REFINEMENT OF STEREO IMAGE ANALYSIS USING PHOTOMETRIC SHAPE RECOVERY AS AN ALTERNATIVE TO BUNDLE ADJUSTMENT

A. Grumpe^a, C. Schröer^a, S. Kauffmann^a, T. Fricke^a, C. Wöhler^a, U. Mall^b

^a Image Analysis Group, TU Dortmund, Otto-Hahn-Str. 4, D–44227 Dortmund, Germany arne.grumpe@tu-dortmund.de ^b Max-Planck-Institut für Sonnensystemforschung, Justus-von-Liebig-Weg 3, D–37077 Göttingen, Germany

Theme session ThS8

KEY WORDS: Stereo Image Analysis, Shape from Shading, Bundle Adjustment, DEM, 67P/Churyumov-Gerasimenko

ABSTRACT:

Topographic mapping, e.g. the generation of Digital Elevation Models (DEM), is of general interest to the remote sensing community and scientific research. Commonly, photogrammetric methods, e.g. stereo image analysis methods (SIAM) or bundle adjustment methods (BAM), are applied to derive 3D information based on multiple images of an area. These methods require the detection of control points, i.e. common points within multiple images, which relies on a similarity measure and usually yields a sparse map of 3D points. The full spatial DEM is then obtained by interpolation techniques or imposed restrictions, e.g. smoothness constraints. Since BAM utilizes all images of the area, it is assumed to provide a more accurate DEM than SIAM which utilizes only pairs of images. Intensity-based shape recovery, e.g. shape from shading (SfS), utilizes the reflectance behavior of the object surface and thus provides a dense map of relative height changes, which provide the possibility to refine the photogrammetric DEMs. Based on Rosetta NavCam images of 67P/Churyumov-Gerasimenko we compare intensity-based DEM refinement methods which use DEMs obtained based on SIAM and BAM as a reference. We show that both the SIAM based DEM refinement and the BAM based DEM refinement are of similar quality. It is thus possible to derive DEMs of high lateral resolution by applying the intensity-based refinement to the less complex SIAM.

1. INTRODUCTION

In 2004, the spacecraft Rosetta was launched by the European Space Agency in order to analyze the comet 67P/Churyumov-Gerasimenko (Churyumov, 2005). Among other subjects, the shape of 67P/Churyumov-Gerasimenko has been analyzed by Sierks and 65 coauthors (2015) based on images of the OSIRIS camera (Keller and 68 coauthors, 2007).

Topographic mapping, e.g. the generation of Digital Elevation Models (DEM), is of general interest to the remote sensing community and scientific research. It has been widely applied to planetary bodies (Smith and 23 coauthors, 2001; Smith and 30 coauthors, 2010; Scholten et al., 2012; Jaumann and 42 coauthors, 2012) and is crucial for geologic analysis and mission planning, e.g the selection of possible landing sites (Kirk et al., 2003b). Consequently, a high effective DEM resolution is required.

Commonly, photogrammetric methods, e.g. stereo image analysis methods (SIAM) or bundle adjustment methods (BAM) (Mc-Glone et al., 2004), are applied to derive 3D information based on multiple images of an area. These methods require the detection of control points, i.e. common points within multiple images, which relies on a similarity measure and usually yields a sparse map of 3D points. The full spatial DEM is then obtained by interpolation techniques or imposed restrictions, e.g. smoothness constraints. Since BAM utilizes many images of the area, it is assumed to provide a more accurate DEM than SIAM which utilizes only pairs of images. The DEMs originating from both types of methods, however, have a lower resolution than the underlying image data (Kirk et al., 2003a; Schenk, 2008; Scholten et al., 2012).

In contrast, intensity-based shape recovery, e.g. shape from shading (SfS), utilizes the reflectance behavior of the object surface and thus provides a dense map of relative height changes (Horn, 1990), which provide the possibility to refine the photogrammetric DEMs. The refinement of a-priori known DEMs has been successfully demonstrated (Soderblom and Kirk, 2003; Schenk, 2008; Grumpe and Wöhler, 2014).

Based on Rosetta NavCam images of 67P/Churyumov-Gerasimenko we compare intensity-based DEM refinement methods which use DEMs obtained based on SIAM and BAM as a reference. We show that both the SIAM based DEM refinement and the BAM based DEM refinement are of similar quality. It is thus possible to derive DEMs of high resolution by applying the intensity-based refinement to the less complex SIAM. The position and orientation differences between frames of the Nav-Cam with respect to the surface being analyzed are not known accurately enough for using a calibrated stereo method. Consequently, we apply an uncalibrated stereo image analysis. This SIAM is defined only up to an arbitrary projective transformation (Hartley and Zisserman, 2004) and thus needs to be constrained by the BAM derived DEM, which is defined up to an arbitrary scaling. Although this approach requires both the SIAM and the BAM derived DEMs, it is generally possible to replace the computationally complex BAM with SIAM if the camera positions are calibrated or three images are analyzed.

2. METHODS

The DEMs are generated using two subsequent steps. At first, a DEM is obtained from triangulation based methods outlined in Section 2.1. The resulting DEM is then refined using the photometric techniques given in Section 2.2.

2.1 Triangulation based shape recovery

Triangulation methods are common techniques for the DEM generation from multiple images of the same object. Bundle adjustment methods (BAM) recover the DEM from multiple images at once and thus pose a considerably larger computational burden than stereo image analysis methods (SIAM). Since BAM includes more images, i.e. more information, the obtained DEMs from BAM, however, are supposed to be more accurate. This section reviews the applied triangulation based methods.

2.1.1 Bundle-adjustment methods (BAM) In contrast to SIAM, BAM recover the depth information from more than two images. While it may be feasible to manually select pairs of stereo images, it is certainly infeasible to manually determine a possibly huge sequence of images showing the same surface region. At first, we thus present a simple image selection algorithm. The BAM is then applied to the selected images yielding a point cloud of 3D coordinates. Finally, the possibly sparse point cloud is projected into a virtual camera and interpolated to a dense grid.

Image selection At first, we manually select one target image showing nearly the complete target region. An additional criterion is the position of the camera. The viewing direction should be as orthogonal to the surface as possible to avoid projective distortions during the photometric analysis. The selected image is then cropped to the target region to reduce false matches.

SIFT features (Lowe, 2004) are extracted from the target image, yielding the target descriptors. We then extract SIFT features from each NavCam image, respectively, and compare them to the target descriptors using the Euclidean distance based approach of Lowe (2004). Images showing less than 50 matching descriptors are discarded.

Triangulation BAM generally do not yield a single intersection of all camera rays due to noise and inaccurate point matches. To avoid this problem, the squared re-projection error

$$E_{\text{BAM}} = \sum_{p} \sum_{c} \left\| \mathbf{u}_{p,c} - \Pr\left(\mathbf{x}_{p}, \mathbf{c}_{c} \right) \right\|^{2}, \qquad (1)$$

i.e. the squared Euclidean distance $\|\cdot\|$ of the position $\mathbf{u}_{p,c}$ of the *p*th point in the *c*th camera from the projection $\Pr(\cdot)$ of the 3D coordinates \mathbf{x}_p of the *p*th point using the camera parameters \mathbf{c}_c of the *c*th camera.

In general, Eq. (1) requires a non-linear optimization algorithm and allows for the simultaneous estimation of camera parameters and 3D points. In this study, we use the VisualSFM toolbox (Wu, 2011; Wu et al., 2011) to solve the problem. For all parameters the default values are used. The optimization produces a point cloud which is exported from VisualSFM.

DEM generation All points of the point cloud are projected into the camera, which corresponds to the target image. The resulting DEM z_{proj} is generally not dense and its scaling is unknown. Consequently, we rescale the resulting DEM such that the average depth value \bar{z}_{proj} matches the distance $d_{67\text{P}}$ to the comet that has been supplied with the NavCam image data and is centered around zero, i.e.

$$z_{\rm BAM} = \left(\frac{z_{\rm proj}}{\bar{z}_{\rm proj}} - 1\right) d_{67\rm P}.$$
 (2)

Finally, the DEM is interpolated to a dense grid using a spring metaphor D'Errico (2012).

2.1.2 Stereo image analysis methods (SIAM) In general, the projection of a 3D point into cameras at two different spatial positions results in a displacement of the projected points in the camera coordinate systems. This displacement, which is known as the disparity, shows an inverse proportionality to the distance

from the optical centres of the cameras. The direction of the displacement is called epipolar line and depends on the orientation of the cameras and varies over the local camera coordinates. To reduce the complexity of the problem and introduce parallel epipolar lines, we apply a projective transformation to the image, i.e. a rectification (Hartley, 1999), prior to the estimation of a disparity map using semi-global matching (Hirschmüller, 2005, 2008). Since SIAM requires two images only, we select the second image manually. Finally, the disparity map is converted to a DEM.

Rectification The rectification step consists of a projective transformation that results in parallel epipolar lines. If the camera has been calibrated, i.e. the extrinsic and intrinsic camera parameters are known, the rectification is straightforward (Loop and Zhang, 1999). In case of the Rosetta NavCam, however, the position and orientation of the camera relative to the surface part being analyzed are not known accurately enough for using a calibrated stereo method. We thus apply the rectification algorithm by Hartley (1999) and Hartley and Zisserman (2004). Accordingly, the fundamental matrix of the stereo image pair is estimated based on control points, i.e. points in both images that correspond to the same object point. Both images are then transformed such that the epipolar lines are parallel. Additionally, the transformations include a translation of the images such that the mean squared Euclidean distance between the camera coordinates of the control points is minimized (Hartley, 1999; Bradski and Kaehler, 2008). This translation minimizes the image size which is required to hold the transformed images.

The control points are obtained from SIFT features (Lowe, 2004). SIFT features are extracted for different scales of the image, i.e. octaves, and different levels of low-pass filtering. Possible feature points are derived from corner points. The corner points are obtained using the corner detector of Harris and Stephens (1988). Based on the intensity gradient, each feature point is assigned an orientation. Finally, an array of histograms of oriented gradients are constructed and the histogram counts form the feature descriptor. The descriptors are matched based on the Euclidean distance between different descriptors. In this study, we apply the SIFT feature extractor and the SIFT matching of the VLFeat toolbox (Vedaldi and Fulkerson, 2008) with the default parameters therein.

Since the estimated projective transformations highly depend on the quality of the matched control points, we apply an additional outlier detection step. The resulting epipolar lines are supposed to be parallel and thus the vertical coordinate should not differ for the transformed control points. Consequently, we remove control points that exhibit a vertical difference of more than one pixel and recompute the transformations.

Semi-global matching The semi-global matching algorithm of Hirschmüller (2005) minimizes the error

$$E_{\text{SGM}}(D) = \sum_{\mathbf{p}} \left(C(\mathbf{p}, D_{\mathbf{p}}) + \sum_{\mathbf{q} \in N_p} P_1 T \Big[|D_{\mathbf{p}} - D_{\mathbf{q}}| = 1 \Big] + \sum_{\mathbf{q} \in N_p} P_2 T \Big[|D_{\mathbf{p}} - D_{\mathbf{q}}| > 1 \Big] \right)$$
(3)

where D is the disparity map, \mathbf{p} is an image pixel, \mathbf{q} is a pixel in the neighborhood N_p of \mathbf{p} , $D_{\mathbf{p}}$ and $D_{\mathbf{q}}$ are the disparity values at \mathbf{p} and \mathbf{q} , respectively, and the function $T [\cdot]$ returns the value one if the logical argument is true or zero if the logical argument is false. Consequently, the overall error function consists of a cost function $C(\mathbf{p}, D_{\mathbf{p}})$ measuring the similarity between the pixel \mathbf{p} and the corresponding pixel in the second stereo image and a smoothness constraint. The first smoothness expression punishes small deviations, i.e. one pixel, in the disparity map weighted by P_1 while the second expression punishes larger deviations, i.e. more than one pixel, weighted by P_2 . Accordingly, P_2 should exceed P_1 . We set $P_1 = 1 \cdot 10^{-5}$ and $P_2 = 1 \cdot 10^{-4}$.

Since minimizing $E_{\rm SGM}(D)$ globally is a NP-complete problem, Hirschmüller (2005) propose the aggregation of one-dimensional cost functions. The modified error function

$$S(\mathbf{p}, d) = \sum_{\mathbf{r} \in N_{\mathbf{r}}} L_{\mathbf{r}}(\mathbf{p}, d)$$
(4)

is obtained by traversing $N_{r}\xspace$ paths $r\xspace$ and aggregating the path costs

$$L_{\mathbf{r}}(\mathbf{p}, d) = C(\mathbf{p}, d) + \min(L_{\mathbf{r}}(\mathbf{p} - \mathbf{r}, d),$$

$$L_{\mathbf{r}}(\mathbf{p} - \mathbf{r}, d - 1) + P_1,$$

$$L_{\mathbf{r}}(\mathbf{p} - \mathbf{r}, d + 1) + P_1,$$

$$\min_{i \in N_d} L_{\mathbf{r}}(\mathbf{p} - \mathbf{r}, i) + P_2)$$
(5)

where **r** is the previous pixel in the path direction, d is the disparity at pixel **p**, and min is the minimum operator which returns the smallest argument. The path costs are recursively defined and terminated at the image boundary by setting the path costs to $C(\mathbf{p}, d)$ for the pixels on the image boundary. The path costs are computed for a set of admissible disparities N_d . Additionally, we adopt the 16 path directions proposed by Hirschmüller (2008). The recursion is started at each pixel on the image boundary and the cost function values are aggregated.

As proposed by Hirschmüller (2008), we set the cost function $C(\mathbf{p}, d)$ to be the negative mutual information. Since the mutual information is based on the full image and requires the disparity map, it is generally not possible to compute a value for each pixel. Therefore, the Taylor expansion derived by Kim et al. (2003) is applied to transform the mutual information into a sum over pixels. The rectification step produces an average disparity value which is close to zero. Consequently, we initialize an all-zero disparity map to perform the estimation for three subsequent iterations. Each iteration uses the previous disparity map estimate to compute the mutual information. Finally, the optimal integer disparity value is refined by fitting a parabola through the overall error function for each pixel, respectively. To ensure valid disparity estimates, we independently compute a disparity map that transforms the left image onto the right image and a disparity map that transforms the right image onto the left image. If both estimates differ by more than one pixel, the disparity is invalid and discarded.

DEM generation Since the camera geometry is uncalibrated, the scene may be reconstructed up to an arbitrary 3D projective transformation (Faugeras, 1992; Hartley, 1999; Hartley and Zisserman, 2004). Furthermore, the disparity map is aligned to the rectified image. Therefore, we apply the inverse projective transformation from the rectification step to generate a disparity map that is aligned to the original NavCam image. We then compute the projective transformation that minimizes the squared Euclidean distance between the transformed disparity map and the BAM-derived DEM z_{BAM} . Finally, the DEM is interpolated to a dense grid using a spring metaphor D'Errico (2012) yielding the SIAM-based DEM z_{SIAM} .

In general, this arbitrary projective transformation limits the usefulness of SIAM. It is, however, possible in principle to resolve the ambiguity if the camera positions are calibrated, i.e. the external parameters of the camera are extracted from the spacecraft position and pointing data. Here, this step is omitted and replaced by the mapping onto the BAM derived DEM.

2.2 Reflectance-based DEM refinement

Reflectance-based surface recovery methods estimate the slope of the surface, i.e. its gradient field, based on the known reflectance behavior of the surface. Consequently, a mathematical model of the surface reflectance is required. From the observed shading and the shading predicted by the model, the gradient field of the surface is then estimated. Afterwards, the surface may be recovered through a gradient field integration step. This section reviews the applied methods.

2.2.1 Reflectance model The reflectance-based surface recovery relies on the predicted shading of a reflectance model. Consequently, an adequate model needs to be accurate and computationally feasible. Prominent physically motivated reflectance models, which are commonly applied to planetary bodies, asteroids, and comets are the model by Hapke (1981) and Shkuratov et al. (1999). Both models rely on the scattering behavior of a single particle modeled as a slab. The surface reflectance is then computed by taking into account multiple scattering between particles in a regolith layer. The main difference between the models of Hapke (1981) and Shkuratov et al. (1999) is the treatment of the single-particle scattering function or "phase function" (Poulet et al., 2002).

Both models, however, have quite a few parameters and a high degree of freedom. Since a reliable estimation of all parameters requires a large variety of different phase angles α , i.e. the angle between the viewing direction and the direction of the incident light, it is not possible to apply these models to single images without prior knowledge. Therefore, we apply a less complex empirical reflectance model. McEwen (1991) suggests the "lunar-Lambert" function

$$R = \rho \left(2L \frac{\mu_0}{\mu_0 + \mu} + (1 - L) \,\mu_0 \right), \tag{6}$$

where the cosines of the incidence angle and the emission angle are denoted by μ_0 and μ , respectively, and the weight function Ldepending on the phase angle can be read from the diagrams in McEwen (1991, Fig. 16 therein). We assume that the NavCam image intensities are proportional to the physical reflectance. The proportionality constant $\rho \in [0, \infty[$ thus has no upper boundary and corresponds to the scaling factor from image intensities to reflectance values. The interpolated DEMs obtained by SIAM and BAM generally do not match the image resolution. Thus we apply the method of Grumpe et al. (2015) to compute a locally varying reflectance parameter ρ using a Gaussian low-pass filter of width $\sigma_{\rho} = 21$ pixels.

2.2.2 Shape-from-Shading Horn (1990) provides an extensive survey of the development of reflectance-based methods. Most of them focus on the estimation of gradients from shading information since the proportion of light reflected from a surface is dependent on the surface slope rather than the absolute height. Furthermore, Horn (1990) presents a coupled scheme for simultaneous estimation of surface height and gradients. This method is referred to as "Shape from Shading" (SfS). The coupled scheme alternates between estimating the surface gradient field and the recovery of the surface from the current gradient field estimate. A similar technique is applied by Kirk (1987).

Reflectance-based gradient field estimation Commonly, the gradient field is estimated by describing the surface normal $\mathbf{n} = (p^2 + q^2 + 1)^{-1/2} [-p, -q, 1]^T$ in terms of reflectance-based gradient estimates $p \approx \frac{\partial z}{\partial x}$ and $q \approx \frac{\partial z}{\partial y}$. The reflectance model R(p,q) depends on the illumination geometry and thus on the gradient field given by p and q. The gradient field estimate is then obtained by minimizing the intensity error

$$E_{\rm I} = \frac{1}{2} \int_{x} \int_{y} (I - R(p, q))^2 \, \mathrm{d}x \mathrm{d}y, \tag{7}$$

i.e. the squared difference between the intensity I and the modeled reflectance. The intensity error, however, relates the two partial derivatives to one intensity quantity and thus does not yield a unique solution. Horn (1990) proposes an additional smoothness and/or integrability constraint to regularize the optimization problem.

In this study, we adopt the method by Grumpe et al. (2014) and use the triangulation based DEM as a regularization of the optimization problem by adding the gradient error

$$E_{\text{grad}} = \frac{1}{2} \int_{x} \int_{y} \left(f(p) - \left(\frac{\partial z_{\text{DEM}}}{\partial x}\right) \right)^{2} + \left(f(q) - \left(\frac{\partial z_{\text{DEM}}}{\partial y}\right) \right)^{2} dx dy.$$
(8)

The gradient error measures the deviation of the photometric gradient field estimate [p,q] from the gradient field $\left[\frac{\partial z_{\text{DEM}}}{\partial x}, \frac{\partial z_{\text{DEM}}}{\partial y}\right]$ derived from a known DEM z_{DEM} after a lowpass filter $f(\cdot)$ has been applied. This approach assumes that z_{DEM} is re-sampled to the image but has a lower effective resolution, which is generally the case for triangulation based DEMs.

Gradient field integration Since the reflectance-based gradient field estimate [p, q] is subject to image noise and modeling inaccuracies, the DEM may not be retrieved by a simple integration step. Horn (1990) proposed the minimization of the integrability error

$$E_{\rm int} = \frac{1}{2} \int_{x} \int_{y} \left(p - \frac{\partial z}{\partial x} \right)^2 + \left(q - \frac{\partial z}{\partial y} \right)^2 dx dy, \quad (9)$$

which measures the squared Euclidean distance of the gradient field derived from the DEM z to the reflectance-based gradient field estimate.

To include information from a known but possibly sparse DEM, Shao et al. (1991) propose the simultaneous minimization of the integrability error and the squared deviation of z from the known DEM z_{DEM} . This approach however, does not include the lower lateral resolution of the known DEM. Therefore, we adopt the depth error

$$E_{\text{depth}} = \frac{1}{2} \int_{x} \int_{y} \left(f\left(z\right) - f\left(z_{\text{DEM}}\right) \right)^2 \mathrm{d}x \mathrm{d}y \qquad (10)$$

from Grumpe and Wöhler (2014), with the same low-pass filter $f(\cdot)$ of $E_{\rm grad}.$

Combined height and gradient estimation The combined estimation of the reflectance-based gradient field and the DEM is obtained by minimizing

$$\min_{p,q,z} E_{\rm I} + \delta E_{\rm grad} + \gamma E_{\rm int} + \tau E_{\rm depth}.$$
 (11)



(a) 67P/Churyumov-Gerasimenko



ROS_CAM1_20140826T060854. ROS_CAM1_20140816T110718.

Figure 1: Study area on 67P/Churyumov-Gerasimenko. (a) Regional map of 67P/Churyumov-Gerasimenko. URL: <http://www.esa.int/spaceinimages/Images/2015/ 01/Getting_to_know_Rosetta_s_comet_region_maps>. Image credit: ESA/Rosetta/MPS for OSIRIS Team MPS/UPD/-LAM/IAA/SSO/INTA/UPM/DASP/IDA. (b) The target image of BAM and SIAM. All images are matched with respect to this image. The red rectangle marks the region of interest, which is refined using SfS. (c) The second image of the stereo pair.

Similar to Horn (1990), we use an alternating update scheme, i.e. at each iteration the reflectance-based gradient field is updated and the DEM is then adapted to the updated gradient field. Since $E_{\rm depth}$ does not depend on p and q, the update equations of Grumpe et al. (2014) are applicable. The update of the DEM, in contrast, does not depend on $E_{\rm I}$ and $E_{\rm grad}$. Consequently, we adopt the update equations from Grumpe and Wöhler (2014). All update equations are somewhat lengthy and thus not repeated here. The weights of the error function are set to $\delta = 0.001$, $\gamma = 0.2$, and $\tau = 10$. The low-pass filter $f(\cdot)$ is set to be a Gaussian of width $\sigma = 21$ pixels.

3. RESULTS

The NavCam image ROS_CAM1_20140826T060854 showing the Seth region is selected for this study. Fig. 1 shows a map of 67P/Churyumov-Gerasimenko, the selected NavCam image ROS_CAM1_20140826T060854, and the corresponding stereo image ROS_CAM1_20140816T110718. The image selection algorithm described in Section 2.1.1 selects 47 NavCam images showing the Seth region. The analysis is based on the region of interest (ROI) marked in 1(b). Fig. 2(a) shows a close-up of the ROI, and shaded DEMs obtained by BAM and SIAM are shown in Fig. 2(b) and 2(c), respectively. Since both DEMs show "spikes" that originate from false matches, we apply a median filter of width 15 pixels prior to the computation of the reflectancebased surface refinement. The refined DEMs are shown in Fig. 2(d) and 2(e).



Figure 3: 3D surface plots of the DEMs. Notably, the overlying images are shaded versions of the DEM and thus represent the details of the DEM data. (a) DEM obtained using BAM. (b) DEM obtained using SIAM. (c) SfS result using the BAM-derived DEM. (d) SfS result using the SIAM-derived DEM.



(a) Close-up of the (b) BAM-derived (c) SIAM-derived study region. DEM. DEM.





(d) Refined BAM-derived DEM.

(e) Refined SIAM-derived DEM.

Figure 2: Region of interest. (a) A close-up view of the ROI (b) DEM obtained using BAM. (c) DEM obtained using SIAM. (d) SfS result using the BAM-derived DEM. (e) SfS result using the SIAM-derived DEM.

3.1 Visibility of small details

Fig. 3 shows 3D surface plots of the DEMs. Although the SIAMbased DEM (Fig. 3(b)) appears more sharp than the BAM-based DEM (Fig. 3(a)), the effective resolution is not better since the impression of sharpness arises from edges and most of the edges in the SIAM-based DEM originate from pixel locking, i.e. integer disparity values, and noise captured by the dense semiglobal matching. It does, however, show less interpolation gaps. Both triangulation based DEMs, however, appear rather planar and capture little details that are clearly visible in the image (Fig. 2(a)).

The refined DEMs based on BAM and SIAM are shown in Fig. 3(c) and 3(d), respectively. Both DEMs appear to be very similar and there is basically no visible difference. Very subtle differences may be observed near the boundary of the ROI and the corners. The "spiky" noise artifacts are suppressed and the interpolation gaps are filled with data. Furthermore, both DEMs closely resemble the NavCam image, indicating a strongly increased effective resolution.

3.2 Contours

Twenty equally spaced contours are overlain on the color-coded DEMs in Fig. 4. The pixel locking artifacts are clearly visible in the SIAM-derived DEM (Fig. 4(b)) while the BAM-based DEM (Fig. 4(a)) clearly shows "spikes" and flat areas due to the interpolation routine. In both cases, the contours of the SfS-refined DEMs are smooth and appear to be less noisy. The contours, however, show that there are small differences between the two SfS-based DEMs. The SIAM-based DEM shows an increased height variation, i.e. smaller minimum values and larger maximum values. The hill on the south-east corner seems to be higher in the refined SIAM-based DEM. It is, however, not clear if this behavior is due to missing matches in the BAM-based DEM or

due to noise inherited from the SIAM-based DEM. The differences are very subtle and the overall appearance of both SfS based DEMs is largely the same.

3.3 Depth profiles

The high noise component of the SIAM based DEM is also visible in the height profiles shown in Fig. 5. The BAM-derived DEM is less noisy but, again, shows "spiky" height changes around control points. The refined DEMs show a very similar behavior. In the left profile shown in Fig. 5(b), they resemble the triangulation-based DEMs quite well. Both DEMs show approximately the same values but differ close to the image boundary. In the right profile shown in Fig. 5(c), both SfS-refined DEMs show a larger deviation from the triangulation-based DEMs after the shadowed region has been passed. This behavior originates from the integration of erroneous shading information if no shadow detection is applied. It is well known for shading-based DEM recovery methods to exhibit this kind of errors. The applied SfS algorithm, however, utilizes the triangulation based DEMs to regularize the solution and thus recovers from these errors yielding a locally slightly distorted DEM. The absolute height thus matches again near the end of the profile. While the triangulation-based DEMs barely show the steep rims and crater-like structures, the SfS-based DEMs clearly show that the corresponding height decreases. However, shadows and poor shading information lead to an oversmooth representation of the steep rims.

4. CONCLUSIONS

The results show that triangulation-based DEM recovery methods, i.e. BAM and SIAM, yield DEMs of a lower effective resolution than the underlying images. Although the dense semiglobal matching algorithm produces DEMs that nominally share the image resolution, the resulting DEM shows artifacts such as pixel locking and noise. The BAM-based DEM shows less noise but more "spiky" artifacts originating from false control point matches. In both cases, the SfS method succeeded in increasing the effective resolution of the DEM while the absolute scale was not affected. The refined DEMs derived from BAM and SIAM based DEMs, respectively, showed no significant difference. Consequently, it is possible to replace the computationally complex BAM by a more simple SIAM. Uncalibrated SIAM, however, may be applied to obtain a DEM up to an arbitrary projective transformation. This ambiguity may be resolved by a camera calibration step, by using at least five control points with known 3D coordinates, or by using a trinocular image analysis method.

ACKNOWLEDGMENTS

We acknowledge ESA for permission to work with NavCam images under the Creative Commons Attribution-Share Alike 3.0 IGO (CC BY-SA 3.0 IGO) license.



(c) Refined BAM-derived DEM.



Figure 4: Contour plots of the DEMs. The 20 contours are equally spaced between -650 m and 350 m. The based maps are created by encoding the height in colors and multiplying the intensity of each image by a Lambertian shading of the DEM. Black dashed lines represent the contours. (a) DEM obtained using BAM. (b) DEM obtained using SIAM. (c) SfS result using the BAM-derived DEM. (d) SfS result using the SIAM-derived DEM.



Figure 5: Depth profiles. All profiles run from top to bottom. (a) The location of the left (green) and right (red) profiles are shown on the NavCam image. (b) Left profile. (c) Right profile.

References

Bradski, G. and Kaehler, A., 2008. Learning OpenCV. 1st edn, O'Reilly Media, Inc., Sebastopol, CA, US.

Churyumov, K., 2005. Discovery and study of Comet 67P/Churyumov-Gerasimenko, the main target of the Rosetta space mission. Kinematika i Fizika Nebesnykh Tel Supplement 5, pp. 487–490.

D'Errico, J., 2012. inpaint_nans. Version 1.1, accessed 28 October 2015, < http://de.mathworks.com/matlabcentral/fileexchange/4551-inpaint-nans>.

Faugeras, O. D., 1992. What can be seen in three dimensions with an uncalibrated stereo rig? In: G. Sandini (ed.), Proceedings of Second European Conference on Computer Vision (ECCV) 1992, Lecture Notes in Computer Science, Vol. 588, Springer, Berlin, Heidelberg, pp. 563–578.

Grumpe, A. and Wöhler, C., 2014. Recovery of elevation from estimated gradient fields constrained by digital elevation maps of lower lateral resolution. ISPRS Journal of Photogrammetry and Remote Sensing 94, pp. 37–54.

Grumpe, A., Belkhir, F. and Wöhler, C., 2014. Construction of lunar DEMs based on reflectance modelling. Advances in Space Research 53(12), pp. 1735–1767.

Grumpe, A., Zirin, V. and Wöhler, C., 2015. A normalisation framework for (hyper-)spectral imagery. Planetary and Space Science 111, pp. 1–33.

Hapke, B., 1981. Bidirectional reflectance spectroscopy. 1. Theory. Journal of Geophysical Research 86(B4), pp. 3039–3054.

Harris, C. and Stephens, M., 1988. A Combined Corner and Edge Detector. Proceedings of the 4th Alvey Vision Conference pp. 147–152.

Hartley, R. I., 1999. Theory and Practice of Projective Rectification. International Journal of Computer Vision 35, pp. 115–127.

Hartley, R. I. and Zisserman, A., 2004. Multiple View Geometry in Computer Vision. Second edn, Cambridge University Press, Cambridge, New York.

Hirschmüller, H., 2005. Accurate and efficient stereo processing by semiglobal matching and mutual information. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR) 2005 2, pp. 807–814.

Hirschmüller, H., 2008. Stereo Processing by Semiglobal Matching and Mutual Information. IEEE Transactions on Pattern Analysis and Machine Intelligence 30, pp. 328–341.

Horn, B. K. P., 1990. Height and Gradient from Shading. International Journal of Computer Vision 5(1), pp. 37–75.

Jaumann, R. and 42 coauthors, 2012. Vesta's Shape and Morphology. Science 336(6082), pp. 687–690.

Keller, H. U. and 68 coauthors, 2007. OSIRIS – The Scientific Camera System Onboard Rosetta. Space Science Reviews 128(1), pp. 433–506.

Kim, J., Kolmogorov, V. and Zabih, R., 2003. Visual Correspondence Using Energy Minimization and Mutual Information. Proceedings of the ninth IEEE International Conference on Computer Vision (ICCV) 2003 pp. 1033–1040.

Kirk, R. L., 1987. III A fast finite element algorithm for two-dimensional photoclinometry. Ph.D. Thesis (unpubl.), California Intstitute of Technology, Pasadena.

Kirk, R. L., Barrett, J. M. and Soderblom, L. A., 2003a. Photoclinometry Made Simple...? In: ISPRS Working Group IV/9 Workshop "Advances in Planetary Mapping".

Kirk, R. L., Howington-Kraus, E., Redding, B., Galuszka, D., Hare, T. M., Archinal, B. A., Soderblom, L. A. and Barrett, J. M., 2003b. Highresolution topomapping of candidate MER landing sites with Mars Orbiter Camera narrow-angle images. Journal of Geophysical Research.

Loop, C. and Zhang, Z., 1999. Computing rectifying homographies for stereo vision. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR) 1999 1, pp. 125–131.

Lowe, D. G., 2004. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision 60, pp. 91–110.

McEwen, A. S., 1991. Photometric functions for photoclinometry and other applications. Icarus 92(2), pp. 298–311.

McGlone, J. C., Mikhail, E. M. and Bethel, J. S. (eds), 2004. Manual of photogrammetry. fifth edition edn, American Society for Photogrammetry and Remote Sensing.

Poulet, F., Cuzzi, J. N., Cruikshank, D. P., Roush, T. and Dalle Ore, C. M., 2002. Comparison between the Shkuratov and Hapke Scattering Theories for Solid Planetary Surfaces: Application to the Surface Composition of Two Centaurs. Icarus 160(2), pp. 313–324.

Schenk, P. M., 2008. Cartographic and Topographic Mapping of the Icy Satellites of The Outer Solar System. ISPRS Archives XXXVII(B4), pp. 967–972.

Scholten, F., Oberst, J., Matz, K.-D., Roatsch, T., Wählisch, M., Speyerer, E. J. and Robinson, M. S., 2012. GLD100: The near-global lunar 100 m raster DTM from LROC WAC stereo image data. Journal of Geophysical Research 117, pp. E00H17.

Shao, M., Chellappa, R. and Simchony, T., 1991. Reconstructing a 3-D Depth Map from One or More Images. CVGIP: Image Understanding 53(2), pp. 219–226.

Shkuratov, Y., Starukhina, L., Hoffmann, H. and Arnold, G., 1999. A Model of Spectral Albedo of Particulate Surfaces: Implications for Optical Properties of the Moon. Icarus 137(2), pp. 235–246.

Sierks, H. and 65 coauthors, 2015. On the nucleus structure and activity of comet 67P/Churyumov-Gerasimenko. Science.

Smith, D. E. and 23 coauthors, 2001. Mars Orbiter Laser Altimeter: Experiment summary after the first year of global mapping of Mars. Journal of Geophysical Reasearch 106(E10), pp. 23689–23722.

Smith, D. E. and 30 coauthors, 2010. The Lunar Orbiter Laser Altimeter Investigation on the Lunar Reconnaissance Orbiter Mission. Space Science Reviews 150, pp. 209–241.

Soderblom, L. A. and Kirk, R. L., 2003. Meter-Scale 3-D Models of the Martian Surface From Combining MOC and MOLA Data. In: Lunar and Planetary Institute Science Conference Abstracts. Abstract #1730.

Vedaldi, A. and Fulkerson, B., 2008. VLFeat: An Open and Portable Library of Computer Vision Algorithms. Version 0.9.20, accessed 26 October 2015, <http://www.vlfeat.org>.

Wu, C., 2011. VisualSFM: A Visual Structure from Motion System. Version 0.5, accessed 18 November 2015, < http://ccwu.me/vsfm/>.

Wu, C., Agarwal, S., Curless, B. and Seitz, S. M., 2011. Multicore bundle adjustment. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2011 pp. 3057–3064.