PHOTOCLINOMETRY MADE SIMPLE...?

Randolph L. Kirk*, Janet M. Barrett, and Laurence A. Soderblom
Astrogeology Team, U.S. Geological Survey, Flagstaff, Arizona (rkirk@usgs.gov)

Commission IV, Working Group IV/9

KEY WORDS: photoclinometry, shape-from-shading, topographic mapping extraterrestrial mapping, Mars

ABSTRACT
We describe a series of programs and scripts under development for the USGS digital cartography and image processing software system ISIS that will make it possible for users to create digital elevation models with single-pixel resolution from images of interest. The key program is an implementation of the two-dimensional photoclinometry (shape-from-shading) algorithm of Kirk (1987) with a graphical user interface designed to reduce the difficulty of operation substantially. Supporting programs and scripts can be used to estimate the surface and atmospheric photometric parameters used in the photoclinometric model, in order to maximize the quantitative accuracy of the results.

1. Introduction
Photoclinometry (more descriptively shape-from-shading) is a cluster of related techniques for estimating topography that have been exploited by planetary scientists for more than 50 years (van Diggelen 1951). Most of the interest in these methods is stimulated by their complementarity to stereo toponmapping. Whereas stereomatching compares finite areas to estimate heights no closer than every 3–5 pixels, photoclinometry generates accurate relative heights at single-pixel resolution, and of course can do so with only a single image. The fundamental taxonomy of photoclinometry (PC) methods addresses the dimensionality of the region over which height information is sought. One-dimensional methods (e.g., Davis and Soderblom 1984) determine slopes from image brightness (radiance) along a line or curve and integrate these into an elevation profile. If only the slopes are of interest, these can be estimated and studied point-by-point via "zero-dimensional" photoclinometry (Byer and McEwen 2002; Byer et al. 2003). These methods require supplemental assumptions about the direction of slope in order to recover the two independent slope components from a single radiance observation at each pixel. Casting the PC problem in terms of a two-dimensional domain removes the need for ad hoc assumptions, because the continuity of the surface supplies a second constraint on the slopes (Wildey 1975; Horn and Brooks 1990). Two-dimensional (2D) PC methods thus build up a digital elevation model (DEM), i.e., a full three-dimensional model of the surface, in which height relations in the sun direction come mainly from radiance information and those in the cross-sun direction mainly from continuity. 2D PC is naturally of greater interest for mapping applications than approaches that produce only isolated topographic profiles, but their increased usefulness comes at a cost of increased software complexity and increased effort by both computer and human. Not only is the amount of data to be processed for a DEM much greater than that for a profile, the calculation itself becomes iterative and much more complex because of the non-local effects that must be dealt with. This complexity has kept 2D PC from being widely applied. Our purpose in this abstract is to describe and demonstrate new software tools that are intended to make the method accessible to a relatively wide audience of planetary mappers and geologists, while improving the accuracy of the results by allowing users to determine appropriate values of the photometric parameters needed as inputs.

2. Two-Dimensional Photoclinometry
Our new software uses the methods developed by Kirk (1987), in which the PC problem is formulated as one of least-squares fitting of a synthetic image calculated from the desired DEM to the actual observed image. The method is doubly iterative: the least-squares problem is nonlinear in the DEM elevations and must be solved by repeated linearizations, for each of which the resulting matrix equations are sparse and are solved by iterative techniques such as successive over-relaxation (SOR; Ortega 1970; Press et al. 1986). These methods resolve local details of the DEM much more rapidly than longer-wavelength topography, so to speed overall convergence, the software incorporates multigridging (Brandt 1977), which is the ability to adjust long-wavelength components of the topography by working on grids 2, 4, 8, … times coarser than the input image. The resulting path of iteration can be very complex, and a large amount of experience unfortunately shows that control of the calculation cannot be automated successfully in the general case. The user must decide interactively when to change the working grid resolution and what SOR parameters to set. The SOR parameters present the main challenge in getting 2D PC to work. If the SOR weight w is too low or the number of SOR steps per linearization is small, convergence can be very slow. Increasing these parameters speeds convergence up to a point, but setting them too high can cause the method to diverge rapidly. The practical limit on the "aggressiveness" of the SOR parameters depends sensitively on the data (low-contrast images of smooth surfaces tolerate more aggressive iteration) and can even change as iteration proceeds. In some cases a near-optimal level for the SOR parameters can be found; for very rough surfaces it may be necessary to alternate between aggressive overrelaxation (w>1), which converges the solution at most DEM points rapidly, and underrelaxation (w<1), which repairs localized divergence occurring during the overrelaxation steps. Perhaps the best news about this complex process is that, provided divergence is avoided, the choice of SOR parameters does not affect the final DEM, only the amount of effort needed to produce it.

* Correspondence author.
Figure 1. Graphical user interface (GUI) for two-dimensional photoclinometry provided by ISIS program pc2d. Clockwise from upper left: Main window with displays for image, DEM, synthetic image, and DEM increment (any 2 displayed, panned, and zoomed simultaneously), widgets to control iteration, and other information; plot of error statistics vs. SOR step within a Newton step; plot of error statistics vs. Newton step, with offsets where resolution was changed; log window giving error statistics in tabular form. Image is MOC M09-03811, in martian south polar cap.

To control the 2D PC algorithm effectively, the user must have access to a large amount of information about the parameter settings, the evolving state of the DEM, and the remaining errors. We have therefore created a new photoclinometry program called pc2d within the USGS ISIS image processing system (Eliason 1997; Gaddis et al. 1997; Torson and Becker 1997), with a graphical user interface that presents this information efficiently (Fig. 1). The input image and DEM are displayed alongside one another (the image modeled from the DEM, and the most recent increment to the DEM can also be displayed in these windows) and can be zoomed and scrolled in unison; statistics describing the progress of iteration are presented both as plots and in tabular form. All of this information is continuously updated, and buttons and slide controls are provided so the user can start and stop iteration, change the working grid resolution, and change the parameters for the SOR method as needed. From this description it should be apparent that we have not actually made photoclinometry simple. Training is needed merely to learn the function of the many displays and controls in pc2d, and more training is needed to interpret the displays in terms of the best settings to complete the calculation. As discussed below, care and experience are also required to evaluate the reliability of the result based on the illumination geometry, data available for photometric calibration, and evidence for local albedo variations. Nevertheless, this process is considerably simpler than controlling the same calculations with the software originally developed by Kirk (1987), which has a harder-to-learn text interface and presents the user with less information. We believe that, distributed as part of ISIS and with adequate documentation, pc2d will put two-dimensional photoclinometry methods within reach of a relatively wide audience.

A second ISIS PC program is truly simple in that it operates without user supervision, performing a fixed number of full-resolution iterations with fixed parameters. Such iteration suffices to add local details to a starting DEM (e.g., from stereo or altimetry) that is already accurate at long spatial wavelengths (Giese et al. 1996; Soderblom and Kirk 2003), hence the name pcsi for “smart interpolation” of the input DEM to higher effective resolution. A second application of pcsi is to derive DEMs for comparatively small image areas (~100 pixels on a side or less). For such images, the linearized equations can be solved directly rather than by SOR, so the unsupervised calculation does not require a low-resolution DEM to start with. The latter mode of operation is useful in calibrating PC to other, more accurate sources of topography as described below. The amount of memory required by pc2d and pcsi depends on the image size, method of solution (i.e., direct or SOR), and details of the illumination geometry and photometric function. Program pcinfo is provided to calculate the amount of memory that should be allocated in running pc2d or pcsi with a given image.

Unlike our older software (Kirk 1987), pc2d and pcsi obtain the geometric information needed for PC automatically by accessing the labels included in the ISIS input image. For images in the native geometry of the camera (sometimes called “Level 1” images), the program can read both the original ISIS label design and the newer standard that was developed to support pushbroom scanning imagers beginning with the Mars Global Surveyor MOC (Malin et al. 1992; Malin and Edgett 2001). Most framing cameras, including those of Viking Orbiter and Voyager, are supported by both variants of ISIS, but a few can currently be used only with programs written to the earlier standard. In addition to Level 1 images, the PC programs can also operate on image data that have been mapprojected (referred to as “Level 2”), subject to certain common-sense restrictions. Processing of Level 2 data will work if the projection is geometrically similar to an image (e.g., an Orthographic projection centered near the subspacecraft point of the data is ideal) but not if the projection is highly distorted (e.g., Sinusoidal projection, away from the equator and central meridian). In addition, the illumination and viewing geometry in a Level 2 dataset must be nearly uniform: a single mapprojected image or a mosaic of images taken in rapid succession can be processed, but a mosaic of images taken at different seasons and times cannot. The newer ISIS projection program lev1tolev2 automatically propagates the information about illumination and viewing geometry needed for photoclinometry to the Level 2 image labels, but mosaic erases this information when combining images. We have created a new program, lev1prop, that will add the illumination and viewing information from a single Level 1 image (chosen by the user to be representative) to the labels of a Level 2 mosaic.

We find the ability to work with such mosaics to be extremely valuable in light of “Murphy’s law of cartography” that the features of greatest interest are always on the boundary between images (e.g. Figueredo et al. 2002).
3. Surface Photometric Modeling

In addition to making 2D PC simpler, we have also been developing tools and procedures to make the process more accurate by deriving appropriate photometric parameters for use in the calculation. Both the surface photometry (Kirk et al. 2000) and (for Mars) that of the atmosphere (Kirk et al. 2001b) are of concern, as described by Jankowski and Squyres (1991). The surface photometric function is an important determinant of the amount of contrast (shading) for given topography. Equivalently, if the wrong photometric function is used, the amplitude of topography inferred will be wrong, and, to a lesser extent, the DEM will be distorted. The iterative nature of the PC algorithm makes it desirable to use a simple, empirical photometric function for the surface, rather than a more complex, physically motivated model. McEwen (1991) pointed out that either Minnaert's (1941) law or a lunar-Lambert photometric function provides a reasonable working approximation to Hapke's (1981; 1984; 1986) function, provided the limb-darkening parameter (k or L respectively) and overall brightness are chosen appropriately as a function of phase angle. ISIS program \textit{pho\_emp\_global} fits Minnaert or lunar-Lambert models to a given Hapke model over nearly the entire visible hemisphere much as McEwen did, yielding a table of parameter versus phase angle that can be used for photometric normalization of images and mosaics with program \textit{photomet} (Kirk et al. 2000b; 2001b). The program \textit{pho\_emp\_local} is similar but intended specifically for determining model parameters for PC. This program performs a fit over a distribution of surface slopes about a mean plane that has specified incidence and emission as well as phase angles. It thus indicates the best empirical photometric parameters for use in processing a given image. Both \textit{pho\_emp\_global} and \textit{pho\_emp\_local} rely on Hapke parameter estimates from the literature, turning these into parameter values for the computationally simpler empirical functions; Fig. 2 shows results applicable to Mars. It is also possible to determine the best-fitting empirical model de novo, by comparing image data with a coregistered DEM (Carr et al. 1994; Sullivan et al. 1996; Giese et al. 1996). We plan to provide ISIS tools for such fitting in the future. Empirical function fits to data at multiple phase angles can then be used to estimate or at least constrain the physically more meaningful Hapke parameters (Kirk et al. 2003b).

4. Atmospheric Photometric Modeling

For Mars, the photometric behavior of the atmosphere is as or more important for PC as that of the surface. Although the photometric behavior of atmospheric particulates is reasonably well constrained (Tomasko et al. 1999; Markiewicz et al. 1999), the amount of such material varies with place and time, affecting the relative contrast of topographic shading in an essentially unpredictable way. If the vertical scale is uncertain, the results of PC may still useful for visualization of the shapes of surface features, but for quantitative studies (e.g., Herkenhoff and Kirk 2001; Herkenhoff et al. 2002; Kirk et al. 2001a; Kirk et al. 2002a; 2002b; 2003a) a reliable calibration of scale and hence an accurate contrast correction is needed. A well-established approach to atmospheric correction is to find a shadow, measure its radiance as an estimate of the radiance contributed by the atmosphere everywhere in the image (the "haze"), and subtract this value from all pixels (e.g., Davis and Soderblom 1984; Jankowski and Squyres 1991). This method is not ideal because different shadows receiving different amounts of sky illumination and hence having different radiances will yield corrections of varying accuracy. In addition, shadows may not be available if the incidence angle or the local slopes are small. We have therefore developed alternate methods based on the photometric result to an independent source of topographic data. These methods are similar, but one involves a comparison in the image (radiance) domain and the other in the DEM (elevation) domain. Which is more appropriate for a given image depends on the resolution of the \textit{a priori} DEM and the illumination geometry, in relation to the scale and arrangement of topographic features.

Haze estimation in the image domain (forward fitting method) is a noniterative process: model radiances are calculated from a \textit{a priori} topographic data, based on a realistic surface photometric function but with no assumed atmospheric scattering. A linear regression of the observed image radiances on the model radiances yields the haze as its intercept value. We initially demonstrated this approach by manually measuring slopes at two locations on the profile of MOLA altimetry (Zuber et al. 1992; Smith et al. 2001) acquired simultaneously with a MOC image of interest and evaluating the photometric model for these slopes to compare with point measurements of image radiance. The haze estimated in this way, when used for PC, yielded a DEM consistent with the MOLA data (Herkenhoff and Kirk 2001; Herkenhoff et al. 2002). More recently, we have used Kirk's (1987) interactive software system both to shade an \textit{a priori} DEM with the surface photometric function and to perform a regression of the image data on the shaded model. In order to automate the process and make it widely available, we are developing an ISIS script \textit{pho\_fit\_forward} that will carry out the operation as shown in Fig. 3a. The linear regression program \textit{linfit} used by this script exists, and the shading program \textit{shade} is currently being developed; both programs have many other potential uses. Our experience with mapping candidate landing sites for the Mars Exploration Rovers (Kirk et al. 2001a; 2002a; 2002b; 2003a) indicates that the forward-fitting approach must be applied with care, particularly when contrast in the image comes from small topographic features that are barely resolved.
in the *a priori* DEM. In such cases, it is necessary to perform the fit over a region in which the shaded DEM best captures the morphology of the features seen in the image. It is also useful to verify whether the fit is reasonable by examining a scatterplot (two-dimensional histogram) of observed versus model radiances (Fig. 4). ISIS program `hist2d` provides this capability. If the topographic features are broad and well resolved in the DEM, however, selection of the domain of the haze fit is less critical.

Fitting in the DEM domain (inverse fitting method) must be performed iteratively. A trial value of the haze is subtracted from the image and the topography is estimated by PC. The results are compared with the *a priori* DEM. The haze estimate and an overall brightness normalization of the image are then adjusted until good agreement is obtained. The normalization mainly affects the tilt of the clinometric topography toward or away from the sun, whereas the haze controls the amplitude of relief, so the two effects can be disentangled if the DEM contains recognizable local features.

We initially developed the inverse fitting method by implementing a simple one-dimensional PC calculation as a spreadsheet (Soderblom et al. 2002). Given a list of radiances extracted from a column of an image, the spreadsheet produces a clinometric elevation profile and compares it to a collocated MOLA profile. The fit is determined interactively by adjusting the haze and normalization values stored in the spreadsheet. This process is somewhat tedious but works well, provided that the MOLA profile clearly resolves topographic features at a variety of slopes, that these slopes cross the profile at a fairly consistent angle, and that the illumination is neither along-strike nor at right angles to the profile (either of which would eliminate the relation between profile slope and shading). The applicability of this one-dimensional approach can be increased by using profiles interpolated from gridded data (a DEM from MOLA or from stereo), which are not limited in their orientation as are individual MOLA orbit tracks. The ISIS program `cube2asci` is used to extract profiles from both image files and DEMs, while `mocmola` finds particular MOLA profile acquired simultaneously with a given MOC image.

A DEM derived by 2D PC can also be compared with an *a priori* DEM to produce an improved haze estimate. A single iteration of this process provides enough information to calculate a very good haze estimate, because the DEM scale depends on haze in a known way (to a good approximation it is inversely proportional to the mean image radiance minus assumed haze). To date, we have carried out this process manually, by using the interactive PC software, collecting spot elevations of matching points in the clinometric and *a priori* DEMs, and calculating the regression relation between these spot heights (Kirk et al. 2001a; 2002a; 2002b; 2003a). This manual process has one advantage in some situations: the use of carefully selected spot heights can lead to a more reliable correlation between the DEMs if the *a priori* dataset does not resolve small features well. If the *a priori* DEM is sufficiently detailed, however (as with the forward method, it is highly advisable to confirm this by examining the cross-histogram of the *a priori* and clinometric DEMs), then the comparison process can be automated. Fig. 3b shows the structure of a script `pho_fit_inverse` currently being developed, that uses `pcsi` to perform trial PC on a small image patch, `linfit` to regress the entire resulting DEM on an *a priori* model, and a single loop to adjust both haze and normalization until the best fit is attained.

We note that either the forward or inverse atmospheric fitting methods can be applied to multiple images and even to successive segments of a MOC or other scanner images to produce maps of the spatial and temporal variation of atmospheric haze. We are currently engaged in a closely related analysis of MOC Wide Angle images from the Geodesy campaign (Caplinger and Malin 2001). These images cover such large areas that planetary curvature plays the role of an *a priori* DEM, and we are able to use a script similar to those in Fig. 3 to adjust the atmospheric optical depth for each image so that the photometric correction program `photomet` removes the shading associated with this curvature. The primary result intended be a global color image mosaic (Kirk et al. 1999) with improved uniformity, but the potential of all these fitting methods to provide useful information for atmospheric studies as well as for mapping should be evident.

### 5. Spatially Varying Albedo

The photoclinometry software described here makes the crucial simplifying assumption that the photometric properties of the surface are uniform over the area being modeled. (It is certainly possible in principle for photoclinometric methods to solve simultaneously for topography and spatially varying photometric parameters, but for each added parameter to be mapped an additional input image will be required as a constraint.) Photometric variations, in particular variations in net surface reflectivity (loosely speaking, “albedo”), will be misinterpreted by the existing single-image algorithm as slope-related shading, leading to a distinctive pattern of artifacts in the DEM. For example, an intrinsically darker patch will be modeled as a slope away from the sun and will be flanked by a ridge on the upsun side and a trough on the downsun side. The presence of roughly sun-aligned troughs and ridges in the DEM is a sensitive indicator that can indicate the presence of albedo variations even when they are too subtle for the human visual apparatus to pick out or are hidden in high contrast parts of the image (more severe albedo variations will be visible to the eye...
and a signal that the image is probably not suitable for PC analysis unless the albedo can be corrected for). The trough-and-ridge pattern of the albedo-related artifacts also makes it possible to suppress them very effectively by digital filtering of the DEM (Fig. 5). The model must first be rotated so that the streaks are in the sample direction; enlarging the DEM at the same time minimizes loss of resolution due to the resampling. The streaks are aligned with the direction in which the dependence of brightness on slope is strongest (sometimes called the direction of "characteristic strips"; Horn 1970) which is close to the sun direction but depends on details of the image geometry and photometric function. The photoclinometric programs (including pcinfo) calculate and report this direction. Once the DEM is rotated, a three-step process implemented in the ISIS procedure dstripe suppresses stripes in a given size range. First, the DEM is lowpass filtered with a boxcar one pixel in width (across the stripes) and the minimum length of stripes to be removed. Second, the results of the lowpass filter are highpass filtered with a boxcar equal in width to the stripes and one pixel in length. The minimum stripe length is typically set to about 3 times the maximum width. The second filter yields an estimate of the "stripe component" of the DEM that is then subtracted from the original. The process can then be repeated with a larger (typically 3x as great) stripe width and length to remove larger stripe features caused by larger albedo patterns. Once filtering is complete, the DEM must be rotated and scaled back to its original orientation and size.

If an estimate of the spatially varying albedo across the image can be obtained, then the image can be corrected by dividing the image by this albedo map after subtracting the haze and topographic shading, and the photometric behavior of the surface must not include contrast reversals or even substantial contrast changes with phase angle (an especially severe problem for Io; see Simonelli et al. 1997). In principle, such an image can be used to correct small-scale as well as broad albedo variations. In practice (in particular, for images from flyby encounters), low-incidence images often have relatively poor resolution, and the accuracy with which the low- and high-incidence images can be coregistered also limits the ability to correct for small albedo features. Despite these limitations, we have found that dividing out a low-incidence image can dramatically improve the uniformity of surface albedo as reflected by both visual appearance of the image and artifacts in the clinometric DEM (Fig. 5; see Figueredo et al. 2002). Our ability to perform photoclinometry in a map-projected coordinate system is extremely useful for performing photoclinometry on the ratio of two coregistered images.

6. References


