 2023.

Key words: Planetary Exploration, Vision based localisation, Absolute localisation.

1. INTRODUCTION

Over the past two decades, significant advancements have been made in enhancing the autonomy of space exploratory rovers, particularly through improvements in localization processes. In 1997, the Mars Sojourner rover of the Pathfinder mission demonstrated limited autonomy of only 2-3 meters per sol, relying solely on wheel encoders and a sun sensor. The Mars Exploration Rovers (MER) Spirit and Opportunity (2004-2018) improved their autonomy by incorporating Visual Odometry (VO) and Inertial Measurement Unit (IMU) measurements to address potential slippages. The rover MSL Curiosity (2012-) has achieved remarkable navigation milestones, covering over 20 kilometers in 2450 sols, thanks to the integration of VO technology. However, these achievements, while commendable, still lag behind the advancements in terrestrial robotics. Notably, the utilization of absolute localization through GNSS has significantly expanded the autonomy of rovers and autonomous vehicles on Earth. The future aspirations of exploratory missions, like the Mars Sample Return (MRS) campaign, underscore the necessity for developing alternative techniques for extraterrestrial applications. Given the approximately 20-minute one-way signal delay to Mars, it's evident that non-autonomous rover navigation imposes substantial limitations on efficient exploration of the red planet. Hence, contemporary and forthcoming space rovers are designed with long traverses in mind. However, the current relative localization approaches pose challenges due to unbounded error growth during rover traversals. This work focuses on absolute localization systems, which are capable of rectifying dead-reckoning pose estimates as they progressively lose reliability. The main obstacles associated with this type of approach revolve around constrained computational resources and the substantial disparity in perspective between rover acquisitions and the orbital data utilized for their localization within a global frame of reference. Several complementary approaches were implemented and evaluated during ALPER to assess these challenges, providing an increasing level of autonomy. Their principal functions were benchmarked on a quad-core LEON4 processor, showcasing their applicability for embedded space applications. The manuscript is structured as follows: Section 2 categorizes and surveys various approaches that utilize information from aerial and orbital imagery to localize rovers on the Martian surface. Section 3 outlines a proposed assessment methodology for orbital-based localization techniques. The results of the Monte Carlo campaign conducted on a representative simulated environment and the live demonstration in the Bardenas Reales are presented in Section 4. Finally, Section 5 provides concluding remarks.

2. STATE OF THE ART

The main solutions that are being considered are map-based absolute localisations using HiRISE stereo pair images, the latter having a resolution of 25cm/px - 1m/px for the associated orthoimages and DTM. The idea is to match landmarks in the global map with the ones in a local map built by the rover's sensors. According to [1], the current techniques employed in the localisation of planetary exploration rovers based on orbital imaging can be categorized into three main classes of approach:

- Skyline matching
- Interest point matching
- Dense image/terrain matching

On-board sun sensors can also be used in complement to these methods to estimate the absolute orientation of the rover. Both MER, MSL as well as ExoMars integrate such capabilities. Although an accuracy of a few degrees can be achieved, long integration time is needed (from 15min to 1 h), and the estimation can be degraded when the sun is at high elevations [2]. Skyline Matching techniques do not achieve a sufficient level of accuracy for the intended application and are burdened by high disk storage usage [3].

2.1. Interest points matching

An interest point is a small region of an image/DTM with distinctive characteristics, such as high contrast/elevation. In the case of the Martian terrain, interest points can typically be rocks, craters, ridges or outcrops. Such features can be detected on rover acquisitions and orbital data. The basic principle of the interest points matching algorithms is to identify corresponding pairs of orbital and local features and to estimate the transformation between the two sets of features using methods such as least square fitting. The pose of the rover is directly computed from the result of this step. The main challenge inherent to this type of approach is thus to reliably find a correct match between an orbital and a local feature. Two different ways to tackle the issue are mentioned in the literature.

2.1.1. Manual pairing

The articles [4] and [5] discuss manual efforts undertaken by JPL to achieve absolute localization of the Curiosity rover within a reference framework. This involved the "geo-referencing" of local orthorectified panoramas in ArcMap using orbital images. This process entailed identifying corresponding pairs of landmarks in both the orthoimage to be "geo-referenced" and the reference orbital orthoimage. An affine transformation is then calculated using a least squares fitting algorithm. In [5], the authors routinely employed Spirit's PanCam and NavCam to generate DEMs and orthoimages of the rover's immediate surroundings. Various local landmarks like rocks, craters, outcrops, and ridges were manually pinpointed on these topographic outputs before being aligned with HiRISE orthoimages. The rover's position, acting as the imaging "center" on the Navcam orthoimage, could then be established on the HiRISE orthoimage.

2.1.2. Constellation matching

As opposed with the previous approach where every features are individually described, constellation matching search for similar feature constellations, where the spacing between features now effectively acts as the descriptor. Chen [7] developed a framework for this approach known as DARCES. This algorithm was used in numerous publications to address computational challenges and resolution and perspective disparity. In [8], Carle et al. presents an algorithms based on peak extraction from local 3D LiDAR maps. In [9], Hwangbo et al. proposes a similar approach using stereo cameras instead of LiDAR. The features are rocks extracted on the rover DEM based geometrical consideration (height of the peaks, area, etc). In the system introduced by Boukas, Gasteratos, and Visentin in [10], candidate ROIs are extracted from orbital and rover orthoimages based on entropy and Hessian matrix eigenvalues, and visual odometry and IMU fusion are used for relative localization to construct a local ROI networks. The latter is periodically paired with the global ROI network through the DARCES algorithm.

2.2. Dense image/terrain matching

The previous methods have to deal with the intricate problem of pairing points of interest. The first method struggles with automatic extraction of corresponding pairs of visual landmarks, necessitating manual intervention, while the second method relies on the presence of rocks to produce a pose estimate. The technique in this section aims to overcome these limitations by utilizing the entire information of onboard and orbital orthoimage for rover localization. This involves building a fused orthoimage mosaic along the rover's path and performing dense co-registration between this mosaic and orbital images. In [6], Y. Tao et al. detail a system that orthorectifies rover NavCam panoramas and constructs an orthomosaic by adjusting individual orthomosaics using a graph-SLAM approach (SIFT feature detection and matching, followed by bundle adjustment). The local orthomosaic and orbital orthoimage undergo preprocessing (resampling, Sobel filtering), and they are matched using Mutual Information within a gradient descent optimization scheme. A simpler approach is outlined in [11], where orthomosaics are generated solely with the rover's visual odometry (without bundle adjustment), using ZNCC as a similarity function. Promising results are demonstrated in real MSL data, both for estimation accuracy and execution time on LEON4 hardware. There are also some attempts to employ DTMs instead of orthoimages. In [11], authors obtained poor results using DTMs in the same system as with the orthoimages. Additionally, Van Pham et al. propose a system in [12] based on a particle filter estimation method to refine the rover position during traversal, achieving satisfactory outcomes, particularly in terrains with significant relief.

3. TECHNICAL DESCRIPTION

As previously discussed, no single method proves effective across all potential mission environments. Therefore, the ALPER project developed three distinct approaches based on complementary modalities. Considering the current state-of-the-art, the project opted to concentrate on interest point matching (manual pairing and constellation matching) as well as dense image matching techniques. Automated feature pairing, and terrain matching methods either lack the necessary reliability or fail to provide estimations meeting the required accuracy threshold (less than 5 HiRise pixels, equivalent to 1.25 meters). These three algorithms demonstrate increasing levels of autonomy. The initial algorithm (TPT) integrates and extends the manual interest point matching approach introduced in [4]. The second one (CM) employs constellation matching with a local network containing rocks extracted from a local DEM. For orbital rocks, manual selection by an operator is performed before the traverse. The final algorithm (DICOR) involves registering local orthomosaics with orbital orthoimages, requiring no operator intervention.

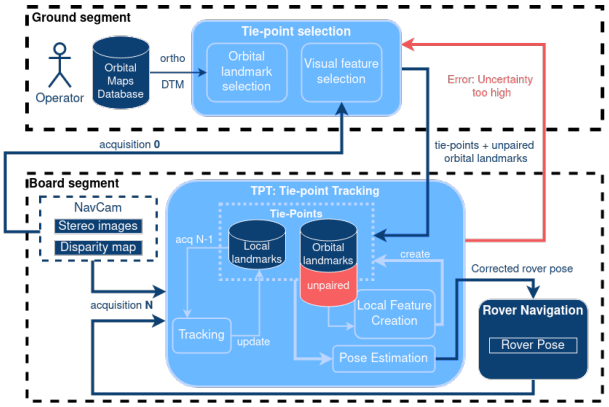

Figure 1. architecture of the TPT component

Figure 2. TPT - initial pose estimation. Left: rover images with visual features. Right: orbital orthoimage with tie-points.

3.1. TPT: Tie-Point Tracking

The Tie-Point Tracking (TPT) algorithm's principle is to identify and visually track local landmarks that were selected on HiRISE orbital images by an Operator, and to estimate the rover pose by computing a local-to-orbital transformation that minimises the positioning errors of these landmarks. It is composed of a ground segment and a board segment, as illustrated in Figure 1. The ground segment encompasses the user interfaces and the tools to select tie-points and isolated orbital landmarks. The board segment implements the core functions in charge of estimating the rover's absolute pose. It is composed of three main submodules: tracking, local feature creation and pose estimation.

Tracking: When a new acquisition is available, the Tracking component updates the position of the local visual features and associated local landmarks by locating them in the new rover image and disparity map. It is based on the algorithm described in [13], developed to track a visual target on MER. The algorithm takes into account the scale and illumination changes of the target by introducing a template image magnification factor and prior to a brute force window ZNCC matching with pyramidal image reduction. In TPT, a multi-target capability was introduced.

Pose estimation: At a given frequency, the Pose Estimation component computes the best transformation between the orbital and local landmarks via a least square fitting approach (Figure 2). The result of this points fitting step is used to produce an estimation of the 2D rover pose in an absolute frame of reference (x, y and yaw).

Local Feature Creation: When an unpaired orbital landmark enters in the rover field of view, a new local landmark is created. To do so, the orbital landmark is projected into the rover's last NavCam acquisition. The resulting pixel positions are used as starting position to detect a new visual features. Upon success, the new visual feature is projected back in the local reference frame using the disparity map to create a local landmark. The resulting tie-point is then ready to be tracked at the next iteration. The detection algorithm finds the most salient area in a search window based on contrast. The size and shape of the search window is adapted with the current rover pose uncertainty.

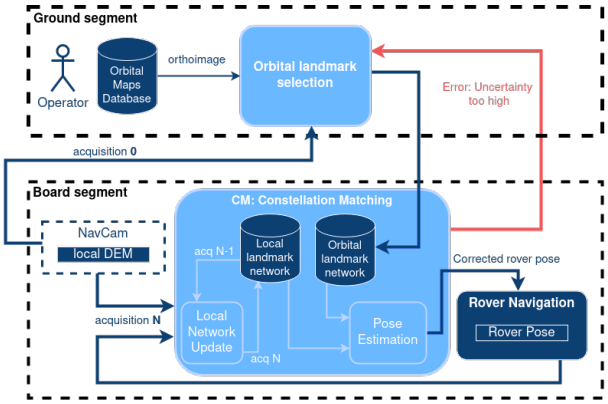

Figure 3. architecture of the CM component

3.2. CM: Constellation Matching

The core principle of the Constellation Matching (CM) system is to estimate the pose of the rover by matching an orbital landmark network – created by the operator-, with a local landmark network – build autonomously by the rover during its traverse. The considered landmarks are rocks that are both visible on the orbital orthoimages and on the rover DEM (size of 30cm-1m). The global architecture of the component is illustrated in Figure 3. The ground segment enables the operator to select the orbital landmarks. The board segment is composed of two main functions: local network update and pose estimation.

Local network update: When a new acquisition is available, the Local Network Update component updates the position of the existing local landmarks as well as their uncertainty relatively to the new rover position. New local landmarks are extracted from the rover DEM and landmarks with a two high relative uncertainty are remove from the local network. The **rock extraction process** (Figure 4) is the key element of the component. First, the local DEM is divided into roughly 1-2m cells using a grid pattern. Within each cell, a plane equation is computed using a RANSAC plane-fitting algorithm. A height map is filled with all points that deviate from this fitted plane (distance to the fitted plane). A binary mask is then computed on the entire grid by thresholding the height map and morphological operations are applied to filter out the outliers. Clusters are then identified on the mask and characterized (area, shape) to form areas denoting landmarks candidates. The candidates that meet the geometric requirements are added to the local network.

Pose Estimation: When the CM system is ready for new pose estimation (enough local landmarks detected since last matching), the Pose Estimation component matches the local and orbital landmark network and estimates the

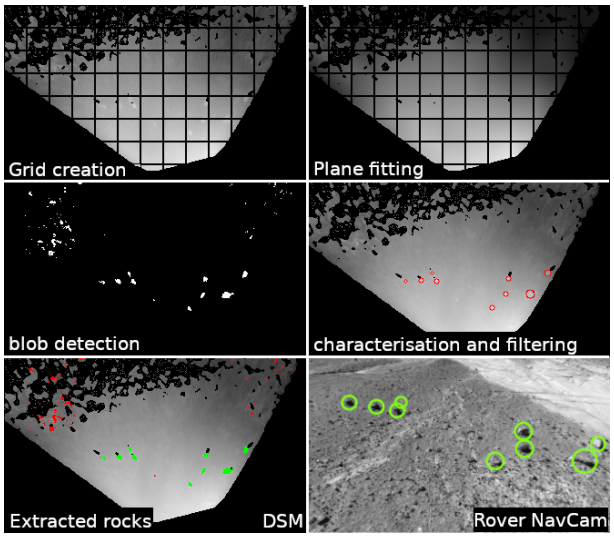

Figure 4. Rock extraction process

transformation between the two networks. Landmarks from the local network are updated based on the results of the matching step. The 2D pose of the rover in the global frame of reference is then estimated (x, y and yaw). If the current input pose is too uncertain. The networks are matched using a modified version of the DARCES **constellation matching** algorithm: A "triangle" of local landmarks is matched with a triplet of orbital landmarks taking into account the uncertainties of the local landmarks (both absolute and relative to the rover position). A transformation is computed from this match and the whole network is transformed to the same reference system as the orbital network. A matching score is computed using the entire local network. If the score is below a threshold, the match is a success, otherwise the process is repeated with another local feature "triangle", using a RANSAC-like approach.

3.3. DICOR: Dense Image Co-Registration

The Dense Image Co-Registration (DICOR) component estimate the pose of the rover by matching an orbital orthoimage with a local orthomosaic built autonomously by the rover along its traverse. Contrary to TPT and CM, it does not require any operator intervention and is thus only composed of a board segment. The global approach is similar to [11], i.e. mosaic built using only the input rover pose (no bundle adjustment) and matching done with ZNCC score function. The system is composed of three main functions: acquisition buffer update, orthomosaic building and pose estimation.

Acquisition buffer update: The acquisition buffer stores the rover acquisitions of the last 10-15 meters traveled, with a step of 1-1.5 meter. When a new acquisition is available, it is added to the buffer of local acquisition if the zone covered by the rover image is sufficiently different from the previous update. Acquisitions too distant from the current one get removed from the buffer. Consequently, the buffer consistently maintains about 10 acquisition entries.

Orthomosaic building: Before performing pose estimation, all acquisitions currently in the local acquisition buffer undergo orthorectification, pre-processing, and are then combined to create a local orthomosaic. These local orthoimages are generated at the orbital orthoimage resolution, typically 25cm per pixel for HiRISE images. The assembly process employs a straightforward blending technique. Orthoimages are positioned relative to each other based on the rover's input pose, and radiometric data from the most recent orthoimages is overlaid onto the mosaic when available. This approach ensures a fast orthomosaic assembly and prioritizes the incorporation of the most recent acquisitions during the matching process. The primary aim here is to maximize the influence of the most recent acquisitions in order to accurately estimate the rover's current pose, even in cases where there might be some relative localization drift.

Pose Estimation: Once the local orthomosaic is built, the local orthomosaic is matched with the orbital orthoimage and the 2D translation between them is estimated. The 2D position of the rover in the global frame of reference is then estimated based on this result. The matching process is based on a two-level pyramidal grid search approach. A sample-based ZNCC matching score is computed for each pose within the search window at the higher pyramid level. Several candidates are selected among the pixels with the best matching score. To ensure accuracy, a search at full resolution is then conducted within a small window surrounding these chosen candidates. This is necessary because the best score may correspond to a different local minimum between the downscale and full-scale search. In this context, a sample refers to a small rectangular portion of the local orthomosaic where the ZNCC score is computed. The global matching score is determined by averaging the scores of all these samples. These samples are positioned on the orthoimage using a simple grid scheme. This approach helps minimize the influence of local exposure or illumination artifacts that might be visible in the local orthoimages, such as reflections, shadows, radiometric variations between acquisitions, etc.

4. EVALUATION

4.1. Simulated environments

The simulator used during the project is based on Blender, which proved to be a powerful tool for generating highly detailed Martian-like terrain, allowing precise control over environmental characteristics. This capability is essential for Monte-Carlo campaigns where various terrain parameters, including rock density, soil texture, and illumination conditions, need meticulous adjustment to ensure statistical relevance. The simulator enables the creation of terrain at user-specified dimensions, manual modeling of relief, and the addition of procedural textures with two key advantages: preventing pattern repetition and incorporating realistic local relief. Objects like rocks and outcrops can be automatically placed on the terrain with adjustable size and density. The simulator also allows manipulation of the artificial sun's settings, trajectory, and internal characteristics. Users can generate orbital orthoimages and create custom trajectories (Figure 5). Ground truth and odometry files are automatically generated, and stereo-bench configurations can be customized for data acquisition. Finally, Magellium's algorithms are employed to generate disparity maps and

Table 1. Variability in Simulated Terrain Characteristics

Terrain characteristics	Variations		
sun inclination	20°	45°	zenith
Mean rock-to-rock distance	2m	4m	6m
Terrain relief	flat	hilly	cratered

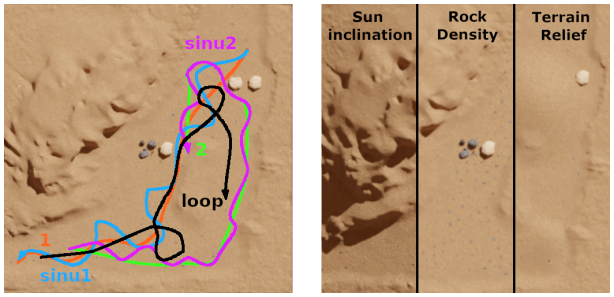

Figure 5. Simulated trajectories and terrains

DEMs from the acquired stereo images. With this simulator, 17 datasets were generated, each of them with different combination of trajectory characteristics, sun inclination, rock density and terrain relief. Different types of rocks populates the simulated environment, from large boulders to smaller gravel. Those impacted by the "mean distance between rocks" parameter fall within the size range of 20cm to 1m in width. The different terrain configurations are summarised in Table 1.

4.2. Monte Carlo evaluation campaign

The Monte-Carlo test campaign is designed to assess and validate the performance of localization algorithms, determining their operational capabilities across various scenario characteristics and noise levels in data inputs. It encompasses a substantial number of test datasets based on the simulated trajectories presented in the previous section, as well as some real case datasets taken from the Erfoud dataset [14]. Different noise level configurations are defined to artificially degrade algorithm input poses, simulating varying odometry quality in terms of drift and initial offset. Randomly generated noise rolls from these configurations are introduced during tests, with each noise roll applied to a random sub-segment of a trajectory constituting a test run. In the following section, the success rate is defined as the proportion of runs where the estimation error is less than or equal to 1.25 meters (equivalent to 5 HiRISE pixels). The estimation rate, on the other hand, represents the proportion of runs in which the algorithm conducted an estimation, regardless of whether it was correct or incorrect. In the simulated dataset, the step between acquisitions is approximately 30cm.

4.2.1. TPT

A run of the TPT Monte Carlo campaign is defined by the relief type, the sun inclination, the trajectory, the starting point on this trajectory and by the noise level of the input pose and of the operator selected visual features. The campaign is composed of 80,244 runs in total. The

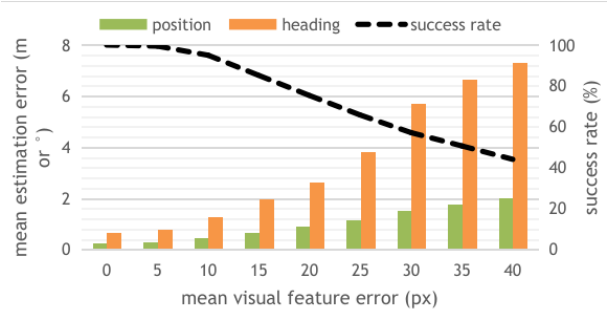

Figure 6. TPT: First estimation error vs. visual feature error

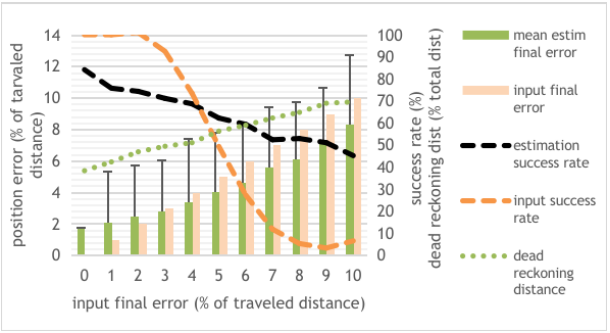

Figure 7. TPT: Estimation error vs. input final error

first part of the campaign focuses on the impact of several parameters on the initial pose estimation – such as number of tie-points, their distribution on the image and the accuracy of their selection by the operator. The initial phase assesses parameters impacting initial pose estimation (Figure 6). Reliable estimates require at least three tie-points, resulting in an average error reduction from 1 meter and 7.5 degrees with two tie-points to less than 0.5 meters and 1.5 degrees with three or more. Adding more than three tie-points doesn't significantly enhance results. Regarding visual feature noise, TPT delivers good results, even with an average error of up to 10-15 pixels on 640x512 images. This provides operators with a reasonable margin for feature selection.

The second part of the campaign focuses on the performances of the autonomous tracking and feature reprojection stage. During this phase, the input pose is assumed to have no initial error. If the tracking phase produces satisfactory results -regardless of the terrain characteristics-, the reprojection step falls short. Visual features are often inaccurately selected, leading to pose estimations not significantly better than the input pose. Figure 7 illustrates that with an accurate input pose (less than 2% error), TPT tends to worsen the estimation. However, when the input error exceeds 3%, TPT yields slightly better results. Notably, the success rate remains above 50% until the input error reaches 9%, compared to 5% without TPT. However, the dead reckoning distance increases, indicating that the algorithm becomes ineffective on average after 25 meters with a 2% input error, as opposed to 15 meters with a 9% input error.

4.2.2. CM

For the CM campaign, three noise sources were evaluated: the initial pose error, the relative error (both posi-

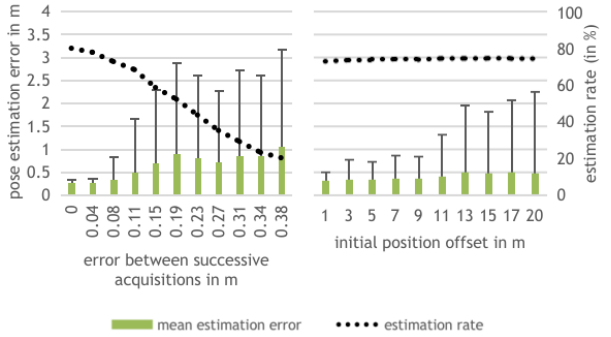

Figure 8. CM – Nominal conditions - pose estimation error vs. relative error between successive acquisitions and input inital error

tion and heading) between successive acquisitions. The first challenge is the algorithm's ability to provide a good pose estimation with a large search window, and the second put to test its robustness to errors in the construction of its local feature network. A total of 255,280 runs compose the full campaign. A run stops at the first pose estimation outputted by the algorithm. As can be seen on Figure 8, CM performs well, even with strong noise on input poses the estimated poses remain on average below 1m, regardless of the input noise. It is especially robust to initial pose offset: both the average error, the estimation and succes rate are not significantly impacted by the level of noise. Nevertheless, it can be noted that the increase of the error standard deviation (black 'T' on Figure 8) can be explained by the growing number of estimation with an error above 10m (up to 1% of the runs with an initial offset of 20m), while the success rate is always above 98% (c.f. Figure 9).

As it has been observed, the terrain characteristic that impact the most CM is logically the rock density. With a low rock density (less than one rock every 6m), the algorithm is less robust to strong input noise, with a success rate dropping to 86% with an error between successive acquisitions of 15cm and to 72% with an input offset of 15m. During the campaign, all runs were performed using the same set of parameters. In the case of low rock density, a better success rate could be achieved by employing more conservative parameters, albeit at the expense of the estimation rate. In contrast, for all other terrain characteristics, such as relief and sun inclination, CM gave satisfactory results, similar to those in the nominal case. It also showcased strong robustness against input heading errors. It consistently delivers a heading estimation error averaging less than 2 degrees, even when the input error ranges up to 20 degrees.

4.2.3. DICOR

Similarly to the CM campaign, the initial position error and the relative error between successive acquisitions are evaluated during the DICOR campaign. The relative error tests the robustness of the matching function to a distorted local orthomosaic. Contrary to CM, no noise was applied to the heading, since DICOR does not estimate it. The full campaign consist in 348015 runs. In the nominal case, DICOR performs well with various input noises (c.f. Figure 10). Specifically, the success rate remains

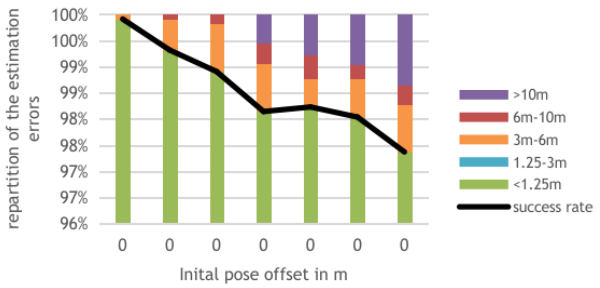

Figure 9. CM: Nominal conditions - Repartition of the pose estimation errors

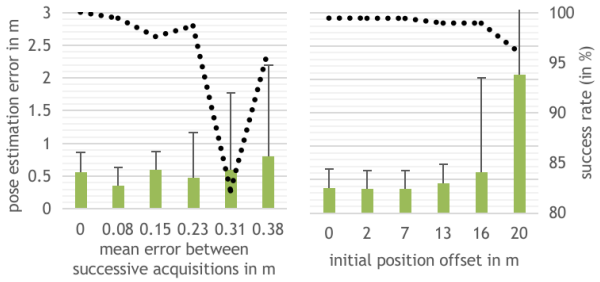

Figure 10. DICOR – Nominal conditions - pose estimation error vs. relative error between successive acquisitions and input initial error

above 98% for an initial position error of up to 16m and an error of 27cm between successive acquisitions (with a 30cm acquisition step). The slight decrease in the success rate observed with an input error of 31cm is primarily due to 18% of "small" estimation errors (less than 3 meters). On the other hand, the drop in the success rate associated with a 20-meter input offset is attributed to 3% of "large" estimation errors (exceeding 10 meters).

Contrary to CM, the sun inclination is the terrain characteristic that has the most negative impact on the performances of DICOR (c.f. Figure 11). This can be explained by the radiometric differences (illumination, shadows and sun reflection) between the rover images (taken with a sun inclination of 20° or at the zenith) and the orthoimages (taken with a sun inclination of 45°). A strong relief also degrade the results for a large initial offset, notably due to the fact that the orthoimages are shrinked because of the terrain obstacles, when compared to the nominal case.

It has to be noted that the results obtained in the Bardenas are overly pessimistic due to inaccuracies present in the ground truth data collected during the field testing, especially in the "orbital" orthoimage produced by the drone. However, they are in line with what could be expected from a terrain with relief variations and differences in sun inclination different between the orbital orthoimage and the rover images, which was the case here.

4.3. Field tests

Three sites with varying geological features were chosen in Bardenas Reales. A Parrot ANAFI Ai drone captured aerial images at three different times of the day: morning (with long shadows), when the sun was at its zenith, and at dusk (with no shadows). These images

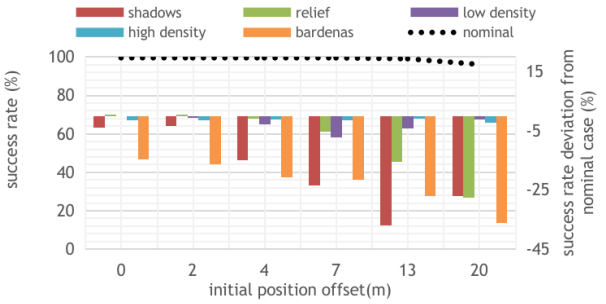

Figure 11. DICOR: success rate deviation from nominal scenario vs. initial position offset

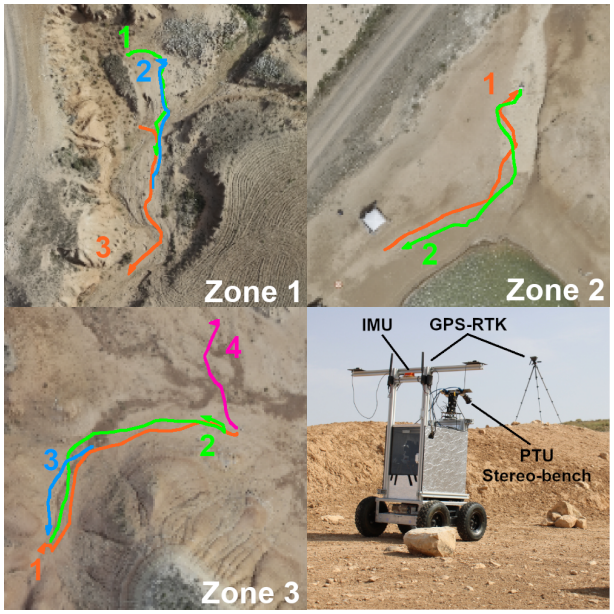

Figure 12. Bardenas field test trajectories and Magellium's robot

were used to create downsampled orthoimages at 25cm per pixel resolution and DTMs at 1m per pixel resolution. On each site, multiple short trajectories were recorded using Magellium's rover equipped with a PTU-mounted stereo-bench, IMU, and centimetre-accurate RTK GPS for precise ground truth positioning (c.f. Figure 12). In total, 19 trajectories were conducted, spanning distances from 15m to 40m, and involving two acquisition scenarios: 1) the rover, with a fixed PTU, acquired data every 50cm, and 2) at intervals of 1.5-2m, the rover captured three images to create panoramas with a 150° FOV. A live demonstration of the three algorithms (TPT, CM and DICOR) was conducted in the presence of an ESA observer in Zone 3.

4.3.1. TPT

As could be observed during the Monte-Carlo campaign, TPT exhibited precise initial pose estimation with an error of only 0.5 meters. However, it was challenging to determine a parameter set that struck the right balance between offering sufficient constraints for effective tracking and maintaining flexibility for feature reprojection. With the selected parameter set, which prioritized feature reprojection, TPT successfully provided 17 pose estimations over an 18-meter-long trajectory. These estimations

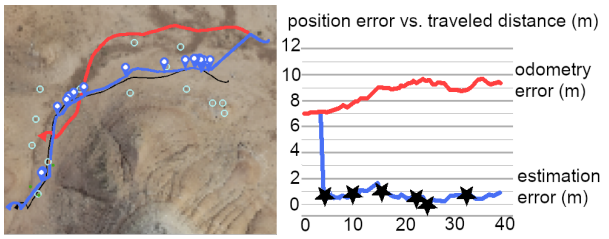

Figure 13. CM: results of the demonstration trajectory

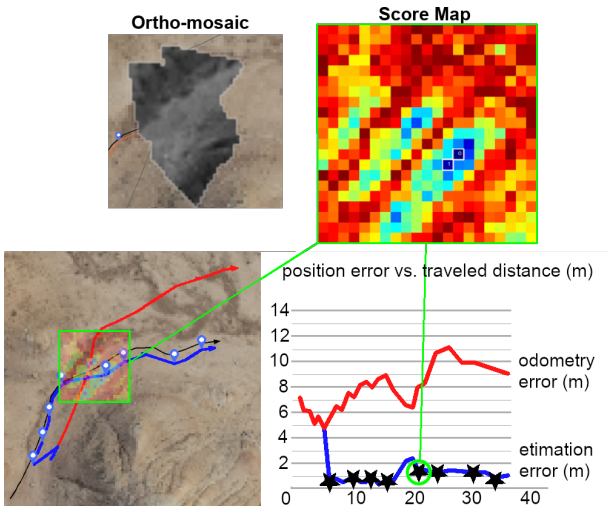

Figure 14. DICOR: results of the demonstration trajectory

had an average error of 0.77 meters and a maximum dead reckoning distance of 8 meters. Towards the end of the trajectory, TPT displayed an error of 3.2% relative to the traveled distance, outperforming odometry, which had an error of 5.5%.

4.3.2. CM

CM exhibited strong performance in the demonstration trajectory, achieving an average estimation error of 0.63m across 19 estimations. This was achieved with an input odometry error amounting to 6.6% of the traveled distance and an initial offset of 7 meters (with a maximum error of 9.8m). The evolution of the estimation error and the complete trajectory can be observed in Figure 13.

4.3.3. DICOR

DICOR performed well in the beginning of the trajectory with 4 estimations with an error bellow 2px of the orbital image. A series of 3 estimations with an error around 4-5px of the orbital image (1-1.3m) were produced in the middle of the trajectory. As can be seen in Figure 14, those degraded results can be explained by the presence of a strong and contrasted linear feature (i.e. a small river bed) in the rover images. This translates into a score map exhibiting a spread optimum region, instead of a well defined peak. This issue could easily be fixed by adding a score map analysis step in the overall process to reject such uncertain estimations. Overall, the estimation results are very satisfying, with an average estimation error of 0.79m for the entire 36m-long trajectory with an input odometry relative error between 7 and 15% of the traveled distance and an initial offset of 7m.

Table 2. Computation time of a CM cycle on the LEON4 processor

Main function	Mean time (ms)	%
Rock extraction	545	40.5
Grid creation	12	0.9
Plane fitting	511	38.9
Blob detection	13	1.0
Network update	10	0.7
Pose estimation	800	59.5
Full cycle	1345	100

4.4. Benchmark on Leon4

The benchmark's primary goal is to evaluate the computational efficiency of core algorithm functions using applicable hardware. This was achieved by integrating benchmarking capabilities into the CM library, which was then deployed on the GR-CPCI-GR740 development board. Notably, this board is equipped with a radiation-tolerant quad-core 32-bit LEON4FT processor. Results from the evaluation on synthetic data indicate that a full estimation cycle, excluding DEM computation, runs in slightly more than one second (c.f. Table 2).

5. CONCLUSION

The operator-guided visual tie-point tracking algorithm demonstrated robust and accurate localization performances as long as the selected visual features are visible to the rover. However, the reprojection method proved to be too simplistic to be reliable over longer traverse, highlighting the need more sophisticated matching techniques like constellation matching. By showcasing a dependable localization accuracy of less than 5 HiRISE pixels (1.25 meters) across diverse terrains and input data noises, the two automated absolute localisation modules, namely Constellation Matching and Dense Image Co-Registration, underscored their robustness and pertinence for missions related to planetary exploration.

6. ACKNOWLEDGEMENT

The work presented in this paper has been funded through ESA GSTP Element 1 Activity under contract No. 4000136724/21/NL/GLC/zk. The authors want to thank Gianfranco Visentin and Martin Azkarate at ESA for the follow up of the activity and their advices, as well as Michel Delpech and Irene Valenzuela-Molina at CNES for their support.

REFERENCES

[1] E. Boukas, A. Gasteratos et G. Visentin, "Localization of Planetary Exploration Rovers with Orbital Imaging: a survey of approaches," *IEEE International Conference on Robotics and Automation*, Hong Kong, China, 2014.

[2] P. Furgale, J. Enright et T. Barfoot, "Sun Sensor Navigation for Planetary Rovers: Theory and Field Testing," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, n°13, pp. 1631-1647, 2011.

[3] F. Cozman, E. Krotov et C. Guestrin, "Outdoor Visual Position Estimation for Planetary Rovers," *Autonomous Robots*, vol. 9, pp. 135-150, 2000.

[4] T. J. Parker, M. C. Malin, F. Calef et R. G. Deen, "Localization and 'Contextualisation' of Curiosity in Gale Crater, and Other Landed Mars Missions," chez 44th Lunar and Planetary Science Conference, The Woodlands, texas, 2013.

[5] R. Li, S. He, Y. Chen, M. Tang, P. Tang, K. Di, L. Matthies, R. E. Arvidson, S. W. Squyres, L. S. Crumpler, T. Parker et M. Sims, "MER Spirit rover localization: Comparison of ground image– and orbital image–based methods and science applications," *Journal of Geophysical Research*, vol. 116, 2011.

[6] Y. Tao, J.-P. Muller et W. Poole, "Automated localisation of Mars rovers using co-registered HiRISE-CTX-HRSC orthorectified images and wide baseline Navcam orthorectified mosaics," *Icarus*, vol. 280, pp. 139-157, 2016.

[7] C.-S. Chen, Y.-P. Hung et J.-B. Cheng, "RANSAC-based DARCES: a new approach to fast automatic registration of partially overlapping range images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, n°111, pp. 1229-1234, 1999.

[8] P. J. Carle, P. Furgale et T. D. Barfoot, "Long-range rover localization by matching LIDAR scans to orbital elevation maps," *Journal of Field Robotics*, vol. 27, n°13, pp. 344-370, 2010.

[9] J. W. Hwangbo, K. Di et R. Li, "Integration of Orbital and Ground Image Networks for the Automation of Rover Localization," *ASPRS 2009 Annual Conference*, Baltimore, Maryland, 2009.

[10] E. Boukas, A. Gasteratos et G. Visentin, "Introducing a globally consistent orbital-based localization system," *Journal of Field Robotics*, vol. 35, n°12, pp. 275-298, 2018.

[11] M. Dinsdale, et al., "Absolute Localisation by Map Matching for Sample Fetch Rover," *Advanced Space Technologies in Robotics and Automation*, Noordwijk, the Netherlands, 2022.

[12] B. Van Pham, A. Maligo et S. Lacroix, "Absolute Map-Based Localization for a Planetary Rover," *12th Sympusium on Advanced Space Technologies and Automation in Robotics*, Noordwijk, the Netherlands, 2013.

[13] W. S. Kim, J. J. Biesiadecki et K. S. Ali, "Visual Target Tracking on the Mars Exploration Rovers," *International Symposium on Artificial Intelligence: Robotics and Automation in Space*, Los Angeles, California, 2008.

[14] S. Lacroix, et al. "The Erfoud Dataset: A Comprehensive Multi-Camera and Lidar Data Collection for Planetary Exploration," 2018.