A METHOD FOR SHAPE-FROM-SHADING USING MULTIPLE IMAGES ACQUIRED UNDER DIFFERENT VIEWING AND LIGHTING CONDITIONS

Keith Hartt and Mark Carlotto
The Analytic Sciences Corporation
55 Walkers Brook Drive
Reading, MA 01867
(617) 942-2000

ABSTRACT

A new formulation for shape-from-shading from multiple images acquired under different viewing and lighting conditions is presented. The method is based on using an explicit image formation model to create renditions of the surface being estimated, which are synthetic versions of the observed images. It is applicable in a variety of imaging situations, including those involving unknown non-uniform albedo. A probabilistic model is developed based on typical characteristics of the surface and minimizing the difference between the synthetic and observed images. This model is used to arrive at a Bayesian formulation of the shape-from-shading problem. Techniques are presented to compute an estimate that is statistically optimal in the sense that it is the expected value of the surface, given the set of observations derived from it. The method is applied to Viking imagery of Mars.

1. INTRODUCTION

A variety of "shape from ..." methods have been developed for estimating the relief of three-dimensional surfaces under a variety of surface, viewing, and lighting assumptions. Two methods which have been used in terrestrial and planetary remote sensing are shape-from-shading and shape-from stereo, or stereoscopy. Shape-from-shading techniques usually involve the determination of surface gradients from shading information in one or more images (the latter includes the method of photometric stereo). Elevations are then determined by integrating the gradients. In stereopsis, the heights of objects are determined by computing the disparity of the object between the left and right image and relating disparity to height via the camera model. Ideally, in shape-from-shading the image(s) should be acquired at nadir, or at least from the same perspective so that there is no geometrical distortion due to relief. For stereo, the images should be acquired under similar lighting conditions so as to minimize confusion in the matching process. Under typical imaging situations neither condition is perfectly satisfied.

The above motivates a new method which integrates shape-from-shading and stereoscopy. Others have proposed methods for combining the two approaches. Ikeuchi [1] proposed that shading information be used instead of interpolation to fill in elevations derived from feature-based stereo. More recently, Marroquin et al [2] have suggested the possibility of taking a probabilistic approach to a cooperative solution of perceptual problems. They also suggest an optimization criterion for single-image uniform-albedo shape-from-shading, but it differs from ours as a special case. Most notably, it is based on the extraction of surface gradients as opposed to direct determination of elevations. To the best of our knowledge, no physically-based, systematic approach to the solution of the integrated problem has been developed.

The method presented in this paper is appropriate for the extraction of shape information from a set of observed images in situations when an image-formation model and any necessary ancillary information is available or can be derived, so that the rendering of a synthetic version of each member in the observed-image set from the surface is possible. Attention is concentrated on planetary remote-sensing, where certain assumptions are valid, and lighting and viewing conditions for each image are known. Elevation information is conveyed by the observed images through shading and parallax.

The basic idea behind the method is to compute a surface $\mathbf{z}$ for which the following cost or energy functional is small:

$$U(\mathbf{z}) = \int\int \sum_k | I_k(x,y) - I_k^{syn}(x,y) | + \lambda |\nabla z(x,y)|^2 \, dx \, dy \quad (1)$$

Each term in the summation is the absolute difference between an observed image $\mathbf{I}_k$ and a synthetic image $\mathbf{I}_k^{syn}$ rendered from the elevation surface $\mathbf{z}$ under the lighting and viewing conditions for the k^{th} sensor. A constant of proportionality λ weights a term which dictates that, loosely speaking, the elevation surface be smooth. More precisely, it represents a prior assumption about the characteristics of the elevation surface.

A flavor of the computational technique follows that is used to arrive at an estimate of the elevation surface which is statistically optimal in the sense that it is the expected value of the elevation surface, given the set of observed images formed from it. The surface is repeatedly swept, with a change proposed at each pixel site. The change is accepted with a probability which increases with the merit of the change, based on a local measure derived from the above energy functional. The convergence behavior of the local updates is such that the elevation states satisfy the energy functional *globally*, in the sense that favorable low-energy states are highly probable. These states are then averaged to obtain an approximation of the optimal surface estimate.

The organization of the remainder of the paper is as follows. Section 2 discusses the effects of illumination, sensor geometry, and surface reflectance on the image formation process. Section 3 presents the probabilistic aspects of the model in the context of Markov random fields. The optimization criterion is constructed using a Bayesian formulation. Section 4 presents computational methods and considerations used to implement the method. Section 5 presents experimental results. Synthetic imagery is used to test the technique in a controlled environment. Initial results obtained by applying the technique to Viking Orbiter imagery of Mars are also presented.

2. IMAGE FORMATION MODEL

In order to recover the shape of imaged surfaces via the creation of synthetic renditions of the surface, an explicit image formation model is developed. The surface is described by a function z(x,y); i.e., it is assumed single-valued. Surface orientation at (x,y) is described through the partial derivatives which make up the gradient, $\partial z(x,y)/\partial x$ and $\partial z(x,y)/\partial y$, denoted p(x,y) and q(x,y), respectively. The image formation process is derived by first calculating the surface, or scene radiance, and then the irradiance falling on the image plane, through the appropriate projection.

Planetary Remote Sensing Model

An image-formation model valid for planetary remote sensing, the application area of the experiments presented here, is given by the following, and is depicted in Fig. 1. The illuminant, the sun, is assumed to be a point source located at infinity. Its position in gradient space for image k is (p_k, q_k), where $p_k=\tan\phi_k\cos\theta_k$ and $q_k=\tan\phi_k\sin\theta_k$, with ϕ the zenith angle, and θ the azimuth angle with respect to the x axis. If the atmosphere is modeled as an optically thin, semi-infinite, plane-parallel, horizontally homogeneous air mass, and if the variation in elevation is small relative to the depth of the atmosphere over the field of view, the scene radiance can be approximated by [3]

$$L_k(x,y) = \left(\frac{\rho}{\pi}\right) T_u \{ I_o T_d R_k[p(x,y),q(x,y)] + I_s r[p(x,y),q(x,y)] \} + L_p \qquad (2)$$

where ρ is the albedo, I_o the extra-terrestrial solar irradiance, T_u and T_d atmospheric path transmittances, I_s the sky irradiance, and L_p the atmospheric path radiance. A reflectance map, or function, relates surface orientation and brightness. $R_k(.)$ is the reflectance map for directly illuminated areas, r(.) for shadowed areas, which are lit by ambient light.

Horn [4] reviews reflectance functions which have been used for creating shaded renditions of surface topography. In planetary remote sensing, the Minnaert function $\cos^{\kappa} i \cos^{\kappa-1} e$ is often used, where incidence angle i is the angle between the surface normal and a vector in the direction of the light source, the emittance angle e the angle between the surface normal and a vector in the direction of the sensor focal point. The Minnaert function with $\kappa=1/2$ has been shown to be a good model for the maria of the moon. In fact, the earliest application of shape-from-shading (also known as photoclinometry) was in lunar mapping [5]. The Minnaert function is widely used in remote sensing studies of Mars (see, for example [6]). With $\kappa=1$, the function is equivalent to the Lambertian reflectance function, which has been used to reconstruct isolated surface features on Mars [7].

In shadowed areas the first term in (2) is zero. A surface patch is shadowed if the incidence angle i of the light-source is greater than 180°, or most generally, if a ray between the patch and the sun intersects the surface (see [8], e.g., for hidden surface removal algorithms). Since $I_o T_d \gg I_s$ typically, in directly lit regions the ambient term in (2) can be ignored.

Lambertian Model

A Lambertian surface radiates uniformly in all directions. Here, reflectance maps for point-source and ambient illumination in (2) are presented. Directly lit portions of the surface are affected by point-source illumination; the reflectance map for a Lambertian surface directly illuminated by a point light-source at (p_k,q_k) is given by

$$R_k(p,q) = \frac{(pp_k+qq_k+1)}{(p^2+q^2+1)^{1/2}(p_k^2+q_k^2+1)^{1/2}} \qquad (3a)$$

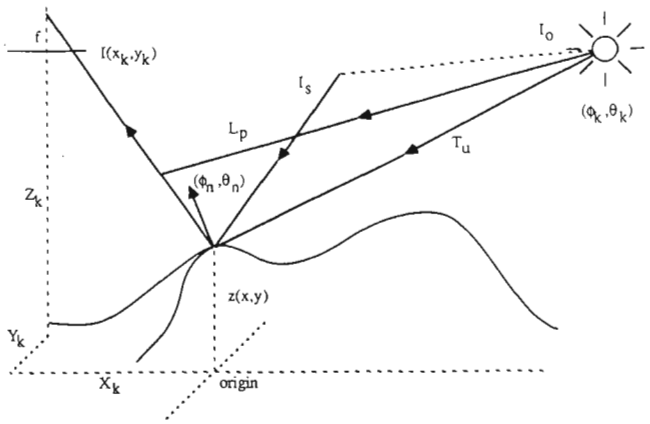


Fig. 1. Planetary image-formation model.

where the notation giving the dependence of p and q on (x,y) has been suppressed. The point-source reflectance map is a normalized inner product of the surface normal vector and a vector in the direction of the light source. $R_k(p,q)$ is maximum at (p_k,q_k), which corresponds to a patch oriented orthogonally to the light source.

In shadowed areas, the surface is lit by ambient light. The reflectance map for a Lambertian surface illuminated by a uniform hemispherical source is

$$r(p,q) = (1+(p^2+q^2+1)^{-1/2}) / 2 \qquad (3b)$$

Note that because the source is circularly symmetric, the mapping depends on only the magnitude of the gradient.

Image-Plane Projection

The model allows viewing the surface from an arbitrary perspective, which involves projecting the surface onto the image plane of the sensor as depicted in Fig. 1. Assume a standard down-looking frame camera with focal length f and position (X_k, Y_k, Z_k); points on the surface z(x,y) are mapped onto the image plane via a perspective transformation [8]. For the sensor geometry in Fig. 1

$$x_k = \frac{f\,(x-X_k)}{[Z_k-z(x,y)]} \;;\; y_k = \frac{f\,(y-Y_k)}{[Z_k-z(x,y)]} \qquad (4)$$

In general, a point's position in (x_k, y_k) depends on x, y, and z; however, if the relief is small relative to the height of the sensor (i.e., $|\Delta z| \ll Z_k$), and the field of view small relative to the sensor position (i.e., $|\Delta x| \ll X_k$ and $|\Delta y| \ll Y_k$), the coordinate transformation is given by

$$x_k = \frac{f}{Z_k}\left(x - \frac{X_k}{Z_k}z(x,y)\right) \;;\; y_k = \frac{f}{Z_k}\left(y - \frac{Y_k}{Z_k}z(x,y)\right) \qquad (5)$$

The above simplifications amount to assuming that a parallel projection holds, which is not unrealistic in remote sensing. In (5), the surface at (x,y) is simply shifted in the image by a disparity proportional to z(x,y).

Image-Formation Inversion

It is worthwhile at this point to examine the benefit of using multiple images in recovering shape information. Note that the problem is an inversion problem; i.e., the image formation processes associated with the observed images are inverted to derive a surface which is consistent with the observations. If the mapping from the surface to an observed image is one-to-one and uncorrupted, a single image is adequate; however, this situation is not realistic. As an example, consider that an image of a surface region which is shadowed and not lit by ambient light conveys no shape information. If another image is available in which this region is lit, it contains the missing information.

Additionally, for some regions of the domains of the reflectance maps described above, the functions are relatively constant, so little information is conveyed. Generally, it makes sense to utilize all available information in reconstructing the surface.

Non-Uniform Albedo

Thus far, homogeneous reflectivity characteristics of the surface have been assumed, i.e., constant albedo is assumed in (2). However, the method is also applicable when albedo is spatially varying. Consider the following image-formation model:

$$I_k(x,y) = \rho(x,y)\, R_k(p(x,y),q(x,y)) \qquad (6)$$

The reflectance map $R_k(.)$ relates surface orientation and image intensity for a patch having unit albedo under the lighting conditions for image k. All shape information, which is manifested as shading, is contained in $R_k(.)$; the albedo $\rho(.)$ contains information about the reflectivity of the surface.

If the albedo is known, it is simply used in the rendering process. If it is unknown, the problem is to find an elevation surface and an albedo field which are consistent with the observed image intensities, through comparison of the observations and synthetic renditions. The albedo is incorporated by associating with each state of the elevation surface an albedo image which makes the comparison most favorable; e.g., the albedo at (x,y) is related to the surface by

$$\rho(x,y) = \arg\min \sum_k [I_k(x,y) - \rho(x,y) R_k(p(x,y),q(x,y))]^2 \qquad (7)$$

which is a simple least-squares optimization. Note that at least two observed images are required in order for the method in this situation to yield meaningful results; a single image can be explained perfectly on the basis of albedo alone, independent of surface shape. Other possibilities include constraining the albedo field to consist of one of a few values, if appropriate. The basic idea is to use an albedo consistent with the shading and observations. The energy functional (1) is then still small for favorable elevation states.

3. MARKOV RANDOM FIELD MODEL

In this section, a probabilistic model is developed which reflects realistic assumptions about the observed image(s) and the underlying elevation. Background on Markov random fields is provided as a context for presentation of the model. The distribution of the elevation surface is given, and the derivation of relationships to well-known processes and the effect of lattice approximation is presented. The distribution of the observed image is then derived based on the elevation model. These distributions are then used in a Bayesian formulation of the optimization problem.

Background on Markov Random Fields

Markov random fields, described through Gibbs distributions, have become popular for image modeling [9], [10], [11]. The main utility of the Gibbs distribution lies in the fact that a random field can be described in a tractable manner in terms of local interactions which are in accord with desired behavior. Gibbs models allow a rich, robust description of interesting classes of images within practical computational constraints.

The Markov property in a multi-dimensional setting on a lattice is defined with respect to a *neighborhood system* $\eta=\{\eta_{ij}\}$, where η_{ij} denotes the neighborhood of pixel site (i,j). The Markov property is given by

$$p(x_{ij} \mid x_{kl},\ (k,l)\neq(i,j)) = p(x_{ij} \mid x_{kl},\ (k,l)\in\eta_{ij}) \qquad (8)$$

Here, the neighborhood consisting of the eight nearest neighbors is adequate. A Markov random field can always be expressed in the form of a Gibbs distribution [9]:

$$p(\mathbf{x}) = \frac{1}{Z}\exp\{-U(\mathbf{x})\};\ U(\mathbf{x}) = \sum_{c\in C} V_c(\mathbf{x}) \qquad (9)$$

where c is a set of pixels that are neighbors of each other, called a *clique*; C is the set of all cliques; $U(\mathbf{x})$ is an *energy functional*; $V_c(\mathbf{x})$ is a *clique potential* which describes the interaction among members of clique c; Z is a normalizing constant. The relationship between the Gibbs distribution and the Markov property is demonstrated by the following. Consider

$$p(x_{ij} \mid x_{kl},\ (k,l)\neq(i,j)) = \frac{p(\mathbf{x})}{\int p(\mathbf{x})\, d(x_{ij})} \qquad (10)$$

where $p(\mathbf{x})$ is given by (9). Note that the normalizing constant and all clique potentials cancel with the exception of the potentials involving cliques of which (i,j) is a member. This gives rise to the Markov property (8).

Distribution of the Elevation Surface

In this section the model for the elevation surface is presented, which yields the prior distribution in the context of a Bayesian formulation. The model is based on using a membrane smoothing constraint as the energy in a Gibbs distribution, which is a special case of a formulation used by Marroquin et al [2]. The resulting process has some interesting properties which are derived here. Most notably, the process described by the Gibbs distribution is the well-known Wiener process, also known as the Wiener-Levy process or Brownian motion. Fractional Brownian motion [12], of which Brownian motion is a special case, has been used previously to model natural surfaces [13] (usually, fractional Brownian motion with a fractal dimension smaller than that of Brownian motion is used, but in this case the field is not Markov). Discrete versions of the process are of obvious practical relevance. In one-dimension, the discrete process is Markov, and is equivalent in distribution to a sampled continuous process. However, the lack of a natural ordering on a lattice prevents this from carrying over fully to two dimensions. These ideas are developed in the following discussion.

Consider first the Wiener process in one dimension, which has the following representation [14]:

$$w(t) = \int_0^t v(s)\, ds \qquad (11)$$

where v(.) is zero-mean normal white noise. There is a minor problem with initial values in this representation; assume that the process is actually a family parameterized by an additive offset, and exists over an adequately sized window. Additionally, assume the existence of appropriate stochastic integration and differentiation operators. The process has independent increments: for $t_0<t_1\le t_2<t_3$

$$E\{[w(t_3)-(t_2)][w(t_1)-w(t_0)]\} = 0 \qquad (12)$$

The variances of the increments are given by

$$E\{[w(t_1)-w(t_0)]^2\} \sim |t_1-t_0| \qquad (13)$$

Now, a discrete, finite-support process is formed by sampling the Wiener process within a window. The joint distribution of the samples, due to the samples having independent normal increments, can be expressed as a product of normal distributions. That product has a Gibbs form with energy functional:

$$U(\mathbf{w}) \sim \sum_i \left\{ \frac{w(iT)-w((i-1)T)}{T} \right\}^2 \qquad (14)$$

where $\mathbf{w}$ is a vector of samples; and T is the sampling interval. Note that each element in the sum corresponds to a clique potential; the following Markov property is then evident (the sampling notation is hereafter abandoned):

$$p(w_i \mid w_k,\ k\neq i) = p(w_i \mid w_{i-1}, w_{i+1}) \qquad (15)$$

Furthermore, letting $T\rightarrow 0$ in (14),

$$U(w(\cdot)) \sim \int \left[\frac{d}{dt}w(t)\right]^2 dt \qquad (16)$$

A Gibbs distribution having this energy describes the Wiener process since the distribution dictates that infinitesimal increments are independent, which is evident from the above representation (11).

The above shows that a 1D discrete process based on approximating the derivative with a first difference is equivalent in distribution to a sampled continuous process. In two dimensions the Wiener process becomes

$$U(w(\cdot,\cdot)) \sim \int |\nabla w(x,y)|^2 \, dx \, dy \qquad (17)$$

From the properties of gradients, w(x,y) is circularly symmetric in the plane, so that the process along any line in the plane is a 1D Wiener process and equivalent in distribution to the process along all other lines, independent of orientation. Therefore, in 2D, (13) becomes

$$E\{[w(x_1,y_1)-w(x_0,y_0)]^2\} \sim [(x_1-x_0)^2+(y_1-y_0)^2]^{1/2} \qquad (18)$$

The 2D discrete process which is used in the experiments is now examined, which is a lattice approximation to the 2D Wiener process, and has energy

$$U(\mathbf{z}) \sim \sum \{(z_{ij} - z_{i-1,j})^2 + (z_{ij} - z_{i,j-1})^2\} \qquad (19)$$

The differences approximate the partial derivatives in (17). The process has the 2D Markov property

$$p(z_{ij} \mid z_{kl}, (k,l)\neq(i,j)) = p(z_{ij} \mid z_{i-1,j}, z_{i+1,j}, z_{i,j-1}, z_{i,j+1}) \qquad (20)$$

However, unlike the 1D case, this process is *not* the one obtained by sampling the 2D continuous process. This can be shown by a simple contradiction. By the above (19),

$$E\{(z_{ij} - z_{i-1,j})(z_{ij} - z_{i,j-1})\} = 0 \qquad (21)$$

this implies

$$E\{(z_{i,j-1} - z_{i-1,j})^2\} = 2E\{(z_{ij} - z_{i-1,j})^2\} \qquad (22)$$

however, by (18),

$$E\{(w_{i,j-1}-w_{i-1,j})^2\} = \sqrt{2}E\{(w_{ij}-w_{i-1,j})^2\} \qquad (23)$$

where E{} denotes an expectation. The 2D Wiener process is such that only increments along a line are independent. Therefore, in the sampled process, the x-direction first-difference is not independent of the y-direction first-difference, which is the basis for the contradiction. The 2D discrete process **z** is not isotropic in the sense that increment variances along other than lines along x or y do not match those of the sampled isotropic, continuous process. In any event, the 2D discrete process is at least an approximation of the sampled process, and has a valid probability distribution because it has a Gibbs form. A realization is shown in Fig. 2.

The above shows the utility of the Gibbs distribution in deriving Markov properties, and of particular relevance here, the effect of discrete approximation of derivatives.

Distribution of the Observations

In this section the distribution of the observed image intensities is derived based on the elevation distribution and a simple point-source illuminated Lambertian surface. The intensity distribution conditioned on a known elevation surface is then provided for the most general image-formation model, which is the distribution directly used in the processing. For the simple model, because the intensity at a site is related to the elevation surface by the point-source reflectance mapping, one can derive the marginal distribution of the intensity using methods for determining the distribution of a function of a random variable having a known distribution [14]. Because the distribution of the elevation surface is circularly symmetric, the orientation of the coordinate system is arbitrary with respect to the (x,y) plane. Choose an orientation such that the derivative in the elevation plane, in the direction of the light source, is along x. The normalized (unit-albedo) intensity I is then given by the point-source reflectance mapping (3a) in this coordinate system

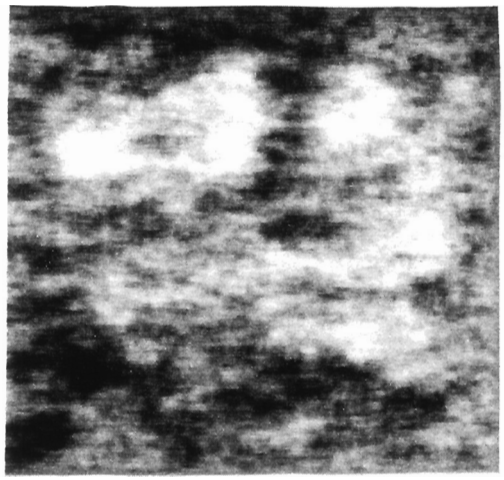


Fig. 2. Realization of the elevation model.

$$I = R(p,q) = \frac{(pp_o+1)}{(p^2+q^2+1)(p_o^2+1)} \qquad (24)$$

where $(p_o,0)$ is the orientation in gradient space of the light source. A plot of R(p,0) for various values of p_o is given in Fig. 3. Assume an off-nadir light-source position and a gradient distribution with a small variance, which constrains the domain of R(.) to a region about the origin in the gradient space (these are valid assumptions for the Martian imagery). The partial derivative in the light source direction then dominates the behaviour of the reflectance map; i.e.,
$R(p,q) \approx R(p,0)$. The p.d.f. for the intensity is given by

$$f_I(I) \approx \sum_i \frac{f_p(p^{(i)})}{|R'(p^{(i)},0)|} \qquad (25)$$

where R'(.) is the derivative with respect to p; and $\{p^{(i)}\}$ are the roots of R(p,0). These are obtained by solving for p in

$$((1+p_o^2)I^2-p_o^2)p^2 - 2p_op - (1+p_o^2)I^2 - 1 = 0 \qquad (26)$$

The second-order polynomial is obtained by manipulating the expression for R(p,0). Note that the intensity is squared in the polynomial so that roots must be checked for the sign of the intensity. There are either one or two roots for each positive I; in the case there is only one, the other root of the polynomial yields -I, which corresponds to a shaded patch, so the term for that root in the sum in (25) contributes to the area of an impulse in $f_I(.)$ at I=0. The distributions for $\sigma_p = \sigma_q = 0.125$ and various values of p_0 are shown in Fig. 4.

Recall that the gradient distribution is normal, and note in Fig. 4 that R(.,0) is close to linear about p=0 for large p_o (off-nadir light sources). For these reasons, the intensity distributions are nearly normal themselves, because R'(.,0) is constant in the p.d.f. expression (25). This implies that a first-order Taylor series expansion at p=0 of R(.,0) is a good approximation under the above assumptions, which results in significant computational savings. This linear relationship is in agreement with the computer graphics trick of simply applying a derivative operator and then scaling and offsetting to arrive at a realistic-looking shaded rendition from an elevation surface.

The intensity distribution derived above is a marginal distribution based on normally distributed gradients. However, in the absence of strict assumptions, the distribution is robust in the sense that it is adequately valid for any elevation surface in which gradients on the surface are distributed about the origin in gradient space. Figure 5 shows the agreement of the theoretical distribution with histograms of portions of the Martian imagery.

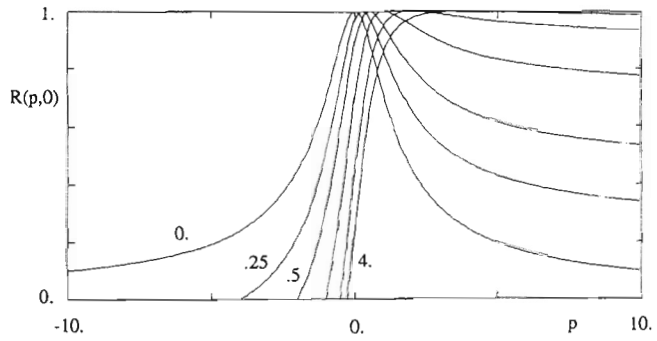


Fig. 3 . R(p,0) for various values of p_o.

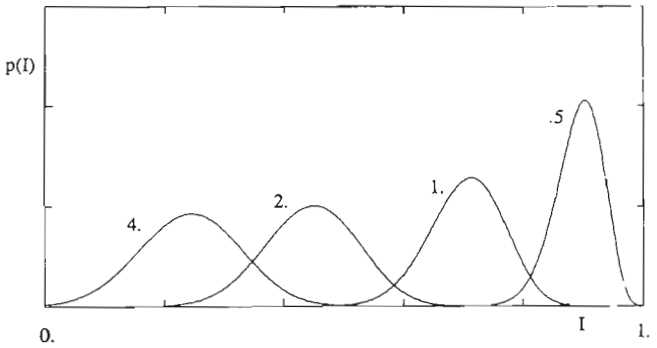


Fig. 4. Intensity distributions for various values of p_o.

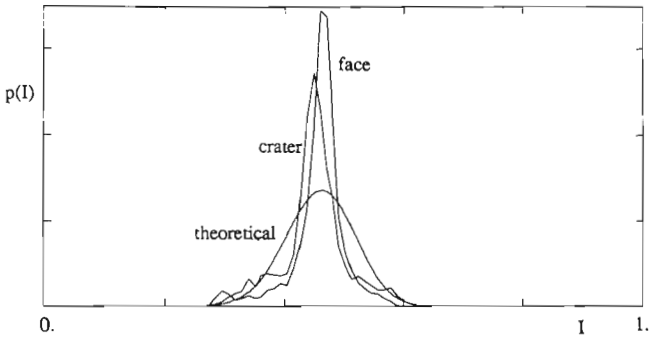


Fig. 5. Martian and theoretical intensity distributions.

Knowledge of the theoretical distribution also allows inferences about the image formation process; e.g., parameters of the reflectance mapping may be derived.

Necessary for the Bayes formulation is the distribution of the observed images, conditioned on the elevation surface, which has a Gibbs form with energy

$$U(\{\mathbf{I}_k\} \mid \mathbf{z}) = c_I \sum_{i,j} \sum_{k} | I_k[i,j] - I_k^{syn}[i,j] | \qquad (27)$$

where c_I is a constant; $\{\mathbf{I}_k\}$ is the set of observed images; $\{\mathbf{I}_k^{syn}\}$ is a set of renditions of the elevation surface obtained through appropriate reflectance mapping; and (i,j) are mapped from (i_k,j_k) via the sensor geometry model as described in Section 2; i.e., the observed images are mapped to a common projection. The distribution is such that an observed image which matches a rendition of the surface is highly probable. The incorporation of the absolute difference in the model as opposed to enforcing strict correspondence allows for three real-world considerations: 1) noisy data; 2) distortion due to the absence of strict model validity; 3) a smooth posterior distribution (Section 4), resulting in better convergence properties of the estimation procedure. The conditional distribution at a single site is a two-sided exponential distribution, chosen for its robust nature, centered on $I_k^{syn}[i,j]$. Note that the rendered intensity at (i,j) depends on the set $\{z_{ij}, z_{i-1,j}, z_{i,j-1}\}$ which is used in the local slope approximation. This set is a three-pixel clique in the Gibbs distribution.

Optimization Criterion

The optimal estimate of the elevation surface is given by

$$E\{\mathbf{z} \mid \{\mathbf{I}_k\}\} = \int \mathbf{z}\, p(\mathbf{z} \mid \{\mathbf{I}_k\})\, d\mathbf{z} \qquad (28)$$

the expected value of the elevation surface conditioned on the set of observed images. The posterior distribution is given by Bayes' theorem:

$$p(\mathbf{z} \mid \{\mathbf{I}_k\}) = \frac{[\Pi\, p(\mathbf{I}_k \mid \mathbf{z})]\, p(\mathbf{z})}{\int [\Pi\, p(\mathbf{I}_k \mid \mathbf{z})]\, p(\mathbf{z})\, d\mathbf{z}} \qquad (29)$$

This distribution is also Gibbsian, with energy

$$U(\mathbf{z} \mid \{\mathbf{I}_k\}) = \sum U(\mathbf{I}_k \mid \mathbf{z}) + U(\mathbf{z}) \qquad (30)$$

Note the implicit assumption that degradations associated with each observed image are independent.

4. OPTIMIZATION METHOD

In this section the computational aspects of the method are presented. The Metropolis algorithm [15] is presented, which allows Monte Carlo approximation [16] of the integral in (28) to obtain the optimal estimate of the surface. This procedure is augmented by a multi-resolution approach framed in the context of scale-space continuation [17].

Metropolis Algorithm

The above optimal estimate (28) is obtained by sampling from the posterior distribution, and then averaging these samples to obtain a Monte Carlo approximation of the expected value of the elevation surface, given all available observations. The Metropolis Algorithm is the procedure by which realizations from the posterior distribution are generated (a different sampling algorithm is given in [18]). The procedure is given as follows. Given the present state $\mathbf{x}$ of the system, one randomly choses another configuration $\mathbf{x}'$. The ratio

$$r = p(\mathbf{x}')/p(\mathbf{x}) \qquad (31)$$

is evaluated. If $r>1$, $\mathbf{x}'$ becomes the new system state. If $r<1$, $\mathbf{x}'$ becomes the new system state with probability r. As the total number of proposed changes gets large, the system states become samples of the joint distribution. When this ocurrs, the system is said to be in *equilibrium*. In precise terms, the procedure generates a Markov chain (in time) which has a limiting distribution $p(\mathbf{x})$.

This technique is applied to the posterior distribution, so that the state of the elevation surface evolves through time, guided probabilistically by the conditioning observations and the characteristics of typical surfaces. The change proposals are chosen so that they differ from the current state at only one site. Because the posterior distribution is Markov with respect to the eight nearest neighbors, the ratio computation (31) involves a small set of intensity and elevation values. To avoid imposing the time causality on the lattice, the change sites are chosen according to the *coding* method [9].

Scale-Space Continuation

It is obvious that a state proposal is not independent of the current state. As noted in [16], it is possible that the system, dependent on the initial configuration and the nature of the posterior distribution, finds itself in a region of the state space for which the transition probability to a more representative region is very small, so that the transition does not occur in a practical amount of time. The system gets stuck in a locally

probable region. Although the potential for disaster is not as severe as with a deterministic update algorithm, which is the basis for simulated annealing [10], some protective measures are in order.

Scale-space continuation was proposed by Witkin, et al [17] in a deterministic setting to combat the local-minimum problem in their solution of the signal-matching problem. The idea of a continuation method [19] is to embed a problem in a parameterized family of problems in which the parameter is related to the ease with which a member of the family is solved. A hierarchy of problems are solved in which a solution at one level serves as the starting point for the next, more difficult level. In scale-space continuation, the parameter is the standard deviation σ of a Gaussian-shaped smoothing kernel which is applied to the input signals. The smoothed input gives rise to well-behaved optimization criteria in the sense of avoiding getting stuck in locally optimal states. Here, a family of posterior distributions is created through smoothing the conditioning intensity images. The hierarchy is formed by smoothing the full-resolution observed images with a kernel with $\sigma=T$ (sampling interval). The Fourier transform of the kernel is itself a Gaussian kernel in frequency with $\sigma_f = 1/(2\pi T)$. The smoothed data are therefore downsampled by a factor of two without excessive aliasing. Lower levels are created from their parents in the same way. Samples (in the statistical sense) are probabilistically tracked through scale by applying the sampling algorithm to the lowest-level posterior distribution first. When the system reaches equilibrium, the elevation surface is upsampled and serves as the initial configuration for the system at the higher level. Conveniently, the relationship among first-order differences among elevation surfaces at different levels is given by (18). The hierarchical scheme generally is known as a coarse-to-fine or multi-resolution approach, which has been widely taken in image processing [20].

The optimization strategy combines probabilistic updating and scale-space continuation. The ideal scenario is one in which at the highest level the states of the system at equilibrium are in the most probable region of the state space. States within this region are averaged to form the optimal estimate. The technique through experimentation has proved to be robust and consistent with the design philosophy of achieving a good solution while expending a minimum of resources.

5. EXPERIMENTAL RESULTS

In this section initial experimental results are presented. At this time we have no results on the simultaneous estimation of albedo and shape; it is an area of future work. The image formation model used here assumes constant albedo, and a point-source Lambertian reflectance function, which was not approximated. The multi-resolution strategy was used with ~200 sweeps of the image performed at each resolution.

Figure 6 shows results which demonstrate using multiple images. The renditions shown in (a) and (b) were synthetically generated from natural terrain, given by a digital elevation map of an area in upstate New York. Oblique sun angles were chosen, a) 80°, b) 75°, to test the technique in a shadowed, non-linear environment. The image (a) light -source is to the left of the page, image (b) from the bottom, as is the view. The view angles were a) 0°, b) 70°. The same renditions based on elevation derived using the shape-from-shading method with only image (a) are shown in (c) and (d). The renditions using both the nadir and the oblique view are shown in (e) and (f).

Next, the technique was applied to actual Viking orbiter imagery of Mars. We assume negligable atmospheric scattering and absorption effects, and negligable ambient light. The albedo ρ, which scales the Lambertian point-source reflectance mapping, was estimated as $\rho=\mu/R(0,0)$, where μ is the mean of the data histogram. The sun angle is 63°. Fig. 7 (a) shows a 60x60 area containing a crater. The estimated z-map is depicted in (b). Another example containing the famous Martian "face" is shown in Fig. 8. Again, the z-map is depicted in (b). In (c), a reconstruction of the image is created by synthetically rendering the extracted elevation surface under the same lighting conditions as in the formation of the Viking image. In (d) a synthetic view generated under different lighting conditions is shown.

6. SUMMARY

A new technique that combines shading and parallax information from multiple images for 3D surface reconstruction is presented. It is based on an explicit model which represents the characterisics of the surface and includes the image formation process. The method essentially inverts the image formation processes, specifically, computes the shape of a surface from one or more images acquired under different viewing and lighting conditions. The inversion is cast as an optimization problem that involves computing the expected value of the elevation surface given the observed images. Experimental results demonstrate the efficacy of the method using both synthetic and actual imagery. Specific contributions are: 1) the derivation of the relationship of a membrane smoothing constraint in a probabilistic setting with well-known processes, and the effect of lattice approximations; 2) a probabilistic optimization strategy that incorporates scale-space continuation and Monte Carlo estimation, which represents an alternative to simulated annealing.

REFERENCES

[1] K. Ikeuchi, "Constructing a depth map from images," AI Memo No. 744, MIT Artificial Intelligence Lab., Cambridge, MA, 1983.

[2] J. Marroquin, S. Mitter, and T. Poggio, "Probabilistic solution of ill-posed problems in computer vision," Journal of the American Statistical Assoc., Vol. 82, No. 397, March 1987.

[3] R. W. Sjoberg, "Atmospheric effects in satellite imaging of mountainous terrain," TR-688, MIT Artificial Intelligence Lab., Cambridge, MA, Sept. 1982.

[4] B. K. P. Horn, "Hill shading and the reflectance map," Proc. IEEE, Vol. 69, No. 1, Jan. 1981.

[5] T. Rindfleisch, "Photometric method for lunar topography," Photogrammetric Eng., Vol. 32, No. 2, March 1966.

[6] R. L. Wildey, "Generalized photoclinometry for Mariner 9," ICARUS, Vol. 25, p 613-626, 1975.

[7] M. J. Carlotto, "Digital imagery analysis of unusual Martian surface features," Applied Optics, Vol. 27, pp 1926-1933, May 1988.

[8] J. D. Foley and A. Van Dam, Fundamentals of Interactive Compter Graphics, Addison-Wesley, Reading, MA, 1983.

[9] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," J. Roy. Stat. Soc. B, Vol. 36, pp. 538-544, 1965.

[10] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," IEEE Trans. Pattern Anal. Machine Intell., Vol. PAMI-6, pp. 721-741, Nov. 1984.

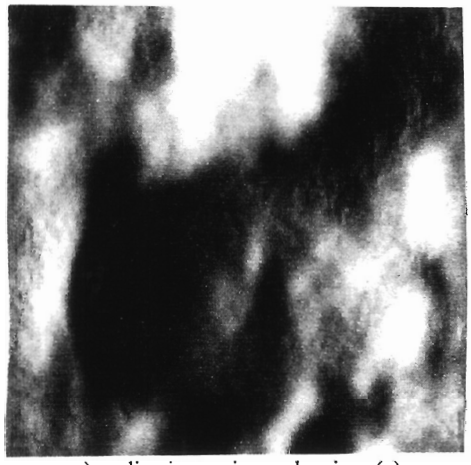
[11] H. Derin and H. Elliott, "Modeling and segmentation of noisy and textured images using Gibbs random fields," IEEE Trans. Pattern Anal. Machine Intell., Vol. PAMI-9, No. 1, Jan. 1987.

[12] B.B. Mandelbrot and J. W. Van Ness, " Fractional Brownian motions, fractional noises and applications," SIAM Review, Vol. 10, No. 4, pp. 422-437, Oct. 1968.

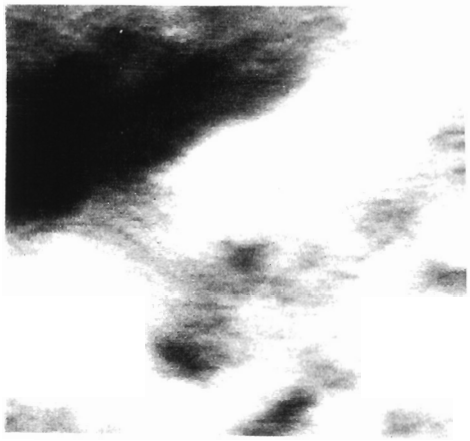
[13] B. B. Mandelbrot, The Fractal Geometry of Nature, Freeman, New York, 1982.

a) nadir view

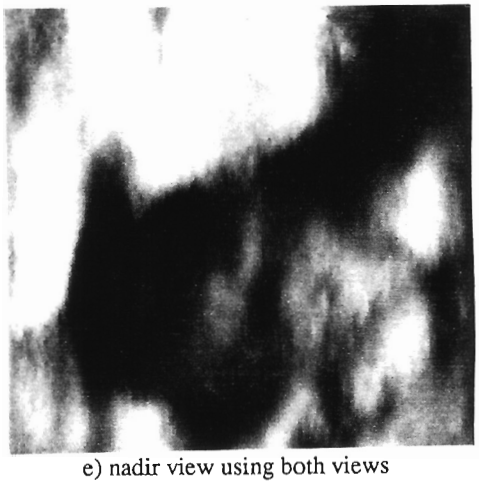
b) oblique view



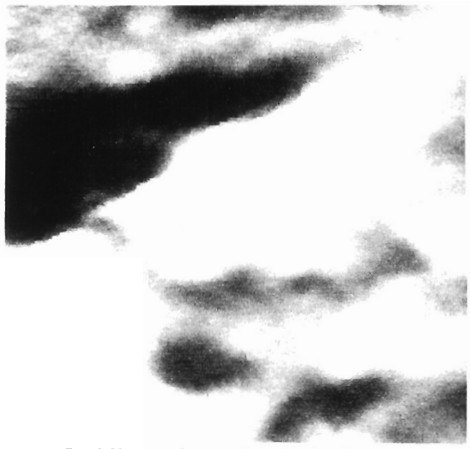
c) nadir view using only view (a)



d) oblique view using only view (a)



e) nadir view using both views



f) oblique view using both views

Fig. 6. Result of application of method to synthetic imagery.

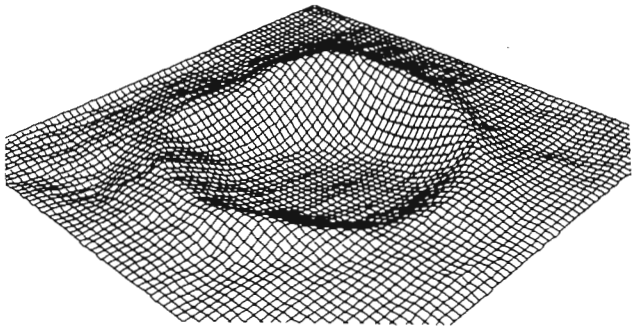
a) original image

a) original image

c) reconstructed image

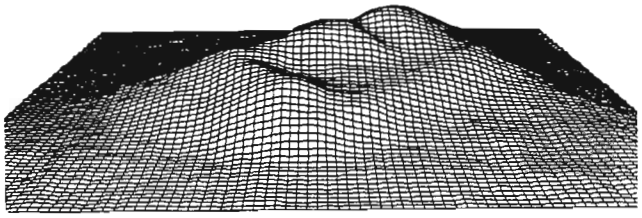
[14] A. Papoulis, Probability, Random Variables, and Stochastic Processes, McGraw Hill, New York, 1984.

[15] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, " Equations of state calculations by fast computing machines," J. Chem. Phys., Vol. 21, No. 6, pp. 1087-1092, June 1953.

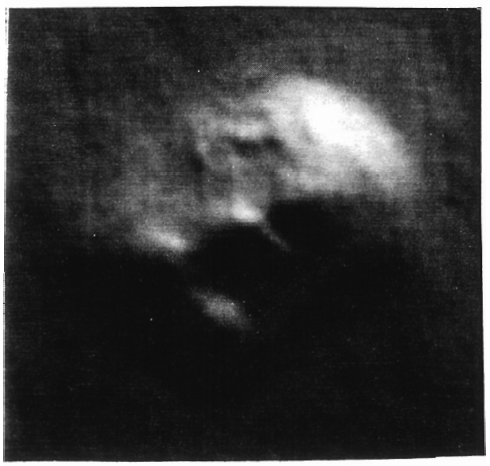


b) extracted elevation

Fig. 7. Result of application of method to Martian crater.



b) extracted elevation



d) synthetic view

Fig. 8. Result of application of method to Martian face.

[16] J. M. Hammersley and D. C. Handscomb, Monte Carlo Methods, John Wiley & Sons, New York, 1964.

[17] A. Witkin, D. Terzopoulos, and M. Kass, "Signal Matching through scale space," Int. J. Comp. Vision, pp. 133-144, 1987.

[18] M. Creutz, " Microcanonical Monte Carlo Simulation," Phys. Rev. Letters, Vol. 50, No. 19, pp. 1411-1414, May 1983.

[19] G. Dahlquist and A. Bjork, Numerical Methods, N. Anderson (trans.), Prentice-Hall, Englewood Cliffs, NJ, 1974.

[20] A. Rosenfeld (Ed.), Multiresolution Image Processing and Analysis, Springer,Verlag, Berlin, 1984.