Derivation of planetary topography using multi-image shape-from-shading

Volker Lohse\textsuperscript{a},*, Christian Heipke\textsuperscript{a}, Randolph L. Kirk\textsuperscript{b}

\textsuperscript{a}Institute of Photogrammetry and GeoInformation (IPI), University of Hannover, Nienburger Strasse 1, 30167 Hannover, Germany  
\textsuperscript{b}United States Geological Survey, Astrogeology Team, 2255 N. Gemini Dr. Flagstaff, AZ 86001, USA

Received 28 January 2005; received in revised form 20 March 2006; accepted 22 March 2006

Abstract

In many cases, the derivation of high-resolution digital terrain models (DTMs) from planetary surfaces using conventional digital image matching is a problem. The matching methods need at least one stereo pair of images with sufficient texture. However, many space missions provide only a few stereo images and planetary surfaces often possess insufficient texture.

This paper describes a method for the generation of high-resolution DTMs from planetary surfaces, which has the potential to overcome the described problem. The suggested method, developed by our group, is based on shape-from-shading using an arbitrary number of digital optical images, and is termed “multi-image shape-from-shading” (MI-SFS). The paper contains an explanation of the theory of MI-SFS, followed by a presentation of current results, which were obtained using images from NASA’s lunar mission Clementine, and constitute the first practical application with our method using extraterrestrial imagery. The lunar surface is reconstructed under the assumption of different kinds of reflectance models (e.g. Lommel–Seeliger and Lambert). The represented results show that the derivation of a high-resolution DTM of real digital planetary images by means of MI-SFS is feasible.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Image processing; Shape-from-shading; Terrestrial planets; Moon surface; Digital terrain models

1. Introduction

High-resolution digital terrain models (DTMs) are an important information source for many applications in planetary sciences, such as the selection of future landing sites, the description of local and regional topographic features, slopes, discontinuities of the surface and thus possible flow directions of liquid material and isostatic considerations to name only a few. On Earth, such DTMs can normally be generated by means of conventional photogrammetry including digital image matching. But planetary missions in general are not topographic missions, and therefore usually only a few stereoscopic images are available. Furthermore, often the few available stereo pairs are acquired with a time lag. Therefore, it is possible that the observed area has changed due to external influences, e.g. by large-scale atmospheric changes and variable surface features. Additionally, many planetary stereo pairs have a disadvantageous camera configuration (e.g. a poor base-to-height ratio or different image resolutions) and in many cases planetary images have poor image texture, which is a further obstacle to automatic matching methods. For these reasons, in many regions no complete high-resolution DTM of planetary bodies can be made available by means of conventional photogrammetric methods.

Besides photogrammetry, there are other methods to derive the topography of planetary bodies. One example is laser scanning as employed by the “Mars Orbiter Laser Altimeter” (MOLA) flown on the Mars Global Surveyor mission (Smith et al., 2001). MOLA acquired high-precision height information along one-dimensional tracks, but despite simultaneous processing of multiple orbits, the horizontal resolution and accuracy of the resulting data set is rather limited.
For these reasons, there has been ongoing interest among planetary scientists in shape-from-shading (SFS, also commonly called photoclinometry in the planetary community) techniques that infer surface shape from image grey values. A variety of SFS algorithms have been described in the literature, most of which require only one image as input. Single-image approaches allow mapping in the absence of overlapping image coverage, but they are based on the assumption that the surface albedo is uniform and will yield erroneous results where the albedo varies significantly. In this paper we describe an approach for multi-image shape-from-shading (MI-SFS), which takes an arbitrary number of images as input. Where multiple images are available, the method can be extended to solve for spatially variable albedo in addition to topography (although the examples presented here assume a constant albedo), but the same approach can also be applied to a single image when needed. As applied to multiple images, the method can be seen as a framework for combining the approaches of SFS and image matching. There are numerous potential benefits to such a combination. At a minimum, the use of multiple images reduces the sensitivity to measurement noise. More importantly, the SFS component strengthens the topographic solution in unfavourable situations for stereo (i.e. poor base-to-height ratio or significantly different viewing directions), and can fill in the shape of areas of modest size that have insufficient texture for image matching. Conversely, image matching constrains absolute elevations, which SFS cannot. The spatial resolution properties of the SFS and stereo components of the algorithm are also highly complementary. Image matching produces independent position estimates only every few pixels, but SFS can, in principle, achieve a lateral height resolution equal to that of the images (i.e. single-pixel). The vertical resolution of SFS can be a small fraction of the pixel width, limited by the relative radiometric resolution of the images. Finally, on bodies such as Mars with a significant atmosphere, SFS will miss-estimate the amplitude of relief unless atmospheric scattering of light (which can vary unpredictably from one image to another) is accounted for. The image-matching part of our algorithm provides the independent constraint on elevations needed to make such an atmospheric correction. The MI-SFS algorithm provides all these potential benefits where multiple images are available, while still permitting mapping in areas of single image coverage and constant albedo.

This paper is organised as follows. Section 2 gives a review of the state-of-the-art of the derivation of topographic information of planetary surfaces by means of SFS methods. Section 3 describes the MI-SFS approach developed by our group. Different reflectance models are presented and the used radiometric and geometric surface model are introduced and discussed in detail. MI-SFS is cast into an optimisation approach, which is solved using least-squares adjustment. Section 4 deals with the investigations of MI-SFS using Clementine images. After describing the input imagery, the test design and the obtained results are presented and discussed. Finally, Section 5 gives conclusions drawn from the described work and an outlook on future activities.

2. Related work

2.1. General considerations

For solving the DTM reconstruction problem by SFS, the image formation process has to be modelled and inverted with respect to the parameters describing the DTM. SFS is based on the fact that surface patches, having different inclination relative to a light source, are imaged with different brightness. SFS uses these variations in the grey values of the digital images to reconstruct the topography of the observed surface.

Using a single digital image to reconstruct a DTM with SFS, the result is ambiguous: the inclination of a surface patch is determined by two components (viz. the slopes $p$ in $x$ and $q$ in $y$ direction) while only one observation, namely the grey value of the patch, is available. For continuous surfaces the integrability constraint (i.e. the derivative of $p$ with respect to $y$ equals derivative of $q$ with respect to $x$) must be fulfilled, and thus corresponding system of equations can be solved. However, when integrating slope to compute heights, the boundary conditions of the integration remain undetermined. In order to overcome this problem, absolute heights at the border of the surface or alternatively the absolute average height of the area can be used.

For SFS it is generally assumed that the albedo is constant within the observed area, because the method does not distinguish between albedo and topographic variations as a reason for grey value changes in image space. As discussed above, the opportunity to relax this assumption of constant albedo is one of the main motivations for multi-image approaches to SFS.

2.2. Summary of existing approaches

The roots of SFS can be found in planetary science (Rindfleisch, 1966) where the measurement of topography from shading was also referred to as photoclinometry, and in the computer vision community (Horn, 1970). These two areas were developed independently of each other for a number of years.

In computer vision, SFS was mostly considered under theoretical aspects, such as the existence and uniqueness of a solution under various constraints, and different iterative schemes for solving the problem were suggested (Horn and Brooks, 1989). Other methods have been developed for the reconstruction of surfaces in terrestrial and close-range surroundings (Fua, 1997; Lee and Kuo, 1997).

In planetary science the main research interest in SFS lies in the geometric surface reconstruction. In the beginning work concentrated on first deriving surface profiles, which
were connected in a separate step in order to generate a two-dimensional surface (Rindfleisch, 1966; Davis and Soderblom, 1984; Schenk, 2002). Already Wildey (1975) and later Kirk (1987), Giese et al. (1996), Kirk et al. (2003) and others derive a two-dimensional surface in one step.

A survey of SFS as applied in computer vision is contained in Zhang et al. (1999), where the authors subdivide the SFS algorithms published in the literature into minimisation, propagation, local and linear approaches, and investigate six different algorithms in more detail. All algorithms are based on orthographic projection from image to object space and make use of one image only. Furthermore they are based on Lambert reflection, and assume a single point light source at an infinite distance. Tests were performed using synthetic and real images. According to the authors the algorithms do not meet the expectations in terms of accuracy and predictability of the results. Reasons given for these disappointing results include inappropriate geometric and photometric modelling, and the lack of redundant information.

Recent work about SFS-based reconstruction of parts of the Earth surface is described by Liu (2003). Linear SFS and minimisation SFS algorithms are applied to a panchromatic image of the Ross Ice Shelf of Antarctica acquired by the SPOT 3 satellite. The terrain surface in this region is covered by permanent ice and snow. Liu shows that the two implemented SFS algorithms can produce a dense, consistent elevation grid using only a single image. Levin et al. (2003) develop and apply a SFS method for extracting height information of sand dunes in Israel from shading effects as seen on Landsat images.

Another recently published approach deals with separating brightness changes due to shape from those caused by variable albedo (Wöhler, 2004). Surface reconstruction of regions on the lunar surface is carried out with images acquired with Earth-based telescopes equipped with CCD sensors. The approach needs at least two shadow-free images of one scene acquired under different illumination conditions and assumes that the observed surface has the properties of Lambert reflectance. Dorrer (2002) and Dorrer et al. (2004) present investigations for the generation of highly accurate DTMs of the Martian surface with Viking Orbiter Imagery and those from the High Resolution Stereo Camera (HRSC) orbiting planet Mars onboard the current ESA mission Mars-Express. The developed approach called “de-re-shading” (DRS)—which contains SFS—removes the actual shading in the original image and then re-shades the image with an artificial light source to produce orthophoto mosaics from images taken at different illumination directions and to refine the DTM. DRS is based on Lambert’s reflectance model. Kirk et al. (2004) show an approach to produce high-resolution DTMs of the Mars surface by photoclinometry and stereoanalysis using Super-Resolution Channel (SRC) images of HRSC, together with Mars Orbiter Camera images. The presented approach can use different reflectance functions, and it is able to fuse matching and photoclinometry results in a consistent manner (see also Kirk et al., 2003). Another interesting method is the one for integrating spacecraft navigation and the determination of small body dynamics, shape, and high-resolution topography by Gaskell (2003). In this approach multiple image stereography and photoclinometry are used to construct high-resolution topography and albedo maps. These maps can be re-illuminated and correlated with images to act as body-fixed navigation tie-points. Their limb projection can be compared with observed limb profiles to better fix their locations and the spacecraft orientation and position.

Gelli and Vitulano (2004) describe a method which increases the accuracy of a surface obtained by SFS by exploiting the light direction information. Edges of concave regions are split in projecting and projected points. The geometric relation between these points allows for the introduction of an additional constraint into the SFS algorithm.

The final goal of many suggested approaches is to develop a combined reconstruction method integrating the two complementary methods image matching and MI-SFS, as already suggested by Bülthoff and Mallot (1988), Heipke (1992a) and Fua and Leclerc (1995). Similar ideas to combine stereoscopic methods and SFS methods, however in the area of archaeology, are investigated by Danzl et al. (2004). If an archaeological object can or shall not be removed from the site where it was found the archaeologist sometimes generates a digital surface model of this object. Because the objects are often rather textureless conventional matching methods reach their limits. In order to solve this problem the researchers propose to combine the stereoscopic methods with SFS.

2.3. Conclusions

In the field of SFS a lot of research deals with the reconstruction of object surfaces in close-range, with the derivation of DTMs of the Earth using aerial images and with the determination of 1D-profiles of planetary surfaces. Only the work of Dorrer, Kirk, Schenk and a limited number of other colleagues deal with the generation of high-resolution DTMs of planetary surfaces. Nevertheless, in the absence of sufficient image texture—a common case for planetary images—most alternative schemes, and in particular image matching algorithms fail. Another important aspect for further intensive investigations in the area of the derivation of planetary surfaces using SFS is the increasing interest of the NASA and the ESA to explore the planets of our solar system, and to design topographic mapping missions such as HRSC (Neukum et al., 2004) on Mars-Express.

Some of the shortcomings of currently existing algorithms include the nearly exclusive use of orthographic as opposed to perspective projection, and of the Lambert reflectance model (the latter only in computer vision), the lack of modelling shadows, occlusions, and breaklines, and the disability to simultaneously and synchronously use
multiple images to improve the reliability of the computed result. Also, most algorithms are not capable to derive absolute heights instead height differences are computed only.

With a view to the upcoming topographic mapping missions, it thus makes sense to develop a method which derives high-resolution planetary DTMs from multiple, low texture images, to take into account occlusions, shadows, and breaklines of the terrain, and to embed this method into a combined approach dealing with image matching and SFS. Our approach described in the Section 3 is a step into this direction.

3. Multi-image shape-from-shading

In contrast to classical SFS methods, our MI-SFS method can deal with an arbitrary number of images and spectral bands. In principle, it can also take into account different types of geometric transformations between image space and object space1 such as perspective or push-broom, however our current implementation is exclusively based on the perspective transformation. MI-SFS relates directly the recorded grey values to the heights of a two-dimensional DTM and the parameters of a radiometric model, which describes the surface reflectance behaviour. The DTM heights as well as the parameters of the radiometric model are estimated from the image grey values in a least-squares adjustment.

The surface reconstruction method MI-SFS has been developed and studied intensively by our group over the last several years. Detailed descriptions and analysis with simulated and aerial images from a desert area on Earth are given in Heipke (1992a), Heipke and Piechullek (1994), Piechullek (2000) and Heipke et al. (2001).

3.1. Reflectance models for planetary surfaces

Like other SFS approaches, we need to make assumptions about the illumination and reflectance properties of the investigated surface. The illumination is assumed to come from a distant point light source (the sun). For perspective images the direction of illumination, which we denote by the vector s, is considered to be known and constant since the image is captured in a fraction of a second. For push-broom imagery s is assumed to be known as a function of time. As already mentioned, the Lambert law is one of the simplest reflectance models. It is given by

\[ r_L(i) = A_L(X, Y) \cdot \cos(i). \]  

\[ A_L(X, Y) \] stands for the Lambert surface albedo, which is assumed to be space-variant, i for the incidence angle (the angle between the direction of illumination s and the surface normal n, see Fig. 1), and \( r_L \) is the bi-directional reflectance (BDR). The BDR2 is the ratio between the irradiance \( I_S \) scattered towards the sensor and the scene irradiance \( E_S \), with which the surface is illuminated. More information of the Lambert law and its connection to SFS can be found, e.g. in Horn (1986), Hapke (1993) and Zhang et al. (1999).

The Lambert law describes a surface, which emits the incoming irradiance uniformly in all directions. This means, that a surface looks equally bright from every viewing direction. The larger the incidence angle \( i \) the bigger the irradiated area on the surface becomes, and therefore, less brightness can be reflected per surface element towards the sensor (Fig. 2). The Lambert model characterises the reflectance from bright surfaces very well.

In order to extend the assumption that light reflection occurs at the boundary surface between two media only, the Lommel-Seeliger law was derived by Seeliger (Horn and Brooks, 1989; Hapke, 1993; Rebhan, 1993). The model derivation is based on the probability with which a photon is scattered within a surface layer (rather than on the boundary between the atmosphere and the surface, as is the case when using the Lambert law), and then captured by the sensor. The Lommel-Seeliger law (Eq. (2)) does not only contain the incidence angle \( i \) but also the emittance angle \( e \), the angle between viewing direction \( v \) and the surface normal \( n \) (Fig. 1).

\[ r_{LS}(i, e) = A_{LS}(X, Y) \cdot \frac{\cos(i)}{\cos(i) + \cos(e)}. \]  

\( r_{LS}(i, e) \) stands for the Lambert-Seeliger reflectance (

1We model the image space by two dimensions and denote the coordinate axes with lower case letters (\( x', y' \)). The object space coordinate system \((X, Y, Z)\) we use is Cartesian, a map projection of any kind is not further considered.

2The BDR should not be mixed up with the bi-directional reflectance distribution function (BRDF). For a Lambertian surface BRDF is constant, while BDR depends on the incidence angle: \( \text{BDR} = \text{BRDF} \cdot \cos(i) \).
A$_{LS}$ is the space-variant Lommel–Seeliger albedo and $r_{LS}$ the corresponding BDR. The significant increase in brightness for large $e$ is due to the fact that with increasing $e$ the area of the imaged surface also increases, and consequently a greater part of the surface layer contributes to the brightness observed in the sensor (Fig. 3). Although at first sight the Lommel–Seeliger law seems to contradict intuition, because a surface patch appears brighter with increasing viewing angle rather than darker, the law has been widely and successfully employed in planetary photoclinometry as part of the Lunar-Lambert law (McEwen, 1986, 1991). In contrast to the Lambert law, the Lommel–Seeliger law describes dark surfaces better.

It should be noted that while both, the Lambert and the Lommel–Seeliger law are being used in SFS, a photometric function in general also depends on the phase angle $\alpha$, the angle between $s$ and $v$ (see again Fig. 1). This can be shown rather easily: imagine that all three vectors ($s$, $n$, $v$) lie in a vertical plane. For given values of $s$ and $n$ and thus a constant incidence angle, there are two possibilities for $v$ with identical emittance angle but with different phase angle, the image taken into the light and the one taken with the light. Experience tells us that the two resulting images look significantly different. Thus, the phase angle must also enter into a general model for photometric reflection. Considerable effort has been devoted to developing models of the photometric behaviour of planetary surfaces that are physically motivated and have parameters that can be physically interpreted (e.g. Hapke, 1981, 1984, 1986). These and similar models are routinely used for studying surface properties and for removing the effects of varying photometric angles from image mosaics. Unfortunately, physically motivated photometric models are generally too complex to be used in SFS calculations.

We therefore take the approach suggested by McEwen (see above) of using a Lunar-Lambert photometric function as a computationally efficient empirical approximation of more complex photometric behaviour. In this model the reflectance function is a linear combination of the Lambert and the Lommel–Seeliger model

$$r_{LL}(i, e, \alpha) = A_N(X, Y, \alpha) \cdot \left[ \frac{A(z)}{\cos(i) + \cos(e)} + (1 - A(z)) \cdot \cos(i) \right].$$

(3)

$A_N$ is the so-called normal albedo and is also allowed to vary with surface position. It is in general set up as a function of the phase angle $\alpha$ in order to model the phase curve dependence of the reflected intensity on the phase angle and thus the normal albedo is different for each image. $A_N$ is equivalent to the BDR of a surface with vertical illumination and viewing direction ($i = e = 0^\circ$) and describes the overall brightness of the surface. When comparing Eq. (3) with Eqs. (1) and (2), and Fig. 2 with Fig. 3 it can be seen that for a vertical image of a horizontal surface patch ($e = 0^\circ$) with illumination directly from above ($i = 0^\circ$) the BDR of the Lommel–Seeliger model $r_{LS}$ is only half of the value of that for the Lambert model $r_{L}$. This is the reason why in the Lunar-Lambert model the Lommel–Seeliger term is multiplied by 2.

The parameter $A(z)$ is known as the limb-darkening parameter and controls the influence of the Lambert term relative to the Lommel–Seeliger term, it also depends on the phase angle $\alpha$. $A(z)$ influences the contrast of shading on topographic slopes and, to a lesser extent, the direction relative to the sun and the observer, in which slopes have the maximum contrast. Values of $A(z)$ reported in the literature, based on prior photometric observations, may be interpolated to the appropriate phase angle. Alternatively,
the Lunar-Lambert function may be fitted to another reported function (e.g. Hapke’s) at the desired phase angle, with \( A(z) \) adjusted to optimise the fit (McEwen, 1991). For the latter case a 3rd-order polynomial with constant parameters \( A, B \) and \( C \) derived for the Moon by McEwen (1996) can be used

\[
A(z) = 1.0 + A \cdot z + B \cdot z^2 + C \cdot z^3
\]

with \( A = -0.019; \quad B = 0.242E - 3; \quad C = -1.46E - 6. \) (4)

3.2. Camera equation

In this section we derive the formulae for describing the so-called model grey value of a pixel \( G(x', y') \) as a function of the slope and the photometric properties of a surface. Using the Lunar-Lambert law, the model grey value \( G(x', y') \) in image space can be formulated based on the well-known camera equation (e.g. Horn, 1986) under the assumption of a point light source at infinite distance and known illumination direction. We extend this equation in order to take into account offset and gain of the camera as well as two space-invariant parameters \( T \) and \( H \) for describing the effects of the atmosphere

\[
G(x', y') = O(x', y') + k(x', y') \cdot \frac{\pi}{4} \cdot \cos^2(\gamma)
\]

\[
\left( \frac{d}{r} \right)^2 \cdot E_S \cdot \left\{ A_N(X, Y, z) A(z) \left( \frac{2 \cdot \cos(i)}{\cos(i) + \cos(e)} \right) \right.
\]

\[
+ \left. (1 - A(z)) \cdot \cos(i) \right) \cdot T + H \}
\]

(5)

with

\[
\cos(e) = \frac{n \cdot v}{|n| \cdot |v|}; \quad \cos(i) = \frac{n \cdot s}{|n| \cdot |s|}
\]

where \( G(x', y') \) is the model grey value at image point \( P' \), \( x' \) and \( y' \) the image coordinates of \( P' \) (projection of point \( P \) into image space), \( O(x', y') \) the camera offset (includes dark current), \( k(x', y') \) the camera gain factor (allowed to vary from pixel to pixel), \( n \) the exponent of light fall off, \( \gamma \) the angle between optical axis and the ray through \( P \) and \( P' \), \( d \) the aperture of optical lens, \( f \) the focal length of optical lens, \( E_S \) the scene irradiance, \( A_N(X, Y, z) \) the normal albedo of the object surface at point \( P(X, Y, Z) \), \( z \) the phase angle, \( A(z) \) the limb-darkening parameter (assumed to be space-invariant), \( T \) the factor for modelling multiplicative effect of atmospheric transmission, \( H \) the additive contribution of light scattered off the atmospheric haze, \( n \) the normal vector of the object surface at point \( P(X, Y, Z) \), \( s \) the vector in illumination direction at point \( P(X, Y, Z) \) and \( v \) the vector in viewing direction at point \( P(X, Y, Z) \).

In planetary science (as in other fields using CCD cameras) it is common practice to obtain an estimate of the offset \( O(x', y') \) and to use the so-called flat-field observations to determine the pixel-to-pixel variations of relative sensitivity as well as the absolute sensitivity of the camera. Then, a radiometrically calibrated image may be constructed as

\[
G_c(x', y') = \frac{(G(x', y') - O(x', y'))}{k(x', y') \cdot \frac{\pi}{4} \cdot \cos^2(\gamma) \cdot \left( \frac{d}{r} \right)^2 \cdot E_S}
\]

\[
= A_N(X, Y, z) \left( A(z) \frac{2 \cdot \cos(i)}{\cos(i) + \cos(e)} \right) \left( T + H \right).
\]

(6)

\( G_c(x', y') \) is the model grey value of the radiometrically calibrated image at image point \( P' \).

Clearly, \( T \) can be grouped with the unknown normal albedo \( A_N \), while \( H \) could be solved for during the calculation, with the information from stereo matching providing the constraint needed to determine it. The examples in the paper, however, are for the Moon, which is airless, so that \( T = 1, H = 0 \). Furthermore, we assume for the present work a constant normal albedo. Hence the model becomes

\[
G_c(x', y') =
\]

\[
A_N(z) \left( A(z) \frac{2 \cdot \cos(i)}{\cos(i) + \cos(e)} + (1 - A(z)) \cdot \cos(i) \right), \quad (7)
\]

Thus, the model grey value \( G_c \) of the radiometrically calibrated image depends on \( A_N, A, s, v \) and \( n \).

3.3. Geometric and radiometric surface model

For the purpose of object surface description, a geometric and a radiometric surface model are introduced (Fig. 4). The
mathematical description of the geometric model is given by means of a DTM with a simple grid structure, which is defined in the XY-plane of the object space. The roughness of the terrain is the decisive point for the choice of the DTM grid size. An independent height \( Z(x_0, y_0) \) is assigned to each grid point \((x_0, y_0)\) of the DTM. A height \( Z \) at an arbitrary point is interpolated from the neighbouring grid heights, e.g. by bilinear interpolation. At each point of the object surface, \( n \) and thus the angles \( i \) and \( e \) become a function of the neighbouring \( Z_{k,l} \).

A radiometric surface model is introduced to establish the connection between the geometric surface model and the reflectance behaviour of the surface. Each DTM grid cell is divided into several object surface elements of constant size. The size is chosen approximately equal to the pixel size multiplied by the average image scale factor. The radiometry of the surface is then described by one albedo \( A_o(n) \) per image valid for the whole surface together with the described Lunar-Lambert reflection model. Thus, our current model does not allow for any space-variant albedo.

### 3.4. DTM height estimation by least-squares adjustment

We can now put all the pieces together and formulate an estimation procedure to compute the DTM heights from the observed (and radiometrically calibrated) grey values of one or more images. In contrast to the calibrated model grey values \( G_i(x', y') \) the observed and radiometrically calibrated grey values \( g(x', y') \) are denoted by lower case. For a given object surface element we first compute \( G_i(x', y') \) according to Eq. (7) using approximate object space knowledge, we then project these model grey values into image space, and compare them with the observed grey values \( g(x', y') \) of the recorded images. By iteratively refining the values for the initially unknown parameters we improve the correspondence between the model grey values and the observed grey values, until we reach an optimum solution in the sense of least-squares adjustment.

In more detail, we assume all parameters of interior and exterior orientation of the images to be known, e.g. from a geometric and a radiometric camera calibration and a previously carried out bundle adjustment. Thus, for each object surface element, the vectors of viewing directions \( v \), where the subscript \( i \) stands for the considered image, are given. The illumination vector \( s \) (the direction to the sun) can be computed from the time of image acquisition. Note that while \( v \) varies across the surface, \( s \) is considered to be constant due to the large distance to the illumination source (the sun). Therefore, also the phase angle \( z \) is known, and \( A(z) \) can be computed as described above. The only remaining unknowns for the computation of \( G_i(x', y') \) (Eq. (7)) are the DTM heights \( Z_{k,l} \) and the normal albedo \( A_N(z) \).

For each image \( j \) every considered object surface element can now be projected into image space using the well-known collinearity equations (Eq. (8)) of central perspective projection (for push-broom imagery, the appropriate transformations have to be used, they are iterative in nature)

\[
\begin{align*}
x' &= x_0 - c \left( r_{11} \cdot (X - X_0) + r_{21} \cdot (Y - Y_0) + r_{31} \cdot (Z - Z_0) \right) + dx', \\
y' &= y_0 - c \left( r_{12} \cdot (X - X_0) + r_{22} \cdot (Y - Y_0) + r_{32} \cdot (Z - Z_0) \right) + dy',
\end{align*}
\]

where \( x' \) and \( y' \) are the image coordinates of \( P' \) (projection of point \( P \) into image space), \( x_0 \) and \( y_0 \) the image coordinates of the principal point, \( c \) the calibrated focal length, \( dx' \) and \( dy' \) the parameters of geometric distortion, \( X_0, Y_0, Z_0 \) the object coordinates of the projection centre, \( r_{11}, r_{12}, \ldots, r_{33} \) the elements of the rotation matrix of exterior orientation and \( X, Y, Z \) the object coordinates of point \( P \).

At the resulting position \( P_j' (x', y') \) the image grey value \( g_j(x', y') \) is resampled from the recorded grey values, e.g. by bilinear interpolation. The \( g_j \) are considered as observations in our least-squares adjustment. The corresponding observation equations are (see also Eq. (7))

\[
v_j(x', y') = G_i \left( \hat{Z}_{k,l}, \hat{A}_N(z_j) \right) - g_j(x', y') \left( \hat{Z}_{k,l} \right),
\]

\[
= \hat{A}_N(z_j) \left( A(z_j) \frac{2 \cdot \cos(i(\hat{Z}_{k,l}))}{\cos(i(...))} + (1 - A(z_j)) \cdot \cos(i(...)) \right)
\]

\[
- g_j(x', y') \left( \hat{Z}_{k,l} \right),
\]

where \( v_j(x', y') \) are the residuals of observed grey value \( g_j \) in image \( j \), \( \hat{Z}_{k,l} \) the DTM-heights of the geometric surface model (unknown), \( A_N(z_j) \) the normal albedo (unknown) and \( g_j \) the grey value in image \( j \) (observation).

Since Eq. (9) is non-linear with regard to \( Z_{k,l} \), an iterative computational scheme is needed, and initial values for the unknown \( Z_{k,l} \) must be available for carrying out the least-squares adjustment. \( A_N(z_j) \) only appears in linear form in Eq. (9), thus no initial value is required. The solution is then computed using standard formulae of least squares calculus (e.g. Mikhail and Ackermann, 1982).

A solution can already be obtained with only one image, although from one image, only height differences rather than absolute heights can be derived, because from a single image an image scale factor cannot be determined. If the height of a DTM point changes, so does the slope of the related object surface element and also its position in image space (since we use the perspective transformation between object and image space). The first effect leads to a change
in the model grey value, whereas the second results in a different grey value observation to be resampled. Both effects have an influence on the numerical stability of the algorithm: obviously, the larger the differences between model grey value and observed grey value become as a consequence of a height change, the more stable the algorithm is. Using the same argument, it is also clear that the reconstruction of a DTM from vertical images becomes more unstable, especially if the area to be reconstructed lies directly under the camera: in this case a change in height results in a very small displacement in image space only.\(^4\) It is numerically much more stable to use oblique images, because then a height change results in a larger image space displacement and thus potentially in a larger grey value difference. If two or more images taken from different viewpoints are given, stereoscopic correspondence between DTM grid cells projected into the images is implicitly exploited, and therefore absolute heights can be computed. Also, the described approach can be extended by simultaneously solving for a value of \(A(g)\) for each image along with the other unknowns of the adjustment, rather than in a pre-processing step. On the other hand, photometric modelling becomes more critical in the case of multiple images: The photometric functions capture in general only an approximation of the complex reflection behaviour of a real surface, and while in the case of a single image, the effects of a somewhat incorrect photometric function distort the DTM, for two or more images modelled with a wrong photometric function, each would correspond to a surface that is distorted differently. Thus, unless the photometric functions are chosen carefully, there will be no solution consistent with all the images, and the algorithm may fail to converge.

4. MI-SFS investigations

In order to investigate the proposed method with real extraterrestrial imagery we have selected suitable overlapping images from NASA’s 1994 scientific lunar mission Clementine. A detailed description of the Clementine mission is published by Nozette et al. (1994).

After presenting the input data in more detail, we discuss the case of a single-image analysis. Subsequently, we present results of multiple-image analysis and determine its radius of convergence.

4.1. Input data

For the reconstruction of a surface by means of MI-SFS it is necessary that the following information is available:

- one or more radiometrically calibrated digital images;
- illumination direction during data acquisition for each image;
- initial values for the unknown heights \(Z_{k,l}\).

For our investigations we selected images from the Ultraviolet/Visible (UV/Vis) digital frame camera, a medium resolution camera based on CCD-technology. We chose two images (Table 1 and Fig. 5). The selected area of the moon is part of the “Northern Mare Orientale Basin” (between 16.3° and 14.3° South and 87.3° and 90.9° West) and is depicted as the white rectangles in the two images (see Fig. 5). In the area a number of small craters and small valleys along with some larger craters are visible. Apart from some brightness variations in the walls of the bigger craters and some isolated bright spots, no major albedo variations seem to be present in the area, making it possible to use the images in our approach.

Radiometric calibration of the images was carried out using the ISIS software routine UVVISICAL (USGS, 2003). UVVISICAL performs dark current and readout correction, non-linearity and flat-field correction, and returns radiance values\(^5\) in (W sr\(^{-1}\) m\(^{-2}\)). The complete radiometric calibration algorithm is described in detail in McEwen et al. (1998).

The interior orientation of the images was taken from the general Clementine mission information, standard values for the calibrated focal length and the pixel size were used, distortion values were not available. The exterior orientation was obtained through bundle adjustment carried out at the German Aerospace Center, Berlin (DLR; Zhang et al., 1996; Wählisch, personal communication, 2005). We work in a local 3D object coordinate system with the origin in the centre of the area under investigation. The exterior orientation parameters provided in Table 1 refer to this coordinate system.

The illumination conditions for both images are approximately equal, the average incidence angles as derived from the Planetary Data System (PDS) amount to 34° and 26°. Image no. 334 has been taken from a more or less vertical position and has an average emission angle of approximately 13°, therefore some numerical instabilities may arise in the reconstruction process using only image 334 (see above). Image no. 338 is an oblique image with an average emission angle of about 47°. The average phase angles of both images (47° and 70°) are sufficiently large to allow the images to be considered for SFS using the Lunar-Lambert model,\(^6\) the difference of the phase angle, however, is also somewhat large. Some problems

\(^4\)The described numerical problem can be circumvented by a suitable selection of the object space coordinate system, namely by ensuring that none of the object space axes is approximately parallel to the optical axis; for more details see Heipke et al. (2001).

\(^5\)Note that the calibrated grey values lie in the interval of [3.27–6.08], not in the usual one for grey values of [0–255].

\(^6\)The photometric behaviour of most bodies is more complex at low phase angles (<30° and especially <10°) than is captured by the Lunar-Lambert model (McEwen, 1991), so low-phase images are unlikely to be suitable for MI-SFS. For the Moon in particular, the “albedo” variations between Maria and highland disappear for phase angles beyond about 100° (McEwen, 1996) so that images at very high phase angles should also be excluded.
resulting from this difference must therefore be expected. The limb-darkening parameters $A(z)$ for the two images were determined according to Eq. (4), they amount to 0.489 for image no. 334 and to 0.353 for image no. 338. The average spacecraft altitude was about 550 km. The geometric resolution of the images differs somewhat, the mean geometric resolution of one pixel is about 180 m. Due to the obliqueness of image no. 338, the pixel footprints are not square, but since we work from object to image space and resample the grey values $g_j$ from the actually recorded ones, this effect is implicitly taken care of. The two images were recorded with a time difference of about 20 h. We assume that during this period of time no relevant changes occurred in the observed lunar area.

In the overlapping part of the two images an area with a size of $24.3 \times 24.3$ km$^2$ was chosen for the investigations. The terrain can be described as somewhat rolling with some topographic variations, but without occlusions, the maximum height difference is about 1.3 km. The area is divided into $54 \times 54$ DTM grid cells with a grid size of 450 m. Thus, in total there are $55 \times 55 = 3025$ DTM heights. Each grid cell consisted of $3 \times 3$ object surface elements with a size of $150 \times 150$ m$^2$ each. The mean displacement in image space of one pixel at a pixel size of 23 µm translates to a height change in the DTM of approximately 360 m.

In order to create reference information against which to compare our results, we manually generated a DTM on a digital photogrammetric workstation using interactive stereo viewing. We carried out this process several times by different human operators, the accuracy of the derived DTM turned out to be in the order of 67 m or 0.2 pixels in height. Obviously, this value influences the comparison with the automatically derived results. The manually measured DTM is shown in Fig. 6. It should be noted that while this DTM can be used to assess the automatically derived results as a whole, it must be kept in mind that the potential of MI-SFS to derive high-resolution height information at a pixel level can be best assessed by visually comparing the original images with an artificially shaded representation of the derived DTM.

### 4.2. Single-image analysis

In this section we report on results of investigations with one image only in order to study the capabilities of MI-SFS using a single image.

We introduced the manually measured DTM (with one of the heights held constant to determine the scale factor, see above) as initial height information into our MI-SFS algorithm. For the evaluation of the results we derived

<table>
<thead>
<tr>
<th>Image no.</th>
<th>File name</th>
<th>$X_0$ (m)</th>
<th>$Y_0$ (m)</th>
<th>$Z_0$ (m)</th>
<th>$\varphi$ (deg)</th>
<th>$\omega$ (deg)</th>
<th>$\kappa$ (deg)</th>
<th>Direction of illumination</th>
<th>Image resolution</th>
<th>Average incidence angle $i$ (deg)</th>
<th>Average emission angle $e$ (deg)</th>
<th>Average phase angle $\alpha$ (deg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>334</td>
<td>LUC1941H.334</td>
<td>-119057.3</td>
<td>-32283.7</td>
<td>555435.2</td>
<td>3.89</td>
<td>-11.63</td>
<td>-179.82</td>
<td>Horizontal/vertical (deg)</td>
<td>Row/col (m/pixel)</td>
<td>145.0/148.0</td>
<td>12.80</td>
<td>47.13</td>
</tr>
<tr>
<td>338</td>
<td>LUC1941H.338</td>
<td>-531971.4</td>
<td>-26245.5</td>
<td>489371.1</td>
<td>-46.77</td>
<td>2.29</td>
<td>-177.19</td>
<td></td>
<td>183.0/270.0</td>
<td>35.9056/64.07</td>
<td>47.20</td>
<td>70.25</td>
</tr>
</tbody>
</table>

---

Fig. 5. Selected images with the investigated area (white rectangles): no. 334 (left) and no. 338 (right).
values for offset and standard deviation between the manually measured and the automatically computed DTM. The results are presented in Table 2. For each run we present the number of iterations, the offset $Z_0$ and the standard deviation $s$. We also show the computed normal albedo $A_N$.

Note that the reported results have to be interpreted with a mean height shift of about 360 m/pixel and the accuracy of the manually measured DTM of about 67 m in mind. When comparing the obtained standard deviation $s$ between the two DTMs to the average height shift per pixel, one can further conclude that the experiment has yielded sub-pixel accuracy for the heights. The standard deviation for the oblique image no. 338 is somewhat smaller than for the more vertical image no. 334.

As expected, the results also show that the surface exhibits a different brightness in the different images, which can be attributed to the different phase angles: when viewed from the two emission angles, the normal albedo $A_N$ for the two images is significantly different (4.9 for image no. 334 vs. 4.1 for image no. 338).

We now turn to a visual assessment of the results using image no. 338 (the corresponding figures for image no. 334 are very similar). Fig. 7 shows the DTM$_{338}$. The similarity with the manually measured DTM (see Fig. 6) is well visible, and as expected Fig. 7 shows more detail. A comparison of the observed grey values, the model grey values (both computed with the manually measured DTM), and the model grey values, computed with DTM$_{338}$, is shown in Fig. 8 from left to right.

In the centre image striping effects along the measuring directions of the human operators (top to bottom) can be seen. This visualises the accuracy limits of the manually measured DTM. The model grey values of the resulting DTM$_{338}$ (Fig. 8 right) correspond much better to the observed grey values (Fig. 8 left). This observation documents the refinement of the DTM achievable through MI-SFS. However, some small regular grid structures, which are not part of the DTM can also be observed, see Fig. 9 for an enlargement. These structures can be interpreted as an indication that the chosen grid size of the geometric surface model is too large, and the potential of MI-SFS for resolving fine height details goes beyond the level of $3 \times 3$ pixels.

The next step in our investigations was to demonstrate that MI-SFS also works for one image with less accurate initial height information. In order to do so we introduced a horizontal plane at the average height of the terrain as an initial DTM and show that our approach returns the correct terrain shape. The results are given in Table 3. Although the results of the single image analysis using a more precise initial DTM (see Table 2) are something better, it can be seen that the computations using a horizontal plane as initial DTM converge to a correct topography. The results still lie in the area of sub-pixel accuracy.

In summary, based on the obtained results we can say that

- The reflectance function of the surface under consideration can be approximated by the Lunar-Lambert law for both images.
- MI-SFS can yield correct results starting from a horizontal plane at an average height of the investigated area.

<table>
<thead>
<tr>
<th>Image</th>
<th>$A(z)$</th>
<th>Number of iterations</th>
<th>$Z_0$ (m)</th>
<th>$s$ (m)</th>
<th>$A_N$ (W sr$^{-1}$ m$^{-2}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>334</td>
<td>0.489</td>
<td>22</td>
<td>-80.7</td>
<td>249.7</td>
<td>4.9</td>
</tr>
<tr>
<td>338</td>
<td>0.353</td>
<td>17</td>
<td>-10.5</td>
<td>197.9</td>
<td>4.1</td>
</tr>
</tbody>
</table>

Fig. 6. Manually measured DTM.
As expressed in the standard deviation $s$ (see Tables 2 and 3), the accuracy potential of MI-SFS lies clearly in the sub-pixel domain and reaches between $\frac{1}{2}$ and $\frac{3}{2}$ of the pixel size (when considering that a height shift of approximately 360 m amounts to a shift of one pixel in image space, and neglecting for reasons of simplicity the accuracy of the manually determined DTM). The
accuracy for oblique images seems to be somewhat smaller than for more vertical images.

4.3. Multi-image analysis

The advantage of applications using multiple images with different exterior orientations is that an additional geometric stabilisation constraint, the correspondence between the images, is added to the determination of the surface. In addition, further images provide independent grey value information for the reconstruction of the unknown DTM heights. Compared to single-image analysis, it is unnecessary to introduce a known height as a scale factor, because homologous image rays intersect in the appropriate object point. In this way, the computation of absolute heights is achieved.

For the multiple image analysis we also use the manually measured DTM as initial values. The results for the multi-image case are depicted in Table 4.

The computation yield a correct result and converge in only 8 iterations in comparison to 17 and 22 iterations in the case of the single-image analyses (Table 2). The accuracy lies in the same range as for the oblique image no. 338. The results can be interpreted as a signal that a multi-image analysis is more robust than a single-image analysis, while retaining the same accuracy potential as that achievable with one image in stable geometric configurations. The resulting DTM\(_{334,338}\) is depicted in Fig. 10, it shares a similar form with the DTM derived from image 338 (see Fig. 7).

4.4. Radius of convergence

Finally, we investigated the radius of convergence of MI-SFS. For this purpose we used both images and inserted different initial DTMs into the algorithm. These DTMs differed from the manually measured DTM by an offset \(a_Z\) and a scale factor \(m_Z\) (Eq. (10)). The height differences were chosen in a way that the mean displacement of a surface element in the two images is a multiple of the pixel size. As mentioned above, a pixel change in the images conforms to a height change of about 360 m. If the scale factor \(m_Z\) takes the value 1.0 the initial DTM has the same shape as the manually measured DTM and if \(m_Z\) is equal 0.0, the initial DTM is a horizontal plane.

\[
Z_{Xi} = \bar{Z}_M + a_Z + m_Z \cdot (Z_{Mi} - \bar{Z}_M),
\]

(10)

where \(\bar{Z}_M\) is the mean height of the manually measured DTM, \(Z_{Mi}\) the DTM-height no. \(i\) of the manually measured DTM, \(Z_{Xi}\) the DTM-height no. \(i\) of the destination DTM, \(m_Z\) the scale factor and \(a_Z\) the height offset.

The results of the two-image analysis computed with different initial DTMs, are presented in Table 5. The numerical results show, that in our experiments the radius of convergence of MI-SFS amounted to approximately four pixels, which is equivalent to an offset \(a_Z\) of about 1440 m. For larger differences between the initial height information and the actual terrain height, the computations were trapped in a local minimum, or no convergence was reached. It should be noted, that the algorithm

![Fig. 10. Resulting DTM\(_{334,338}\).](image-url)
produced a correct result also starting from a horizontal plane. This means that for the reconstruction of a surface by means of MI-SFS it is enough to know an approximate average surface height. On the other hand, a poor selection of the direction of the object space coordinate axes may lead to a smaller radius of convergence (see also footnote 4).

5. Conclusions

In this paper we have presented a new method called multi-image shape-from-shading (MI-SFS) for the generation of digital terrain models from multiple, overlapping planetary images. We presented experimental investigations in single-image and multiple-image SFS using imagery from the lunar mission Clementine. The obtained results show that MI-SFS is a method that can refine a DTM, which was generated with some other technique, to a level of detail of a few pixels, and to sub-pixel accuracy in height. Results using multiple images have been found to be superior to those obtained with single images. However, the method also produces good results, if only one image is available. In this case, vertical images should be avoided for numerical reasons, and an average height must be available, since only shape, but no absolute heights can be generated from a single image. Furthermore, in our investigations we obtained a radius of convergence of about four pixels.

Limits of our technique lie in the fact that we cannot model shadows or occlusions, and our surface model does not contain breaklines. We plan to improve our model to be able to handle such cases. We also work on extending our method to be able to allow also for spatially variable albedo. A reasonable possibility is to split the normal albedo into two factors, one taking care of the phase angle dependence and thus varying from image to image (but not in space), and another one taking into account the variations in space (but being constant across the images).

This is equivalent to assuming that the albedo of the surface varies, but the whole surface darkens with increasing phase angle at an equivalent rate.

We will also try to increase the geometric level of detail by introducing more sophisticated object surface models, i.e. DTM with a grid size equal to the pixel size with appropriate smoothness and curvature constraints. Finally, we plan to integrate line sensor geometry and thus push-broom camera models into the algorithm in order to be able to use other planetary data, e.g. HRSC data of ESA’s Mars-Express mission (Neukum et al., 2004). While we expect the geometric part of this extension to be relatively straightforward, we will need to carefully revisit our photometric models and the related assumptions in order to adapt them to imagery from Mars.

Acknowledgement

This work was developed within the priority program “Mars and the terrestrial planets” financed by the German Science Foundation (Deutsche Forschungsgemeinschaft - DFG) under the project number HE 1822/10. The support is gratefully acknowledged. Thanks also go to Prof. Egon Dorrer, University of the Federal Armed Forces, Munich, for many valuable suggestions to improve the method and the manuscript, to the German Aerospace Center (Deutsches Zentrum für Luft- und Raumfahrt—DLR) for providing and helping with the Clementine data, and to the ISIS Support Center at USGS for their assistance in the radiometric calibration of the Clementine images. Finally, we express our thanks to the reviewers for valuable comments, which have considerably improved the paper.

References


Table 5

<table>
<thead>
<tr>
<th>DTM parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale factor $m_Z$</td>
<td>Offset $a_Z$ (m/pixel)</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>1.0</td>
<td>0.0/0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0/0</td>
</tr>
<tr>
<td>1.0</td>
<td>360.0/1</td>
</tr>
<tr>
<td>0.0</td>
<td>360.0/1</td>
</tr>
<tr>
<td>1.0</td>
<td>720.0/2</td>
</tr>
<tr>
<td>0.0</td>
<td>720.0/2</td>
</tr>
<tr>
<td>1.0</td>
<td>1080.0/3</td>
</tr>
<tr>
<td>0.0</td>
<td>1080.0/3</td>
</tr>
<tr>
<td>1.0</td>
<td>1440.0/4</td>
</tr>
<tr>
<td>0.0</td>
<td>1440.0/4</td>
</tr>
<tr>
<td>1.0</td>
<td>1800.0/5</td>
</tr>
<tr>
<td>0.0</td>
<td>1800.0/5</td>
</tr>
</tbody>
</table>


