Localization of Planetary Exploration Rovers with Orbital Imaging: a survey of approaches

Evangelos Boukas¹, Antonios Gasteratos¹ and Gianfranco Visentin²

Abstract—Owing to the rovers exploring the surface of Mars being assigned with ever more complex tasks, it is the autonomy of such operations that enables their effective planetary activities. An indicative case of the escalation in requirements is the upcoming Mars Sample Return mission (MSR), which is a joint effort of the National Aeronautics and Space Administration (NASA) and the European Space Agency (ESA). The mission involves a Sample Fetching Rover (SFR) to gather a previously deposited cache of soil that, eventually, will be directed towards Earth. With the aim to retrieve the cache and rendezvous with the Mars Ascent Vehicle (MAV), the rover should be apt to globally localize itself on the Martian surface. Due to the absence of Global Navigation Satellite System (GNSS) on Mars, the most suited approach is the blending of information stemming from ground rovers and orbital imagery. The scope of the paper in hand is to summarize the work delivered so far on the localization of space exploratory rovers based on such information.

I. INTRODUCTION

Upon a successful probe landing along the robotic Mars exploration missions, one of the most important tasks is the accurate global localization of the rover on inertial and fixed coordinate systems, such as the Mars Mean Equator and IAU vector of J2000 frame and Mars body-fixed rotating frame [1], [2]. The localization of a space exploratory rover on a georeferenced orbital image is equivalent to global localization and, hence, is considered sufficient. During the last decade, some approaches have been proposed employing urban structures earthly observed by the robot as patterns for orbital recognition. For example, the authors in [3], [4] employ the skyline perceived by a fisheye camera in a city block, in order to match it with the 3D models of the city buildings. Other approaches utilize stereoscopic techniques to create top views of either the fully perceived field of view [5], [6] or of prominent obstacles [7] which are then matched to aerial equivalents. However, such dense, prominent regions do not exist in the Martian surface and, therefore, these techniques are of limited interest for Mars exploration missions.

Bearing in mind that an one-way signal to Mars requires approximately 20 min, it is apparent that the non autonomous navigation of rovers poses a noticeable overhead to the exploration of the red planet. Therefore, recent and future space exploratory rovers are designed for long traverses [8], [9]. This is the reason why the rovers are equipped with several cameras. An example of camera setup design for such a specific application is reported in [10], where a camera system for localization and mapping of space exploratory rovers is proposed. The algorithms required for the long range autonomous navigation are 3D reconstruction, mapping [11], [12], localization and path planning. The algorithms that are currently implemented in the space exploratory rovers can be found in [13], whilst a detailed review of visual odometry methods is available in [14].

Moreover, notwithstanding the advanced and noteworthy mechanical and electrical design of space exploratory rovers, their capabilities in terms of computational and power needs are modest, especially compared to contemporary robots. The algorithms for the localization, as well as for any other task, should be as lightweight as possible. Toward this end, the authors in [15] have implemented a Visual Odometry (VO) algorithm that is capable of achieving state of the art results, while being of low computational cost. In terms of power consumption, the authors in [16] have implemented the localization and mapping algorithms on FPGA devices for ESA’s future space exploratory rovers. The gain of this implementation is twofold: Firstly the cost, in power, of the execution of these algorithms is diminished and, moreover, the parallel implementation on the FPGA speeds up the algorithms to frequencies higher than 1Hz.

The manuscript is organized as follows: Section II categorizes and surveys all the approaches that employ information stemming from aerial and orbital imagery in order to localize rovers operating on the Martian surface. A proposal for the assessment of orbital based localization techniques is described in Section III and, last, conclusions are drawn in Section IV.

II. CATEGORIZATION OF METHODS

An attempt to categorize the techniques employed in the localization of planetary exploration rovers based on orbital imaging results into three major classes of approaches. These are separated according to specific characteristics summarized as follows:

• employment of integrated descent imagery to capture the Martian surface (Subsection: II-A)
• registration of the skyline acquired from rover on the Orbital terrain model (Subsection: II-B)
• seeking for common interest areas on both rover and orbital imagery to accomplish localization; this category is further distinguished into methods utilizing points of

¹Evangelos Boukas and Antonios Gasteratos are with the School of Engineering, Democritus University of Thrace, Vass. Sophia 12, GR-67100 Xanthi, Greece. evanbouk@pme.duth.gr, agaster@pme.duth.gr
²Gianfranco Visentin is the Head of Automation and Robotics Section (TEC-MMA), European Space Agency, The Netherlands. Gianfranco.Visentin@esa.int
The techniques being reviewed in the context of the paper in hand are graphically assembled in Fig. 1.

A. LOCALIZATION WITH DESCENT IMAGERY

Following the design and implementation of NASA’s first Mars rover, namely the Pathfinder’s Sojourner [17], the utilization of descent imaging for the localization of the robot near its landing position has been examined. The authors in [18], [19] presented a detailed explanation of state of the art descent and landing approaches followed by NASA and ESA. Matthies et al. [20] proposed the employment of Digital Elevation Maps (DEM), computed from consequent descent images, as prior information to the rover’s localization system. The authors theoretically proved the possibility to create DEMs by applying cross-correlation to match common surface points appearing on descent images at different scales. They also identified three approaches for the localization of rovers:

- terrain matching among ground and orbital deriving 3D models
- skyline identification on rover stemming images
- pairing of structures of interest, such as ridges, ravines and rocks on descent and rover imagery

1) DEM Construction from Descent Images: The automatic generation of a hierarchical DEM from consequent descent images is explained in detail in [21]. At the descent image “timeline” every image is acquired at a half distance from the previous one (starting from 5000 m). Three steps comprise the overall methodology:

- initial one-meter spacing DEM
- refinement of the one-meter spacing DEM
- generation of the hierarchical DEM

Firstly, Ground Control Points (GCPs) are used to acquire the exterior orientation. Then, a six-point based affine transform is performed to register each image to the previous one. A cross-correlation metric is used to find the correspondences between consecutive descent frames. The descent images present a vertical baseline and the 3D position of each matched point is inserted in the ground coordinate system by triangulation, forming the initial 1m resolution DEM grid. Since the lower images have greater resolution, the initial DEM grid is refined by a bottom-down and bottom-up methodology that performs bundle adjustment to ensure that each point has acquired the most accurate elevation. Lastly, the DEM is formed into a hierarchical layered grid with different resolutions ranging from 0.1-1m.

2) Incremental Localization: In [22] the localization of the rovers position is performed at each time step, by an incremental bundle adjustment. The difference of the implementation presented in [23] is that for each observation only the current and previous state set of variables are preserved and not the entire set of system variables. In order to build a bundle adjustment that would take into consideration both descent and rover image data, features appearing in both kinds of images were manually selected by an operator to serve as tie points [24], [25]. For a detailed analysis of the incremental bundle adjustment that includes both descent and rover features, the reader should refer to [23]. The performance possibilities of the image network of descent and rover imagery is proved in [26]. Li et al. [22] presented an astonishing overall localization error of less than 0.1% on a 500 meters course.

It is noteworthy to mention that the aforementioned approaches suffer from some crucial drawbacks. The first and foremost is that the working area is limited due to the coverage and the resolution of the descent images and the produced DEMs, as explained in [20]. Moreover the localization of the rovers (without the existence of GCPs, which are unfeasible on Mars) is performed by employing a local coordinate frame. Therefore, the global location - a condicio sine qua non in advanced space applications - remains unknown. Lastly, all of the approaches require a human operator to select strong tie points, due to the fact that the high difference in scale and orientation between the rover and descent images does not permit robust extraction and matching of common features. In some situations, even a manual selection of such points is difficult, resulting in a deterioration of localization accuracy up to 300% [23].

B. SKYLINE MATCHING

The skyline is a curve that passes through the edge of the horizon. Contrary to other approaches, the horizon matching does not aim to accurately localize a rover, but to calculate a preliminary uncertainty location area. The approach can be considered as an extension to the sun detection methods [27], where the localization produces an area of uncertainty in the order of 20 km. Some approaches relying on the horizon matching for rover localization exist, which assume accurate rover orientation based on sun sensor as the one presented in [28]. The most influencing and informative approaches are the one introduced by Stein and Medioni [29], [30].

![Fig. 1. The Venn diagram for the categorization of the approaches considered in this review.](image-url)
and the one presented by Cozman and Krotkov [31], [32], which is known as the VIsual Position Estimation for Rovers (VIPER). The main parts of both methodologies are summarized as follows:

- skyline detection from rover images
- skyline detection on DEMs
- feature extraction on skylines (optional)
- search and locate rover skyline on DEM images
- pose estimation based on the skyline location

The main approach is relatively easy to understand. The rover is considered to be at an unknown position, also known as the “lost in space” or “drop off” problem, laying inside an area that has been mapped, by means of orbital imaging. Also, the robot is considered to hold precise information about its orientation, as explained previously. The rover is equipped with panoramic cameras, or normal cameras on pan units, capable of capturing full circle images of the horizon. Then, the skyline is detected employing image segmentation techniques. Assuming a DEM covering the area of uncertainty, within which the robot lays and given the orientation of the rover, a “simulated” skyline is rendered at each and every point of the DEM. A search follows to match the rover’s skyline to the rendered ones.

1) Skyline Detection and Feature Extraction on Skyline:

There are two approaches of including the skyline into the system: (i) feature based and (ii) signal based ones. Talluri and Aggarwal [33] as well as Stein and Medioni [30] and Cozman in [32] utilized a signal-based approach, as the skyline is the elevation of the highest observed obstacle form the horizontal plane. On the contrary, the authors in [31], [33], [34] utilize a feature vector of the aforementioned signal. The vectors of rover and orbital stemming skylines are computed in the same manner.

2) Skyline detection on DEMs: Such approaches require the computation of skylines for each and every point in the DEM, through a procedure called, “skyline rendering”. Assuming an interval $\phi$, which represents an azimuth orientation and ranges into $[0 : N : 359]$, where $N$ is the required resolution, a skyline is predicted at each point of the DEM. The calculation of the skyline employs the separate line scanning and sampling on the DEM at each $\phi$ and due to the discreteness of the DEM requires an interpolation. An analytic explanation of the actual VIPER implementation can be found in [32]. According to the authors in [35], the computation of all those possible skylines is computational expensive, as it might “take some days for each DEM” but it is only executed once.

3) Rover Position Estimation: Given a full set of possible locations in a DEM and the corresponding skylines, the next step comprises the matching with the rover’s actual skyline. This is performed by a Bayesian posterior estimator that provides the probability of a rover to be at each point on the map.

Extensive testing has been performed by Furgale et al. [35], proving that although the VIPER algorithm is able to provide adequate results in some circumstances, a lot of conditions exist that result in software unresponsiveness or inaccurate estimation. The skyline approaches are prone to specific area formations, such as planar, repetitive scenery or even to a close obstacle that can cause occlusion [30]. An additional issue with such methods is the extend of the mapped area taken into consideration. The extend of the DEM should be large enough, to include all the areas that appear in the rovers horizon, but in the same time to be sufficiently small so as to allow adequate resolution and, hence, accuracy.

C. LOCALIZATION WITH ORBITAL-ROVER IMAGERY INTEGRATION

The global localization of Mars rovers at their landing site has been employed by Direct to Earth (DTE) radio signals, via two-way Doppler tracking [36] or via descent image analysis, as reviewed in Section II-A. The Mapping and GIS laboratory of the Ohio State University (OSU) has extensively analyzed the localization of the MER rovers, employing DEMs either from descent images or from the orbiting imagery [37]. Nevertheless, the majority of approaches comprise steps that require the manual selection of common points by a human operator. Some of the approaches that appear in the literature point towards the automation of the global localization, yet without any of them reaching the state of completeness that would allow its selection for realization.

1) MER Rovers Localisation Employing Orbital Imagery: The authors in [38] describe the initial approaches for the localization of MER rovers. The OSU Mapping and GIS Laboratory performed incremental bundle adjustment on the rover image network and on the orbital images, separately, in order to assist the operations of MER [38], [39], [37]. Although the inter-stereo tie points were computed automatically by 90%, the cross-site ties were computed manually. Nevertheless, the orbital imaging proved that the mapping procedures of the rover level were performing adequately. Li et al. [40] noted that with the aim to produce high quality localization estimates, orbital and rover image networks should be connected through tie points. Moreover, the authors marked the difficulty of extracting such features due to great differences in the scaling and viewing angle. In order to practically prove the necessity of the integration of rover and orbital imagery, the authors in [41] compared the bundle adjustment VO with a traverse that was corrected utilizing orbital images. They overcome the barrier of ground points extraction by manually selecting the common points. The comparison results in a deviation of 1.5% among the two approaches, which can be corrected with the automation of the tie-points selection.

2) Terrain Matching: Van Pham et al. [42] employed a Bayesian recursive algorithm, namely a modified particle filter able to retrieve the location of the rover on a global DEM. The rover was equipped with a stereo camera used to create a local DEM. The state of the particle was chosen to be: $X = [x, y, \theta]$, where $x, y$ are the spatial coordinates and $\theta$ is the orientation. This simplified approach was selected because of the global DEM formation. The particles were resampled into a discrete grid with the same spatial
resolution of the global DEM. The experiments were carried out in a sand quarry and the global DEM was created by means of a UAV, while ground control points were used for georeference. The robot employed was Ral Space’s Rimmer, as described in [43]. The resulting accuracy, which is confirmed by DGPS, is 1m, equal to the resolution of the DEM and the convergence distance is at 58-78m. The approach seems reasonable and well documented, however the examined area is quite limited. The uncertainty of rover location may reach up to 10ths of km, a range within which no convergence is guaranteed. Finally, the computational burden of the initialization of every point of DEM will lead to an unfeasible implementation.

3) Interest Point Matching: Di et al. [44] propose a novel approach for the “orbital-based rover localization” problem, by incorporating both rock detection/matching and Scale Invariant Feature Transform (SIFT) feature detection/matching in a system that handles outliers by employing a procedure similar to RANDom SAMple Consensus (RANSAC). The rock detection on the rover field of view is an intuitive algorithm that represents a 3D morphological filter, which is executed incrementally and leads to extraction of rock peak points [45]. The rock detection on the orbital image is a statistical intensity filter bearing constraints that ban shadows from being detected as a rock. The matching of the rover and orbital rocks employs a RANSAC-based algorithm, which through a random initial sample, is able to: (i) transform the rover rocks on the orbital image coordinate system and (ii) select the matches with a threshold over the Euclidean distance. Moreover, the system involves SIFT detection and matching among the orbital image and an orthophoria, which is created via the rover’s stereo images. A similar to the aforementioned RANSAC based outlier detection algorithm is applied to the matched points. The remaining inliers are used to calculate the position of the rover inside the orbital image. The employed dataset does not contain a groundtruth as it includes real Mars rover images.

Carle et al. [46] proposed the employment of a LIght Detection And Ranging (LIDAR) device in a system that is able to globally localize a rover laying within the imaging region of a georeferenced DEM. The approach follows a feature based approach, with features being the prominent peaks that appear on local and global DEMs. The local DEM is produced with a transformation on a single LIDAR scan. The extraction of the features, a peak detection in particular, is performed on the images, by filtering them with the external morphological gradient, i.e. the difference between the dilated image and the original one. The feature extraction is the same on orbital and rover DEMs. The correspondences between the rover and ground features are computed utilising a method named “Data-Aligned Rigidity-Constrained Exhaustive Search” (DARCES) [47]. Firstly, several groups of three matched points are considered and the group the resulting 2D transformation of which offers the least error is selected. This estimation is then fed into a “Simultaneous Localization and Mapping” (SLAM) module, namely the MOGA which includes a RANSAC outlier detection part. The performance of the system is noteworthy with a 22.1 m error at a range of 1.5 km with a DEM resolution of 13m×13m. The authors compare their system with VIPER (see Section II-B), which they outperform. The most significant issue of this work is the infeasibility of the LIDAR integration on a Mars Rover due to both heaviness and power consumption.

The last approach being presented here is the one developed by Hwangbo et al. [48]. The distinctiveness of this approach lays on the the employment of both full terrain and rock matching to hierarchically trace correspondences between orbital and rover images. Firstly, the 3D terrain matching is utilized to locate the region of interest in the global DEM that corresponds to the rover position. The rover DEM is created from stereo imagery and both DEMs are treated as images. With the purpose of having the same scale in elevation, the mean elevation is subtracted from both DEMs. Then, the matching is performed by identifying the region of the maximum weighted correlation of values and slopes within the global DEM. As soon as the region of interest is located the rock extraction and pattern matching is performed. The rock detection on rover imagery includes the removal of ground points and the identification of peaks, i.e. the highest points in an area having elevation of more than 25cm. The rocks on orbital imagery are identified via an intensity thresholds technique. Furthermore the morphology of the rocks is examined, discarding those rocks with long axis longer than 2m and short axis shorter than 0.2m. The rotation angles are considered to be zero and, thereby, the translation vector which produces the most matches is selected. Yet, it should be mentioned that this method is prone to local minimum on the terrain matching, which result in false localization estimation.

III. ASSESSMENT

As rovers are called to fulfill more complicated functions the need for higher accuracy and autonomy correspondingly emerges. During the last years there has been an increasing interest in approaches for the localization of planetary rovers utilizing orbital imagery. Howbeit, a standard procedure for the assessment and comparison of a new reported method does not exist yet and neither does a standard dataset upon which the method can be valorized. Thus, taking an interest in robustly solving the localization problem, it is our belief that a benchmark framework should be setup with a view to define a common base-line for the coming out approaches. The main points of such a framework can be summarized as follows:

- Real World Relevance: In order to prove the relevance of a newly proposed method, one should test his/her approaches on specific and real world datasets. According to the authors, the most sufficient datasets are the ones produced by the ESA on the Chilean Atacama Desert (the ones provided by SEEKER [43] and SAFER activities). The selected areas in the Chilean desert are considered to be the most Mars-like regions on Earth.
Section of orbital image

Three categories, based on the type of information used for rover localization. We have classified the approaches into:

- **Openness:** The performance should be tested publicly.
- **Repeatability:** The method should be tested at different scenarios with different trajectories and surrounding environments. This may be accomplished by the concurrent usage of simulators, such as the Pangu and 3D Rover, [50], [51], for instance.
- **Feasibility:** The approach should be feasible, taking into consideration current space rover apparatus. Any addition to the current equipment should be accompanied by a feasibility assessment or, at least, a reasonable prediction of such future space qualified hardware.
- **Openness:** The performance should be tested publicly similar to the successful Middlebury evaluation for stereo correspondence algorithms [52], [53].

(Figure 2). Furthermore, these datasets contain both rover stereo images, registered with DGPS measurements for groundtruth and georeferenced aerial images, sampled appropriately to resemble the orbital imagery on Mars.

Another source of data could be the actual MER and MSL programs, but since the stereo images are sparse (at least sparser compared the actual rover’s frame-rate) and no groundtruth exists, then the Atacama datasets are the most appropriate.

- **Accuracy:** The accuracy requirements of space rovers localization should be similar to the requirements on Earth. The accuracy should be measured both by: (i) the accumulated signed error, expressed as the difference between the measured and the desired position at the end point and (ii) the evolution of the error in all 6 Degrees of Freedom (DOF) of the robot’s pose along its course. More specifically, in the case of utilization of orbital data, it is the the resolution of the respective imagery that defines the maximum resolution that can be achieved. Nowadays, the imagery of High Resolution Imaging Science Experiment (HiRISE) [49] is able to produce orthorectified images of 0.25\(m\) spatial resolution and DTMs with 1\(m\) resolution.

- **Repeatability:** The method should be tested at different scenarios with different trajectories and surrounding environments. This may be accomplished by the concurrent usage of simulators, such as the Pangu and 3D Rover, [50], [51], for instance.

- **Feasibility:** The approach should be feasible, taking into consideration current space rover apparatus. Any addition to the current equipment should be accompanied by a feasibility assessment or, at least, a reasonable prediction of such future space qualified hardware.

- **Openness:** The performance should be tested publicly similar to the successful Middlebury evaluation for stereo correspondence algorithms [52], [53].

IV. DISCUSSION

We have presented the approaches that have been proposed during last decades for the problem orbital-based rover localization. We have classified the approaches into three categories, based on the type of information used for localization: Descent Imagery, Skyline and Orbital Imagery ones. Along the last years, there has been observed an increasing attention to the problem and it is anticipated that more and more of novel approaches will emerge in the near future. Therefore, we have proposed a benchmark framework for the reliable and undisputed evaluation of such methods. One of the issues that arise from our study is the lack of robust techniques to overcome the manual tie point selection among the orbital and rover imagery. Towards this end the authors in [54] are examining the extraction of such commonly observed regions of interest.

ACKNOWLEDGMENT

This work was funded by the European Space Agency via the Network/Partnering Initiative under the contract No. 4000109064/13/NL/PA.

REFERENCES


