

## **OPTICAL TERRAIN NAVIGATION FOR PINPOINT LANDING: IMAGE SCALE AND POSITION-GUIDED LANDMARK MATCHING**

**Jeff Delaune,<sup>\*</sup> Guy Le Besnerais,<sup>†</sup> Martial Sanfourche,<sup>‡</sup> Thomas Voirin,<sup>§</sup>  
Clément Bourdarias,<sup>\*\*</sup> and Jean-Loup Farges<sup>††</sup>**

Map-based accurate navigation is a requirement for future planetary missions requiring a pinpoint landing capability. Based on an Extended Kalman Filter tight fusion architecture between a camera and inertial sensors, this paper proposes a new method to match descent image landmarks with orbital ones down to very low altitude while being robust to 3D terrain topography and illumination changes. It is based on the idea that image scale can be used as an additional description for a landmark. Filter estimates are used to predict the research area and scale of the landmarks in the descent image. A Monte Carlo analysis using software simulation shows its compliance with pinpoint landing requirements.

### **INTRODUCTION**

Autonomous absolute navigation is the capability of a vehicle to locate itself with respect to an existing map of its environment without any external help. Mission objectives are defined in terms of elements of this map, e.g. “reach this point” or “survey this area”. The navigation system estimates the position, speed and attitude of a frame tied to the vehicle with respect to a reference one tied to the map. For most outdoor terrestrial flight applications, satellite or radio signals are used to determine the position of the vehicle on a world map. This process is only semi-autonomous though since it relies on external radio beacons or satellites for which the positions on the map must be known. When full autonomy is necessary, embedded terrain sensors such as altimeters, lidars or cameras are the only alternatives for accurate navigation and are reviewed in (Reference 7).

On Earth, this need for autonomy becomes apparent for Unmanned Air Vehicles (UAVs) in military operations since radio and satellite can suffer from jamming. In urban environment, multipath propagation also alters navigation accuracy from these signals and thus places a need for terrain-relative navigation.

For planetary exploration landers, full autonomy is mandatory because of the unavailability of any radio or satellite navigation network. Teleoperation is also ruled out because the communica-

---

<sup>\*</sup> PhD Student, Systems Control and Flight Dynamics, ONERA, 2 av. Édouard Belin - 31000 Toulouse - France.

<sup>†</sup> Research Engineer, Modeling and Information Processing, ONERA, Chemin de la Hunière - 91123 Palaiseau- France.

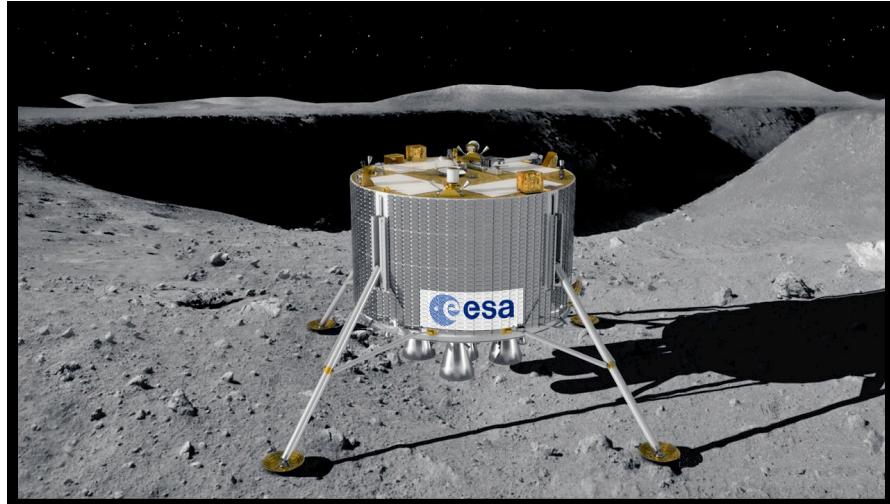
<sup>‡</sup> Research Engineer, Modeling and Information Processing, ONERA, Chemin de la Hunière - 91123 Palaiseau- France.

<sup>§</sup> GNC System Engineer, TEC-ECN, ESA-ESTEC, Keplerlaan 1 - 2200 AG Noordwijk - The Netherlands.

<sup>\*\*</sup> Navigation Engineer, Mission Analysis and Navigation Systems, Astrium ST, 66 route de Verneuil - 78133 Les Mureaux - France.

<sup>††</sup> Research Engineer, Systems Control and Flight Dynamics, ONERA, 2 av. Édouard Belin - 31000 Toulouse - France.

tion delay due to long distance from Earth is simply not compatible with the tight time schedule of landing operations. Nevertheless, some future missions will need landing errors within 100 m of a point designated on a map of the target planet. This map is expected to be made from previous orbital observation data. This pinpoint landing capability is required to reach astronauts or assets previously landed, a small area of great scientific interest, or to meet technical constraints about the landing site. For instance, the European Space Agency's (ESA) lunar lander mission planned for 2018 and shown in Figure 1 is aimed for areas of extent as limited as 200-m at the lunar south pole.<sup>17</sup> Indeed, such places offer better illumination condition, and thus a better solar power access for mission operations of up to several months.



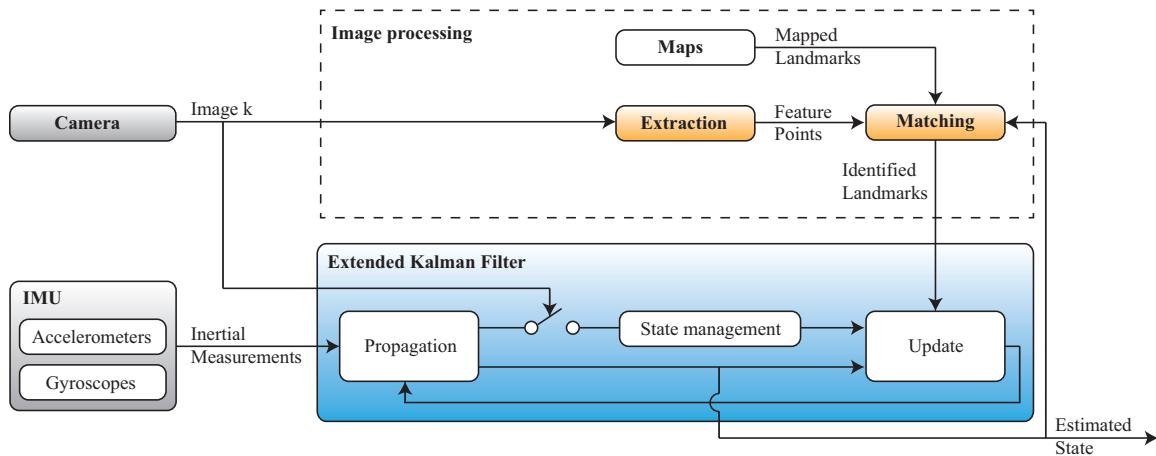
**Figure 1. ESA's lunar lander (Credit: ESA).**

This work is focused on the design of a terrain-relative navigation system to achieve pinpoint planetary landing. Among all possible terrain sensors, the authors decided to go for a simple optical camera since those are lightweight, cheap, and have already flown in space. These qualities make them relatively easy to implement on any spacecraft design. Because they are passive devices, they can operate from any distance to the terrain, from orbit to touchdown. The ground must be clearly visible though, e.g. not covered by clouds or in shadow. There are three main challenges in using optical cameras for planetary landing. The first is to match descent images taken over a significant range of altitudes with an orbital image. The second is to deal with variation in illumination conditions at the time of landing. And eventually the last difficulty is to work over rugged terrains with highly-3D topography. This capability becomes even more interesting now that orbital images and Digital Elevation Models (DEM) with resolutions of the order of the meter are available for other planets.<sup>10</sup> Indeed, these allow to make maps which can be used down to a few hundred meters of altitude for absolute navigation in order to improve landing accuracy.\* However at such low altitudes, the 3D effects of most terrains become even more important.\* These effects can be noticed when flying over the succession of peaks, valleys and crater ridges at the Lunar south pole, when approaching the bumpy shape of an asteroid, or over Martian canyons. The navigation system must be able to understand this 3D information with a single passive 2D camera only.

---

\* Especially if the camera is not pointing nadir.

The aim of this study is thus to propose navigation system for pinpoint planetary landing able to work with any type of terrain topography and robust to illumination changes between the map and descent images. The architecture of this system is illustrated on Figure 2. It is based on an Extended Kalman Filter (EKF) which fuses inertial and optical measurements. An Inertial Measurement Unit (IMU) is present on every space mission for orbital navigation. It is also relevant for landing in order to get high-frequency measurements to feed the control loop and correct sudden motion but also to keep the navigation system running when the information coming from the camera is not reliable, for instance over a shadowed area. The non-gravitational acceleration and angular rate measurements of the IMU are used in the propagation part of the filter. The propagated estimates are initialized from the navigation of the orbital phase.\* The camera descent images are processed to extract feature point landmarks which are to be matched with those from orbital images of the area. The EKF measurements are the image positions of the landmarks which can be predicted using their world coordinates, the a priori state estimate and the calibrated camera model. Image processing delays are taken into account through the use of a state management block which adds a state for the camera at the time of image acquisition and maintains it correlated with current state with the help of inertial measurements. All the details of the filter can be found in (Reference 3).



**Figure 2. Navigation system architecture.**

This paper starts by reviewing the state of the art to match feature points between a descent image and a map. The image scale of the features is shown to be of particular importance in this matter and we give some background and theory about it in the third part. Subsequently, the scale and position-guided landmark search algorithm is presented. And lastly, some simulation results show the performance of the whole filter on a lunar landing approach trajectory.

## LANDMARK MATCHING REVIEW

The visual updates to the filter are based on matching descent image feature points with mapped landmarks rather than on a whole image-to-map correspondence. There are two main advantages in doing this. First it requires less on-board memory space for map storage since only

---

\* For most exploration missions, the standard orbital navigation technique is inertial propagation with corrections coming from a star-tracker in attitude and from Earth-based radio-tracking in position.

information from small image patches around the features of the orbital image needs to be saved and not all of it. Secondly it is more adapted to 3D terrains since 2D image measurements and the associated planarity assumption only concern the local terrain neighborhood of each landmark. The computer vision community makes great use of feature detection and matching to retrieve the geometry of a scene from multiple views.<sup>6</sup> The recursive tight filtering scheme shown in Figure 2 allows to get measurements even from only one landmark and offers a finer modeling of the measurement noise. Different landmark-matching techniques proposed for planetary landing using a camera and considered compatible with real-time hardware constraints are now reviewed.

## State of the Art

One landmark can be described uniquely by the pixel values in a neighborhood big enough on an orbital image of the area. Mourikis et al. extract small descent image patches around intensity corner features extracted through the Harris transform and rectify them onto the horizontal plane using attitude and altitude knowledge from the filter.<sup>5, 12</sup> Landmark matches are the peaks of pseudonormalized cross correlation between the descent patches and the orbital image. Lowe also proposed a very popular Scale-Invariant Feature Transform (SIFT) to match two images with invariance in scale (or resolution) and rotation.<sup>9</sup> SIFT features are selected at a specific image location and scale. They are oriented and described by the orientation and magnitude of the intensity gradients in their neighborhood. This information is stored in histograms which are then compared between two images for matching. Intensity-based signatures are very popular in the computer vision community but they are discarded here because of their sensitivity of the description to illumination changes.

As an alternative, landmarks can also be described by their geometric arrangement. These methods are usually more interesting in terms of space saving in the on-board memory. Cheng et al. identify craters two by two by computing a parameter for a pair of coplanar conic invariant under a viewpoint change.<sup>1</sup> Singh et al. minimize the cumulated square distances between the set of detected craters and a virtual set projected from the a priori state estimates of the filter and the 3D crater map.<sup>15</sup> These two techniques only work over terrains which have craters though, but they cannot be found everywhere. Pham et al. solve this problem by working with simple image feature points since they can be detected over any textured terrain. They use a Shape Context description to characterize them by the pixel distance and angular distribution of their neighbors on the orbital image plane.<sup>14</sup> Because of the planar assumption though, this last method is not compatible with rugged terrains having 3D topography. This is the innovation proposed by the current study.

## Issues Faced with Geometrical Descriptors on Feature Points

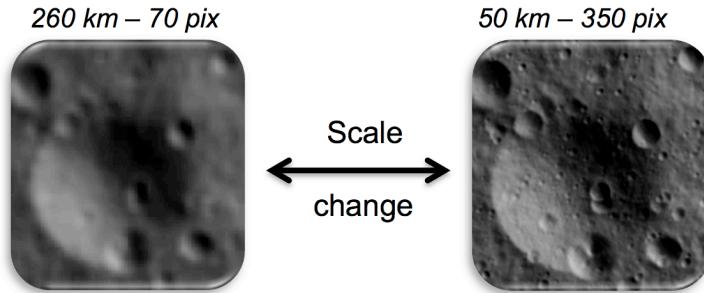
In a previous work, landmarks are extracted using the Harris operator on an orbital image.<sup>3</sup> The 3D world coordinates of the landmarks are then interpolated offline from a DEM of the same area to create the map. During the descent, mapped landmarks are virtually projected on the focal plane of the camera based on the a priori state estimates of the filter. These predicted points are to be matched with the Harris features extracted on the descent image during the landing. Two matching methods were tested: a first one based on Shape Context description (SC), another one computing a General Hough Transform (GHT) to find a global image translation between the two sets of points in the focal plane. Both of them face one major issue: the feature points extracted on the descent image are not necessarily repeated on the orbital image. There are some non-detections, but also some false alarms with descent points which were not detected in the orbital image. These repeatability failures appear even under similar illumination conditions and are due to the differences in distance to the ground, in viewpoint, or in sensor resolution along with other factors such camera sensor noise, impingement of the engine plume, or again the dust cloud it

whips up near the surface. They lead SC to fail to give enough correct matches for the filter to converge. GHT does a lot better because it is a voting process and it is designed to cope with lower feature point repeatability rate. In a 200-run Monte Carlo simulation, 93% of the runs are converging and had a landing error compatible with pinpoint landing. Nevertheless, 7% of the runs of the GHT are still diverging even after a RANSAC check to look for outliers.

To remove these divergent cases, the first educated guess is to work on improving the repeatability rate of selected features by testing different extraction methods than the simple Harris operator. Structured features like craters are interesting because they are not only characterized by their terrain location but also by other geometric parameters like their radius for instance. It provides additional information to detect and match them. But craters are ruled out because they cannot be found on every planet and this study aims at being generic for any type of planetary surface. We investigated if no useful additional information, equivalent to crater radius, could also be found for point features like Harris or SIFT. These can indeed be found on any non-uniform images at locations of local intensity gradients. They are assumed to be due either to terrain topography elements or changes in surface albedo that can be repeated on different images.\* A deeper look at their extraction process highlights a fundamental extraction parameter which appears relevant for landmark matching purposes: the image scale.

## IMAGE FEATURE SCALE

The image scale of an object in the observed scene can be defined as its apparent size in pixels. For a given object size, it is determined by the distance between the camera and the object, the focal length of the lens, and the sensor resolution.<sup>†</sup> Figure 3 shows a typical scaling issue with a crater being imaged by the same camera but at two different altitudes. It illustrates the fact that from an image point of view, a change in scale is equivalent to a change in pixel resolution. The same crater has a scale of 50 pixels on the higher image shot at 260 km of altitude, compared to 250 pixels at 50 km. The level of detail is totally different from one image to the other. The small-scale structures appearing on the 50-km image cannot be seen any more on the higher image.



**Figure 3. Scale change illustration: the same crater is imaged by the same camera at 260 km of altitude (left), or 50 km (right).**

---

\* This is only an assumption since the magnitude of intensity gradients can also become important along shadow edges, which are usually not repeated under different illumination conditions. The system must be robust to this.

<sup>†</sup> Scale can thus be predicted from an estimation of those acquisition parameters in the filter.

## Scale-Space Modeling

Figure 3 underlines the fact that changing the image resolution is equivalent to smoothing it. A smoothing can be modeled by a convolution operation of the reference image array  $\mathbf{I}$  with a Gaussian kernel  $\mathbf{G}$  as shown in Equation (1). Lindeberg formalized the set of images  $\mathbf{L}$  obtained by smoothing  $\mathbf{I}$  under the notion of scale-space.<sup>8</sup>

$$\mathbf{L}(x, y, \sigma) = \mathbf{G}(x, y, \sigma)^* \mathbf{I}(x, y) \quad (1)$$

$x$  and  $y$  are the 2D array coordinates. The Gaussian kernel is reminded in Equation (2).

$$\mathbf{G}(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/\sigma^2} \quad (2)$$

The new image scale  $\sigma$  is defined as the standard deviation of the Gaussian function.

## Scale-Invariant Extractors

Scale-invariant feature detectors select not only a location but also a characteristic scale for each feature. By browsing the scale space for a pair of images of the same scene, they aim at matching common features between them, even when they were taken with different cameras and from a different distance.

Lowe uses a Difference-of-Gaussian (DoG) detector to select SIFT features.<sup>9</sup> The DoG operator basically computes the difference between two smoothed images at nearby scales separated by a multiplicative constant  $k$  as shown in Equation (3).

$$\mathbf{D}(x, y, \sigma) = (\mathbf{G}(x, y, k\sigma) - \mathbf{G}(x, y, \sigma))^* \mathbf{I}(x, y) \quad (3)$$

SIFT features are the local extrema of the  $D$  function within its 3-dimensional definition domain. They have a low computational cost since only a subtraction is needed in addition to the required Gaussian smoothing.

Mikolajczyk et al. proposed the Harris-Laplace detector, inspired from the classical Harris detector, to detect image corners at different scales.<sup>11</sup> It is based on the scale-adapted second moment matrix  $\boldsymbol{\mu}$ . If  $L_i(x, y, \sigma_D)$  is the image derivative in direction  $i$  after a Gaussian filtering at scale  $\sigma_D$ , then this matrix has the form shown on Equation (4). It represents the local gradient distribution after averaging the derivatives through another Gaussian kernel at scale  $\sigma_I$ .\*

$$\boldsymbol{\mu}(x, y, \sigma_I, \sigma_D) = \sigma_D^2 \mathbf{G}(x, y, \sigma_I)^* \begin{bmatrix} L_x^2(x, y, \sigma_D) & L_x L_y(x, y, \sigma_D) \\ L_x L_y(x, y, \sigma_D) & L_y^2(x, y, \sigma_D) \end{bmatrix} \quad (4)$$

A corner is defined as a strong intensity change in two directions. From Equation (4), image point of coordinates  $(x, y)$  is said to be a corner at scale  $\sigma_I$  if  $\boldsymbol{\mu}$  has two large eigenvalues of the same order of magnitude. Avoiding eigenvalue computation, the Harris detector actually selects the local maxima of the cornerness measure of Equation (5) in an image neighborhood sized according to scale.

---

\*  $\sigma_I$  and  $\sigma_D$  are usually set such that  $\sigma_D = 0.7 * \sigma_I$ .

$$C(x, y, \sigma_I, \sigma_D) = \det(\mu(x, y, \sigma_I, \sigma_D)) - \alpha \operatorname{tr}^2(\mu(x, y, \sigma_I, \sigma_D)) \quad (5)$$

$\alpha$  is usually experimentally set in the interval [0.04;0.15]. Note that in the original Harris extractor,  $\sigma_I$  is simply chosen so as to smooth image noise.<sup>5</sup> The Harris-Laplace extractor computes Harris corners over a chosen range of scales but subsequently keeps only those corresponding to an extremum of the Laplacian-of-Gaussian (LoG) operator shown in Equation (6) over scale space. The identified maximum sets the characteristic image scale of the feature when several Harris corners are identified within the same image neighborhood at different scales.

$$|\text{LoG}(x, y, \sigma_I)| = \sigma_I^2 |L_{xx}(x, y, \sigma_I) + L_{yy}(x, y, \sigma_I)| \quad (6)$$

Mikolajczyk et al. tested the repeatability rates of both DoG and Harris-Laplace extractors introduced earlier under scale change.<sup>11</sup> The Harris-Laplace operator was shown to have better performance with a computation time of the same order of magnitude. It is chosen as the baseline scale-invariant extractor for the rest of this study.

### Scale Reprojection Formula

Image features can be associated to actual features in the scene that the camera is observing. In the case of planetary landing, the scene features are called landmarks and they are selected from the orbital image. With scale-invariant extractors, a characteristic terrain scale  $\sigma$  can thus be defined for any landmark based on its characteristic scale on the orbital image  $\sigma_{orb}$ . By projecting it on the image plane of the descent camera, the characteristic scale of the landmark in the descent image  $\sigma_{reproj}$  is given by Equation (7).

$$\sigma_{reproj} = \sigma_{orb} \cdot \frac{d_{orb}}{d_{dsc}} \cdot \frac{\tan \beta_{orb}/2}{\tan \beta_{dsc}/2} \cdot \frac{S_{dsc}}{S_{orb}} \cdot \frac{\cos \alpha_{orb}}{\cos \alpha_{dsc}} \quad (7)$$

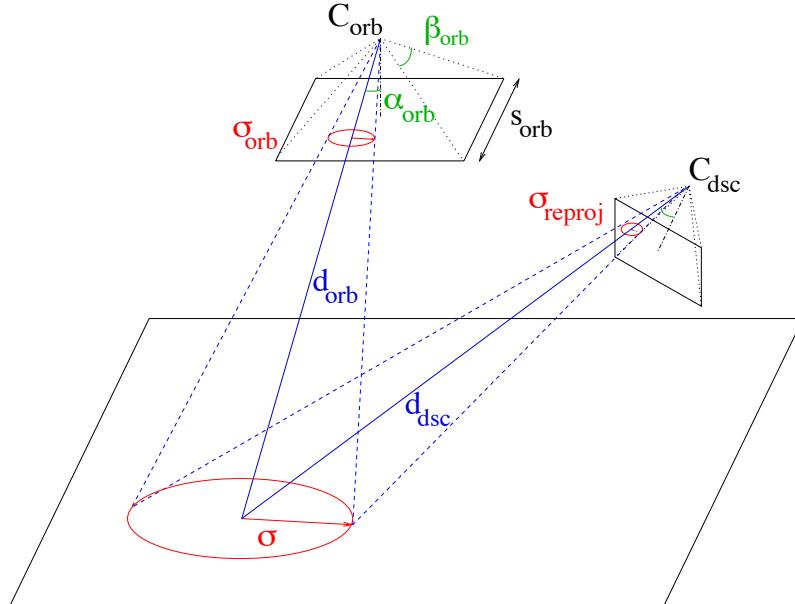


Figure 4. Scale reprojection geometry.

For camera  $i$ ,  $d_i$  is the physical distance between the optical center  $C_i$  and the landmark,  $\beta_i$  is the field of view,  $s_i$  is the sensor resolution in pixel,  $\alpha_i$  is the offset angle from the optical center between the landmark and the optical axis. The geometry of the scene is illustrated in Figure 4.

The formula in Equation (7) can be demonstrated simply using Thales' theorem for each camera within the plane containing the landmark and the optical axis. It is based on a few assumptions though. They are explained below though.

*Isotropic orbital feature gradient.* A corner is defined as a strong and equivalent intensity gradient in at least two directions. Nevertheless, we assume this gradient is strong and the same in every direction so that it can be simply modeled by a circle of radius  $\sigma_{orb}$  in the orbital image.

*Flat level world at feature scale.* The terrain is assumed to be flat and level within the back-projected landmark scale footprint. This flatness assumption is only local around the landmark and did not happen to be a problem during our tests over terrains with strong 3D topography.\*

*Orbital camera orientation.* The orbital camera is assumed to be pointing nadir since most orbital images are taken this way. The landmark scale circle on the orbital image can then be projected onto a characteristic terrain circle of radius  $\sigma$ .

*No affine distortion on descent camera at feature scale.* Affine distortion is a simpler model of image perspective distortion often at feature level in computer vision. An affine transform turns the characteristic landmark terrain circle into an ellipse in the descent image. Equation (7) actually gives the semi-major axis length  $\sigma_{reproj}$  of this ellipse, which defines the reprojected descent scale of the landmark. The affine transform alters the intensity gradient distribution around the landmark in such a way that a corner in the orbital image is not necessarily one in the descent image, even with scale correction. But since a significant part of the descent image is seen with a low incidence angle, affine distortion is neglected and we assume a corner in the orbital corresponds to one in the descent image. This is equivalent to assuming the characteristic landmark terrain circle is projected not into an ellipse but into a characteristic circle of radius  $\sigma_{reproj}$  in the descent image.

## Feature Repeatability with Scale

The orbital vs. descent repeatability rates are compared with and without using a scale-invariant technique for image feature extraction.

*Test setup.* Based on simulated lunar images, two navigation maps were created.<sup>†</sup> The first one uses the settings presented in (Reference 3) to extract Harris corners on an orbital image of the area which are then interpolated with a DEM for 3D world coordinates. The second map is made the same way but with Harris-Laplace features instead and it stores their characteristic scales too. The descent sequence contains 80 images of a lunar approach taken between 2000 and 10 m of altitude, with a 20-deg illumination direction difference in azimuth. Image acquisition rate is 1 Hz. With the Harris map, descent images are processed with the same Harris extractor as the orbital image. With the Harris-Laplace map, each descent image is processed with a multi-scale Harris extractor. It extracts Harris features over the whole reprojected scale range of landmarks computed from the scale reprojection formula of Equation (7). The scale range is discretized by a multiplicative constant  $\sigma_{step} = 1.4$ , such that  $\sigma_n = \sigma_{step}^n \sigma_0$ .

\* Should it become one in other tests, knowledge of the DEM could still be used to get rid of this assumption, but with greater complexity.

<sup>†</sup> For further information about the simulation environment, please refer to Simulation Results.

*Definitions.* On a simulated descent image sequence for which the true camera trajectory is known, the exact image positions of mapped landmarks are computed to assess repeatability of the descent image feature extractor. One landmark is said to be repeated if a descent image feature is found within 10 pixels.\* One Harris-Laplace landmark is said to be repeated with scale if a descent image feature is found within 10 pixels and at a similar scale. A reprojected landmark scale  $\sigma_{\text{reproj}}$  said to be similar to a descent feature scale  $\sigma_{\text{dsc}}$  if it satisfies Inequation (8) which characterizes the discretization domains in scale space.

$$\text{Similarity} \Leftrightarrow \frac{\max(\sigma_{\text{reproj}}, \sigma_{\text{dsc}})}{\min(\sigma_{\text{reproj}}, \sigma_{\text{dsc}})} \leq \sqrt{\sigma_{\text{step}}} \quad (8)$$

Lastly, the orbital repeatability rate is defined as the ratio of the number of repeated landmarks over the total number of landmarks within the field of view.

**Table 1. Repeatability rate comparison between the single-scale feature extractor Harris and the multi-scale Harris-Laplace.**

# Image	Altitude (m)	Harris repeatability (# points, ratio)	Harris-Laplace repeatability / repeatability with scale (# points, ratio)
1	1990	267, 15.4%	(340, 85.9%) / (282, 71.2%)
10	1500	176, 29.6%	(350, 90.2%) / (300, 77.3%)
20	1055	75, 38.5%	(267, 84.5%) / (216, 68.4%)
30	685	271, 13.5%	(138, 95.8%) / (118, 81.9%)
40	395	168, 27.8%	(38, 55.1%) / (22, 31.9%)
50	195	60, 33.9%	(6, 28.6%) / (2, 9.5%)
60	80	7, 25.0%	(0, 0%) / (0, 0%)
70	26	0, N/A	N/A

*Results.* Table 1 compares the orbital repeatability rates between the classical Harris and the Harris-Laplace approaches every 10 images of the descent sequence. The increase of the number of repeated points between image 20 and 30 for the Harris extractor is due to the switch to a higher density map of orbital landmarks as explained in (Reference 3). In both cases, repeatability is not defined after image 70 because the 26-m altitude is so low that no mapped landmark is seen by the camera any more.<sup>†</sup> The Harris-Laplace operator achieves a repeatability rate better or equivalent to Harris all along the descent down to image 50 at 195 m of altitude. At image 30 for instance, repeatability is 95.8% with Harris-Laplace, while only 13.5% with Harris. At image 60 though, Harris-Laplace does not repeat landmarks any more when Harris still does. But results in

---

\* This 10-pixel value was empirically set for 1024x1024 orbital and descent images.

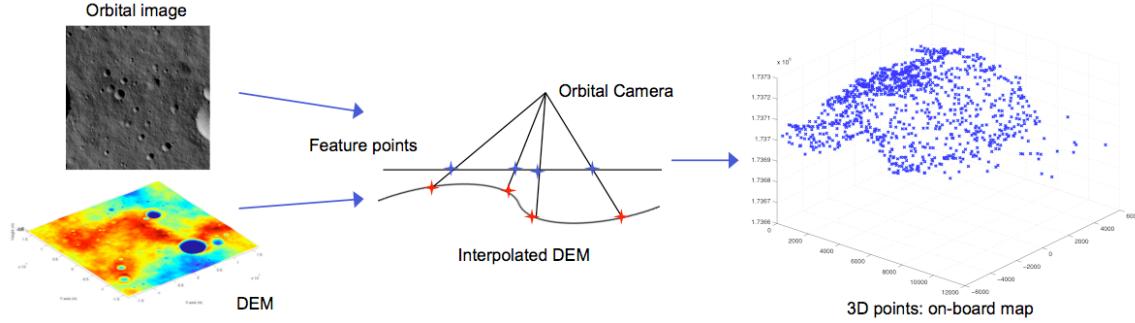
<sup>†</sup> A higher orbital extraction density would solve this problem if needed.

the last images should be read cautiously since the Harris version used had an enforced image extraction density that was not implemented in the Harris-Laplace one. The most interesting result in that table is that up to image 40 more than half of the repeated landmarks are so at the scale computed by reprojecting that from the orbital image with Equation (7), and still at a better rate than regular Harris. This means that by using the a priori state estimate of the filter, the descent image scale of most landmarks can be predicted and thus used as an additional description for matching. The next part introduces a new landmark matching method which takes this landmark scale information into account.

## SCALE AND POSITION-GUIDED LANDMARK SEARCH

A new landmark matching algorithm is described in this part. It uses a priori EKF state estimates to predict the scale and the research area for each landmark in the descent image, in order to select only those which are suitable for matching.

Landmarks are extracted using the Harris-Laplace method on orbital images of the landing area. Since the orbital camera parameters are known, the 3D positions of the landmarks can be interpolated from a DEM of the area. The map is eventually a Nx5 matrix which contains for each landmark its 3D position, the characteristic orbital image scale, and the Harris cornerness score at that scale from Equation (5). It is generated before the mission and loaded in the on-board memory. The overall map generation process is illustrated in Figure 5.



**Figure 5. Map generation process.**

The outline of the online algorithm is divided in 4 main steps which detailed below:

1. Landmark prediction,
2. Landmark selection,
3. Descent image measurement,
4. Outlier removal.

*Landmark prediction.* The position of the landmark in the descent image, its image scale, and research area to look for it, are computed from the EKF estimates of the state of the lander. The

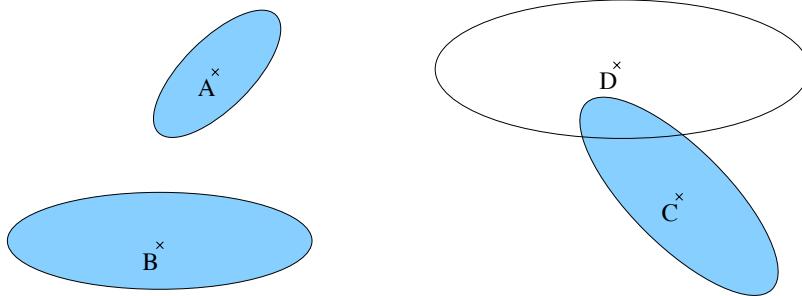
image position  $z_j$  of landmark  $j$  is predicted using the associated projection function  $h_j$  for the state  $x$  as shown in Equation (9).\*

$$\mathbf{z}_j = h_j(\mathbf{x}) = \frac{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{p}_j}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{p}_j} \quad (9)$$

$\mathbf{p}_j$  is the 3D coordinate vector of the landmark in the camera frame. The predicted descent scale  $\sigma_{reproj}$  is computed from Equation (7) using the known orbital parameters. At first order, the descent image position covariance matrix  $\Sigma_j^I$  is derived from the state covariance matrix  $\Sigma$  propagated in the filter by Equation (10).

$$\Sigma_j^I = \mathbf{J}_j \Sigma^T \mathbf{J}_j + \mathbf{Q} \quad (10)$$

$\mathbf{J}_j$  is the Jacobian matrix of the landmark projection function and  $\mathbf{Q}$  is the covariance of the image extraction noise. Equation (10) allows to compute an elliptical area in the image within which the landmark is located with a Gaussian probability of more than 99%.†



**Figure 6. Landmark selection:** Research ellipses of landmarks A and B are not overlapped, they are both selected. Landmark C and D overlap each other, only C is selected since it has the highest cornerness score.

*Landmark selection.* Not all the landmarks visible on the descent are considered suitable for matching. If the elliptical research area of a landmark is not overlapped by another, then it is considered for matching. However when several ellipses cross each other, only that corresponding to the landmark with the higher Harris characteristic score is selected.‡ This selection process is illustrated in Figure 6.

*Descent image measurement.* For each landmark selected, the Harris cornerness criterion of Equation (5) is calculated within the research ellipse with  $\sigma_I = \sigma_{reproj}$  the predicted reprojection scale. The position of the pixel with the maximum value is chosen as the 2D image match for the landmark.

\* This equation is also the measurement model of the EKF.

† Equation (10) and Gaussian distribution of the image position of the landmark are based on a linearization of the projection function which can be assumed when the EFK estimation error is not important.

‡ The characteristic score of a Harris-Laplace feature is defined as the cornerness score at its characteristic scale.

*Outlier removal.* The set of map-to-image landmark matches is verified to eliminate outliers, i.e. mismatches. To do so, the consistency of the matches with the camera projection matrix model of Equation (9) is checked using the RANSAC algorithm.<sup>4</sup> RANSAC is a proven technique in real-time visual odometry terrestrial robotics. It randomly selects sets of 3 match pairs from which a camera model can be computed, counts the number of other matches which agree with that model, and repeats this process a sufficient number of times to ensure with a probability of more than 99% that the model with most inliers is the correct one. Associated inliers form the final set of matches which is fed as a measurement to the navigation filter.

## SIMULATION RESULTS

*Software simulation setup.* Based on the guidance scheme of the Apollo Lunar Module, a lunar landing simulation software was designed.<sup>16</sup> An approach phase scenario starting at a 2-km altitude, ending 10 m above the landing site, and lasting 80 seconds, was implemented on it to test low-altitude navigation capability.<sup>2</sup> Inertial data were generated through an IMU model calibrated to match performances expected from the hardware used on this kind of mission.

Descent trajectory images have been rendered in the PANGU planetary scene generator.<sup>13</sup> A NASA-LRO DEM was scaled to obtain height variations of 500 m and test robustness to rugged terrain. An image example from the descent sequence is shown in Figure 7. It covers a field of view of 70 degrees with a 1024x1024-pixels 8-bit sensor considered realistic of landing camera. The simulated orbital images were taken from a 50-km altitude with 1024x1024 sensor behind a lens providing a 2.85-degree field of view. The elevation of the Sun above the horizon at the targeted landing site was 15 degrees in both the orbital and descent images, but a 20-deg azimuth difference was implemented. Pixel values in descent images were degraded with an Gaussian noise of zero mean and a standard deviation equal to one intensity level.

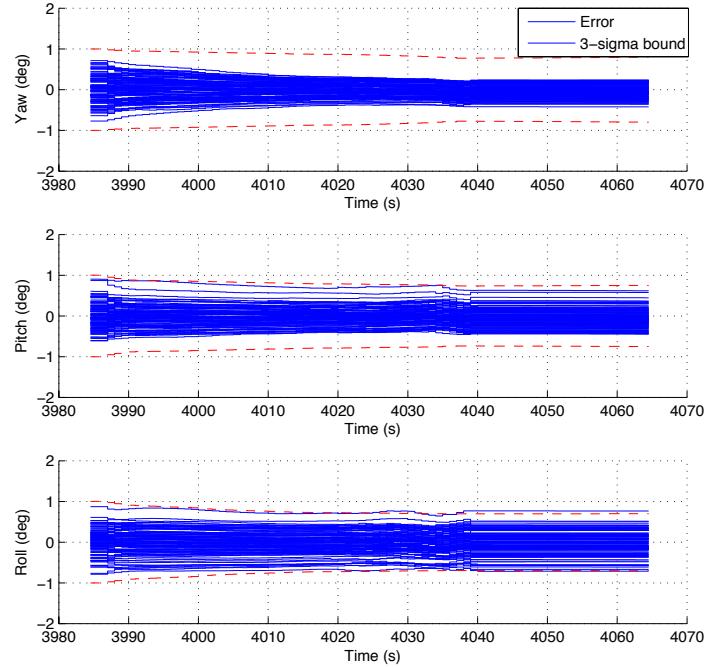


**Figure 7. Descent image generated with PANGU.**

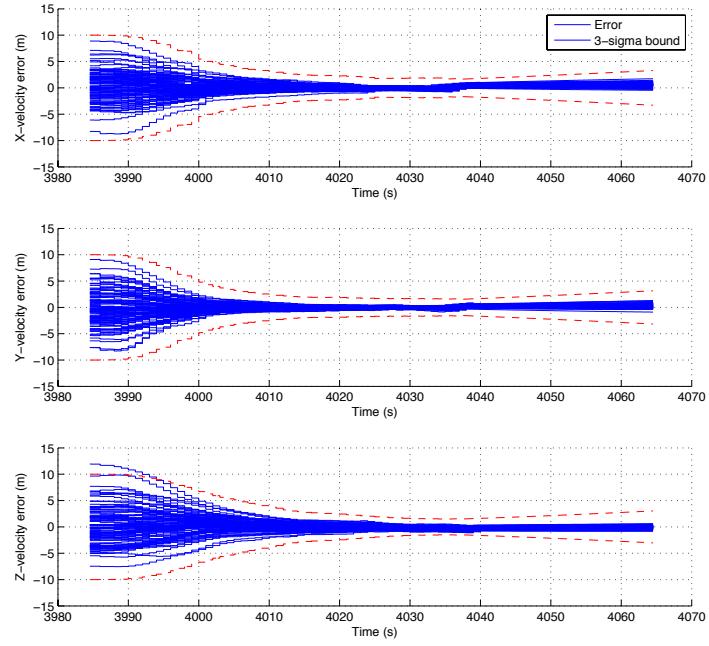
*Performance evaluation.* A preliminary Monte Carlo analysis over 100 runs of the whole navigation using the landmark matching method presented in this paper was performed. The initial

navigation errors were according to independent zero-mean Gaussian distribution on each axis, with  $3\sigma$  values equal to 1 degree in attitude, 10 m/s in velocity and 100 m in position. Such values were considered as representative of expected navigation uncertainties at this stage of the descent. The Harris-Laplace extraction on the orbital images browsed through 25 different scales between 1 and 80 pixels with a multiplicative factor of 1.2 between each. Only Harris corners with scores above 0.2 that of the image maximum were selected at each scale.

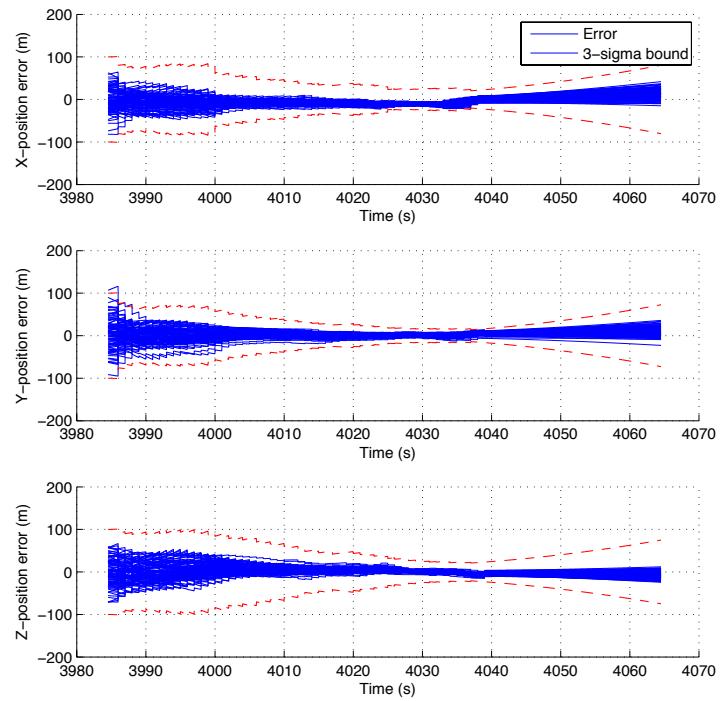
Figure 8, Figure 9, Figure 10 and show the results of the estimation error and the largest  $3\sigma$  bound from the filter for each axis respectively for attitude, velocity and position. The most significant result is that all the runs are now converging and the 7% divergent cases previously found with the GHT method are now eliminated.



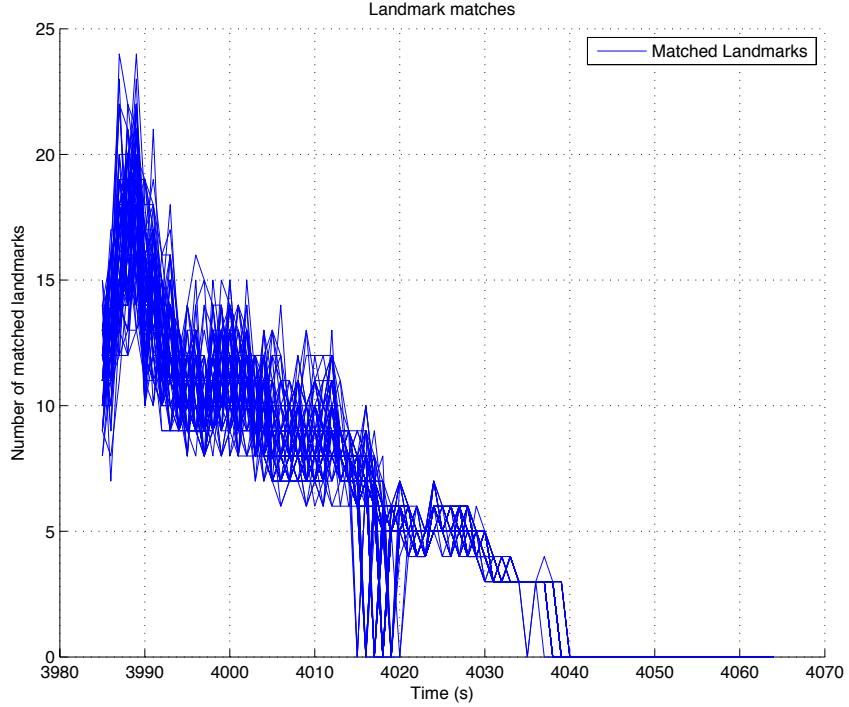
**Figure 8. Attitude errors Monte Carlo results.**



**Figure 9. Velocity errors Monte Carlo results**



**Figure 10. Position errors Monte Carlo results.**



**Figure 11. Number of matches Monte Carlo results.**

Figure 11 shows the evolution of the number of matches fed to the filter with time. Because of the limited number of landmarks on the map, less and less of them appear in the field of view of the descent camera as the lander progresses down its descent trajectory. Thus at  $t=4040$  s with the current map extraction density settings, no landmark can be identified by the system any more at an altitude of 80 m. Up to  $t = 4040$  s is the so-called visual phase, when image measurements are processed by the filter. After  $t = 4040$  s is the inertial phase when navigation estimates are propagated using only IMU data which suffer drift that increase the error. Statistics for the Monte Carlo runs are given in Table 2.

**Table 2.  $3\sigma$  uncertainties at the end of the visual phase (V) and at touchdown (TD).**

Variable	$3\sigma_V$	$3\sigma_{TD}$
Attitude / Yaw (deg)	0.5	0.5
Attitude / Pitch (deg)	0.7	0.7
Attitude / Roll (deg)	0.9	0.9
Velocity / X-axis (m/s)	0.7	1.2
Velocity / Y-axis (m/s)	0.8	1.2
Velocity / Z-axis (m/s)	0.9	0.9
Velocity norm (m/s)	1.0	1.1
Position / X-axis (m)	12.9	35.3
Position / Y-axis (m/s)	11.8	32.9
Position / Z-axis (m)	10.1	23.9
Position norm (m)	12.0	36.1

The norm of the position error at touchdown is 36.1 m which largely meets the 100-m requirement for pinpoint landing. This error is even smaller at the moment of the last landmark matching since it falls down to 12.0 m before inertial drift occurs.

## CONCLUSION

A new mapped landmark matching method has been proposed to tackle challenges associated to planetary landing. It was designed to be compatible with 3D terrain topography likely to be met at low altitude, and tested from 2000 m of altitude to touchdown on a terrain with 500-m height variation. It is also robust to the 20-degree illumination changes between the map and descent images. It uses the notion of image scale as an additional description for matching Harris-Laplace feature points from an orbital image to descent one. Performance found through Monte Carlo software simulation show it is so far compatible with pinpoint landing requirements.

In the future, we plan to take affine distortions of the descent camera into account to compare the performance with the current assumption that they can be neglected. Further simulation testing is also needed with different trajectories, more noisy image conditions, a broader range of illumination conditions and with different terrain topography. A vision lunar test bench is also currently under design at ESA-ESTEC to test all the trajectory from orbit to touchdown with hardware in the loop.

## ACKNOWLEDGMENTS

The authors would like to thank ONERA, the European Space Agency, and EADS Astrium Space Transportation for their technical and financial support to this study.

## REFERENCES

- <sup>1</sup> Y. Cheng and A. Ansar, "A Landmark Based Position Estimation for Pinpoint Landing on Mars" *International Conference on Robotics and Automation*. 2005, pp. 1573-1578.
- <sup>2</sup> J. Delaune, D. De Rosa et al., "Guidance and Control system design for lunar Descent and Landing" *AIAA Guidance, Navigation and Control Conference and Exhibit*. Toronto, Ontario, Canada, 2010.
- <sup>3</sup> J. Delaune, G. Le Besnerais et al., "Camera-Aided Inertial Navigation for Pinpoint Planetary Landing on Rugged Terrains" *8<sup>th</sup> International Planetary Probe Workshop*. Portsmouth, VA, USA, 2011.
- <sup>4</sup> M. A. Fischler and R. C. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography" *Communications of the ACM*. Vol. 24, No. 6, 1981, pp. 91-110.
- <sup>5</sup> C. Harris and M. Stephens, "A Combined Corner and Edge Detector" *4<sup>th</sup> Alvey Vision Conference*. 1988.
- <sup>6</sup> R. Hartley and A. Zisserman, *Multiple View Geometry in computer vision*. Cambridge University Press, Second Edition, 2003.
- <sup>7</sup> A. E. Johnson and J. F. Montgomery, "Overview of Terrain Relative Approaches for Precise Lunar Landing" *IEEE Aerospace Conference*. Big Sky, MT, USA, 2008.
- <sup>8</sup> T. Lindeberg, "Scale-space theory: A basic tool for analyzing structures at different scales" *Journal of Applied Statistics*. Vol. 21, No. 2, 1994, pp. 225-270.
- <sup>9</sup> D. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints" *International Journal of Computer Vision*. Vol. 60, No. 2, 2004, pp. 91-110.
- <sup>10</sup> E. Mazarico, D. Rowlands et al., "Orbit determination of the Lunar Reconnaissance Orbiter" *Journal of Geodesy*. Published online: 8 September 2011.
- <sup>11</sup> K. Mikolajczyk and C. Schmid, "Scale & Affine Invariant Interest Point Detectors" *International Journal of Computer Vision*. Vol. 60, No. 1, 2004, pp. 63-86.
- <sup>12</sup> A. I. Mourikis, N. Trawny et al., "Vision-Aided Inertial Navigation for Spacecraft Entry, Descent and Landing" *IEEE Transactions on Robotics*. Vol. 25, No. 2, 2009, pp. 264-280.
- <sup>13</sup> S. M. Parkes, I. Martin et al., "Planet Surface Simulation with PANGU" *8<sup>th</sup> International Conference on Space Operations*. 2004.
- <sup>14</sup> B. V. Pham, S. Lacroix et al., "Visual Landmark Constellation matching for spacecraft pinpoint landing" *AIAA Guidance, Navigation and Control Conference and Exhibit*. Chicago, IL, USA, 2009.
- <sup>15</sup> L. Singh and S. Lim, "On Lunar on-orbit Vision-Based Navigation: Terrain Mapping, Feature Tracking driven EKF" *AIAA Guidance, Navigation and Control Conference and Exhibit*. Honolulu, HI, USA, 2008.
- <sup>16</sup> R. R. Sostaric and J. R. Rea, "Powered Descent Guidance for the Moon and Mars" *AIAA Guidance, Navigation and Control Conference and Exhibit*. San Francisco, CA, USA, 2005.
- <sup>17</sup> B. Vanoutryve, D. De Rosa et al., "An Analysis of Illumination and Communication Conditions near Lunar South Pole based on Kaguya Data" *7<sup>th</sup> International Planetary Probe Workshop*. Barcelona, Spain, 2010.