A Vision-Based Navigation Facility for Planetary Entry Descent Landing

Piergiorgio Lanza[1], Nicoletta Noceti[2], Corrado Maddaleno[1], Antonio Toma[3], Luca Zini[2], and Francesca Odone[2]

[1] Thales Alenia Space Italy
[2] DIBRIS - Università degli Studi di Genova, Italy
[3] Politecnico di Torino, Italy

Abstract. This paper describes a facility set up as a test bed and a proof of concept to study open issues of future space missions. The final goal of such studies is to increase the on board autonomy, of primary importance for missions covering very high distances. We refer in particular to vision-based modules, in charge of acquiring and processing images during the Entry Descent and Landing (EDL) phases of a Lander, and contributing to a precise localization of the landing region and a safe landing. We will describe the vision-based algorithms already implemented on the facility, and a preliminary experimental analysis which allowed us to validate the approaches and provided very promising results.

1 Introduction

The key factor for the future space robotic missions is the *on board autonomy*, particularly relevant for missions towards planets at high distances from Earth, where a robotic closed loop control in real time is not possible — as an example, the travel time of a radiofrequency signal requires from 4 to 22 minutes covering the distance between the Earth and Mars.

In this general and ambitious goal, Computer Vision algorithms may help obtaining information about the 3D environment of space missions to verify the correct execution of robotic commands. Unfortunately, the difficulty of gathering real data and experimenting on the field is a real limitation for these approaches, and indeed most of the Computer Vision research for space applications is carried out by very few well known premises — such as the NASA JPL. With the exception of some early works (see e.g. [1,2]), DIMES (Descent Image Motion Estimation System) [3], developed by NASA, can be considered as the first historical attempt to exploit Computer Vision to control a spacecraft during the EDL phase. More recently, thanks to the successful application to the Mars Exploration Rovers (MER) landings, NASA invested in the last years in designing and developing vision-based solutions for EDL. The works in [4,5], describe improvements of DIMES technology for what concerns the problems of slope estimation and hazard detection on the planet surface.

This paper reports a work in progress in the design and the development of a facility – the Vision Based Navigational Facility (VBNF) – for the benchmarking

A. Fusiello et al. (Eds.): ECCV 2012 Ws/Demos, Part II, LNCS 7584, pp. 546–555, 2012.

of software modules and sensors to be used during planetary Entry Descent and Landing (EDL). This facility will allow us to generate a realistic environment for testing Computer Vision algorithms against various degrees of complexity, and acquire and distribute such data to the benefit of other research groups.

We will start with a description of the facility layout, with particular emphasis on technical characteristics and operability. Then, we will present the vision-based software modules tested and installed so far, having in mind two main requests: *landing precision* – that is, a small error between actual and expected landing coordinates – and *safety* – that is, an appropriate landing spot in terms of sunlight illumination, absence of rocks and limited surface slope. The remainder of the paper is organized as follows. Sec. 2 discusses the requirements of landing approaches. Sec. 3 is devoted to a detailed description of the facility. The software validation plan and the vision-based algorithms are described in Sec. 4 and 5. Sec. 6 is left to a final discussion.

2 Lander Approach Main Requirements

This section summarizes state of the art and challenges of planetary EDL, with a specific reference to vision-based approaches. Two common requirements for future scientific missions are the exploitation of a *more precise* and *safer* landing.

Over the last decades, within the Viking1 and Viking2 lander missions (1976), Mars Path Finder (1997), Mars Exploration Rover (Spirit and Opportunity) (2004), Phoenix (2008), and Mars Science Laboratory (2009), the landing major axis ellipse has been reduced from 300 km to 20 km [6]. The current aim is to further reduce this dispersion ellipse in order to achieve, within few years, ellipse landing of hundreds of meters. This precision is strictly necessary to reach particular interesting landing zones, as the so called Peaks of Eternal Light (PELs) – zones where the sunlight is almost always guaranteed, found on Mercury and on the Moon. Depending on the distance from the landing surface, a combination of sensors, including video cameras and altimeters, may be adopted.

A second driving criterion, beyond the landing precision, is the definition of safe or unsafe landing zones. The evaluation of the safeness of a zone can be performed by means of image processing algorithms applied to images acquired during the last lander descent phase (lander altitude from 3000 m to 100 m). More specifically, according to the European Space Agency (ESA) recommendations [7], the parameters taken into account are the following:

Level of illumination. There is usually a minimum level of illumination (defined by the producer) for the solar panels to operate.

Slope. Terrains within a 10 deg of maximum slope can be considered as safe for landing.

Rock height. Rocks with height up to 0.5 meter are considered not dangerous for the landing.

Terrain morphology. Unfortunately no objective criteria exist in the definition of safe zones with respect to the morphology. So far, only human judgment with respect to a ground truth has been employed. Sharp cliffs, big

rocks, and high craters rims are considered critical zones. It can be observed that they are characterized by distinctive local textures.

Finally, the use of image processing techniques calls for a further requirement: the capability of acquiring and possibly storing images during the mission approach to the target. The acquired images can be directly sent to the Earth Ground Station in real time, or stored on dedicated mass memory and transferred afterwards.

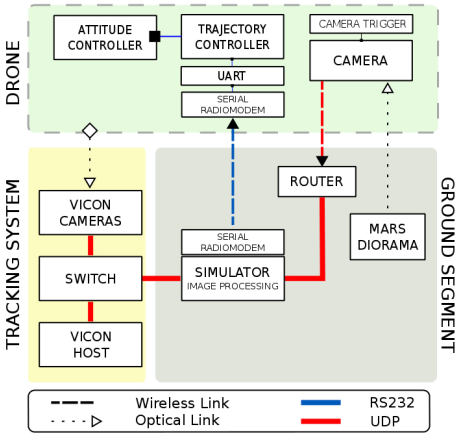


Fig. 1. The functional Layout of the VNF

3 Vision-Based Navigation Facility

The VBNF (Vision-based Navigation Facility) installed at the Thales Alenia Space Italy premises is a platform to be used as a test bed for innovative vision-based technologies and solutions. It is suited to validate them and to analyse their performances in operative situations similar to real ones. In fact, the VBNF simulates EDL operations of a laboratory *lander* represented by a drone equipped with a video camera.

From a functional point of view the facility is composed by three main components (see Fig. 1): (1) the Tracking System, (2) the quadrotor Drone, (3) the Diorama and the Simulator/Image Processing workstation.

The *tracking system* is based on the Vicon Motion Capture System and is composed by 13 Vicon Bonita infrared cameras, connected together in a PoE (Power over Ethernet) network hosted by a dedicated workstation on which runs the proprietary Vicon Tracker software. The infrared cameras are attached on a cube-shaped aluminum structure with 9 meters long sides (Fig. 2, left). They acquire 0.3 Megapixel images up to 240Hz and can track a single marker with a precision of 1 mm. The markers are small spheres covered with an infrared-reflective coating that are fixed on the body to track. The purpose of the tracking

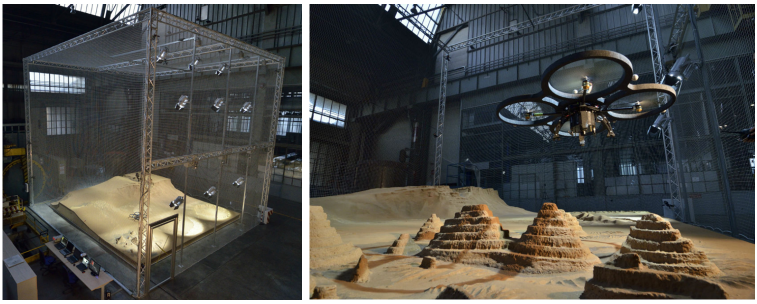
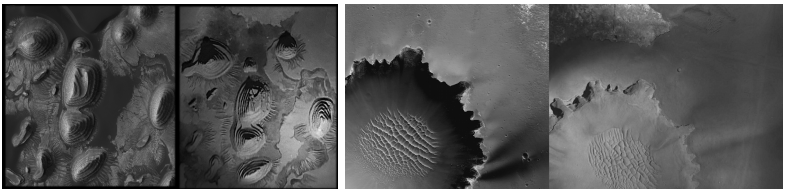


Fig. 2. Left: the Vision-based Navigation Facility. Right: the drone flying (note the camera facing down observing the diorama).



(a) (b)

Fig. 3. (a) Left: part of Western Arabia Terra. Right: the same portion of Arabia Terra on the diorama. (b) Left: the Victoria Crater. Right: the Victoria Crater on the diorama.

system is to provide to the trajectory controller the position and the attitude angles of the quadrotor while it flies inside the tracked volume.

The *quadrotor drone* is a rotorcraft propelled by 4 rotors (Fig. 2, right). The model used in the VBNF has a lightweight carbon fiber frame and can bring a payload up to 1 Kg heavy with an endurance of about 15 minutes. The payload is constituted by the camera, the trajectory control board and the serial radio modem, for an in-flight weight of the drone of 2.7 Kg. While the flight stability of the quadrotor is guaranteed by a commercial autopilot (Mikrokopter) that accurately balances the roll and pitch angles of the drone in order to keep it hovering in the air, the trajectory controller is implemented on an additional board running 4 parallel PID (Proportional-Integral-Derivative) controllers, one for each of the remaining degrees of freedom (x, y, z and yaw angle). This board receives on a dedicated wireless-serial link the feedback obtained from the UDP (User Datagram Protocol) Vicon data stream and sends the commands to the Mikrokopter board using a second UART (Universal Asynchronous Receiver-Transmitter) port. Finally it also takes care of triggering the camera when a new picture is requested by the Image Processing Workstation. The video camera is a CMOS gray level camera (1Kb x 1Kb, 8 bits per pixel) with a rolling shutter. These characteristics are comparable to the already space qualified camera based on STAR1000 sensor used for space applications.

The 1:300 scale *diorama* (8m × 8m, 1.5m of maximum relief height) accurately reproduces some geographic peculiarity of the Mars surface: the Nili Fossae, the Victoria Crater, the Xanthe Terra, the Arabia Terra, and the Dilly Crater. Fig. 3(a) and Fig. 3(b) show, on the left, images of two real Mars surfaces, and on the right the diorama corresponding portions. As it can be observed, the representation is very accurate.

4 Software Validation Plan

All vision-based algorithms to be installed on the Image Processing Workstation(see Sec. 5) are tested in accordance to a *validation plan*. It can be summarized according to the following pipeline:

Preliminary tests (few tests). They include evaluations of the consistency for each software module on a small set of data. The goal is to verify and validate the algorithm and all its modules.

First tests campaign (100 tests). The objective is twofold: first, optimizing the algorithms and, second, characterizing the method and determining their performances.

Second tests campaign (> 10000 tests). It aims at evaluating the robustness of the selected methods with respect to scenarios with variable clutter, noise, angle of the sunlight, texture.

Full chain tests validation (> 65000 tests). The goal here is to evaluate the precision of the whole algorithmic chains.

Datasets include synthetic data, created with a viewer of 3D terrain models developed by ESA (PANGU [8]). Real data include public data from NASA and NASDA and a dataset of video sequences from the VBNF facility.

5 Vision-Based Algorithms

In the following we describe the main Computer Vision modules currently included in the facility. They all have been thoroughly examined following the validation plan (Sec. 4) with the exception of the slope estimation, which is under characterization (Second Tests Validation Campaign). All feature based methods make use of an optimized implementation of SURF features [9] chosen as a trade-off between efficiency and accuracy, although other features (such as, SIFT and Harris corners) are available.

5.1 Relative and Absolute Pose Estimation

The lander pose estimation, which controls the landing position, is based on two different algorithmic chains (see Fig. 4), that are described in the following.

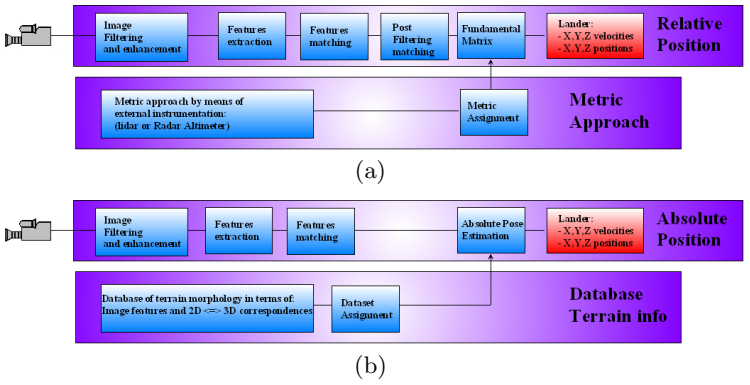


Fig. 4. Flowchart of the main steps of Relative and Absolute Pose determination

Relative Pose Estimation. Relative pose estimation is evaluated between two consecutive images of the sequence acquired during the descent phase. At each time instant, we extract SURFs and compute sparse correspondences between adjacent frames [10,11]. Such correspondences are used to estimate the relative geometry between the two views, that is the *fundamental matrix* F. To this purpose we adopt an approach proposed in [12], where the authors exploit a prior knowledge on the observed scene which appears to be quasi-planar from a large distance (see Fig. 3(b)).

In [12] the authors propose an efficient implementation of the *plane + parallax model* originally proposed in [13] and based on the geometrical model $F = [\mathbf{e}']_{\times} H$.

An external instrument (Lidar or an altimeter in a real space implementation of the algorithm, the tracking system in the VBNF) delivers the altitude information needed to associate a correct metric to the relative pose estimation. The obtained estimate can finally be applied to infer the spacecraft position with respect to an estimated entry point. This approach can advantageously be used when no 3D landing model terrain is available. Tab. 1 reports average relative errors between adjacent frames w.r.t. the ground truth on both synthetic and real data, at the end of the validation pipeline: $RelativeError = \frac{||TrueStep - EstimatedStep||}{||TrueStep||}$.

Absolute Pose Estimation. The absolute pose is determined by Perspective-n-Point (PnP) [14] algorithm which evaluates the matching correspondences between the features extracted from an input image and their equivalent 3D points measured from the 3D terrain model of the landing zone. Thus, this requires such 3D points to be known in advance. For instance, they may correspond to relevant elements of the planet surfaces, as craters or rock structures. Within this context, we adopted a set of labeled correspondences, and obtained very promising results in terms of estimation accuracy, with an error limited to the 3-4 % with respect to the ground truth.

Table 1. Relative pose estimation results on the full chain validation

Sequences type	Trajectory type	Relative error
Synthetic	linear	0.08 ± 0.01
	polynomial	0.3 ± 0.05
Real	linear	0.2 ± 0.18
	polynomial	0.4 ± 0.14

5.2 Hazard Map Determination

In this section we review the main steps performed to derive an hazard map of the surface. The map combines danger levels produced by the various algorithms running in parallel at different resolution levels (patch level for illumination and texture, variable region size level for slope).

Analysis of the Illumination. The algorithm we adopt relies on well known techniques. A simple thresholding is performed on the values of the image pixels. The threshold should correspond to the minimum acceptable level of illuminance required for an exploration mission to operate autonomously.

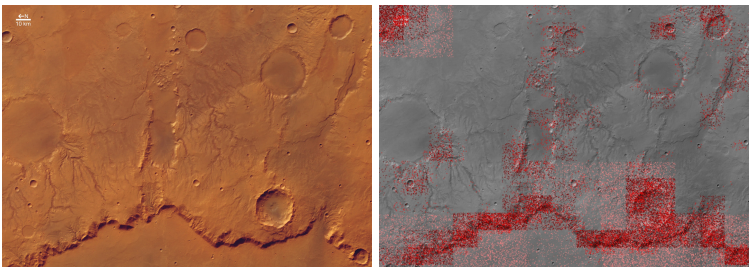


Fig. 5. Left: a real image of a Mars terrain. Right: the output of Hazard Map algorithm. The dangerous areas have been highlighted with colors of increasing intensity depending on the dangerousness.

Texture Analysis. Texture is an important element to human vision and people also tend to relate texture elements of varying size to a plausible 3D surface. The algorithm we adopt is based on the local evaluation of statistical properties of an image, considering local regions of 16x16 pixels. The statistical approach is based on Intensity Standard Deviation (ISD) developed at JPL [4]. The output of the procedure is an evaluation of the danger level (safe, low, medium, high dangerous). Fig. 5 reports a visual representation of such levels of dangerousness, for an example image of variable complexity.

Slope Determination. The slope of the planetary surface can be computed exploiting the relationship between two adjacent images of the planet surface acquired during descent (see the examples in Fig. 6). This relationship is defined by a homography, since the surface is quasi-planar because of the relative distance between the spacecraft and the surface. It is well known, the homography

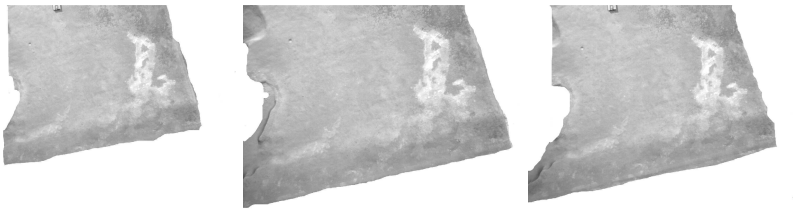


Fig. 6. Examples of images, acquired by the VBNF, of a terrain of slope 15°

H encodes information on the lander relative motion between poses (rotation matrix R and translation vector t) and on the slope of the observed landing site [13]. It holds that $H = R - \mathbf{T}\frac{\mathbf{n}_0^\top}{\mathbf{d}}$ where $\mathbf{n}_0$ is the normal of the surface and $\mathbf{d}$ is the distance between the plane and the camera center. This can be seen a special case of Structure from Motion, when the scene is planar.

In our work we start from a set of feature correspondences and estimate the homography H with RANSAC [15]. Since the homography H can be recovered only up to a scale factor, the equation becomes $H \approx R + \mathbf{T}\mathbf{n}^\top$ where $n = -\frac{\mathbf{n}_0^\top}{\mathbf{d}}$. The obtained solution carries a two-fold ambiguity which can be solved by looking at a further view (that is, combining the solutions from another pair of views).

A previous work in the literature disregarded this ambiguity [4] proposing a solution based on Levenberg-Marquardt, which could lead to one of the possible solutions, not necessarily the right one. Instead, we adopt a spectral approach proposed in [16] (see also [17]) and based on computing a SVD of H and extracting the unknowns from the factorization. We obtain 4 possible solutions, 2 of which may be discarded by geometrical constraints. A unique solution may be obtained by estimating the most stable solution from consecutive image pairs.

Finally, the slope λ of the surface is simply computed as the angle between the estimated normal vector $\mathbf{n}_s$ and the reference normal vector of the surface $\mathbf{n}_r$ in the coordinate system of the camera as $\lambda = \arccos \frac{\mathbf{n}_r^t \mathbf{n}_s}{\|\mathbf{n}_r\| \|\mathbf{n}_s\|}$ (see Fig. 7).

The results obtained at the end of the second campaign on real images representing a variable texture, illumination, and an even distribution of the slope in the range $[0, 30]$ degrees, produced a 1.2% average error (estimated slope diverging from the real one of more than 5 degrees). The performances degrade when the quality of images (and image features) degrades. A set of experiments carried out on the previous set of images, after a low-pass filtering with Gaussian

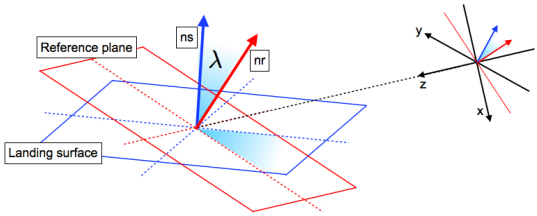


Fig. 7. The slope λ can be recovered computing the angle between ns and nr

filters with standard deviation equal to $\sigma = 3$ and $\sigma = 4$ causes an increasing of the average error to 1.8% and 2.5% respectively.

Preliminary experiments carried out on the full pipeline campaign confirm this analysis. However, they also highlight that a main drawback of real images within the domain of interest is a poor texture and, therefore, a high degree of noise in the estimation of H.

6 Discussion

In this paper we presented a work in progress on the design and the development of the VBNF facility whose main goals are the benchmarking of software modules and sensors to be used during Planetary EDL. We are currently improving both the hardware and the software modules, working on a more stable quadrotor, a more precise camera calibration, and vision-based modules more robust to poorly textured images.

Here we presented very promising results of the vision based algorithms in terms of accuracy with respect to synthetic ground truths. In particular, the error for the Relative and Absolute Pose estimation is within the 3-4 %. By means of the VBNF it will be possible to further verify the algorithm performances against optical deformation introduced by rolling shutter, quadrotor mechanical vibrations and real illumination effects in terms of visible terrain textures and shadows. All these effects are present in a real space mission environment and they shall be taken into account to determine the performances of vision based EDL algorithms.

The facility will also allow us to acquire benchmark datasets, some of which will be shared with the research community.

Acknowledgments. The VBNF facility has been designed and integrated with the support of the STEPS project co-funded by Regione Piemonte (Project co-financed by EC Platform: POR FESR 007/2013).

References

1. Cheng, Y., Johnson, A., Matthies, L., Wolf, A.: Passive imaging based hazard avoidance for spacecraft safe landing. In: Proc. iSAIRAS 2001 (2001)
2. Cheng, Y., Goguen, J., Johnson, A., Leget, C., Matthies, L., Martin, M.S., Willson, R.: The mars exploration rovers descent image motion estimation system. IEEE Intelligent Systems 19, 13–21 (2004)
3. Cheng, Y., Johnson, A., Matthies, L.: Mer-dimes: A planetary landing applications of computer vision. In: IEEE Proc. CVPR (2005)
4. Huertas, A., Cheng, Y., Madison, R.: Passive imaging based multi-cue hazard detection for spacecraft safe landing. In: Aerospace Conf. IEEE (2006)
5. Cheng, Y.: Real time surface slope estimation by homography alignment for spacecraft safe landing. In: ICRA (2010)
6. Braun, R.D., Manning, R.M.: Mars exploration entry, descent and landing challenges. In: IEEE 29th AAS Guidance and Control Conference (2006)
7. ESA: Navigation for planetary approach a general approach and landing (2006)
8. Parkes, S., Martin, I., Dunstan, M.: Planet surface simulation with pangu. In: Int. Conf. on Space Operations (2004)
9. Bay, H., Ess, A., Tuytelaars, T., Van Gool, L.: Surf: Speeded up robust features. CVIU 110, 346–359 (2008)
10. Scott, G.L., Longuet-Higgins, H.C.: An algorithm for associating the features of two images. Proc. of the Royal Society Biological Sciences 244, 21–26 (1991)
11. Delponte, E., Isgrò, F., Odone, F., Verri, A.: Svd-matching using sift features. Graphical Models 68 (2006)
12. Zini, L., Odone, F., Verri, A., Lanza, P., Marcer, A.: Relative Pose Estimation for Planetary Entry Descent Landing. In: Koch, R., Huang, F. (eds.) ACCV 2010 Workshops, Part II. LNCS, vol. 6469, pp. 255–264. Springer, Heidelberg (2011)
13. Hartley, R., Zisserman, A.: Multiple view geometry in computer vision, vol. 2. Cambridge Univ. Press (2000)
14. David, P., Dementhon, D., Duraiswami, R., Samet, H.: Softposit: Simultaneous pose and correspondence determination. Int. J. Comput. Vision 59, 259–284 (2004)
15. Fischler, M., Bolles, R.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM 24, 381–395 (1981)
16. Faugeras, O.: Three-dimensional computer vision: a geometric viewpoint. The MIT Press (1993)
17. Zhang, Z., Hanson, A.R.: 3d reconstruction based on homography mapping. In: ARPA Image Understanding Workshop, pp. 249–399 (1996)