USC-SIPI REPORT # 116

Computational Vision Algorithms for Synthetic Aperture Radar Imagery

by

Robert Thomas Frankot

Research supported in part by the National Science Foundation
Grant No. MIP-84-51010 and by the Hughes Aircraft Company,
IBM, and AT&T .

COMPUTATIONAL VISION ALGORITHMS FOR SYNTHETIC APERTURE RADAR IMAGERY

by

Robert Thomas Frankot

A Dissertation Presented to the
FACULTY OF THE GRADUATE SCHOOL
UNIVERSITY OF SOUTHERN CALIFORNIA
In Partial Fulfillment of the
Requirements for the Degree
DOCTOR OF PHILOSOPHY
(Electrical Engineering)

October 1987

Dedicated to my parents

John and Helen Frankot

Acknowledgments

I would like to thank my thesis committee for their time and effort in reviewing this dissertation. Special thanks are due to Rama Chellappa, my advisor, for introducing me to the field of image modeling and for enduring patience, guidance, and moral support.

A.A. Sawchuk, director of SIPI, has provided a very stimulating research environment and technical guidance as well. Charles Weber, director of the Communication Sciences Institute has given me advice and moral support through the years that have enriched my life as a graduate student. Other USC faculty, Ramakant Nevatia, William Lindsey, Richard Leahy, Solomon Golomb, and Ted Harris have been sources of inspiration.

I would like to thank the graduate students at SIPI, especially Shankar Chatterjee, Zhou Yitong, Surendra Lele, Herb Barad, Tal Simchony, Dan Antzoulatos, and Kung-Shiuh Huang, for their friendship and interesting discussions.

Alan Weber's technical support as the SIPI computer laboratory manager and Ray Schmidt's photographic processing are gratefully acknowledged. Gitta Domik and Franz Leberl of the Vexcel Corporation have provided the imagery and DTMs needed for testing the SAR shape from shading algorithm, and participated in some enlightening technical discussions. Annie Holmes and Martin Ruzek of the Jet Propulsion Laboratory were very helpful in providing SAR imagery for the stochastic modeling experiments.

The consistent financial support of the Hughes fellowship program is gratefully acknowledged. Thanks are due to Sue Baumgarten and Tim Jentes for their continuing support as department managers. My colleagues at Hughes Aircraft Company have contributed to my professional and scholarly growth through the years. Harold Davis provided valuable guidance at the beginning of my Ph.D. studies. Ralph Hudson has been a continuing source of insight and constructive comments. Dan Evans taught me the basics of SAR image simulation. Ted Broida, Bob Latter, Chuck McNary, Larry Rubin, Ray Bergenheier, Pete Demopoulos, Howard Nussbaum, Rich Wojslaw, Tim Putra, and others too numerous to mention deserve thanks.

I would like to acknowledge the partial support of the NSF Grant MIP-84-51010 and matching funds from Hughes Aircraft Company, IBM, and AT&T Information Systems.

The greatest thanks of all go to my parents, my brothers John and Tom, and my fiancée Melinda Melcon.

Contents

List of Figures

List of Tables

Abstract

Two classes of models, computational vision models and stochastic models, are examined for synthetic aperture radar (SAR) images of natural terrain. Algorithms are developed for surface topography estimation, image registration, and texture synthesis.

Shape from shading techniques are used for extracting topographic information. Previous numerical solutions to the shape from shading problem estimated the surface derivatives without ensuring that they are integrable, a serious drawback. The performance of a previously developed shape from shading technique is substantially improved using a fast least-squares algorithm to enforce integrability. The resulting algorithm is then applied to SAR by representing the terrain surface height relative to the "slant plane" (a plane parallel to the line-of-sight) and accounting for the radiometric properties of SAR imagery.

For noisy imagery, such as SAR, low frequency surface information is difficult to recover from a single image. A fast Fourier transform implementation of the integrability projection provides an efficient method for combining low frequency surface information with the shading information. This technique may be suitable for combining the SAR imagery and the low resolution altimetry provided by Magellan to construct high resolution topographic maps of Venus. The resulting algorithm is applied to SIR-B SAR imagery and the surface reconstructions are compared with stereoscopically derived digital terrain models. The use of auxiliary low frequency information is tested, allowing estimation of reflectance map parameters and providing coarse surface structure to complement the surface details obtained from shading information. This simulates the Magellan scenario.

An automatic registration algorithm is used for matching image intensity predictions with the observed images. This registration algorithm matches two

SAR images made from nearly orthogonal flight paths and matches a SAR image with an aerial photograph without detailed a priori knowledge of the terrain, two very difficult problems for images of hilly terrain.

Stochastic models for SAR image intensity based on nonlinear transformations of Gaussian random fields are introduced. Methods for selecting transformations to normality and model order are presented and tested on SAR imagery and synthesis of textures appearing in SAR imagery is demonstrated.

Chapter 1

Introduction

Algorithms developed for application to visual imagery are usually not directly applicable to radar imagery. Yet, many of the paradigms that have evolved in image processing and computer vision research *are* suitable for radar imagery. Translation of paradigms and methodologies into algorithms for a specific application requires a model, either explicitly in a formal derivation or implicitly in the experiences of the algorithm designer. Models previously used in image processing and computer vision research are inappropriate for SAR imagery because of fundamental differences between the physics of SAR image formation and that of conventional images at visible wavelengths. The purpose of this dissertation is to help bridge that gap by formulating models appropriate for SAR imagery and using those models for algorithm development. An algorithm for estimation of surface topography from SAR image intensity is developed. In order to evaluate the surface reconstruction results an algorithm for registration of SAR imagery to terrain elevation data is developed. Synthesis of texture in SAR imagery is also demonstrated using simple extensions of stochastic models used in image processing and image analysis.

1.1 The Role of an Image Model

In both image processing and computer vision two major schools of thought exist—one centered on models of image characteristics and one based on emulation of biological vision systems. The artificiality of this dichotomy is illustrated by contrasting two anecdotes from two different professional communities. Computer vision researchers have referred to vision as "controlled hallucination" [11], describing the underdetermined nature of the vision problem. A similar saying among radar image interpreters, "I wouldn't have seen it if I hadn't believed it", does not state that vision, per se, is a difficult problem. The human visual system (HVS) and cognitive processes can provide very accurate interpretations of radar imagery. However, radar imagery is so different from more familiar visible imagery that radar image interpreters must undergo extensive training to become familiar with the properties of radar imagery and to learn how objects of interest appear in radar imagery. In some sense they must re-learn how to see. Similarly, for machine processing of SAR and other non-visual image data, a thorough understanding of image characteristics is needed, even for applying methods inspired by the HVS. The key issue is not, therefore, which school of thought to follow, rather, it is how to model image data and how to utilize image models.

1.2 Classes of Models

This dissertation examines two kinds of models for SAR imagery—computational vision models and 2-D stochastic models. Computational vision models represent the image in terms of the physical scene being imaged and are useful for image simulation or for inferring object properties from the image. Attention is restricted to scenes composed of continuous surfaces instead of general 3-D scenes. The basic principle is that the observed image is a function of both the orientation of the surface relative to the radar and the intrinsic reflectivity of the materials composing the surface. The relationship between the angle of incidence and image intensity depends, in part, on surface microstructure which can often be characterized by a few unknown parameters. The computational vision models developed for SAR imagery parallel those for visual imagery, but the particular image coordinate system, the relationship between angle of incidence and image intensity, and the noise characteristics are unique to SAR. We have applied computational vision models to estimation of terrain surface elevation from SAR imagery and to image registration.

The 2-D stochastic models represent the image as a random field corrupted by random noise and do not require direct knowledge of scene characteristics. Stochastic models are useful for data compression, image segmentation, object detection, edge detection, and image simulation. Lognormal random field models are proposed as models for the SAR imagery. The lognormal models are treated as specializations of transformed-Gaussian random fields, which are straightforward extensions of the Gaussian random field models often used for visual imagery. After being passed through an invertible point-nonlinearity, the data is modeled as a Gaussian random field with linear spatial interaction. The use of lognormal random field models for radar imagery has been motivated by empirical considerations, based on statistical tests. Modeling the logarithm of SAR image intensity as a random field with linear spatial interaction (as done for lognormal data) has also been motivated by physical considerations, based on the high dynamic range of radar imagery, the multiplicative relationship between illumination, reflectance, and noise components. An additional physiological motivation arises from the approximately linear response of the early HVS to the logarithm of image intensity [141,153]. We have applied lognormal random field models to texture synthesis in radar imagery and developed model selection tests.

The stochastic models presented in this dissertation treat the observed image as a homogeneous random field. They could be extended to handle nonhomogeneous (nonstationary) data. For example, doubly stochastic models where the parameters of a Gaussian random field are piecewise-constant functions of image coordinates, partitioning the image into homogeneous regions, have been considered in the literature. The partitioning process is also a random field such as a Poisson line process [114], or a Markov random field [175,55]. Point processes have been used for modeling transitions between regions and also for modeling the placement of objects on a background [2,135,144,148] and the distribution of reflectivity categories in radar clutter [82]. A second way in which point-processes are useful is for deriving models for signal characteristics [151]. This has been done for deriving speckle statistics [81], and for deriving reflectivity models [9,152].

The choice of a modeling approach depends on several issues, such as

- Sensor characteristics.
- Scene characteristics.
- Applications.

For example, coherent imagery differs from noncoherent imagery both in the presence of speckle and in the point spread function (PSF) [59]. Radar imagery presents different problems than visual imagery, and infrared imagery differs markedly from visual and radar images. All three sensors measure electromagnetic radiation but the different spectral bands measure different physical aspects of the scene. Different type of scenes require different types of models even for identical sensors. For example, SAR images of the barren surface of Venus are as different as night and day from SAR images of a city. In the former case, multiple reflections are probably negligible, the local reflectance mechanism is primarily non-specular on the scale of a resolution element, and the surface conductivity is relatively low and nearly constant. The resulting image intensity is relatively well-behaved. In the latter case, specular reflection dominates, much of the returned energy may be from multiple reflections, and surface conductivity is highly variable. The result is an image with extremely high dynamic range and the breakup of extended features into collections of bright spots. Completely different models would be used for reconstruction of Venus surface topography from SAR images than would be used for delineating man-made structures in SAR images.

Different applications require different models. In computer vision, models of how objects appear in images are useful as a basis for inferring scene properties from image properties. In image processing, models for the image signal and degradations are used for developing restoration, enhancement, and bandwidth compression techniques.

The world is so complex that we cannot possibly model every aspect of it simultaneously. We must choose a few particular facets of reality to model in any given case. The maxim of Occam's razor—that models should be no more complicated than required for explaining the observed data—is followed in order to guide this choice. For example, when detecting small targets on a background, the best approach might be to represent the background as a 2-D random field rather than to use a 3-D computational vision model.

In this thesis, attention is restricted to a few basic models that are immediately useful. Models for discrete clutter such as roads and vehicles are not considered. Only so called diffuse clutter is examined. This is useful in applications where the diffuse component is considered important, as in radargrammetry. It is also useful when the diffuse component is considered to be clutter and objects of interest are detectable as entities not fitting the clutter model.

1.3 Thesis Organization

Chapter 2 develops a $2\frac{1}{2}$-D computational vision model for SAR imagery and reviews the relevant radar literature to support that model. Limitations and possible model extensions are discussed. Chapter 3 reviews shape from shading research and provides a method for enforcing integrability in the surface slopes estimated by shape from shading algorithms. A significant advance over existing algorithms results. Integrability is enforced at each iteration by projecting the possibly nonintegrable surface slopes onto the nearest integrable surface slopes. This projection is easily solved if the surface slopes are represented by a finite set of basis functions that are orthogonal and integrable. A Fourier basis function specialization of that projection is developed which allows efficient processing using fast Fourier transforms (FFTs). It provides a convenient method for including auxiliary low frequency information into the shape from shading solution. In Chapter 4 the computational vision models from Chapter 2 and the shape from shading algorithm from 3 are combined to derive a shape from shading algorithm for SAR imagery. This algorithm is similar to that of Chapter 3 except that the surface height is represented relative to a plane parallel to the line-of-sight (instead of a plane orthogonal to the line-of-sight) and a SAR reflectance model is used. Noise is not considered explicitly, although the noise level is shown to be very high, even after considerable noncoherent integration. The resulting algorithm is tested on SIR-B imagery and compared to independently obtained digital terrain models (DTMs). Chapter 5 applies the computational vision model to registration of images to DTMs and for registration of dissimilar images. The method used for registering SAR imagery to their companion DTMs for evaluating shape from shading results is presented and the use of shading information for registering images of smooth surfaces is demonstrated. Chapter 6 addresses stochastic models where transformed-Gaussian random field models are proposed for SAR imagery. A statistical decision rule for selecting model order simultaneously with the transformation to normality is developed, and the synthesis of textures in SAR imagery using lognormal random fields is demonstrated. The formulation provides a tractable modeling approach suitable for model-based image analysis applications such as edge detection, object detection, texture classification, and texture segmentation. Conclusions are presented in Chapter 7 along with suggestions for future research.

1.4 Thesis Contributions

The following contributions are reported in this dissertation:

- A noniterative method for enforcing integrability in shape from shading algorithms [48,51].
  - An efficient implementation of the integrability projection using FFTs.
  - Application of the integrability projection to improve an existing shape from shading algorithm.
- The development of a shape from shading algorithm for SAR imagery [50].

  - Efficient surface reconstruction and image simulation is obtained by representing surface height relative to the slant plane.
  - Auxiliary low resolution surface information is used to estimate reflectance map parameters.

- The approach of combining low frequency surface information with shading information [51,50].
  - A method for combining shading cues with the low resolution depth information provided by area correlation based stereo matching (as opposed to sparse feature based matches) is given.
  - An algorithm suitable for fusing high resolution SAR image shading information with the lower resolution radar altimetry data from the Magellan Venus Radar Mapper is provided.

- Demonstration of the use of shading information for registration of dissimilar images.

- A transformed-Gaussian random-field model formulation for image processing [47,49].
  - Decision rule for choice of random field models for SAR imagery.
  - Application of lognormal random field models to synthesis of SAR image textures.

Chapter 2

Computational Vision and Statistical Models for SAR Imagery

Image simulation from surface data relies on models that relate surface orientation and intrinsic reflectivity (albedo) to image intensity. Vision is, in some sense, the inverse of image simulation so that the same models used for image simulation are useful for computer vision applications. Accordingly, models that relate image intensity to surface properties are called computational vision models. Such models have been applied to simulation of radar and optical imagery [69,72,173]. With each approach, the simulated image is the product of a known albedo component and a known topography component. The function relating image intensity to the surface slopes is referred to as the reflectance map [71]. Two-dimensional stochastic models are useful for modeling albedo, surface structure, image intensity, and observation noise. In this Chapter we introduce computational vision models and stochastic models for SAR image intensity.

By following a common modeling approach for both radar and visual images it is possible to gain better insight into the *similarities* between, say, SAR images and aerial photographs. By comparing specific models for radar and visual images derived by the same approach it is also possible to gain greater insight into the *differences* between SAR images and aerial photographs. These differences fall into three categories

1. The spatial deformations introduced in projecting from a 3-D scene to 2-D image.
2. Reflectivity characteristics, expressed in terms of both the reflectance map and albedo.
3. Observation noise characteristics.

Each category is discussed below.

2.1 SAR Image Coordinate System

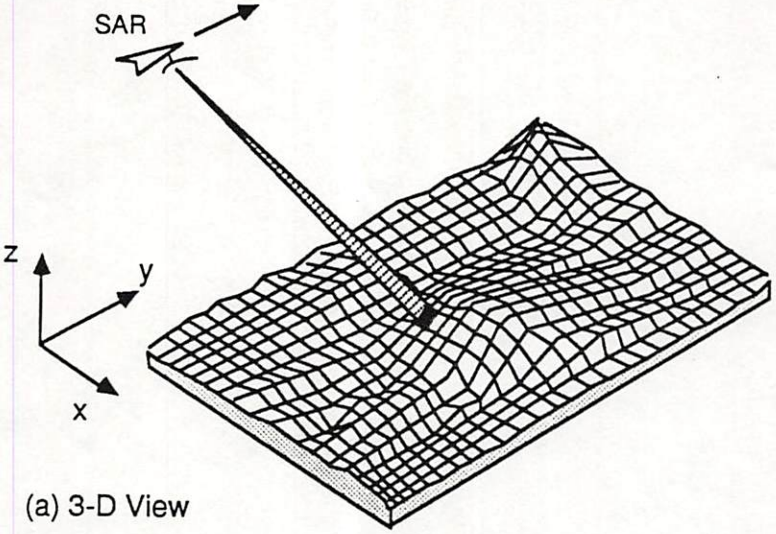
An image can be thought of as the projection of a 3-D scene into a 2-D representation. For an image created using conventional optics the spatial transformation from 3-D to 2-D is given by the perspective projection.

The coordinate system for a SAR strip map is very different, as depicted in Figure 2.1. For simplicity, consider the case of a straight flight path with the antenna main beam orthogonal to the flight path (zero squint angle). Assume that the x-axis is lined up with the r-axis. Then x may be referred to as ground range and is related to slant range r by

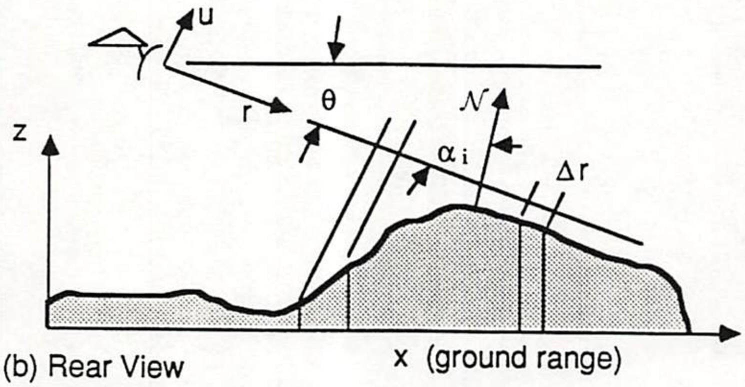
$$r^2 = x^2 + (h - z)^2 \tag{2.1}$$

where h is the radar altitude and the origin is located on the ground directly below the radar. The second SAR strip map axis, by design, represents the along track distance y. It is referred to as cross-range or azimuth. Thus, SAR image coordinates are a projection of the physical scene, $(x, y, z(x, y))$, into the "slant plane" (r, y). One result is foreshortening of the mountainsides that slope upward as r increases. In extreme cases, an increase in ground range x will decrease the slant range so that features in SAR imagery may appear in reversed order. This condition, sometimes referred to as layover, does not occur provided that

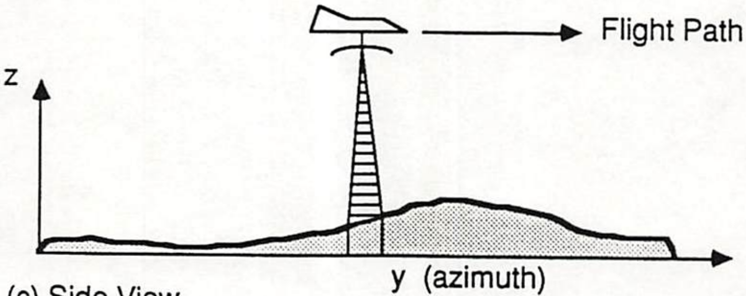
$$z_x \leq \frac{1}{\tan\theta}\,, \tag{2.2}$$



(a) 3-D View



(b) Rear View



(c) Side View

Figure 2.1: SAR Imaging Geometry

where θ is the grazing angle.

Clearly the transformation from surface coordinates (x, y, z) to SAR image coordinates (r, y) is not an orthographic projection. However, it is possible to approximate the SAR coordinates as an orthographic projection of a rotated version of the surface. Suppose that r is large relative to the image size so that arcs of constant range are approximately straight lines over the depression angle subtended by the image. Then the surface height $u(r, y)$ relative to the slant plane is given by the following rotation of $z(x, y)$;

$$\begin{pmatrix} r \\ y \\ u \end{pmatrix} = \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} r_0 \\ y_0 \\ u_0 \end{pmatrix} \tag{2.3}$$

where (r_0, y_0, u_0) is an arbitrary reference point. Here the transformation from slant plane surface coordinates to SAR image coordinates is indeed an orthographic projection. In the absence of layover, $z(x, y)$ being single valued implies that $u(r, y)$ is single valued. If both $z(x, y)$ and $u(r, y)$ are single valued functions then the data structure needed for representing the terrain and the methods used for synthesizing an image are very simple.

A very interesting parallel between the coordinate systems of SAR strip maps and conventional photographs becomes clear at this point. For SAR, the image coordinates can be approximated by an orthographic projection of the surface coordinates relative to a plane *parallel* to the line-of-sight (the slant plane). For conventional photographs the image coordinates can be approximated by an orthographic projection of the surface coordinates relative to a plane *orthogonal* to the line-of-sight. For example, high altitude aerial photographs taken from nadir could by represented closely by an orthographic projection of the surface topography if the field of view is small compared to altitude. This viewing geometry is illustrated in Figure 2.2

Note that the surface height u relative to the slant plane will have a ramp component—a term that increases approximately linearly with increasing range. For computation and storage purposes, the dynamic range of u can be reduced significantly by removing this ramp trend, giving

$$u_1(r, y) = u(r, y) - (r - r_0)\tan\theta \,. \tag{2.4}$$

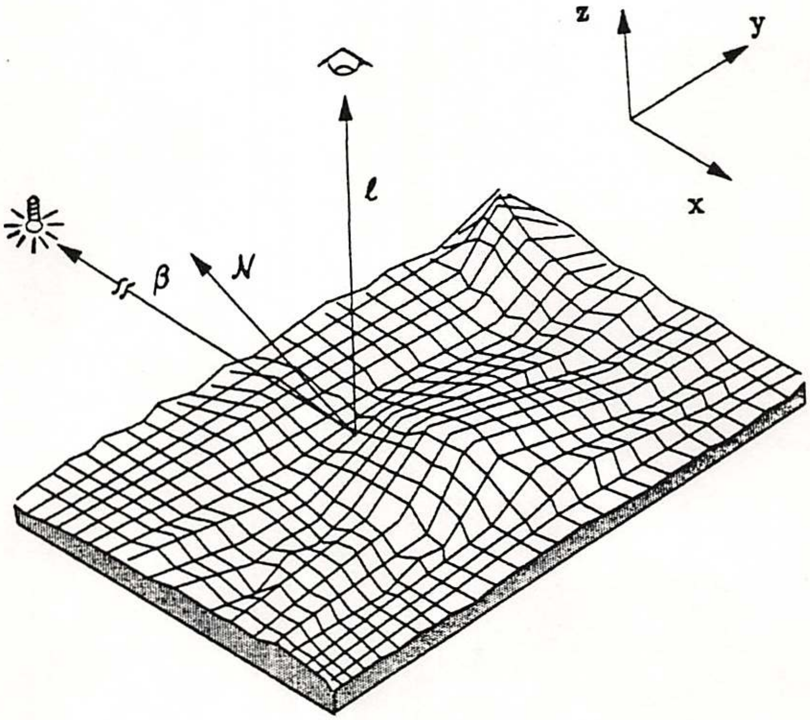


Figure 2.2: Geometry for aerial photograph example.

This transformation is utilized in the shape from shading algorithm of Chapter 4.

To evaluate the reflectance map for surface reconstruction or image synthesis it is necessary to evaluate the angle of incidence α_i or its cosine. The cosine of the angle of incidence is simply the normalized dot-product between the surface normal and the illumination vector,

$$\cos\alpha_i = \frac{\beta \cdot \mathcal{N}}{\| \beta \| \| \mathcal{N} \|} \,. \tag{2.5}$$

This has a particularly simple form when expressed in (r, y, u) coordinates. Here $\beta = (-1, 0, 0)$ and $\mathcal{N} = (-u_r, -u_y, 1)$ so that

$$\cos\alpha_i = \frac{u_r}{\sqrt{u_r^2 + u_y^2 + 1}} \tag{2.6}$$

where

$$u_r = \frac{\partial u}{\partial r} = \frac{\partial u_1}{\partial r} - \tan\theta$$

and

$$u_y = \frac{\partial u}{\partial y} = \frac{\partial u_1}{\partial y}$$

are the partial derivatives of u. Note that a point at which u_r passes through zero while decreasing is a shadow entry point and that $|u_r| \to \infty$ when transitioning to or from an area in layover.

For computer implementation, $u(r, y)$ is represented in a 2-D array with constant sample spacing in r and y as in a SAR strip map. Standard finite difference approximations are used for the partial derivatives. With finite central differences we get

$$u_y(k, l) = \frac{u(k, l+1) - u(k, l-1)}{2\Delta y} \tag{2.7}$$

and

$$u_r(k, l) = \frac{u(k+1, l) - u(k-1, l)}{2\Delta r}$$

or, in terms of the de-ramped version,

$$u_r(k, l) = \frac{u_1(k+1, l) - u_1(k-1, l) - 2\Delta r \tan\theta}{2\Delta r} \tag{2.8}$$

$$= u_{1r}(k, l) - \tan\theta \,,$$

where u_{1_r} is the central difference approximation to $\partial u_1/\partial r$.

As can be seen from (2.3) through (2.8), this slant plane representation allows relatively simple and efficient processing techniques to be applied for SAR image synthesis from terrain elevation data. For the same reasons, the slant plane representation allows fairly straightforward modifications of existing shape from shading techniques to be applied to SAR imagery. Efficient surface reconstruction and image synthesis using the slant plane representation requires a method for resampling a DTM from ground coordinates to slant plane coordinates and vice versa. Such a method was developed which is a variation of the summed area table method reported in the computer graphics literature [68] and is amenable to high speed implementation.

2.2 SAR Reflectance Maps

This section provides a brief review of relevant prior research in radiometry and develops a modeling approach for SAR reflectance maps. Cosgriff, et al. [32] enumerated the following five parameters affecting the characteristics of terrain backscatter for radar:

- Incidence angle
- Surface roughness
- Frequency
- Complex dielectric constant
- Polarization.

We are mainly interested in understanding the effects of incidence angle on backscattered power. That functional relationship depends on the other four parameters. Surface roughness relative to wavelength partly determines the shape of the reflectance map and, hence, the reflectance map is frequency dependent. The complex dielectric constant of the surface materials partially determines albedo but will not be considered explicitly. Polarization effects are also significant. While the interplay between all of these factors is still not completely understood, the utilization of frequency diversity and polarization diversity appear to provide extra constraints needed for pinning down variations in reflectance characteristics in radar imagery.

Radiometry is the microwave counterpart to photometry. In the 1700's Johann Heinrich Lambert developed principles of photometry based on geometrical optics which are still used today. In particular, the concept of the "perfectly diffusing" (or Lambertian) surface and the term albedo were introduced by Lambert [165]. At radar wavelengths, most surfaces are not perfectly diffusing so that the reflectance properties are more highly directional than those often assumed in optics. Even at visible wavelengths, reflection characteristics are very often non-Lambertian [115,158], so that generalizations of early optical reflectivity models have been developed for partially rough surfaces [7,158].

The reflectance map $\mathcal{R}(\cdot)$ is the function which relates local surface orientation to image intensity. In both radiometry and photometry $\mathcal{R}(\cdot)$ is a combination of three effects;

1. The electromagnetic scattering properties of the surface.
2. The geometric properties of the illumination.
3. The geometric properties of image formation.

The reflectance model should be parameterized to retain the distinction between factors unique to SAR image formation and factors due to electromagnetic scattering properties of the surface while explicitly including the dependence on the surface slopes in both the range and azimuth directions. For SAR, the reflectance map $\mathcal{R}(\cdot)$ can be expressed as the product of an area factor A multiplied by the radar cross section (RCS) σ_0 per unit area. Both A and σ_0 are functions of imaging geometry and surface orientation, i.e.

$$\mathcal{R}(u_r, u_y, \beta) = A(u_r, u_y, \beta)\sigma_0(u_r, u_y, \beta) \ . \tag{2.9}$$

This resulting reflectance map is normally much more directional than if the same surface was viewed by a passive imaging system at visible wavelengths. This is for two reasons. First, the area factor is different for SAR imagery than for conventional photographs in that the SAR pixel is formed in slant-range/cross-range coordinates instead of azimuth/elevation. By itself, A is more directional than the reflectance map for conventional photographs of a Lambertian surface used in much of computer vision research. Second, any given surface is much smoother relative to radar wavelengths than for visible wavelengths. Hence, σ_0, by itself, also tends to be more directional than the reflectance map for a visible image of the same surface. Thus, the SAR reflectance map is the product of two factors, both of which can be more directional than typical optical reflectance maps.

A typical SAR reflectance map is plotted in Figure 2.3 along with a reflectance map for a visual image of an ideal Lambertian surface illuminated by a point source. This helps to illustrate the drastic difference between the characteristics of radar imagery and aerial photographs.

A visual comparison of SAR imagery with aerial photographs is shown in Figure 2.4. Figure 2.4 (a) shows a simulated aerial photograph with illumination from the left 60 degrees above the horizon. The Lambertian reflectance model from Figure 2.3 was used with the camera at Nadir. Figure 2.4 (b) shows a simulated SAR image of the same area with the same illumination geometry, i.e. a 60 degree grazing angle, and the SAR reflectance model from Figure 2.3. The effect of the directional reflectance is clearly visible in the higher contrast of the SAR image. A closer examination also reveals the foreshortening of mountain slopes facing the radar.

2.2.1 Area Factor

Two different parameterizations for area/RCS have been used in the radar literature. One parameterization uses the area of a pixel projected onto the surface times the normalized radar cross section. This approach was introduced by Goldstein in 1946 [57]. A second parameterization by Cosgriff, et al. [32] uses the area projected into a plane normal to the illumination. We will refer to the former area factor as "surface area" and the latter as "illumination area". The illumination area parameterization is more compatible with the definition of RCS used by electromagnetic scattering theorists [13].

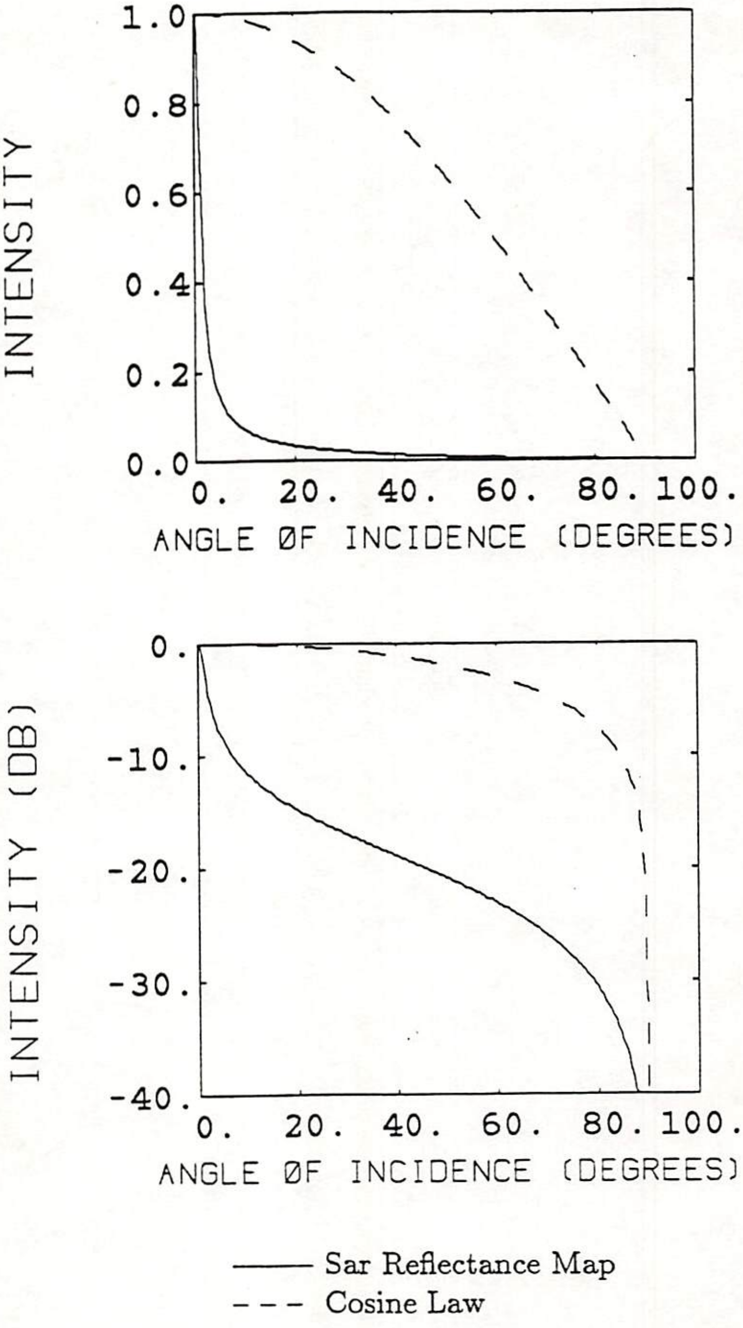


Figure 2.3: A typical SAR image reflectance map and Lambertian reflectance map plotted on the same scale. The top plot uses a linear scale and the bottom plot uses a logarithmic scale

Surface Area

Suppose for now that the pixels can be described by an ideal rectangular point spread function (PSF). The projection in the range direction of such a pixel onto a locally planar surface is given by

$$\Delta L_r = \Delta r \sqrt{1 + u_r^2} \, . \tag{2.10}$$

The projection in azimuth of a pixel onto the surface is

$$\Delta L_y = \Delta y \sqrt{1 + u_y^2} \, . \tag{2.11}$$

The resulting elemental area of a pixel projected onto the surface is

$$A_{\mathrm{S}}(u_r, u_y) = \Delta L_r \Delta L_y = \Delta r \Delta y \sqrt{(1 + u_r^2)(1 + u_y^2)} \, . \tag{2.12}$$

To simplify this expression, note that for small to moderate slopes

$$(1 + u_r^2)(1 + u_y^2) \approx 1 + u_r^2 + u_y^2 \, . \tag{2.13}$$

Substituting (2.13) into (2.12) and normalizing to $\Delta r \Delta y = 1$, we can make the following simplifying approximation

$$A_{\mathrm{S}} \approx \sqrt{1 + u_r^2 + u_y^2} = \frac{1}{\cos \alpha_i} \, . \tag{2.14}$$

Illumination Area

Illumination area A_{I} denotes the area of the incident plane wave that impinges on the projected pixel. For the rectangular PSF this gives

$$A_{\mathrm{I}} = \Delta r \Delta y |u_r| \tag{2.15}$$

providing a very simple expression in terms of the differential geometry of the surface. Equation (2.15) could also be expressed as the tangent of the angle between the range vector and the terrain surface. Figure 2.5 depicts the surface and illumination area in slant plane coordinates for a locally planar surface.

Both forms of area factor have their advantages and disadvantages. The approximation from (2.14) gives surface area as a function of α_i only. If RCS

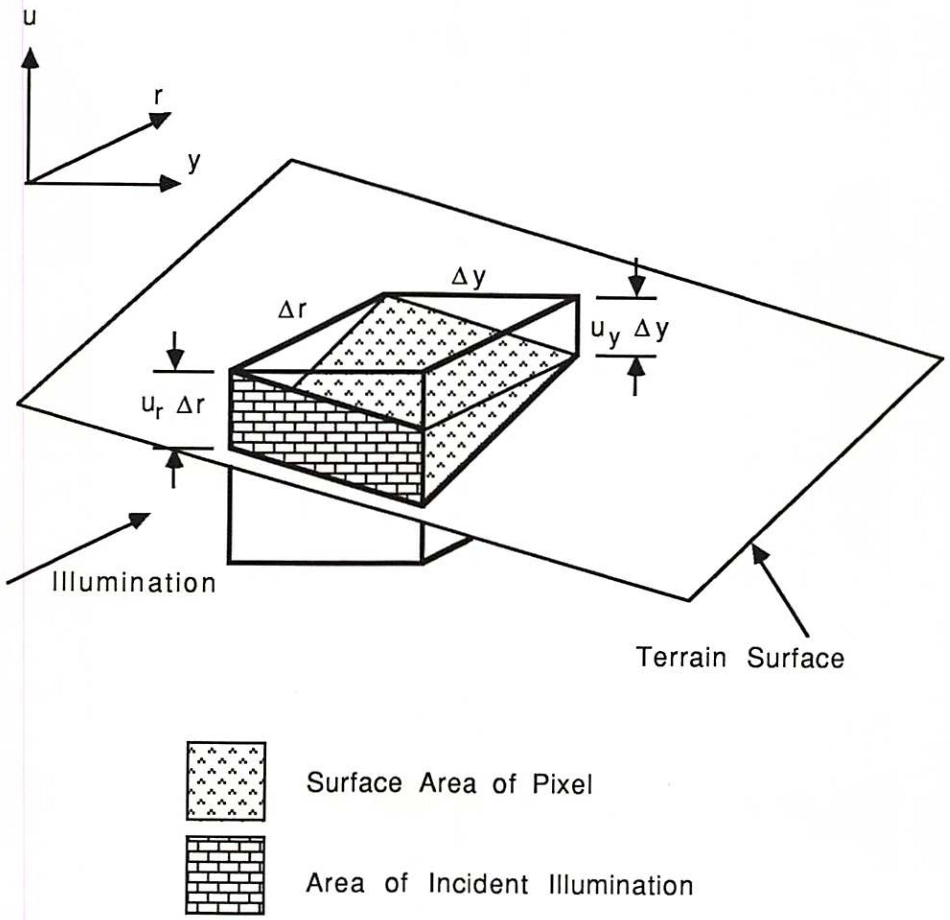


Figure 2.5: Surface and illumination area of a SAR image pixel represented in slant plane coordinates.

is expressed as a function of α_i and implemented in a look-up table then the area factor can be absorbed into the same table, saving computation. While illumination area from (2.15) is not parameterized by α_i, it is also simple and is more physically meaningful than surface area. Experimental results comparing simulated images with real SAR images described in Chapter 4 show the illumination area parameterization to provide a more accurate model.

2.2.2 RCS Models

Many models for the variation of σ_0 as a function of incidence angle are presented in the literature [9,12,106] based on both empirical studies and by modeling the various possible scattering mechanisms that may occur. Interestingly, some of the models developed by the radar community appear to parallel the models developed for quasi-specular reflection in optics [115] and computer graphics. The functional form of σ_0 will depend on surface microstructure relative to the wavelength, and can vary locally in ways that are difficult, if not impossible, to detect from image data alone.

The perfectly diffusing reflectance models were rejected early in the development of radar technology except for some special cases such as fields of grain [31,32]. Clapp considered Lambert's work in photometry as a starting point in developing RCS models. Clapp's first model [106] is given by $\sigma_0 = \cos\alpha_i$ when used with A_I. The relation $\sigma_0 =$ constant, Clapp's third model, results in a simple reflectance map $\mathcal{R}(\alpha_i) = \cot(\alpha_i)$ when combined with the area factor for a flat surface. This was derived by modeling the surface as being composed of many layers of spheres, each absorbing some fraction of the incident power and reradiating the remainder isotropically. Clapp's third model has been found to be in close agreement with measured RCS of some vegetation covered terrain, such as fields of grain.

For better or worse, radar clutter is often more directional than indicated by the above models. The generalized Lambert model, $\sigma_0(\alpha_i) = \cos^k(\alpha_i)$, developed as a remedy for this and has also been considered for computer animation of visual scenes [129]. A further generalization, based on empirical data, is given by [91]

$$\sigma_0(\alpha_i) = \frac{\cos^k \alpha_i}{\sin^l \alpha_i} . \tag{2.16}$$

It has also been suggested that (2.16) with k in the range of 3 to 4 may provide a reasonable model for relatively smooth surfaces such as asphalt [106], and that $l = 0$, with $k \in [1, 2]$ is appropriate for very rough surfaces. In our experiments from Chapter 4, $l = 0$ was appropriate but much higher values were needed for k.

RCS models for partially rough surfaces derived using stochastic models of surface microstructure are discussed below.

Specular Point Models

One characteristic of radar clutter is the relatively frequent occurrence of specularities from smooth surfaces, such as water, lava flows, roads, and other cultural features. An ideal specular surface is one for which all energy is reflected at an angle equal to the angle of incidence, sometimes referred to as Fresnel reflection [147]. Most radar clutter, however, is neither perfectly specular or perfectly diffuse but somewhere in between. Hence, early clutter modeling research attempted to combine the physics of specular reflection with that of reflection by rough surfaces. One example is the specular point theory based on modeling the surface as a collection of many smooth facets, each with well-defined reflectivity properties but oriented at random. Since each facet is a specular reflector the expression for σ_0 is related to the average number of facets oriented towards the radar. This approach was investigated by Katzin [88], and Spetner and Katz [152]. This is an example of using a spatial point process representation for radar returns to derive a practical model. The resulting RCS models have been found to have a close fit to empirical data in some cases. Specular point theory is based on geometrical optics which does not lend itself well to predicting the effects of varying wavelength, but does have the advantage of conceptual simplicity.

Physical Optics Approaches

As of 1963, according to Beckmann and Spizzichino [13], the study of electromagnetic radiation reflected from planar surfaces was well-understood but the more general case of rough surfaces remained an unsolved problem. Although practical advances have been made, electromagnetic scattering from rough surfaces is still an area of active research as evidenced by more recent publications [120,121]. Approaches based on wave theory (of which geometrical optics is the special case with zero wavelength) have been useful when the surface height is modeled as a continuous random field. One elegant result relates physical optics results in the zero wavelength limit to the joint probability density function for the surface slopes [10]. This observation is consistent with earlier work using facet models [152], and physical optics models with finite wavelength [13].

We have considered a class of models derived by a physical optics approach but evaluated in the zero wavelength limit. The surface is modeled as a random process with specular reflection occurring at each point on the surface. The radar cross section depends on the average number of points oriented towards the radar and the average Gaussian curvature at those points, which in turn depends on the root-mean-square (RMS) surface roughness and correlation length. We have concentrated on the model reported by Barrick [9]. For Gaussian distributed surface height the resulting RCS is

$$\sigma_0(\alpha_i) = \frac{K}{s^2 \cos^4 \alpha_i} \exp\{-\frac{\tan^2 \alpha_i}{s^2}\} , \tag{2.17}$$

where s is the RMS surface slope. A different expression is obtained for the case of exponentially distributed surface slopes [9]. It was found in Chapter 4 that, for the available imagery, Barrick's model with Gaussian height provided a better fit than Barrick's exponential slope model or Keydel's empirical model.

Other Factors Influencing RCS

Many real-world surfaces defy description by either geometric or stochastic approaches. For example, ground cover from vegetation is extremely complex. Even if adequate descriptions of their shape could be obtained, the role of depth of penetration and internal reflection of the electromagnetic waves is uncertain

[147], and an item of current research. Further, in shape from shading and even simulation applications we may not know a priori the parameters for well-defined effects such as simple surface roughness.

Past experimental results also indicate a need for better models for $\alpha_i \longrightarrow \pi/2$ (near grazing). This is consistent with our experience in SAR image simulation where the RCS models based on simple models of surface microstructure did not adequately predict image intensity at low grazing angles. One way to look for better RCS models is by using better surface models. Modeling rough surfaces is complicated by the fact that "roughness" of real surfaces depends on the length of the yardstick used. For example surface undulations from multiple physical sources can take on a wide range of scales. This effect has been investigated with both physical and geometric optics approaches by modeling surface height as the sum of independent random fields each with different slope statistics [13,176]. Two-scale models have been developed which seem to better explain the characteristics of radar backscatter near grazing in some cases [176].

Another phenomenon is that the scattering mechanism used for intermediate angles of incidence may not be valid at the extremes near grazing and near vertical incidence. As $\alpha_i \rightarrow \pi/2$ specular reflection takes over since the Rayleigh criterion for a rough surface is [13]

$$s < \text{constant} \times \frac{\text{wavelength}}{\cos \alpha_i}.$$

This shows that for finite surface roughness s that the surface effectively becomes smooth near grazing. This effect is partly responsible for road glare near the horizon observed by travelers and has been observed in radar clutter [106].

Some basic RCS models are summarized in Table 2.1.

2.2.3 Albedo of Radar Returns

Cosgriff, et al. [32] applied the term albedo to radiometry by analogy with photometry to mean that portion of the incident energy which is reradiated by the surface. For convenience, we lump together all multiplicative constants that appear in front of the reflectance map and refer to that product as albedo. Albedo depends on the electrical properties of the surface materials and any

RCS MODEL	DESCRIPTION
$\cos(\alpha_i)$	Clapp's 1st Model (Lambertian)
Constant	Clapp's 3rd Model (Perfectly Diffuse)
$\cos^k(\alpha_i)$	Generalized Lambertian Model
$\cos^k(\alpha_i)/\sin^l(\alpha_i)$	Keydel's Model [91]
$[K\sec^4(\alpha_i)/s^2]\exp\{\tan^2(\alpha_i)/s^2\}$ where s = RMS surface slope	Barrick's Model [9], Based on specular-point model for reflection from rough surface with Gaussian slopes

Table 2.1: Some models for normalized radar cross-section versus angle of incidence. Illumination area parameterization is used.

ground cover. Variations in the median of measured radar albedo for typical clutter spans a wide dynamic range on the order of 60 dB where the extreme lows are for water and the extreme highs are for cities [106].

Albedo can vary significantly in ways that are difficult to predict with changes in moisture. For example, water on asphalt changes both the smoothness and the dielectric constant, lowering the RCS curve by as much as 10 dB [32]. In the same investigation, water on vegetation was observed to have the opposite effect, that of raising RCS on the order of 3 dB [32]. Water *in* vegetation (say due to increasing moisture content from winter to spring) has caused an even larger increase in RCS of 10 dB [32]. The effect of snow depends heavily on snow depth since the electromagnetic waves can penetrate into the snow.

The issue of albedo variation is so complex that we do not attempt to develop a predictive model for it. It does appear that the random field models discussed in Chapter 6 are, in some cases, appropriate models for albedo. Further, frequency diversity [15,41] and polarization diversity provide some information on variations in the electrical properties of the surface being imaged. That, in turn, gives an indication of albedo variations.

2.3 Observation Noise

The most significant source of noise for SAR imagery is speckle, which arises from interference between the returns from multiple scatterers within a resolution element. For complicated surfaces and resolution much larger than a wavelength the complex amplitude of the speckle can be modeled as complex Gaussian noise multiplied by the signal [60]. The correlation length of the speckle is determined by the width of the PSF of the imaging system. Thermal noise from the radar electronics introduces a second noise source with complex amplitude modeled as white complex Gaussian noise added to the image.

A third source of noise arises from the voltage accumulated in the pixel sidelobes. Since SAR image amplitude is formed essentially by calculating a weighted Fourier transform of a finite aperture, sidelobes in the PSF are unavoidable, just as sidelobes are an inherent part of the frequency response of a finite impulse response filter. This voltage accumulated in the sidelobes of any given pixel is the coherent summation of many different weighted voltages, each dependent on the radar backscatter in some area. For natural terrain images, this voltage is approximately independent of the signal in any particular pixel, approximately white and Gaussian with variance dependent on the level of the overall backscatter in a large neighborhood.

It is apparent that speckle, thermal noise, and sidelobe noise have very similar statistical properties, except that speckle is a multiplicative noise source. The combined effect is that the noisy image intensity can be expressed as a noise-free image, combined with a bias term (dependent on thermal and sidelobe noise levels) multiplied by a noise term

$$I(r,y) = \left[G_s \eta \mathcal{R}(u_r, u_y, \beta) + \sigma^2_{\text{bias}}\right] n(r,y) \tag{2.18}$$

where σ^2_{bias} is the combined power of the thermal and sidelobe noise, $n(r,y)$ is unit power random field distributed as chi-squared with two degrees of freedom, and G_s is radar the system gain. It is well-known that a dual additive–multiplicative noise source is equivalent to a single multiplicative noise source with additive bias (as in equation (2.18)). A proof of this assertion may be found in [177].

G_s includes all of the terms of the radar equation [147] not appearing in $\eta\mathcal{R}$, such as wavelength, range attenuation, transmitter power, duty factor, etc., but also includes compensation for range dependence, antenna beam-shape, and calibration factors. One goal in the design of high quality SAR systems is for G_s to be a known constant throughout the image. We will assume that

$$G_s = 1\ . \tag{2.19}$$

For uncalibrated radars system gain is assumed to be an unknown constant.

The noise variance can be reduced at a cost of spatial resolution by averaging neighboring pixels. However, if the thermal noise and sidelobe noise levels are significant, then the bias term will remain and must be accounted for regardless of the amount of spatial averaging. Given enough noncoherent integration one could model image intensity as

$$I(r,y) \approx \eta \mathcal{R}(u_r, u_y, \beta) + \sigma^2_{\text{bias}} \tag{2.20}$$

For example, SEASAT SAR imagery, among the highest quality publicly available at this time, has an integrated sidelobe ratio (ISLR) of only 6 to 20 dB, depending on the processing method used [130]. Since ISLR measures the ratio of mainlobe energy to total sidelobe energy, this means that σ^2_{bias} can be as high as 6 dB below the average value of $\eta\mathcal{R}$ in a region—surely a significant bias level. For the SAR imagery used in Chapter 4, σ^2_{bias} was actually larger than the average value of $\eta\mathcal{R}$. It is, therefore, necessary to incorporate this bias into the reflectance model.

2.4 Stochastic Models for Image Intensity

We now introduce the stochastic modeling approach used for SAR image intensity and examined in more depth in Chapter 6. Let $\sigma^2_{\text{bias}} = 0$ for now and model the observed square-law detected image I as

$$I(r,y) = v(r,y)n(r,y) \ , \tag{2.21}$$

where $v(r,y)$ is the noncoherent image proportional to the ensemble average radar backscatter for the object at the physical location corresponding to (r,y).

Various models have been considered for I and v in the radar literature [12,56,81,147]. The main shortcoming of those models is that they have not devoted attention to spatial correlation in the data. Frost et al. [52] modeled v as a Gaussian random field with causal spatial interaction to account for spatial correlation. We examine transformed Gaussian random fields with possibly noncausal spatial interaction. That is, a transformed version of v, given by

$$w(r,y) = g(v(r,y)) \ , \tag{2.22}$$

is modeled as a Gaussian random field. The function g is an invertible point-nonlinearity. The spatial interaction in the underlying Gaussian random field w is then represented by a noise-driven linear system

$$w(r,y) = \mu * h\ (r,y) \tag{2.23}$$

where μ is the Gaussian driving process and h is a possibly noncausal 2-D PSF. Some specific examples are considered in more detail in Chapter 6.

If $g(\cdot) = \ln(\cdot)$ then v is referred to as a lognormal random field. We concentrate on the lognormal case for SAR images although other transformations are possible. Presence of thermal noise and sidelobe noise requires a three-parameter lognormal distribution model for the first order density of v. The extra parameter is simply the additive bias σ^2_{bias}. More general transformations can be useful, for example, one which is linear at low intensities and logarithmic for high and intermediate intensities, or a piecewise linear transformation, or a cube root. Hence, the general form for g will be retained as much as possible.

Stochastic models are also useful for inferring surface properties. Cox and Munk [33,34] used texture in photographs of sunlight reflected from the ocean surface to estimate surface roughness. More recently, Pentland [125] has attempted to relate surface statistics to image statistics for visible wavelength images of fractal surfaces. It is also possible to develop relations between stochastic models of terrain elevation and stochastic models for the noncoherent component of SAR image intensity. We do not address that problem but proceed directly to reconstruction of deterministic surface topography from image intensity.

Chapter 3

Shape from Shading

3.1 Background

An image is a projection of a three-dimensional (3-D) world into a 2-D signal and vision is, in some sense, the inverse of that projection. Hence, vision is a highly underconstrained problem requiring additional knowledge or simplifying assumptions. In a more restricted domain, shape from shading refers to the problem of inferring the shape of a surface given a single image of that surface. The basic principle is that surface shape information can be obtained by modeling the observed image intensity in terms of the surface orientation and then using that relationship to solve for the surface slopes.

Let $I(x_2, y_2)$ be the observed image intensity and $z(x, y)$ be the unknown surface height above the (x, y) plane. Note the distinction between image coordinates (x_2, y_2) and surface coordinates (x, y, z). The relationship between image intensity and surface slopes can be expressed in the following form [70,134,164],

$$I(x_2, y_2) = \mathcal{R}(z_x, z_y, \beta, \ell, \eta) \tag{3.1}$$

where $z_x = \partial z/\partial x$ and $z_y = \partial z/\partial y$ are the surface slopes, β is the illumination direction vector, ℓ is the vector from the surface to the camera, and η is the albedo or intrinsic reflectivity of the materials composing the surface.

The problem of reconstructing z given I has been expressed as that of solving a nonlinear partial differential equation in x and y [70]. Considerable progress has been made towards solving this in practice assuming that (i) at any point (x, y) the reflectance map $\mathcal{R}$ is a function of the surface slopes and albedo only at (x, y), i.e. multiple reflections are ignored, (ii) the albedo η and the vectors β and ℓ are known over the entire image, (iii) $\mathcal{R}$ is spatially invariant, and (iv) the transformation from surface coordinates to image coordinates is an orthographic projection, specifically

$$(x_2(x, y, z), y_2(x, y, z)) = (x, y) \; . \tag{3.2}$$

Unfortunately, an exact solution to the imaging equation (3.1) does not always exist, or there may be an infinite number of solutions. In practice, modeling errors such as reflectance map mismatch, imperfect knowledge of the light source, spatial and radiometric quantization error, observation noise, and albedo variations are inevitable. Further, boundary conditions are generally not completely known and sometimes may not be available at all. These factors all influence the existence and uniqueness of a solution to (3.1) and the estimation of a good solution in the case that a unique one does not exist. For these reasons shape from shading is a very difficult problem in practice.

It is clear that the shape from shading problem, in its full generality, is an underdetermined problem. Consequently, its solution requires additional constraints. Possible constraints are smoothness and integrability of (z_x, z_y), shadowing cues, auxiliary data from multiple images, altimetry, higher level scene analysis, and simplifying assumptions as discussed above.

One early application of shape from shading techniques was for estimating lunar topography from photographs [134,164] in the 1950's and 1960's. In that work a fundamental difficulty was noted: Even with a known reflectance map and albedo, each image intensity sample is a function of the two surface gradient components, z_x and z_y, at the corresponding location. One can only specify a locus of gradient values capable of producing a particular intensity value so that local measurements, by themselves, do not suffice. A simplification was possible for the surface of the moon due to the particular reflectance model for the lunar surface being constant along straight lines in the gradient space [134], i.e. image intensity was only a function of the gradient in one particular direction. This allowed surface reconstruction along straight lines in the image by 1-D integration, leaving the problem of estimating the relative elevations of each of the integration paths. This was done by assuming the average surface height along each line of the image to be equal. A more accurate reconstruction could have been provided by enforcing integrability of the surface slope estimates [51].

Horn addressed the more general case where reflectivity is not a straight line in gradient space [70]. Direct inversion of the resulting differential equation was accomplished using the characteristic strip expansion, where the 2-D nonlinear differential equation is converted into five ordinary differential equations (one for each of x, y, z_x, z_y, and z). The solution is obtained by integrating along characteristic curves in the image—contours that proceed parallel to the gradient of the reflectance map in image coordinates and parallel to the gradient of image intensity in gradient space. The characteristic curves are data dependent and the solution is sensitive to observation noise and error accumulation. Accordingly, this approach served to demonstrate the concept of shape from shading but does not appear to be appropriate for application to noisy data such as SAR imagery.

Haralick et al [65,66] estimated surface height from LANDSAT imagery by integrating intensity along ridges and valleys. Since the gradient at a ridge or a valley is zero in some direction, it is possible to estimate the gradient orthogonal to that direction unambiguously using image intensity (assuming known albedo). This approach is restrictive in that it only provides reasonable estimates along ridges and valleys: What if there are none? Shape from shading is more useful if it can be applied in areas absent of such salient features so that it may complement other techniques, e.g. stereopsis, that rely more heavily on the presence of significant feature.

Pentland [124] developed an approach which examines local estimates of the Laplacian of image intensity to estimate surface slopes under the assumption that the surface is locally spherical at each point. Lee and Rosenfeld later improved this technique so that it only requires first derivatives [100], reducing its sensitivity to noise. Yet the dependence on accurate estimates of intensity derivatives (even first derivatives) makes it extremely unstable in the presence of noise, such as speckle. Further, the local sphericity assumption is only an approximation. A third difficulty is that no provision is made for enforcing global consistency (e.g. integrability) of the local slopes so that it is difficult to construct a surface from the slope estimates. Hence the local approach is probably not appropriate for SAR imagery.

Iterative search approaches have been developed which cast the problem in a cost minimization framework rather than in a pure inversion framework. Such iterative search techniques are attractive in that they allow for the inevitable uncertainties due to modeling errors and noise. For example, Brooks and Horn [22] proposed the approach of selecting the surface slope estimates, $(\hat{z}_x, \hat{z}_y)$, which minimize the following cost function

$$\epsilon = \int\int (I - \mathcal{R}(\hat{z}_x, \hat{z}_y))^2 + \lambda \cdot (\hat{z}_{xx}^2 + 2\hat{z}_{xy}^2 + \hat{z}_{yy}^2) dx dy \ , \qquad (3.3)$$

subject to known boundary conditions. The first term in the integrand is the squared error between the observed image intensity and the image intensity predicted by substituting the estimates $(\hat{z}_x, \hat{z}_y)$ into (3.1). This mean squared error term allows for modeling errors and noise. The second term in the integrand is a measure of quadratic variation in the surface slopes. This is a smoothness criterion which, under some conditions, assures a unique smooth solution to (3.3) even when (3.1) does not possess a unique solution. It is also interesting that minimization of a quadratic variation of the surface slopes is roughly equivalent to minimization of the potential energy of a thin elastic plate [62]. The constant λ establishes a tradeoff between smoothness of the solution $(\hat{z}_x, \hat{z}_y)$ and the mean-square value of the residuals $I - \mathcal{R}$.

Brooks and Horn [22] developed an algorithm to minimize ϵ in (3.3) subject to the constraint that $(\hat{z}_x, \hat{z}_y)$ satisfy known boundary conditions. Although the Brooks and Horn algorithm may converge to a unique solution of (3.3), that solution is generally not integrable. That is, given the final estimates of the partial derivatives many possible surfaces could be constructed, each dependent upon the path of integration.

Earlier, Ikeuchi and Horn [79] developed a similar algorithm that, instead of using gradient space (z_x, z_y) was parameterized in stereographic coordinates (f, g) where

$$f = \frac{2z_x}{1 + \sqrt{1 + z_x^2 + z_y^2}} \quad \text{and} \quad g = \frac{2z_y}{1 + \sqrt{1 + z_x^2 + z_y^2}} \ . \qquad (3.4)$$

This parameterization eliminates the difficulty in handling occluding boundaries where z_x, z_y, or both are unbounded. To understand the (f, g) coordinates geometrically, note that the locus of all conceivable unit normal vectors defines a unit sphere, called the Gaussian sphere. The (f, g) coordinate system above can be obtained through a central projection from the south pole of the entire Gaussian sphere into a plane tangent at its north pole. The locus of all occluding boundaries correspond to the equator of the Gaussian sphere and a circle of radius 2 in (f, g) space. The algorithm of Ikeuchi and Horn allows f and g to vary independently in the iterative solution and (z_x, z_y) is obtained by a point transformation of (f, g). The result is that Ikeuchi and Horn's algorithm also suffers from the drawback of not producing an integrable solution.

Lee [101] presented a shape from shading algorithm, inspired by that of Ikeuchi and Horn, for which he proves the existence of a unique solution. Further, Lee proves that his algorithm converges to the unique solution if the surface slopes are known on the square image boundary. Unfortunately, that solution is not integrable.

3.2 Enforcing Integrability in Shape from Shading Algorithms

In deriving iterative solutions to (3.1) by the calculus of variations, it appears to be much more straigthforward to solve for surface orientation than to solve directly for z [74]. The question of consistency between z_x and z_y then arises. A reasonable consistency constraint to place on the surface slopes is that they are integrable, where integrability is defined by

$$z_{xy}(x, y) = z_{yx}(x, y) \tag{3.5}$$

for all (x, y) on the support of I. That is they correspond to a surface with second partial derivatives independent of the order of differentiation. This guarantees that the surface height at any particular point is independent of the path of integration so that the integral equation

$$\oint_{\mathcal{C}} (z_x dx + z_y dy) = 0 \tag{3.6}$$

holds for all closed curves $\mathcal{C}$ in the region of interest. Integrability is related to smoothness because a surface with discontinuities violates (3.5). Horn and Brooks assert [74], as do we, that past shape from shading algorithms can be improved considerably by incorporating an integrability constraint.

Both the local shape from shading algorithm and earlier iterative algorithms reported in the literature [22,79,100,101,124,126] do not take into account the interdependence of the two components of surface orientation but allow them to vary independently. If we do allow (z_x, z_y) to vary independently then (3.1) may have an infinite number of solutions even when boundary conditions are completely known and there are no modeling errors or observation noise. Just to illustrate this point, neglect the issue of boundary conditions and consider shape from shading as a discrete problem. Suppose we observe image intensity on a $N \times N$ grid. We have N^2 observations and hope to determine $2N^2$ samples of the partial derivatives of $z(x, y)$. Clearly, there is an infinite number of solutions to (3.1) if the partial derivatives of z are not interdependent. Suppose instead that we simply require that these partial derivatives are integrable in the sense of (3.5). Then the partial derivatives must correspond to only N^2 unknown samples of z so that we are dealing with approximately N^2 equations in N^2 unknowns. Hence, an integrability constraint cuts the number of unknowns by a factor of two, neglecting boundary conditions.

Integrability constraints have been used before in iterative shape from shading algorithms. Smith [149] derived a set of equations relating derivatives of image intensity to surface orientation independent of albedo and illumination angle, and without assuming a particular local characteristic for the surface (such as Pentland's local sphericity assumption). It is interesting to note that integrability was assumed in the formulation and imposed in deriving algorithms for solving the equations. Smith considered the use of a penalty function to encourage integrability during iteration, and a local spatial approach for constructing a surface from nonintegrable slope estimates, although no convergent algorithm was reported.

Strat developed an iterative shape from shading algorithm that minimizes the mean squared image intensity prediction error plus a penalty term that measures the departure from integrability [154]. This penalty is the sum of squares of discrete approximations to the loop integral (3.6) evaluated at each pixel in the image. The contour of integration is a 2×2 pixel block containing the pixel being evaluated.

Recently, Horn and Brooks [74] have developed an approach for deriving iterative shape from shading algorithms using the variational calculus. It was shown that several previously developed algorithms are special cases of the variational calculus approach. A new iterative algorithm which attempts to enforce integrability through the use of a penalty function, $\int\int(\partial\hat{z}_x/\partial y - \partial\hat{z}_y/\partial x)^2 dxdy$, was developed and its discrete implementation was shown to be very similar to that of Strat's algorithm. Problems at occluding boundaries were overcome by employing a unit surface normal parameterization for surface orientation instead of the simple partial derivative form. Horn and Brooks encountered some difficulties in developing a convergent iterative algorithm that strictly enforces integrability. Further, algorithms that were presented do not appear to allow much flexibility for incorporating additional constraints.

We have developed a method for strictly enforcing integrability and for enforcing constraints on the surfaces spectrum. The method for enforcing integrability is to project the possibly nonintegrable surface slope estimates onto the nearest integrable surface slopes [48]. In Section 3.3, this orthogonal projection is solved for the case that the surface slopes are represented by finite sets of orthogonal, integrable basis functions. It is shown that the projection maps closed convex subsets of the space of nonintegrable surface slopes onto closed convex subsets of the space of integrable surface slopes provided that certain conditions are met. Ikeuchi and Horn have previously derived an iterative algorithm, described in [74], that constructs depth information from nonintegrable surface orientation data by iteratively solving the same projection. That approach has the advantage of not requiring a finite basis function representation. Our method has the advantage of being noniterative along with all of the advantages and disadvantages of the basis function representation. In Section 3.4

the projection is presented for the special case of Fourier basis functions. In Section 3.5 the discrete periodic Fourier basis formulation of the projection is applied to extending a shape from shading algorithm of Brooks and Horn [22] so that it converges faster and with much less error than the original version. This extended algorithm includes the integrability constraint, and both the regularization and intensity prediction error penalties from equation (3.3). Simulation results are presented for images of complicated surfaces representative of natural terrain. The technique is appropriate for integrating noisy slope estimates and could, perhaps, be used in improving the results of several other shape from shading algorithms [100,101,124,126,149]. It is this extended algorithm which we have adapted to SAR imagery. Finally, in Section 3.7, possible applications of the integrability projection to other vision problems are considered.

3.3 Enforcing Integrability by Orthogonal Projections

There are many conceivable ways of enforcing (3.5). The key is to develop a computationally efficient method suitable for use with iterative algorithms. We have developed an approach based on projecting the possibly nonintegrable estimated surface slopes $(\hat{z}_x, \hat{z}_y)$ onto a set of integrable surface slopes $(\tilde{z}_x, \tilde{z}_y)$, where

$$\frac{\partial}{\partial y}\tilde{z}_x = \frac{\partial}{\partial x}\tilde{z}_y \tag{3.7}$$

while simultaneously minimizing the following distance measure

$$d\{(\hat{z}_x, \hat{z}_y), (\tilde{z}_x, \tilde{z}_y)\} = \int\!\!\int |\tilde{z}_x - \hat{z}_x|^2 + |\tilde{z}_y - \hat{z}_y|^2 dx dy \ . \tag{3.8}$$

Note that the minimum distance property makes this an orthogonal projection.

The task of finding such a projection is simplified if we can represent the surface slopes by a finite set of integrable basis functions, each satisfying (3.7). Suppose that we represent the surface $z(x, y)$ by the functions $\phi(x, y, \underline{\omega})$ so that

$$z(x, y) = \sum_{\underline{\omega}\in\Omega} C(\underline{\omega})\phi(x, y, \underline{\omega}) \ , \tag{3.9}$$

where $\underline{\omega} = (\omega_x, \omega_y)$ is a two dimensional index, Ω is a finite set of indices, and the members of $\{\phi(x, y, \underline{\omega})\}$ are not necessarily mutually orthogonal. If each $\phi(\underline{\omega})$ satisfies (3.7) then it follows that z does also. Note that the partial derivatives of z can also be expressed in terms of this expansion, giving

$$z_x(x,y) = \sum_{\underline{\omega}\in\Omega} C(\underline{\omega})\phi_x(x,y,\underline{\omega}) \tag{3.10}$$

and

$$z_y(x,y) = \sum_{\underline{\omega}\in\Omega} C(\underline{\omega})\phi_y(x,y,\underline{\omega})$$

where $\phi_x = \frac{\partial\phi}{\partial y}$ and $\phi_y = \frac{\partial\phi}{\partial y}$.

Now a method is presented for finding the expansion coefficients $C(\underline{\omega})$ which minimize (3.8). Suppose that the members of $\{\phi_x(x, y, \underline{\omega})\}$ are mutually orthogonal as well as the members of $\{\phi_y(x, y, \underline{\omega})\}$. Then we can compute the coefficients for the expansions of $\hat{z}_x$ and $\hat{z}_y$ such that

$$\hat{z}_x(x,y) = \sum_{\underline{\omega}\in\Omega} \hat{C}_1(\underline{\omega})\phi_x(x,y,\underline{\omega}) \tag{3.11}$$

and

$$\hat{z}_y(x,y) = \sum_{\underline{\omega}\in\Omega} \hat{C}_2(\underline{\omega})\phi_y(x,y,\underline{\omega}) \; .$$

The orthogonality of the basis functions in (3.11) greatly simplifies the minimization of (3.8).

Proposition 1: The expansion coefficients $\tilde{C}(\underline{\omega})$ in equation (3.9) that minimize (3.8) given a nonintegrable estimate of surface slopes, $\hat{z}_x(x, y)$, $\hat{z}_y(x, y)$, are given by

$$\tilde{C}(\underline{\omega}) = \frac{P_x(\underline{\omega})\hat{C}_1(\underline{\omega}) + P_y(\underline{\omega})\hat{C}_2(\underline{\omega})}{P_x(\underline{\omega}) + P_y(\underline{\omega})} \tag{3.12}$$

for each $\underline{\omega} \in \Omega$, where

$$P_x(\underline{\omega}) = \int\!\!\int |\phi_x(x,y,\underline{\omega})|^2 dxdy$$

and

$$P_y(\underline{\omega}) = \int\!\!\int |\phi_y(x,y,\underline{\omega})|^2 dxdy \; .$$

The integrated surface $\tilde{z}(x, y)$ and integrable surface slopes, $\tilde{z}_x(x, y)$, $\tilde{z}_y(x, y)$ are then obtained by substituting $\tilde{C}(\underline{\omega})$ into the expansions (3.9) and (3.10).

Proof: By substituting the expansion for $\hat{z}_x$, $\hat{z}_y$, $\tilde{z}_x$, and $\tilde{z}_y$ into (3.8) the distance measure becomes

$$d\{(\hat{z}_x,\hat{z}_y),(\tilde{z}_x,\tilde{z}_y)\} = \tag{3.13}$$

$$\int\int \left|\sum_{\underline{\omega}\in\Omega}\tilde{C}\phi_x - \sum_{\underline{\omega}\in\Omega}\hat{C}_1\phi_x\right|^2 + \left|\sum_{\underline{\omega}\in\Omega}\tilde{C}\phi_y - \sum_{\underline{\omega}\in\Omega}\hat{C}_2\phi_y\right|^2 dxdy \ ,$$

which reduces to

$$d\{(\hat{z}_x,\hat{z}_y),(\tilde{z}_x,\tilde{z}_y)\} =$$

$$\int\int \sum_{\underline{\omega}\in\Omega}|\tilde{C}-\hat{C}_1|^2|\phi_x|^2 + \sum_{\underline{\omega}\in\Omega}|\tilde{C}-\hat{C}_2|^2|\phi_y|^2 \, dxdy$$

due to the orthogonality of $\{\phi_x(\underline{\omega}),\ \underline{\omega}\in\Omega\}$ and of $\{\phi_y(\underline{\omega}),\ \underline{\omega}\in\Omega\}$. By interchanging the order of summation and integration the distance in (3.8) reduces to

$$d\{(\hat{z}_x,\hat{z}_y),(\tilde{z}_x,\tilde{z}_y)\} = \tag{3.14}$$

$$\sum_{\underline{\omega}\in\Omega}|\tilde{C}(\underline{\omega})-\hat{C}_1(\underline{\omega})|^2P_x(\underline{\omega}) + |\tilde{C}(\underline{\omega})-\hat{C}_2(\underline{\omega})|^2P_y(\underline{\omega}) \ .$$

Equation (3.14) can be minimized by minimizing each term of the summation individually. By differentiating (3.14) with respect to the real and imaginary parts of $\tilde{C}(\underline{\omega})$ for each $\underline{\omega}$, setting the result to zero, and solving for $\tilde{C}$, (3.12) results.

For clarity in the subsequent proofs, express (3.10) and (3.11) in terms of vectors. First, define the vectors

$$\begin{aligned}
\Phi_x(x,y) &= \mathrm{col}[\phi_x(x,y,\underline{\omega}),\underline{\omega}\in\Omega] \\
\Phi_y(x,y) &= \mathrm{col}[\phi_y(x,y,\underline{\omega}),\underline{\omega}\in\Omega] \\
\underline{0} &= [0,\cdots,0] \\
\mathbf{C} &= \mathrm{col}[\tilde{C}(\underline{\omega}),\underline{\omega}\in\Omega] \\
\mathbf{C_1} &= \mathrm{col}[\hat{C}_1(\underline{\omega}),\underline{\omega}\in\Omega] \\
\mathbf{C_2} &= \mathrm{col}[\hat{C}_2(\underline{\omega}),\underline{\omega}\in\Omega] \ .
\end{aligned}$$

Equation (3.11) can now be written as

$$\begin{bmatrix}\hat{z}_x \\ \hat{z}_y\end{bmatrix} = \begin{bmatrix}\Phi_x^T & \underline{0} \\ \underline{0} & \Phi_y^T\end{bmatrix}\begin{bmatrix}\mathbf{C_1} \\ \mathbf{C_2}\end{bmatrix} , \tag{3.15}$$

with the set of possibly nonintegrable surface slopes expressed as

$$\mathcal{S}_0 = \text{span}\begin{bmatrix} \Phi_x^T & \underline{0} \\ \underline{0} & \Phi_y^T \end{bmatrix} . \tag{3.16}$$

Similarly, (3.10) can be rewritten as

$$\begin{bmatrix} \tilde{z}_x \\ \tilde{z}_y \end{bmatrix} = \begin{bmatrix} \Phi_x^T \\ \Phi_y^T \end{bmatrix} \mathbf{C} , \tag{3.17}$$

with the set of integrable surface slopes given by

$$\mathcal{S} = \text{span}\begin{bmatrix} \Phi_x^T \\ \Phi_y^T \end{bmatrix} \tag{3.18}$$

which is a vector subspace of $\mathcal{S}_0$.

We will denote the mapping of $\mathcal{S}_0$ onto $\mathcal{S}$ using (3.12) as $\mathcal{P}$. This can be written in vector form as

$$\mathbf{C} = [\,\mathbf{P} \quad \mathbf{I} - \mathbf{P}\,] \begin{bmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \end{bmatrix} \tag{3.19}$$

where $\mathbf{P}$ is a diagonal matrix with diagonal elements $P_x/(P_x + P_y)$ and $\mathbf{I}$ is the identity matrix.

Proposition 2: The projection (3.12) maps convex subsets of $\{\hat{z}_x, \hat{z}_y\}$ onto convex subsets of $\{\tilde{z}_x, \tilde{z}_y\}$.

Proof: Convexity follows directly from the observation that (3.19) is a linear mapping. Suppose that $\mathcal{E}$ is a convex subset of $\mathcal{S}_0$. Then $\mathcal{P}(\mathcal{E})$ is a subset of $\mathcal{S}$ and we must now prove that $\mathcal{P}(\mathcal{E})$ is convex. The definition of convexity is that, for all $\mathbf{U}, \mathbf{V} \in \mathcal{E}$ and for all $0 < \mu < 1$ we have $\mu\mathbf{U} + (1-\mu)\mathbf{V} \in \mathcal{E}$. Then $\mathcal{P}(\mu\mathbf{U} + (1-\mu)\mathbf{V}) \in \mathcal{P}(\mathcal{E})$, since $\mathcal{E}$ is convex. $\mathcal{P}$ is linear so that

$$\mathcal{P}(\mu\mathbf{U} + (1-\mu)\mathbf{V}) = \mu\mathcal{P}(\mathbf{U}) + (1-\mu)\mathcal{P}(\mathbf{V}) . \tag{3.20}$$

This proves that the right hand side of (3.20) is an element of $\mathcal{S}$, thus completing the proof.

Proposition 3: The projection (3.12) maps compact subsets of the functions $\{\hat{z}_x, \hat{z}_y\}$ onto compact subsets of $\{\tilde{z}_x, \tilde{z}_y\}$. Since we are working in metric spaces, this is equivalent to saying that closed and bounded sets are mapped onto closed and bounded sets [139].

Proof: From Theorem 4.15 of [139] we know that Proposition 3 holds if $\mathcal{P}$ is a continuous mapping. Since all of the elements of the transformation in (3.19) are finite, it follows that $\mathcal{P}$ is continuous, completing the proof.

It was shown that the integrability constraint represents an orthogonal projection mapping closed convex subsets of its input space onto closed convex subsets of its output space. This is a highly desirable property if the projection is to be applied as a constraint in iterative algorithms [159,178]. The shape from shading problem, in its entirety, does not appear to lend itself well to solution by the method of projection onto convex sets (POCS) since the nonlinearity of the reflectance map makes it difficult to realize an orthogonal projection. However, it is possible to introduce additional constraints, e.g. integrability, into an iterative shape from shading algorithm using POCS as an intermediate step. An example of this is presented in Section 3.5. Note that the methods for enforcing boundary conditions in previous shape from shading algorithms [22,74,79,101] could be expressed as projections onto closed convex sets. This follows from the fact that the operation of setting a function equal to prescribed values over some region of its support (e.g. strictly enforcing boundary conditions) can be expressed as an orthogonal projection mapping closed convex sets onto closed convex sets [178].

This approach for enforcing integrability has a second distinct advantage in that the surface can be reconstructed in one pass using all of the information available in $(\hat{z}_x, \hat{z}_y)$ simply by performing the summation in (3.9) for the values of $\tilde{C}(\underline{\omega})$ computed in (3.12). Thus it may be useful as a more efficient integrator for nonrecursive local shape from shading approaches.

3.4 Integration by Fourier Expansion

The usefulness of the expansion presented in Section 3.3 depends on the specific choice of basis functions. The Fourier basis functions, $\phi(x, y, \underline{\omega}) = \exp(j\omega_x x +$

$j\omega_y y)$, are particularly convenient computationally, and they also form a complete orthornormal basis for z.

Using this basis the surface is represented by

$$\tilde{z}(x,y) = \sum_{\underline{\omega}\in\Omega} \tilde{C}(\underline{\omega}) \exp\{j\underline{\omega}\cdot(x,y)\} \tag{3.21}$$

where $\{\tilde{C}\}$ are the coefficients of the Fourier series expansion of $\tilde{z}$. For images of size N by N, a reasonable choice for Ω would be $(2\pi n, 2\pi m)$, with $n, m \in \{0, 1, \cdots, N-1\}$. Now let $\hat{C}_x$, $\hat{C}_y$, $\tilde{C}_x$, $\tilde{C}_y$ be the Fourier coefficients for $\hat{z}_x$, $\hat{z}_y$, $\tilde{z}_x$, $\tilde{z}_y$ respectively.

The derivatives of the Fourier basis functions possess the following useful properties

$$\phi_x = j\omega_x\phi \quad \text{and} \quad \phi_y = j\omega_y\phi\ , \tag{3.22}$$

$P_x \propto \omega_x^2$, $P_y \propto \omega_y^2$, $\hat{C}_1(\underline{\omega}) = \hat{C}_x(\underline{\omega})/j\omega_x$, and $\hat{C}_2(\underline{\omega}) = \hat{C}_y(\underline{\omega})/j\omega_y$. Then using (3.12) and (3.22) it is straightforward to show that (3.8) is minimized by taking

$$\tilde{C}(\underline{\omega}) = \frac{-j\omega_x\hat{C}_x(\underline{\omega}) - j\omega_y\hat{C}_y(\underline{\omega})}{\omega_x^2 + \omega_y^2} \tag{3.23}$$

with the Fourier coefficients of the constrained surface slopes given by [48]

$$\tilde{C}_x(\underline{\omega}) = j\omega_x\tilde{C}(\underline{\omega}) \text{ and } \tilde{C}_y(\underline{\omega}) = j\omega_y\tilde{C}(\underline{\omega})\ .$$

The expressions above are valid except at the point $\underline{\omega} = (0,0)$ which simply means that we cannot recover the average value of z without some additional information.

Note that

$$\omega_y \to 0 \implies \tilde{C}(\underline{\omega}) \to \hat{C}_x(\underline{\omega})/j\omega_x$$

and

$$\omega_x \to 0 \implies \tilde{C}(\underline{\omega}) \to \hat{C}_y(\underline{\omega})/j\omega_y\ .$$

In general, the Fourier coefficients of the raw slope estimates are combined in proportion to their relative frequencies. For example, if we divide the weight for $\hat{C}_x(\underline{\omega})$ in (3.12) by that of $\hat{C}_y(\underline{\omega})$ we get ω_x/ω_y. The effect is to reduce "random walk" errors that result from amplification of low frequency noise during integration.

This frequency domain interpretation helps in understanding an inherent difficulty in the shape from shading problem. The observed image intensity is a function of the derivatives of the surface. Equation (3.22) shows that the low frequency information for the surface is lost in the image formation process. Hence, the reconstructed surface inevitably suffers from low frequency distortion, the severity depending on observation noise characteristics. The frequency domain formulation of the integrability constraint is appropriate here since the low resolution data could be introduced as a projection constraint in an iterative shape from shading algorithm. This is accomplished by substituting the Fourier coefficients of the low resolution surface in place of the lowest frequency shape from shading results. Experimental results in Figure 3.6 of Section 3.6 demonstrate the improvement that results from using four known low frequency terms in compensating for completely unknown boundary conditions.

It is conceivable that low resolution information could be drawn from other sources in real-world applications of shape from shading techniques. For example, in some remote sensing applications low resolution surface height information is available from DTMs [72,172]. A particular example where low resolution constraints on shape from shading is available and useful is the Magellan project's Venus Radar Mapper which will collect SAR imagery of the surface of Venus along with much lower resolution radar altimetry data [36]. With shape from shading techniques it may be possible to improve upon the resolution of the radar altimetry data by combining it with the higher resolution SAR imagery as discussed in Chapter 4.

3.4.1 Discrete Periodic Formulation

For computer implementation some form of discretization is necessary. Assume, for computational simplicity, that the surface slopes are circularly periodic and evaluate the Fourier coefficients $\hat{C}_x$ and $\hat{C}_y$ using the discrete Fourier transform (DFT). With this discrete periodic formulation (3.8) is minimized by

$$\tilde{C}(\underline{\omega}) = \frac{a_x^*(\omega_x)\hat{C}_x(\underline{\omega}) + a_y^*(\omega_y)\hat{C}_y(\underline{\omega})}{|a_x|^2 + |a_y|^2} \tag{3.24}$$

with

$$\tilde{C}_x(\underline{\omega}) = a_x(\omega_x)\tilde{C}(\underline{\omega}) \quad \text{and} \quad \tilde{C}_y(\underline{\omega}) = a_y(\omega_y)\tilde{C}(\underline{\omega}) \ ,$$

where a_x and a_y are the Fourier coefficients of the appropriate discrete differentiation operator in x and y. Suppose we approximate the derivatives by finite central differences, e.g.

$$z_y(l,m) \approx \frac{1}{2}[z(l+1,m) - z(l-1,m)] \tag{3.25}$$

and similarly for z_x. For the central difference operator above we get $a_y(\omega_y) = \frac{1}{2}\exp\{j\omega_y\} - \frac{1}{2}\exp\{-j\omega_y\} = j\sin(\omega_y)$, and similarly $a_x(\omega_x) = \frac{1}{2}\exp\{j\omega_x\} - \frac{1}{2}\exp\{-j\omega_x\} = j\sin(\omega_x)$.

The estimated surface is constructed by performing the inverse DFT of $\tilde{C}(\underline{\omega})$ from (3.24), so that the integrability projection doubles as an integrator. In all cases the DFT is evaluated using a fast Fourier transform (FFT) algorithm.

Figure 3.1 shows the improvement of the above fast least-squares integration method over simple spatial integration for reconstructing a surface from noisy slope estimates. The top plot, Figure 3.1a, shows the noise-free surface. Forward difference approximations to the surface slopes (z_x, z_y) were calculated and then each sample corrupted by additive white Gaussian noise. The noise standard deviation was 0.25 for each slope component, and the maximum value for the surface slopes was about 0.8. Figure 3.1b shows the reconstruction obtained by simple spatial integration along a single path. Error propagation is apparent from the ridge lines that appear on the reconstructed surface. Figure 3.1c shows the surface reconstruction obtained using the fast least-squares integration described above. Error propagation is greatly reduced and a much better reconstruction results.

The intuitive frequency domain interpretation of the integrability projection is easily extended to the discrete case using the discrete periodic formulation. Alternatives to the discrete periodic formulation are possible, but are not considered here. The Fourier expansion could be formulated on a finite lattice instead of a periodic lattice. The mathematics are somewhat more complicated but FFT calculations are still possible [44] and more careful attention could then be paid to boundary conditions. Other basis functions, such as Chebyshev polynomials could also be useful for finite lattice realizations.

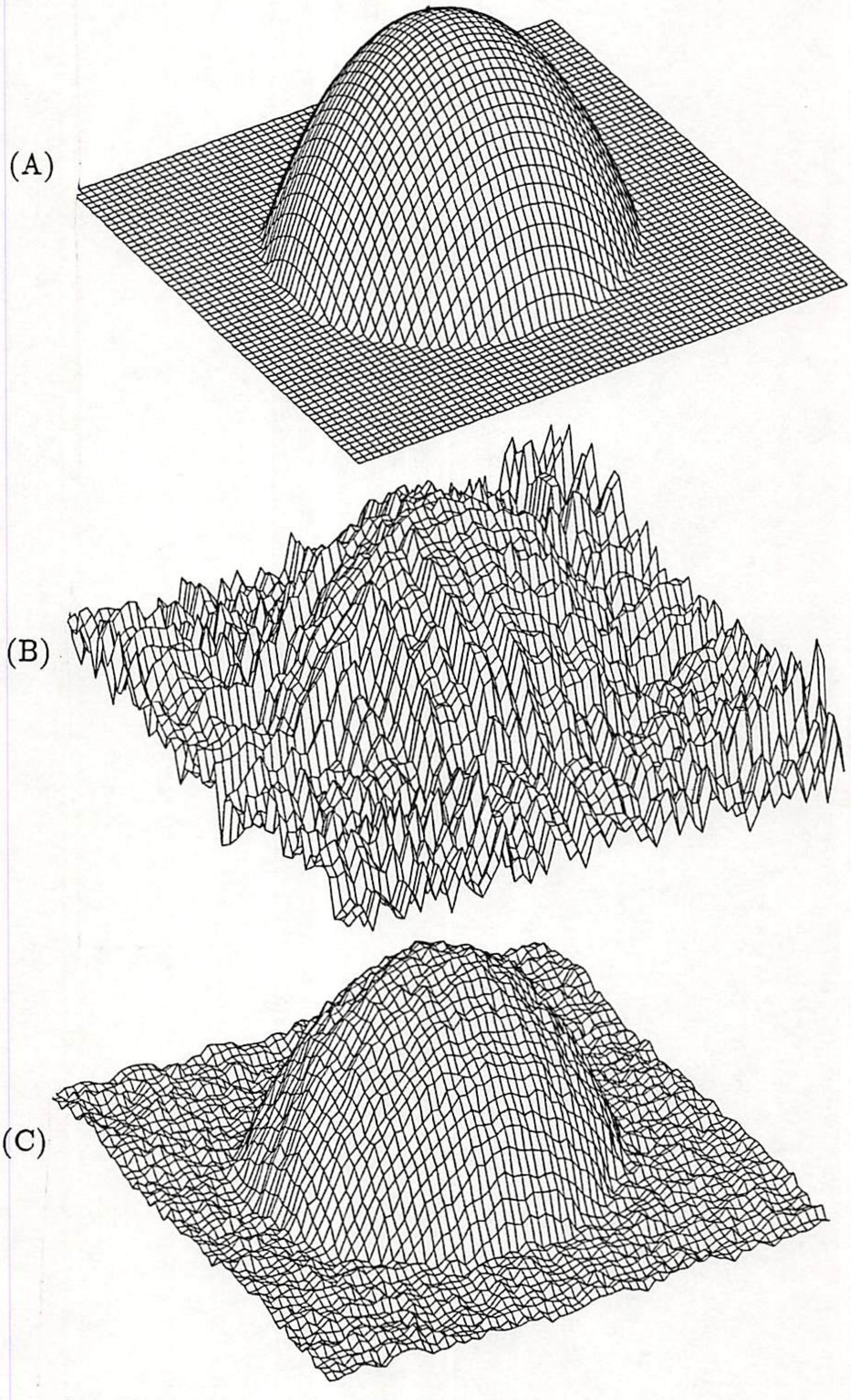


Figure 3.1: (A) Noise-free surface. (B) Reconstruction from noisy slopes obtained by simple spatial integration along a single contour. (C) Reconstruction using least-squares integration.

3.5 An Improved Shape from Shading Algorithm

We have utilized the projection from Section 3.4 to obtain a simple extension of an algorithm presented by Brooks and Horn [22]. The result is an algorithm which minimizes (3.3) while satisfying the discrete form of (3.7) and (3.8) at each iteration. The experimental results presented in Section 3.6 show a marked improvement in performance due to the integrability constraint.

Using the finite difference approximations, the values of $\hat{z}_x(x, y)$ and $\hat{z}_y(x, y)$ which minimize (3.3) are found iteratively by the following recursion for each point (x, y),

$$\begin{bmatrix} \hat{z}_x \\ \hat{z}_y \end{bmatrix}_{k+1} = \begin{bmatrix} \hat{\hat{z}}_x \\ \hat{\hat{z}}_y \end{bmatrix}_k + \lambda_1 (I - \mathcal{R}) \begin{bmatrix} \mathcal{R}_x \\ \mathcal{R}_y \end{bmatrix} \tag{3.26}$$

at the $(k+1)$th iteration, where

$$\mathcal{R}_x = \frac{\partial \mathcal{R}}{\partial z_x} \quad \text{and} \quad \mathcal{R}_y = \frac{\partial \mathcal{R}}{\partial z_y} .$$

$\mathcal{R}$ and its partials are evaluated at $[\hat{\hat{z}}_x, \hat{\hat{z}}_y]_k$, λ_1 is a constant inversely proportional to λ in (3.3), $[\hat{\hat{z}}_x, \hat{\hat{z}}_y]_k$ is a smoothed version of $[\tilde{z}_x, \tilde{z}_y]_k$, and $[\tilde{z}_x, \tilde{z}_y]_k$ is obtained by substituting the raw estimates $[\hat{z}_x, \hat{z}_y]_k$ into the integrability projection (3.24).

The smoothing applied during each iteration is given by [22]

$$\begin{aligned} &\hat{\hat{z}}_y(l, m) = \\ &\tfrac{1}{5}[\tilde{z}_y(l, m+1) + \tilde{z}_y(l, m-1) + \tilde{z}_y(l+1, m) + \tilde{z}_y(l-1, m)] \\ &+\tfrac{1}{20}[\tilde{z}_y(l-1, m-1) + \tilde{z}_y(l-1, m+1) \\ &+\tilde{z}_y(l+1, m+1) + \tilde{z}_y(l+1, m-1)] \end{aligned} \tag{3.27}$$

and similarly for $\hat{\hat{z}}_x$. Note that this is just a discrete approximation to the Laplacian with the center pixel left out. This follows from a manipulation of a discrete form of the Euler equations for (3.3) as discussed in [22,74]. The resulting expression relating λ_1 to λ from (3.3) with this particular version of the Laplacian is [22] $\lambda_1 = \frac{3}{10\lambda}$. The rationale for evaluating $\mathcal{R}$ with the smoothed slope estimates is that it stabilizes (3.26) [79].

The iterative algorithm can be summarized as follows [48];

(i) Smooth the previous slope estimates using (3.27),

(ii) Generate a new set of raw slope estimates using (3.26),

(iii) Project the raw slope estimates onto the nearest integrable solution using (3.24).

The process is repeated until the cost function either stops decreasing or becomes sufficiently small. The surface height is obtained by simply performing the inverse DFT of $\tilde{C}(\underline{\omega})$ after the final iteration.

Parameterization in gradient space (z_x, z_y) rather than stereographic coordinates (f, g) is sometimes criticized because the surface gradient is not well-defined at occluding boundaries where z_x, z_y, or both are infinite. However, for discrete data, z_x and z_y are bounded for all practical purposes. This observation leads to a potential method for utilizing occluding boundaries in the above shape from shading algorithm. Intuitively, it seems that most of the information provided by the occluding boundary comes from knowledge of the location of its contour in the (x, y) plane and the orientation z_x/z_y along its contour. The occluding boundary can be detected by examining zero crossings of the Laplacian of image intensity [62]. The orientation of the occluding boundary z_x/z_y can be found from an analysis of this occluding contour. Then, perhaps, the orientation can be enforced as a constraint in the shape from shading solution while allowing the magnitude $\sqrt{z_x^2 + z_y^2}$ to be controlled by the regularization penalty in (3.3). Alternately, it is possible to apply the integrability constraint in gradient space and the surface slope update in (f, g) coordinates.

3.6 Experimental Results

The above algorithm was tested on synthetic imagery with and without known boundary conditions and was also tested on real imagery. In each case discussed below the images are of size 64 × 64 pixels. Lambertian surfaces illuminated by point sources were assumed in each case, giving the following the reflectance map

$$\mathcal{R}(z_x, z_y, \beta, \ell, \eta) = \eta \frac{\beta_x z_x + \beta_y z_y + \beta_z}{\sqrt{\beta_x^2 + \beta_y^2 + \beta_z^2}\sqrt{1 + z_x^2 + z_y^2}} . \tag{3.28}$$

However, the implemented algorithm used numerical derivatives of the reflectance map so that any reasonably well-behaved known reflectance map could have been applied.

3.6.1 Comparison of Constrained Versus Unconstrained Algorithm

First, a series of experiments was performed to compare the performance of the shape from shading algorithm in Section 3.5 with and without the integrability constraint. An image was simulated for a surface consisting of a partial sphere protruding above a plane.

In the first experiment, the surface slopes were provided where the sphere intersects the plane. No noise was added to the observed intensity, the only source of errors being intensity quantization and finite precision arithmetic. Under these conditions, the algorithm of Brooks and Horn converges to a reasonably accurate solution for the surface slopes. This case was used to get an indication of the impact of enforcing the integrability projection under very benign conditions. Table 3.1 shows the mean and standard deviation of the surface orientation error for both the constrained and unconstrained algorithm after 100 iterations. Note that the application of the integrability constraint reduces the error standard deviation from 2.36 degrees to 0.53 degrees, a factor of 4.5. A second observation is that a minimum value of $\lambda = 7000$ was required for stabilizing the unconstrained algorithm, while the constrained algorithm was stable for a wider range, down to $\lambda = 5000$. Interestingly, the very significant change in λ (from 7000 to 5000) resulted in an insignificant change in the error standard deviation (from 0.53 to 0.56), perhaps indicating that most of the smoothing is coming from the integrability constraint. This appears to corroborate the finding of Horn and Brooks [74] where the regularization penalty from (3.3) was dropped and an integrability penalty was used instead, improving upon the results of Ikeuchi and Horn [79].

The rate of convergence was also compared for the partial sphere case. For the constrained algorithm the orientation error standard deviation was down to 2.7 degrees after only 5 iterations. The unconstrained algorithm took about 90 iterations to reach this same level of convergence. After 8 iterations the constrained algorithm was down to 1.25 degrees error standard deviation, while the unconstrained algorithm took nearly 1000 iterations to reach the same level. This demonstrates that enforcing integrability by the method developed in this chapter can greatly speed up convergence.

Similar experiments were repeated with relaxed boundary conditions, i.e. boundary conditions provided around the square border of the image. The integrability constraint also reduced the error standard deviation by a factor of about 4.5 in this case. The numerical comparison is shown in Table 3.2.

When the unconstrained algorithm was applied with relaxed boundary conditions the resulting solution was nearly locally cylindrical (i.e. $\hat{z}_y \approx 0$) everywhere. This is clearly nonintegrable. The integrability projection was applied to the above erroneous slope estimates after 100 iterations of the unconstrained algorithm. Here the error standard deviation is much higher than for constrained iteration but significantly lower than for unconstrained iteration, as indicated in Table 3.2. This demonstrates the usefulness of the integrability projection as a method for constructing a surface from nonintegrable slope estimates.

Figure 3.2 (A) shows the true partial sphere surface, (B) an image generated from that surface, (C) the surface estimated using the constrained algorithm given both the image intensity and knowledge of the surface slopes around the border of the image (corresponding to Table 3.2), and (D) the surface estimated given only the simulated image without the knowledge of boundary conditions. For the known boundary condition case the reconstructed surface is almost identical to the true surface. For the case of unknown boundary conditions, use of the regularization penalty term in (3.3) combined with the integrability constraint was sufficient to provide convergence to a reasonable solution in all test cases.

The positive results above can be attributed mostly to the integrability constraint, although other issues deserve some attention. Illumination from the

		Orientation Error	(Degrees)
Algorithm	λ	mean	st. dev.
Unconstrained	7000	4.07	2.36
Constrained	7000	0.64	0.56
Constrained	5000	0.61	0.53

Table 3.1: Comparison of surface orientation error after 100 iterations for the unconstrained algorithm versus the constrained algorithm. The boundary slopes are provided at the intersection of the sphere and the plane in these examples.

		Orientation Error	(Degrees)
Algorithm	λ	mean	st. dev.
Unconstrained	7000	6.49	10.33
Unconstrained (until last iteration)	7000	8.45	6.66
Constrained	5000	1.89	2.45

Table 3.2: Comparison of surface orientation error after 100 iterations for the unconstrained algorithm versus the constrained algorithm. In these examples, boundary conditions are supplied around the square border of the image.

side (instead of directly overhead) has a positive impact on the results. Had the sphere been illuminated from directly overhead, one would not be able to distinguish a concave-downward sphere from a concave-upward sphere using shading alone. Some other information would be needed, such as boundary slopes at the joint between the sphere and the plane or low resolution surface information. This did not appear to be a very serious issue in our experiments since a slight deviation of the illumination source from vertical would allow the algorithm to resolve this type of ambiguity. No special measures were taken to avoid wraparound due to the discrete periodic formulation of the integrability projection: $N \times N$ FFTs were used for $N \times N$ images. Similar methods have been used before in solving similar elliptic problems [44]. Periodicity appears to help a little bit for surfaces without discontinuities at the wraparound points, such as in Figure 3.2, but it has a slight negative impact for more complicated surfaces.

3.6.2 Shape from Shading for Complicated Surfaces

A second set of experiments demonstrates the performance of the constrained shape from shading algorithm using images of more complicated surfaces. In Figure 3.3, a picture of the surface of the moon is shown along with a surface estimated based on a guessed light source direction and an assumed reflectance map. Boundary conditions are neither known nor guessed in advance and a reasonable surface estimate is obtained. Although the lunar surface may be non-Lambertian [134], the surface estimate shown in Figure 3.3 was obtained by assuming Lambertian reflectivity. For this moon image, surface estimates consistent with our visual intuition were obtained using a variety of assumed reflectance maps and for a broad range of plausible light source directions.

Images were simulated using a digital terrain model (DTM) of the Laguna Hills in California. Figure 3.4 shows three dimensional (3-D) plots of the true DTM and the results of the constrained shape from shading algorithm after 50 iterations. Comparison of the 3-D plot in Figure 3.4A with those of Figure 3.4B and 3.4C show that the general shape of the surface has been reconstructed both with and without known boundary conditions. The 1-D slices compared in

Figure 3.5 better illustrate how closely the reconstructed surface (dotted line) fits the true surface (solid line) in the known boundary conditions case.

The dashed line in Figure 3.6 shows a 1-D slice of the surface reconstruction obtained with completely unknown boundary conditions. Although the overall shape of the reconstruction is very similar to that of the true surface (solid line) there is some apparent low frequency distortion. This distortion is caused by the loss of low frequency information in the process of image formation, combined with the tendency of the regularization penalty and periodic boundary conditions to flatten the surface. Qualitatively similar low frequency errors appear in the solution for $z(x, y)$ when small amounts of observation noise, or other sources of error are present. This low frequency distortion is an inherent part of the shape from shading problem since we are observing image intensity, which is a function of the partial derivatives of $z(x, y)$. The reconstruction of $z(x, y)$ involves integrating those partial derivatives, amplifying the low frequency portion of the noise spectrum as discussed in Section 3.4. Thus, it is important to be able include low frequency constraints from other sources.

3.6.3 Incorporating Low Resolution Information

The dotted line in Figure 3.6 shows the improvement that results from including a very small amount of low resolution data in the shape from shading solution. This demonstrates that if low resolution information is indeed available in some form other than shading then errors resulting from missing boundary conditions and observation noise can be greatly reduced. The application of the integrability constraint lends itself well to including low resolution information as a constraint in the iterative shape from shading algorithm of Section 3.5. This was accomplished by taking the Fourier transform of the Laguna DTM in Figure 3.4, and saving only the four lowest frequency terms (not including the DC term). The four lowest frequency terms of the true DTM were then substituted for the corresponding part of the Fourier transform of the estimated surface height at each iteration. This did not add significantly to the computational load because the Fourier transform of estimated surface height is obtained as a byproduct of the integrability constraint as can be seen from (3.23). Enforcing

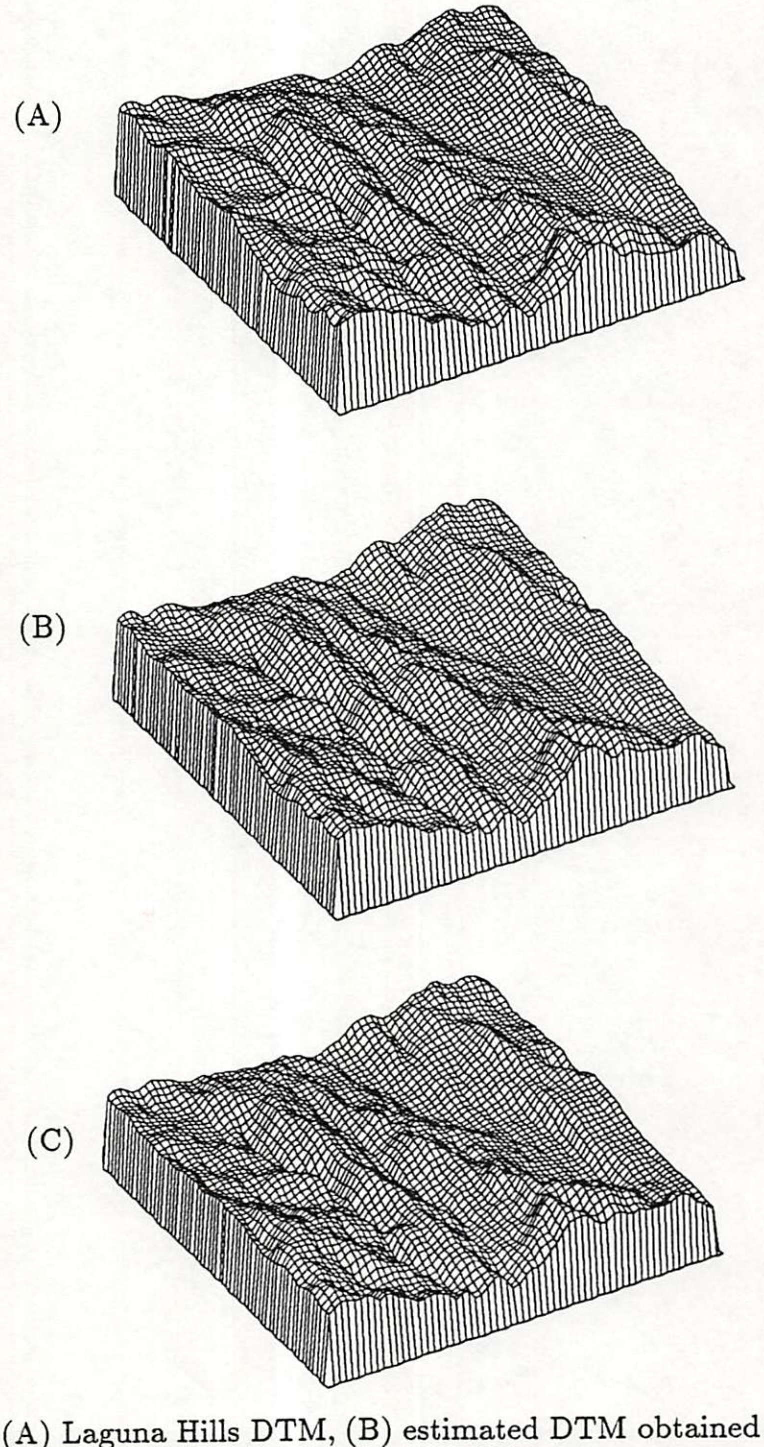


Figure 3.4: (A) Laguna Hills DTM, (B) estimated DTM obtained from a simulated image of that surface, and boundary slopes, (C) estimated surface obtained with unknown boundary slopes.

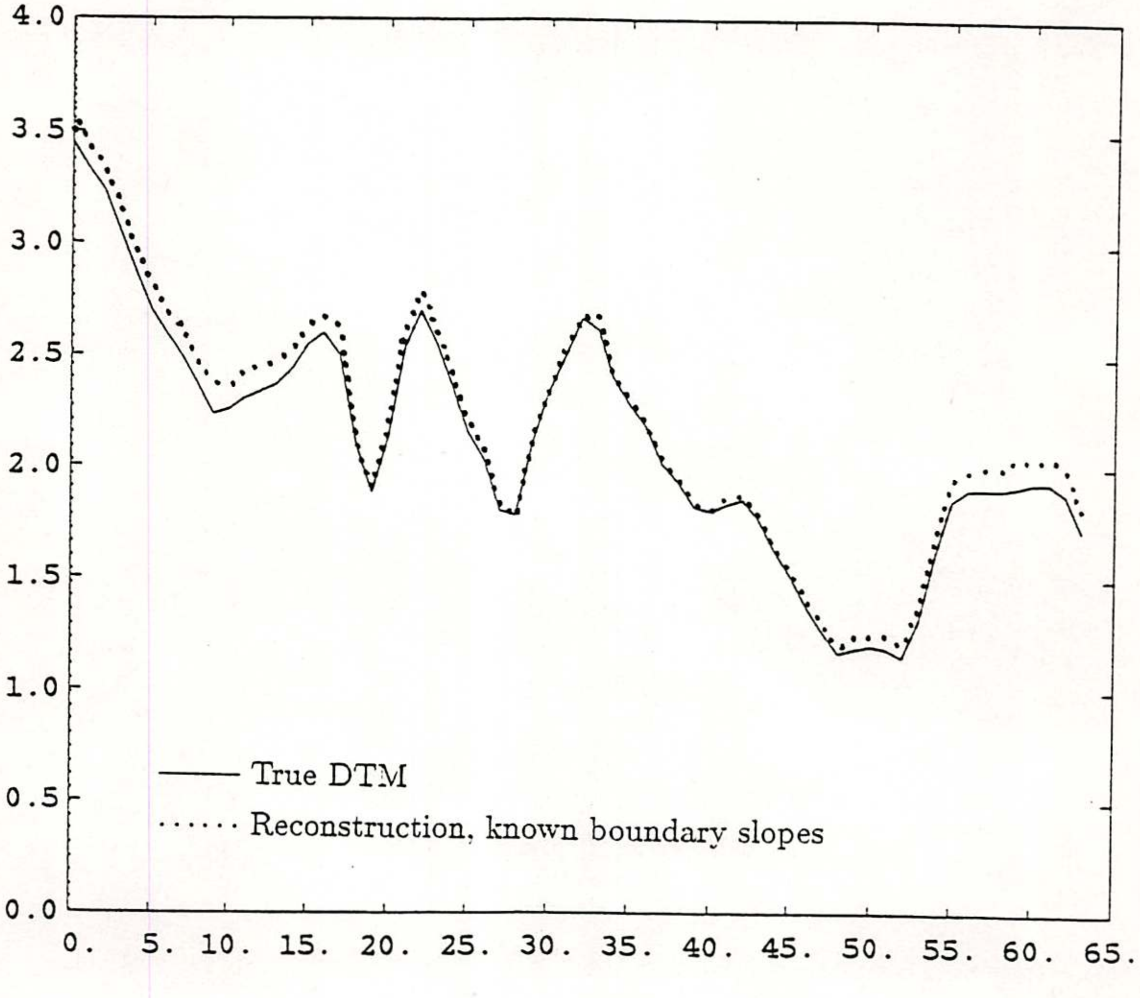


Figure 3.5: One dimensional slice through the DTMs of Figures 3.4A and 3.4B. The horizontal axis corresponds to the 32nd line in the images of Figure 3.8. That is, it forms a line parallel to the x-axis running from left to right through the center of the images. The units for the vertical axis are the same as the image sample spacing in (x, y).

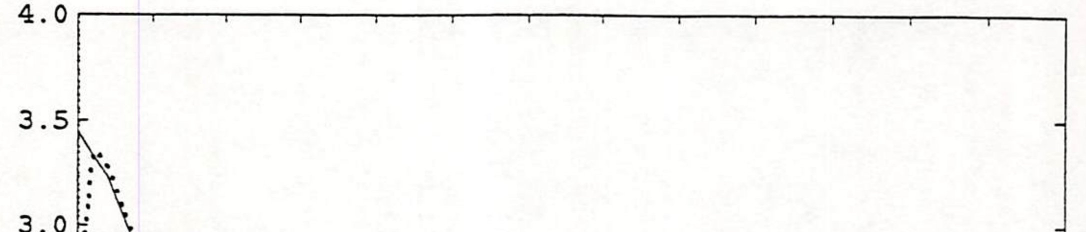


Figure 3.6: One dimensional slice comparing the true DTM, shape from shading results with unknown boundary conditions (Figure 3.4C), and shape from shading with unknown boundary conditions but utilizing low frequency constraints provided by the 0.1 percent lowest frequency Fourier coefficients of the true DTM.

the low frequency data as a projection constraint did not appear to slow down convergence in the cases tested.

3.6.4 Synthesis Results

The shape from shading algorithm was used to obtain a viewpoint - independent representation of an image suitable for predicting image intensity patterns. In Figures 3.7 and 3.8 predicted images are synthesized for various illumination geometries given true surface shapes and also given surface slope estimates obtained using the constrained shape from shading algorithm of Section 3.5.

The top row of Figure 3.7 shows simulated images for the partial sphere from Figure 3.2 created with four different illumination directions. The second row synthesizes images with the same four illumination directions but using the surface reconstruction from Figure 3.2(B), which uses known boundary conditions around the border of the image. The close match between each of the images in the second row and their counterparts in the first row appears to indicate that the shape from shading algorithm is utilizing most of the information available in the observed image intensity. The third row uses the moon surface estimate from Figure 3.3(B). Although ground truth is not available for testing the accuracy of this solution it is encouraging that the surface estimate did produce a set of synthesized images that are consistent with our visual intuition.

Figure 3.7 also reminds us that the human visual system does a good job of evaluating surface shape from image intensity for a wide variety of illumination conditions. Then it is reasonable to ask if the shape from shading algorithm can also produce high quality surface reconstructions without much sensitivity to illumination geometry. In Figure 3.4, simulated images created using four different illumination geometries were used to test the constrained shape from shading algorithm. First, simulated images were created using the DTM from Figure 3.4 given the various illumination directions. This is shown in the top row. Second, the shape from shading algorithm was applied to each of those images. The resulting four surface slope estimates were each used to simulate four images using the same set of four illumination directions. These images are presented in the four subsequent rows. For example, the DTM was reconstructed given the first (leftmost) image in the top row and that surface was used in simulating all of the images in the second row from the top. The surface reconstruction obtained from the second image in the top row was used in forming the four images in the third row, and so on. The purpose of forming the synthesized images is so that we can examine the sensitivity of the observables (image intensity) to errors in the function we want to estimate (surface shape) over a range of operating conditions (illumination geometry in this example).

Inspection of the images in Figure 3.8 for constant illumination geometry, i.e. within the same column, shows very little difference in the synthesized images. The most notable difference is a slight smoothing in the direction orthogonal to the illumination vector β as could be predicted by examining the partial derivatives of the reflectance map in (3.26). Figure 3.8 demonstrates that the constrained shape from shading algorithm obtains results that are not sensitive to the illumination geometry of the observed image, provided that it is known. This indicates that the constrained shape from shading algorithm is robust with respect to illumination geometry.

Notice that a significant portion of the first image in Figure 3.8 is covered by shadows. A reasonable surface estimate was still obtained as indicated by the predicted images in the second row of Figure 3.8. Figure 3.9 shows a representative 1-D cut through the surface reconstructed from the shadowed image as compared with the true surface and a surface estimated from an image free from shadows. It is possible, in principle, to extract powerful surface orientation constraints at shadow entry points and the relative heights between shadow entry and exit points. However, it can be difficult to reliably detect shadows and locate their entry and exit points. The predictive ability of even partial shape from shading results might provide a method of distinguishing between shadows and areas with very low albedo. Hence, we wanted to test the "robustness" of the shape from shading algorithm when confronted with shadowing, while not taking advantage of the constraints that are, in principle, available from shadowing. The only provision made in obtaining shape from shading results for partially shadowed images was to set $\lambda_1 = 0$ in the shadowed regions. This is equivalent to setting the albedo to zero in shadows, in effect ignoring

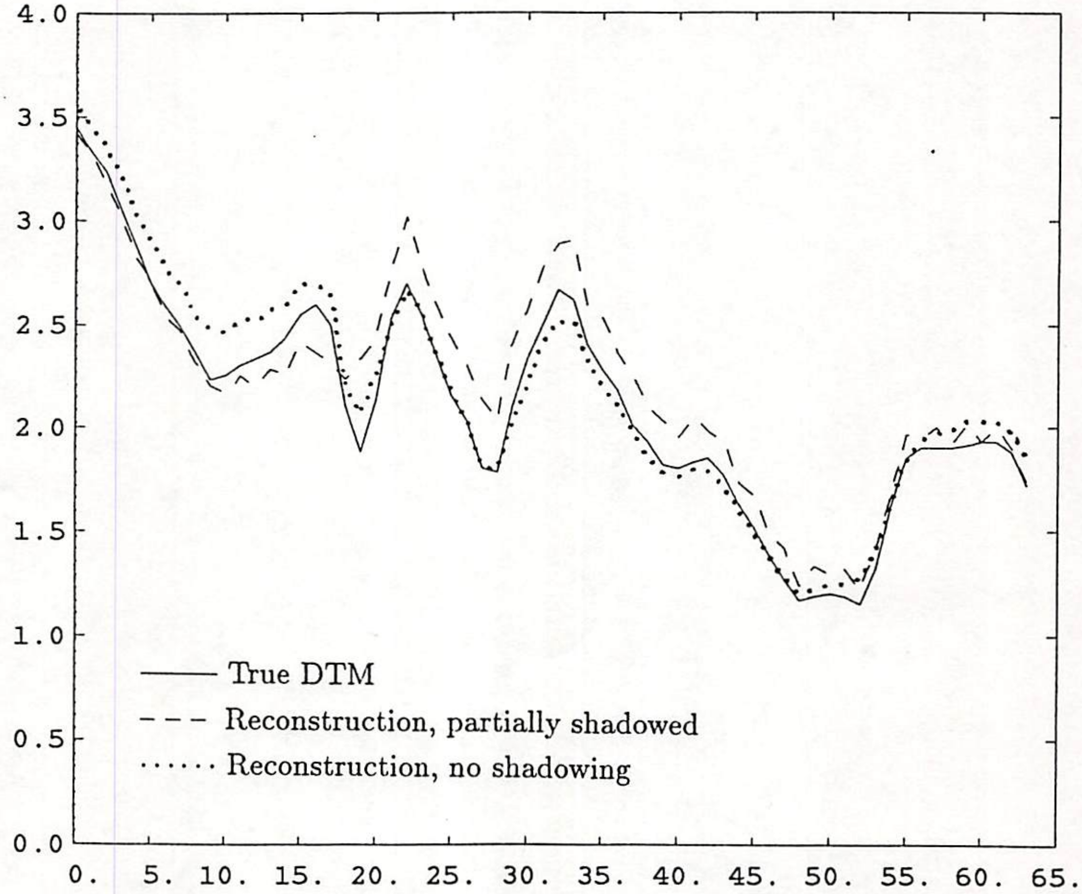


Figure 3.9: One dimensional slice showing true DTM and the estimated surface given a partially shadowed simulated image of the surface compared to an estimate made from an image without shadows. No attempt was made to utilize the relative height information available from the shadow boundaries.

the nonexistent shading information and letting the regularization penalty and integrability projection bridge the gaps left by cast shadows. This worked well with small but significant degrees of shadowing. As shadowing becomes more severe the algorithm becomes more difficult to stabilize.

3.7 Discussion

The projection represented by (3.23) was used to provide a very effective extension to the iterative shape from shading algorithm of Brooks and Horn [22]. The resulting extension converged relatively well for a variety of test cases.

This technique also can be used as an integrator which minimizes the effects of local errors by combining all of the available information in a globally consistent manner. Accordingly, it may be useful for improving other computer vision techniques that obtain possibly erroneous local derivative estimates. Two obvious examples are the shape from texture work reported by Witkin [170] and the technique recently reported by Pentland [126] which attempts to unify shape from texture with shape from shading. In both papers, mention was made of difficulties encountered in reconstructing a depth map by integrating inexact surface orientation estimates.

Shape from shading results were presented in Section 3.6 demonstrating a case where, for all practical purposes, knowledge of the 0.1 percent lowest frequency Fourier coefficients of the surface makes up for the fact that boundary conditions are completely unknown. This is exactly the situation in some remote sensing applications: boundary conditions are not available but low resolution information is. For example, the Magellan Project will provide SAR imagery of the surface of Venus along with much lower resolution radar altimetry data [36]. The algorithm in Section 3.5 easily combines the low resolution altimetry data with high resolution shading information. A simple extension such that the algorithm is suitable for SAR imagery is developed in Chapter 4.

A second application for the low frequency constraint capability is in combining stereo information with shading information. Correspondence between stereo image pairs provides low frequency information not available in shading alone. Conversely, shading provides information not available from either sparse or low resolution stereo correspondences. Further, illumination differences between stereo image pairs (for photographs taken at different times of day or for SAR imagery) may limit the accuracy of stereo image matching. It should be possible, therefore, to use shading and stereo information synergistically.

A similar integrability constraint has been used to unwrap the phase of complex SAR imagery in two dimensions [142]. A fast least squares approach was presented there which removes the inconsistencies in local phase derivative calculations (due to phase ambiguities and noise) to provide a phase image. What is most interesting here in the context of computer vision is this: The phase of a coherent image (e.g. SAR) pixel is related to surface structure. Hence, phase may provide depth information not completely available from image intensity alone. The contribution of phase in determining surface structure is an underlying principle in some applications of speckle interferometry [42], and also was utilized for SAR imagery in one dimension by Harger [67]. It may be possible to combine phase and shading information for surface shape reconstruction. Interestingly, the technique presented in this chapter is applicable to both components of that problem.

Very recently, Simchony and Chellappa [146] have applied orthogonal transforms to directly solve a linearized version of the irradiance equation and to enforce an integrability projection similar to that presented in this chapter. The integrability projection of Simchony and Chellappa was derived using a discrete formulation of the problem while we have discretized the solution to a continuous formulation of the integrability problem. More significantly, they have shown how to efficiently utilize very general boundary conditions, and how to combine sparse stereo depth maps with photometric stereo and monocular shading.

In the next Chapter, we will adapt the shape from shading algorithm from Section 3.5 to SAR imagery. One could follow the same approach to adapt other shape from shading algorithms, such as that of Simchony and Chellappa, to SAR imagery.

Chapter 4

Obtaining Shape from Shading Information in SAR Imagery

Topographic data can be extracted from SAR imagery using at least three kinds of approaches; those that exploit radiometric, geometric, and interferometric information. Radiometric information is useful because the backscattered power is a function of terrain surface slopes. This is the basis of shape from shading techniques. Geometric information enters through the relatively stable relationship between the scene geometry and its projection into a SAR image. This is the basis for stereogrammetry, using multiple images, and for mensuration of discrete objects in a single image. Interferometry uses the relative phase variation betwen the complex amplitudes of two SAR images made from two vertically separated antennas to obtain estimates of the elevation angle to the terrain surface.

It has been suggested that all three approaches are complementary to some extent [39]. It is well-known in the computer vision community that shading of smooth surfaces and stereo parallax provide complementary information: Shading is useful in areas with constant albedo and fairly smooth variations in surface orientation. Conversely, stereo matching is more effective in areas with large albedo variations or sharp surface orientation changes that provide the required intensity features. For SAR imagery, geometric and radiometric information are not only complementary, they are synergistic. Because SAR is an active sensor, both geometric as well as radiometric disparities are created when a SAR stereo image-pair is formed. This limits the accuracy of radar-stereogrammetry [99]. The performance of interferometry depends on the ability to reliably unwrap the phase of noisy image data and correct for a variety of phase errors [104]. The point is that a single information source usually does not yield all of the topographic information available from SAR imagery.

In this Chapter we apply a shape from shading technique to SAR imagery and illustrate some of the strengths and weaknesses of shading as a source of topographic information. SAR image characteristics have not been addressed in any of the previous shape from shading research.

The use of radar backscatter power to extract terrain surface orientation, sometimes referred to as radarclinometry [167], has been considered before. Cosgriff, et al. [32], while concentrating on clutter modeling for system design purposes, discussed the possibility of radarclinometry and suggested a solution method using multiple radar images. Wildey [167,168,169] developed algorithms for reconstructing surface topography from the shading in a single SAR image. Those algorithms directly inverted a differential equation similar to (3.1) for SAR imagery subject to local constraints on the relationship between z_x and z_y. These constraints are similar, although more general, than the local sphericity assumption used by Pentland [125] and also must be inferred from image intensity derivatives.

We have adapted the shape from shading method from Chapter 3 to SAR imagery using an approach that could also be applied to any number of related algorithms from the computer vision literature. This provides more general solutions, more effective methods for applying additional constraints, and more efficient numerical computation than previously available for radarclinometry. In Chapter 2, SAR image characteristics were related to surface properties. Based on those characteristics, a shape from shading algorithm suitable for SAR imagery is presented in this Chapter. The algorithm, its applications, and possible extensions, are discussed in Section 4.2. In Section 4.3, this algorithm is applied to reconstructing surface topography for both simulated and real SAR imagery. Conclusions are summarized in Chapter 7.

4.1 SAR Image Based Shape from Shading Algorithm

As illustrated in Chapter 2, the relationship between image intensity and surface properties for SAR imagery differs markedly from that of, say, aerial photographs. Yet, there are also important *similarities* between the computational vision models for SAR imagery and aerial photographs. It is those similarities which allow us to apply some of the same methodologies utilized in previous shape from shading research.

First, it was pointed out that a SAR strip map can be represented as an orthographic projection of a surface relative to the slant plane. The simplifying assumption of orthographic projection, utilized in many algorithms applied to visible images (including shape from shading), is less restrictive for SAR images. We have represented the surface topography relative to the slant plane as formulated in (2.3). The resulting algorithm reconstructs $u(r, y)$, an orthographic projection of (u, r, y), given observed SAR image intensity.

While the reflectance map models typically used for visual images are not valid for SAR images, a similar computational vision modeling approach has been used quite successfully for radar image simulation [53,69]. The particular reflectance maps are different, but the same methodology is applied. The methodology used for solving the shape from shading problem has been to model the observed image intensity in terms of local surface slopes and develop an algorithm to solve for those surface slopes (or surface height). The algorithms reported by Brooks and Horn [22] and Frankot and Chellappa [48] solve for surface slopes by minimizing mean-squared intensity error criterion balanced against a surface smoothness criterion. The resulting cost function is

$$\epsilon = \int\int [I - \mathcal{R}(\hat{u}_r, \hat{u}_y)]^2 + \lambda \cdot (\hat{u}_{rr}^2 + 2\hat{u}_{ry}^2 + \hat{u}_{yy}^2) dr dy \ . \qquad (4.1)$$

Along with the cost function in (4.1) an integrability criterion is enforced. The algorithm constrains the surface slopes (u_r, u_y) to satisfy

$$\frac{\partial}{\partial y} u_r = \frac{\partial}{\partial r} u_y \ , \qquad (4.2)$$

or $u_{ry} = u_{yr}$. The fast least squares technique presented in Chapter 3 is applicable given the orthographic projection property of $u(r, y)$. Suppose that we are presented with a possibly nonintegrable pair of slope functions $[\hat{u}_r(r, y), \hat{u}_y(r, y)]$ such that any number of surface reconstructions $\hat{u}$ could be obtained depending upon the path of integration. The least squares integrability projection algorithm finds the pair of slope functions $[\tilde{u}_r(r, y), \tilde{u}_y(r, y)]$ that satisfy (4.2) simultaneously with minimizing the distance

$$d\{(\hat{u}_r, \hat{u}_y), (\tilde{u}_r, \tilde{u}_y)\} = \int\int (\hat{u}_r - \tilde{u}_r)^2 K_r + (\hat{u}_y - \tilde{u}_r)^2 K_y dr dy \; , \tag{4.3}$$

where $K_r + K_y = 1$. The constants K_r and K_y are relative weights assigned to u_r and u_y, respectively. It is helpful to set $K_r > K_y$ because the error in u_r is normally smaller than the error in u_y.

The constrained minimization problem defined by (4.1), (4.2), and (4.3) can be solved by extending the iterative algorithm presented in Chapter 3. The resulting algorithm is now summarized. Denote the surface slope estimates after the kth iteration as $[\tilde{u}_r, \tilde{u}_y]_k$. The surface slope estimates for the next iteration are obtained by the following five steps:

1. Obtain a smoothed version, $[\hat{\hat{u}}_r, \hat{\hat{u}}_y]_k$, of the previous iterate using

$$\begin{aligned} &\hat{\hat{u}}_y(l, m) = \\ &\tfrac{1}{5}[\tilde{u}_y(l, m+1) + \tilde{u}_y(l, m-1) + \tilde{u}_y(l+1, m)\tilde{u}_y(l-1, m)] \\ &+\tfrac{1}{20}[\tilde{u}_y(l-1, m-1) + \tilde{u}_y(l-1, m+1) \\ &\quad +\tilde{u}_y(l+1, m+1) + \tilde{u}_y(l+1, m-1)] \end{aligned} \tag{4.4}$$

and similarly for $\hat{\hat{u}}_r$.

2. Simulate the SAR image by evaluating the reflectance map with the smoothed slope estimates, giving $\mathcal{R}(\hat{\hat{u}}_r, \hat{\hat{u}}_y)$ for every range-azimuth sample location, (l, m).

3. Calculate the partial derivatives of the reflectance map with respect to u_r and u_y. The derivatives are evaluated using the smoothed surface slopes from step 1,

$$\mathcal{R}_r(\hat{\hat{u}}_r, \hat{\hat{u}}_y) = \frac{\partial}{\partial a}\mathcal{R}(a, b)\bigg|_{a=\hat{\hat{u}}_r, b=\hat{\hat{u}}_y} \tag{4.5}$$

and

$$\mathcal{R}_y(\hat{\hat{u}}_r, \hat{\hat{u}}_y) = \frac{\partial}{\partial b}\mathcal{R}(a,b)\bigg|_{a=\hat{\hat{u}}_r, b=\hat{\hat{u}}_y} .$$

Numerical derivatives are used to provide greater flexibility, allowing the use of a reflectance map look-up table and, if necessary, an empirical reflectance map. Normally $\mathcal{R}_r \gg \mathcal{R}_y$, explaining why $\hat{u}_r$ has less error than $\hat{u}_y$. This is what Wildey referred to as "poor photometric leverage" for the cross-slopes [167].

4. Update the slope estimates using the recursion [22],

$$\begin{bmatrix} \hat{u}_r \\ \hat{u}_y \end{bmatrix}_{k+1} = \begin{bmatrix} \hat{\hat{u}}_r \\ \hat{\hat{u}}_y \end{bmatrix}_k + \lambda_1 \left(I - \mathcal{R}(\hat{\hat{u}}_r, \hat{\hat{u}}_y)\right) \begin{bmatrix} \mathcal{R}_r(\hat{\hat{u}}_r, \hat{\hat{u}}_y) \\ \mathcal{R}_y(\hat{\hat{u}}_r, \hat{\hat{u}}_y) \end{bmatrix} \tag{4.6}$$

where λ_1 is inversely proportional to λ in (4.1). The new slope iterate obtained from (4.6) is virtually guaranteed to be nonintegrable.

5. Enforce integrability using the following projection constraint [51] on the Fourier series of $[\hat{u}_r, \hat{u}_y]_{k+1}$

$$\tilde{U}(\underline{\omega}) = \frac{K_r a_r^*(\underline{\omega})\hat{U}_r(\underline{\omega}) + K_y a_y^*(\underline{\omega})\hat{U}_y(\underline{\omega})}{K_r|a_r(\underline{\omega})|^2 + K_y|a_y(\underline{\omega})|^2} \tag{4.7}$$

where $\hat{U}_r(\underline{\omega})$ and $\hat{U}_y(\underline{\omega})$ represent the discrete Fourier transforms (DFT) of $[\hat{u}_r, \hat{u}_y]_{k+1}$ respectively, $\tilde{U}(\underline{\omega})$ represents the DFT of the new estimate of surface height $\tilde{u}(r,y)$. The vector $\underline{\omega} = (\omega_r, \omega_y) \in (\frac{2\pi n}{N}, \frac{2\pi m}{N})$, with $n, m \in \{0, 1, \cdots, N-1\}$ for images of size N by N. The coefficients a_r and a_y are given by the frequency response of the partial differentiation operators in r and y, respectively. For example, if we approximate the surface slopes using first order central differencing then the frequency response is $a_y(\underline{\omega}) = j\sin(\omega_y)$ and similarly for a_r. The Fourier transform of the new surface slope iterates are given by

$$\tilde{U}_r(\underline{\omega}) = a_r(\underline{\omega})\tilde{U}(\underline{\omega}) \quad \text{and} \quad \tilde{U}_y(\underline{\omega}) = a_y(\underline{\omega})\tilde{U}(\underline{\omega}) \tag{4.8}$$

with the final step in the iteration that of performing the inverse DFT of (4.8) to get $[\tilde{u}_r, \tilde{u}_y]_{k+1}$. The integrability projection constraint from (4.7) doubles as an efficient 2-D integrator for constructing a surface from noisy slope estimates utilizing all of the available data.

The same process is repeated until the intensity mean squared error is sufficiently small or until a predetermined number of iterations have been completed. The overall procedure follows the structure shown in Figure 4.1.

4.2 Applications in Radargrammetry

The integrability projection used in the shape from shading algorithm has some distinct advantages. The main advantage is that it is noniterative—it constructs a surface from noisy slope estimates in one pass utilizing all of the available data, unlike simple spatial integration. A second advantage is computational efficiency. It utilizes fast Fourier transform (FFT) algorithms routinely used for SAR image formation. Hence the method is efficient in terms of the number of computations required, the ability to utilize high speed FFT processors, and synergy with SAR image formation hardware and software. A third advantage is modularity.

It is possible to apply other constraints in a similar manner. This is represented in the block diagram of Figure 4.1. The integrability constraint is a form of self-consistency. Another form of self-consistency is that shadows appearing in the image should correspond to shadows predicted by surface reconstruction. The second class of constraints comes from auxiliary data sources. This might come from altimetry data, stereopsis, and higher level scene analysis. If auxiliary information is local in nature it can be applied as a constraint directly to the surface. If it is global in nature it is sometimes more conveniently applied in the frequency domain. Altimetry data and stereo data tend to be of lower resolution than SAR imagery obtained by the same radar system and, accordingly, may be useful as global low frequency constraints for surface reconstruction. This is discussed below.

4.2.1 Low Frequency Information

Note that (4.7) is valid only for $\underline{\omega} \neq (0,0)$. This illustrates the fact that low frequency surface height information is lost in the image formation process, since image intensity is a function of the derivatives of surface height. A surface reconstructed based on shape from shading techniques inevitably suffers from low frequency errors, the severity depending upon observation noise characteristics. This also appears to be true of interferometry [61], which relies on noisy and ambiguous observations of phase differences to solve for relative surface height.

An added advantage of the Fourier based integrability constraint is that low frequency information from other sources can easily be included during or after iteration of the shape from shading algorithm. This can be accomplished by performing the DFT of the low resolution surface height information, denoted by $U_L(\underline{\omega})$. This can be precomputed with modest computational cost. The appropriate low frequency terms of U_L are used to replace the low frequency DFT coefficients $\tilde{U}(\underline{\omega})$, in effect treating U_L as a projection constraint. The algorithms of Brooks and Horn [22] and Ikeuchi and Horn [79], which our algorithm is based upon, requires knowledge of boundary conditions. While boundary conditions are generally not available, low frequency information sometimes is. Further, availability of low resolution surface height data can, for all practical purpose, eliminate the need for boundary conditions in addition to replacing lost low frequency information as shown in Chapter 3. The significance of this issue for SAR imagery is illustrated in Section 4.3.2.

4.2.2 Fusion of Stereopsis and Shape from Shading

A second possible source of low resolution information is from stereo image pairs. One computer vision paradigm uses shape from shading to provide an interpolation between sparse height information provided by matching stereo image pairs. That is, the height information provided by the stereo image pair may have very high resolution but is only available at a few points in the image. The sparseness of this stereo information results from the practice of extracting correspondence by matching discrete features, such as edges. A second method, with its own advantages and disadvantages, obtains correspondence using cross-correlation of multiple subareas [132]. The correlation approach does not rely on the presence of sharp changes of intensity but requires fairly large areas to average over in order to obtain reliable matches with SAR imagery [46]. Hence, use of multiple subarea correlation to estimate correspondence between stereo image pairs provides low resolution correspondences, but complete area coverage. Such low resolution stereo data could be combined with shape from shading results using the algorithm presented above.

Conversely, shape from shading can be used improve stereo match results. For SAR imagery, a tradeoff exists between two competing effects: On the one hand, the larger the difference in look angles for the stereo image pair the less sensitive the surface reconstruction is to errors in stereo correspondence [98]. On the other hand, the errors in stereo correspondence grow as the disparity between the look angles increases [99]. In many cases the human eye is not even be able to recognize that two SAR images are of the same terrain if they are made from drasticly different look angles because the change in illumination geometry causes changes in shading. Hence, stereopsis helps shape from shading and shape from shading helps stereopsis.

4.2.3 Venus Radar Mapper

The geology of the planet Venus is of great scientific interest because it is about the same size and density of Earth and it is Earth's closest neighbor [24]. The scientific community relies heavily upon radar observations to provide this information since the Venusian atmosphere is opaque at visible wavelengths. The Magellan project's Venus Radar Mapper will provide high resolution (nominally 150 meter) SAR imagery of the surface of Venus along with much lower resolution (5 to 50 kilometers depending on operating altitude) radar altimetry data [77]. It is believed that the Magellan SAR imagery will provide new information about the geology of Venus, which in turn will create more insight into the evolution of the solar system.

For example, certain significant geological features, such as folds and faults, appear as banded terrain in SAR imagery. Bands from 10 to 20 kilometers wide have been observed in previously obtained radar images of Venus and similar bands as small as 5 kilometers appear in the Appalachian mountains [24]. While the altimetry data might not be able to resolve the bands, the Magellan SAR imagery certainly will.

Using the techniques presented in Section 4.1, the altimetry data and the SAR imagery could be combined to construct high resolution topographic maps of Venus. The Magellan scenario, where SAR imagery is combined with orders of magnitude lower resolution altimetry data for surface reconstruction, is simulated in Section 4.3. The grazing angle for the SAR imagery will be anywhere from 40 to 80 degrees, depending on orbital position. Imagery near 45 degrees grazing should provide the best compromise between shadowing and layover. Experience with Seasat imagery indicates that the higher grazing angle imagery is still useful but the likelihood of encountering layover is much greater. Noncoherent integration of anywhere from 4 to 30 looks (depending on orbital position and the demands placed upon the communication link) will be performed to reduce speckle. The highest possible degree of noncoherent integration is desired to provide an adequate radiometric resolution without compromising spatial resolution. If 30 look imagery is indeed available, then the shape from shading algorithm can probably be applied without any additional noncoherent integration at the full resolution specified for the SAR imagery.

Earlier scientific observations point out some potential difficulties in applying shape from shading to SAR imagery of Venus. For example it has been observed that surface roughness is often higher at high elevation than at low elevation on the surfaces of both Earth and Venus (apparently due to erosion at mountains and sedimentation in valleys) [127]. The reflectance map is a function of surface roughness. Therefore, the impact of this variation on shape from shading might be to produce distortions in the estimated surface which are not easily detected, perhaps making a mountain seem to be less jagged than it really is or making plains appear rougher than in reality. There also appears to be correlation between high albedo and high elevation, perhaps caused by erosion exposing metallic ore at high elevations. Thus, while albedo variations on the surface of Venus appear to be much lower than those typical of Earth, they are significant [127], and their effect on shape from shading is probably correlated with surface structure.

Prior scientific work also provides possible solutions to the above difficulties. Along with low resolution surface height data, it will also be possible to obtain estimates of of the albedo and surface roughness (and therefore the reflectance map) at comparable resolutions. Scattering laws have been derived which provide a familiy of reflectance maps parameterized by RMS surface roughness and albedo [9,63]. Pettengill et al. [127] used a scattering law due to Hagfors [63] to obtain reflectivity and surface roughness estimates from earlier Pioneer Venus radar altimetry data. The observed distribution of echo power versus time delay was first converted to a distribution of echo power versus angle of incidence, which was matched with the scattering law to provide an estimate of the surface roughness and albedo. This information helped scientists to develop theories on the geological make-up of Venus. In the same manner, we can construct practical estimates of albedo and reflectance maps to support the application of shape from shading. The likely absence of cultural features and vegetation may practically eliminate the need for any higher resolution estimates of albedo.

4.2.4 Albedo Variation

Haralick et al. [66] extended a technique due to Eliason et al. [41] to detect albedo variations using multi-spectral Landsat imagery. In a similar manner, it may be possible to detect changes in albedo using multi-frequency radar imagery [15]. It may also be possible to segment areas of drasticly different albedo (e.g. no return versus moderate return versus very high return) prior to application of shape from shading techniques and to develop shape from shading algorithms that are more robust with respect to albedo variations.

There also appears to be some informaton on both albedo and surface structure in SAR image phase [67]. Presently, it is not clear just how much can be gained through the use of phase information, especially with diffuse reflection. Polarization diversity may provide additional information useful for obtaining shape from both shading and phase information. The recent analysis of multi-polarization radar imagery [163] may provide insight into the utilization of phase data and polarization diversity for inferring surface structure. For example, it is possible to distinguish between surface and volume scattering mechanisms, and therefore reflectance map classes, using the phase differences between images sensed with different polarizations.

4.3 Experimental Results

4.3.1 Simulation Results

The shape from shading algorithm presented in Section 4.1 was tested on simulated and real SAR imagery. Figures 4.2 and 4.3 illustrate the simulation results. In Figure 4.2 (a) the original DTM is shown. A simulated noise-free SAR image with a 45 degree grazing angle is shown in Figure 4.2 (b). A simulated speckled SAR image is shown in Figure 4.2 (c). Although the speckle noise is reduced by noncoherent averaging of 16 speckle samples, the noise is still significant both visually and statistically. For this particular case, the standard deviation of the noise alone is about the same as the the standard deviation of the noise-free image from Figure 4.2 (b). Figure 4.2 (d) shows the reconstructed surface obtained from the noisy image. Although no special provisions were made to accommodate the noise, a reasonable reconstruction was still obtained.

A more clear evaluation of the results can be made by examining 1-D slices through the original and reconstructed DTMs. In Figure 4.3 the solid line shows a 1-D slice through the original DTM. The dotted line shows the shape from shading results given the noise-free simulated image and the dashed line shows the results obtained using the noisy image. For the reconstructions shown, boundary values were not provided to the algorithm. Instead, 0.1 percent of the lowest frequency DFT coefficients of the DTM were provided.

While the simulation results are encouraging, they are unrealistic in that no modeling errors are present. A practical test of shape from shading on real SAR imagery is now presented.

4.3.2 Results with SIR-B SAR Imagery

A set of three SIR-B SAR images, an aerial photograph, and a DTM derived stereoscopically from aerial photographs (all for the same location) were used as test cases. These were provided by Dr. Gitta Domik of the Vexcel Corporation. The area coverage was near the town of José de San Martin in Argentina. Figure 4.4 shows a SAR image simulated from the DTM with and without speckle noise,

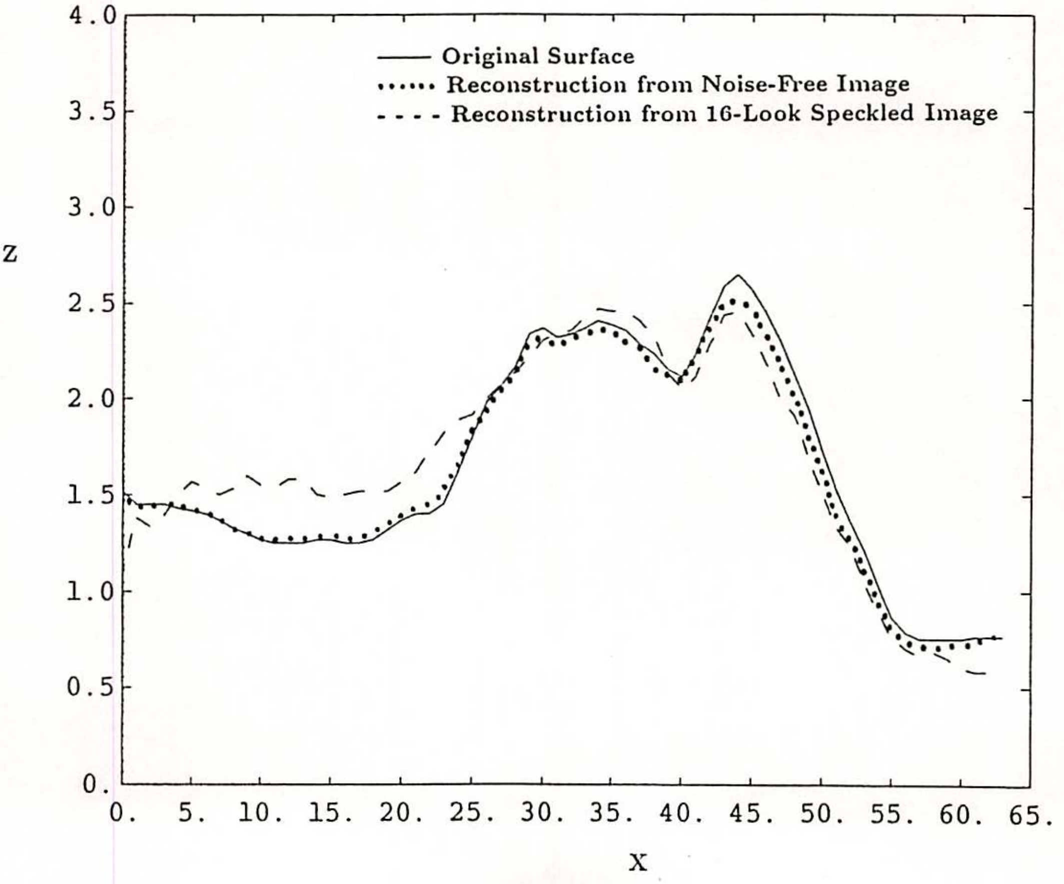


Figure 4.3: 1-D cuts through the original DTM compared to the noise-free reconstruction, and the reconstruction given the noisy image.

one of the SIR-B SAR images, and a registered aerial photograph. Proceeding diagonally across the aerial photograph a riverbed and a roadway are clearly visible, and a great deal of other albedo variations are evident. A small airport can be seen in the upper left hand corner of the aerial photograph. The town of José de San Martin is just to the right of the airport but is not discernable at this resolution.

Comparison of the images with registered terrain elevation data revealed some important characteristics of this data set. First, the empirical reflectance map for the aerial photograph was very close to the often assumed Lambertian characteristic. The reflectance map for the SAR imagery was even more directional than expected. Second, the albedo variations present in the aerial photograph prevented successful application of any current shape from shading technique, while the albedo variation in the SAR imagery was much less severe. The SAR albedo variation was significant enough to be noticed visually and to have a negative impact on the surface reconstruction results but did not cause catastrophic failure.

For each of the three SAR images we used the following experimental procedure:

1. *Reduce SAR image pixel spacing to match the DTM pixel spacing.* The original SIR-B pixel spacing is about 12.5 meter while the DTM sample spacing is 37.5 meters, a factor of 3 reduction. Resolution reduction was performed by averaging 3×3 blocks of pixels. The original SAR image is obtained as a noncoherent average of four looks. The equivalent number of looks for a 37.5 meter pixel is then 36, assuming independent speckle samples. However, SAR imagery is normally slightly oversampled. The equivalent number of looks estimated using the ratio of the square of image mean to image variance in a very low contrast region is about 28 (after compensating for intensity variations due to topography).
2. *Precisely register DTM to SAR image.* This allows us to compare the reflectance parameters for the different images, evaluate signal to

noise ratio, and compare the SAR image based surface reconstruction with the DTM. To estimate the transformation we simulated a SAR image given the DTM and approximate knowledge of imaging geomtery. Next we performed an exhaustive search over translation, rotation, and scale, selecting those parameter values which maximized the normalized cross-correlation. This is very similar to the approach Horn and Bachman used for registering LANDSAT imagery to DTMs [72] and is discussed in Chapter 5. The orthographic projection representation in the slant plane provided an excellent model for the SAR coordinates: Sub-pixel registration accuracy between simulated and observed images was achievable after compensating only for translation, rotation, and scaling of the ground coordinates.

3. *Generate low resolution DTM.* In each case, the Fourier transform of the registered high resolution DTM was obtained, all but the lowest frequency 0.1% of the coefficients were set to zero, and then the inverse Fourier transform of the resulting spectrum was performed. This represents a resolution reduction factor of about 23 in range and azimuth, roughly simulating the Magellan scenario.
4. *Estimate reflectance map parameters using low resolution DTM.* This ensures that the shape from shading test results represent the useful case where the high resolution surface structure is unknown.
5. *Execute shape from shading algorithm given a SAR image and low resolution DTM.* Ten iterations of the algorithm from Section 4.1 where performed in each case. The total execution time for a 256 × 256 image was 45 minutes on an unloaded Micro VAX, including resampling into ground plane coordinates.

Details of these results are presented below.

Reflectance Parameter Estimation

Reflectance models from Chapter 2 were fit to the SAR imagery using both the low resolution and high resolution DTMs. Three types of parameters need to be estimated: Additive noise bias, albedo, and shape parameters. An exhaustive search over the parameter space was used in each case to select those values which minimize the following mean-squared prediction error

$$\epsilon_p = \sum_{(l,m)\in\Omega} \frac{1}{M_\Omega} \left[I(l,m) - \mathcal{R}(u_r(l,m), u_y(l,m))\right]^2 , \tag{4.9}$$

where M_Ω denotes the number of pixels in the domain Ω under consideration. Equation (4.9) is just the sample second moment of the error between the observed image intensity and predicted image intensity. We compared the results using the generalized Lambert model and Barrick's rough surface model with Gaussian and exponential surface slope probability density functions (PDF). This was done for various images, DTM resolutions, and image resolutions. Barrick's model with Gaussian height PDF nearly always resulted in lower mean squared prediction error than the exponential model and the generalized Lambert model, although the visual realism of the image simulation was about the same in each case. Reflectance parameters were fairly stable as a function of image resolution. Each parameter varied slightly as a function of DTM resolution. The error in the parameter estimate obtained using a low resolution DTM appears to be somewhat predictable but no attempt was made to compensate for it in our experiments.

The two different area factor parameterizations, illumination area and surface area, were compared when used in conjunction with Barrick's (Gaussian) RCS model. Use of illumination area resulted in consistently lower mean squared prediction error. Visually, the quality of image simulation did not vary significantly between the two area factor parameterizations.

Figure 4.5 shows a scatter plot of observed SAR image intensity versus $\cos \alpha_i$ for the SAR image of Figure 4.4 (bottom-left), providing an empirical reflectance map. The nominal depression angle is 32.9 degrees so that the value of $\cos \alpha_i$ is approximately 0.84 for flat ground and the surface orientation data centers itself around this value. Notice from Figure 4.5 that the reflectance is a strong function of $\cos \alpha_i$ at high angles of incidence but only a very weak function at low angles of incidence. Accordingly, the up-slopes (areas increasing in height as slant range increases) have much more contrast and higher intensity than the down-slopes. Small-signal suppression due to additive noise and intensity quantization therefore destroys shading information on the down-slopes.

Tables 4.1 through 4.3 show the reflectance map parameter estimates obtained from the three SIR-B SAR images given high and low resolution DTMs. In each case we used Barrick's RCS model for Gaussian surface slopes with the illumination area parameterization. RMS surface roughness determines the shape of the RCS curve with the resulting reflectance map being more directional the lower the roughness. The surface roughness estimates, obtained as a byproduct of reflectance map parameter estimation, were between 17.5 and 24 degrees in each case.

RMS prediction error between real and simulated image intensity is also given in Tables 4.1 through 4.3. The prediction error was calculated in each case using the high resolution DTM to predict image intensity through the reflectance model. In the first column, an indication of prediction error with no modeling error was obtained by simulating a 28-look speckled image with the same reflectance map parameters as the real SAR image. The mean squared error between the simulated speckled image and the simulated image without speckle was then computed. The second prediction error column in Tables 4.1 through 4.3 uses the same reflectance map parameters (estimated using the high resolution DTM) and compares predicted image intensity to the observed SAR image intensity. The third prediction error column uses reflectance map parameter estimates obtained given the low resolution DTM. This is the situation that would be encountered in practice.

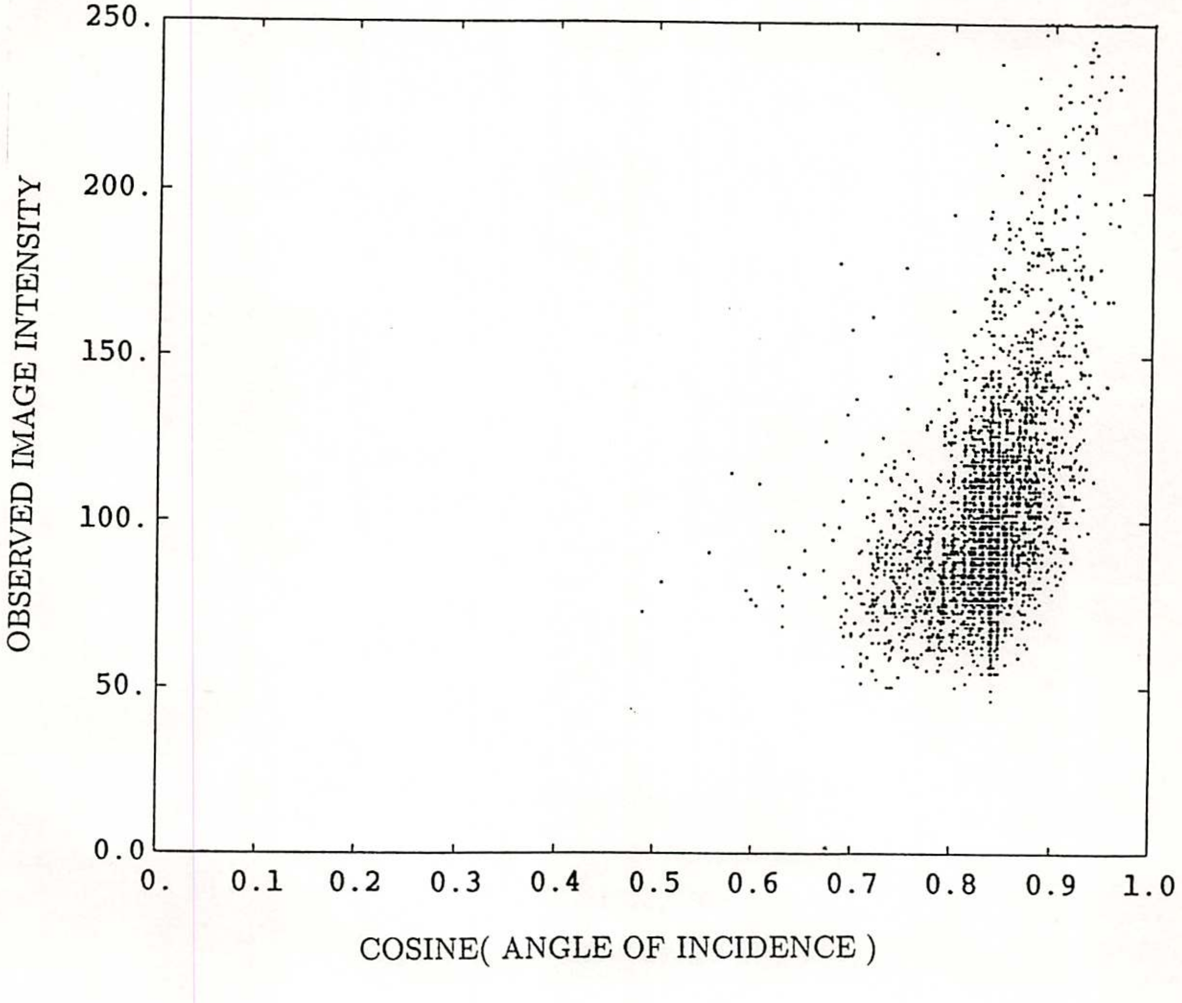


Figure 4.5: Scatter plot of observed SAR image intensity versus cos(angle of incidence), SIR-B image with pixel spacing reduced to 37.5 meters. The nominal depression angle is 32.9 degrees giving a mean $\cos \alpha_i \approx .84$.

	Simulated image with speckle No modeling errors	Parameters estimated using high resolution DTM	Parameters estimated using low resolution DTM
Surface roughness (degrees RMS)	–	19.1	17.5
η	–	12.2	11.6
σ^2_{bias}	–	77.7	84.9
Intensity prediction RMS error	19.46	28.55	28.72
SNR (dB)	1.5	-1.04	-1.16

Table 4.1: Reflectance map parameter estimates, resulting prediction error statistics, and shape from shading SNR measure for SIR-B SAR image in the upper left hand corner of Figure 4.8.

	Simulated image with speckle No modeling errors	Parameters estimated using high resolution DTM	Parameters estimated using low resolution DTM
Surface roughness (degrees RMS)	–	23.5	23.0
η	–	13.6	15.5
σ^2_{bias}	–	38.0	38.8
Intensity prediction RMS error	8.56	10.49	10.57
SNR (dB)	.13	-2.85	-3.0

Table 4.2: Reflectance map parameter estimates, resulting prediction error statistics, and shape from shading SNR measure for SIR-B SAR image in the top row, center column of Figure 4.8.

	Simulated image with speckle No modeling errors	Parameters estimated using high resolution DTM	Parameters estimated using low resolution DTM
Surface roughness (degrees RMS)	–	24.1	20.5
η	–	10.6	9.2
σ^2_{bias}	–	78.1	88.1
Intensity prediction RMS error	19.74	25.27	25.50
SNR (dB)	.01	-1.68	-1.88

Table 4.3: Reflectance map parameter estimates, resulting prediction error statistics, and shape from shading SNR measure for SIR-B SAR image in the top row, right hand column of Figure 4.8.

Signal to Noise Ratio

A measure of signal to noise ratio (SNR) specially defined for shape from shading is presented in Tables 4.1 through 4.3. This measure is given by

$$\frac{\sigma_I^2 - \epsilon_p}{\epsilon_p} , \tag{4.10}$$

where

$$\sigma_I^2 = \frac{1}{M_\Omega} \sum_{(l,m)\in\Omega} I^2(l,m) - \left(\frac{1}{M_\Omega} \sum_{(l,m)\in\Omega} I(l,m) \right)^2 \tag{4.11}$$

is the sample variance of the observed image and ϵ_p is the intensity prediction error defined in equation (4.9). Several points are now evident. First, comparison of SNR for simulated speckle images with that of real images indicates a 2 to 3 dB loss due to modeling errors for these particular cases. We conjecture that spatial variations of reflectance parameters, i.e. albedo, noise bias, and surface roughness account for most of the modeling error. Second, comparison of SNR using high resolution versus low resolution reflectance parameter estimates indicate only an additional 0.2 dB loss for the low resolution estimates. This demonstrates successful reflectance map estimation from low resolution topography. Third, the SNR is very low (even without modeling error), suggesting that future improvements might be obtained through optimal filtering of speckle noise.

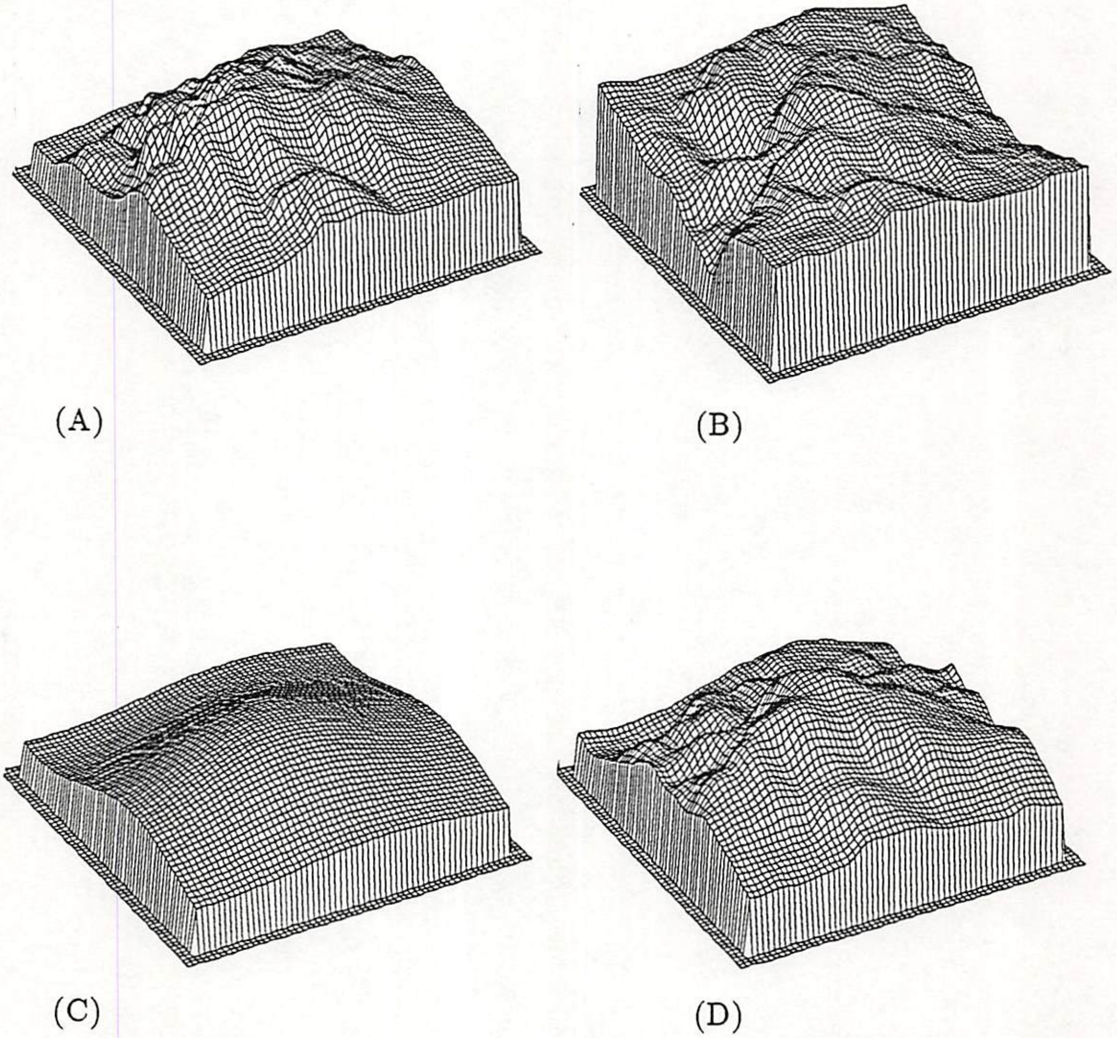
Low Frequency Information

It has been suggested that auxiliary low frequency information be used to compensate for unknown boundary conditions and observation noise [51]. An example using a SIR-B SAR image window of size 64 by 64 pixels with 75 meter pixel spacing is now provided. Figure 4.6 demonstrates the results of shape from shading with and without auxiliary low frequency information. In Figure 4.6A a section from a DTM is shown. The surface reconstruction given only the shading information in a SAR image shown in Figure 4.6B. Notice that the slopes in the range direction are fairly accurate while the slope estimates in the azimuth direction are not. This is partly a consequence of unknown boundary conditions and partly due to the fact that image intensity is a very weak function of the azimuth slopes. A low resolution surface is shown in Figure 4.6C. This was obtained from the five lowest frequency Fourier coefficients of Figure 4.6A. The results of using shape from shading to combine the low resolution surface with high resolution image intensity information is shown in Figure 4.6D. Interestingly, the distorted surface of Figure 4.6B gives a lower value for the cost function in (4.1) than does the true surface. Actually, both the intensity prediction error and the surface smoothness criterion were lower for the distorted surface. Tests on simulated imagery indicate that this is not due solely to modeling error: Similar results were obtained for simulated speckled imagery, but not for speckle-free imagery. The general trend is as expected—as the noise level increases so does the loss of low frequency surface information in the image.

The same kind of approach was applied to the 256 by 256 pixel SAR images corresponding to Tables 4.1 through 4.3. The procedure described in the beginning of Section 4.3.2 was followed. Figure 4.7 shows an image simulated from the low resolution DTM. This illustrates how little low frequency surface information is available in image intensity and also a possible difficulty in estimating reflectance map parameters from low resolution surface data. The higher slope values are not well represented so that the reflectance map is fit mostly to data with angle of incidence near the radar depression angle. This indicates that some degree of extrapolation is required and, therefore, care must be taken in selecting an appropriate reflectance model class.

Surface Reconstructions

Images Predicted from Surface Reconstructions: Figure 4.8 shows three original SAR images and and intensity predictions given shape from shading results. The top row shows the three observed SAR images corresponding to Tables 4.1 through 4.3, respectively. In each case the range direction and the illumination is along the horizontal axis. The first two images are taken from about 12 degrees difference in depression angle and about 1.3 degrees difference in orientation or aspect. The third image has about 81 degrees aspect and illumination difference from the first two images. Surface reconstructions from each of the SAR images were obtained and then each surface reconstruction was



(A) (B)

(C) (D)

Figure 4.6: (A) Section from DTM co-registered with SAR image, (B) surface reconstruction given only the shading information in a SAR image, (C) low resolution surface obtained from the five lowest frequency Fourier coefficients of the DTM, (D) surface reconstruction obtained using shape from shading algorithm to combine the low resolution surface with high resolution image intensity information.

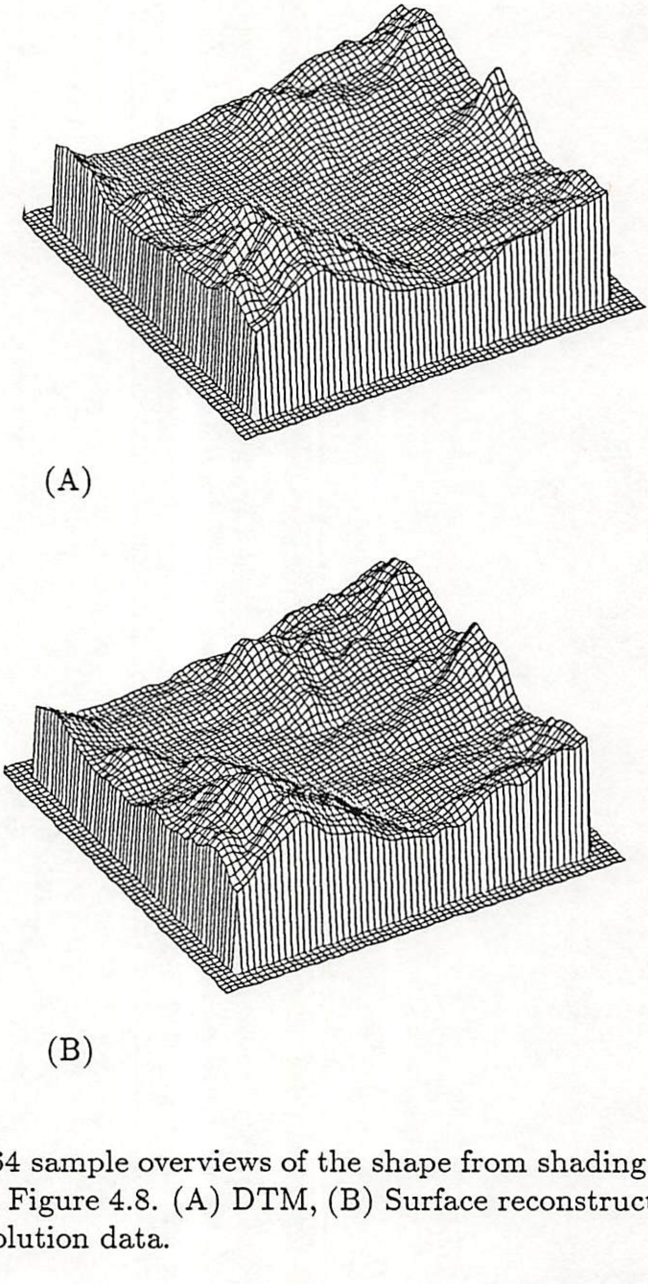
used to predict each of the original SAR images. The second row shows the three predicted SAR images given the shape from shading results of the top-left SAR image. The third row shows predicted images given the shape from shading results of the top-center SAR image, and the fourth row does the same given the top-right SAR image.

Comparison of the observed images with predicted images show excellent intensity predictions for similar imaging geometries, as expected. Prediction of the first two SAR images given the third (and vice versa) provides a worst-case example because of the near orthogonal aspect difference. This means we are using mostly the estimates of azimuth slope for one image to predict the intensity of the other image. The left and center images in the bottom row show very good predictions of the first and second SAR images, respectively. The predictions of the third SAR image given in the rightmost column, second an third rows, are of fair quality. They show some loss of the structures that were in the downslopes of the observed image and some distortions due to albedo variations, yet most of the structure is still recognizable.

The large illumination differences also present some difficulty in deriving stereo correspondences. The predictive ability of shape from shading results might support stereo matching in this case [99].

Comparison with DTM: Actual surface reconstructions are now compared with the "true DTMs". Figures 4.9, 4.10, and 4.11 show 64 by 64 overviews of the shape from shading results showing a favorable comparison between the DTM and the surface reconstruction. The reconstructions correspond to the three SAR images in Figure 4.8. Some apparent differences can be observed along drainage patterns in the surface, perhaps due to albedo variations or surface roughness variations.

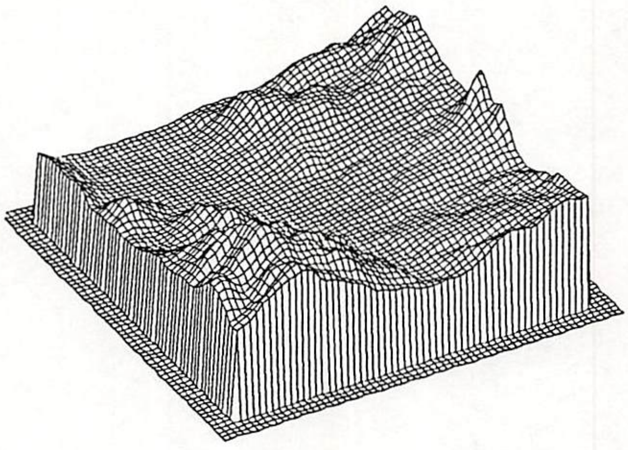
Figure 4.12 shows a full resolution window of one of the mountains from the DTM and its reconstruction from the third SAR image. Finally, Figure 4.13 provide 1-D cuts diagonally across the DTM and surface reconstruction obtained from the first SAR image. Table 4.4 provides RMS error for the surface reconstructions. First the standard deviation of the high resolution DTM is given



(A)

(B)

Figure 4.9: 64 by 64 sample overviews of the shape from shading results for the first SAR image in Figure 4.8. (A) DTM, (B) Surface reconstruction from SAR image and low resolution data.


(A)

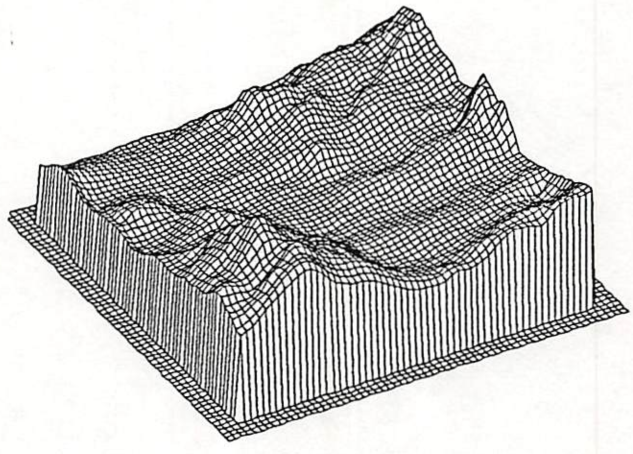
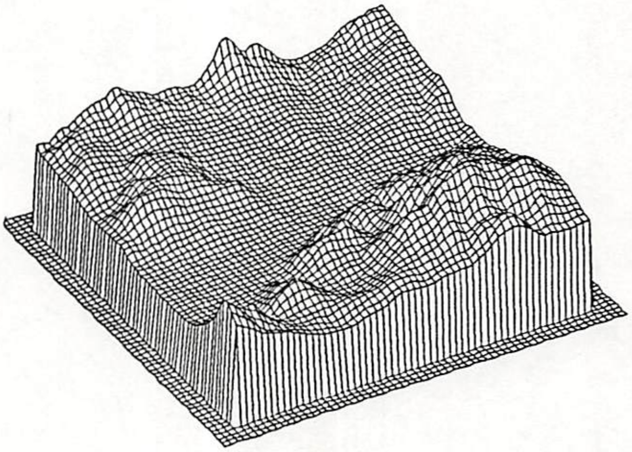
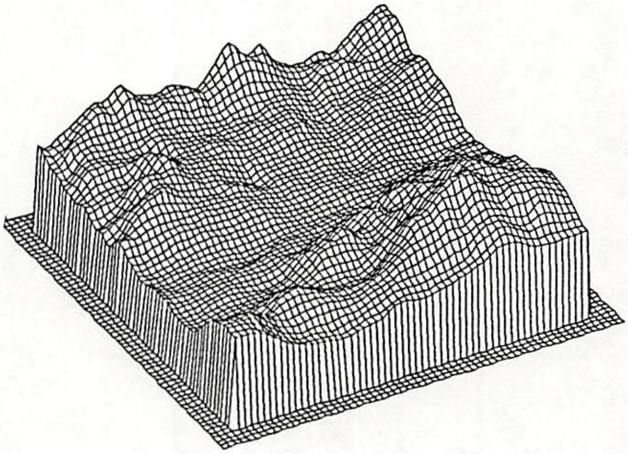

(B)

Figure 4.10: 64 by 64 sample overviews of the shape from shading results for the second SAR image in Figure 4.8. (A) DTM, (B) Surface reconstruction from SAR image and low resolution data.



(A)



(B)

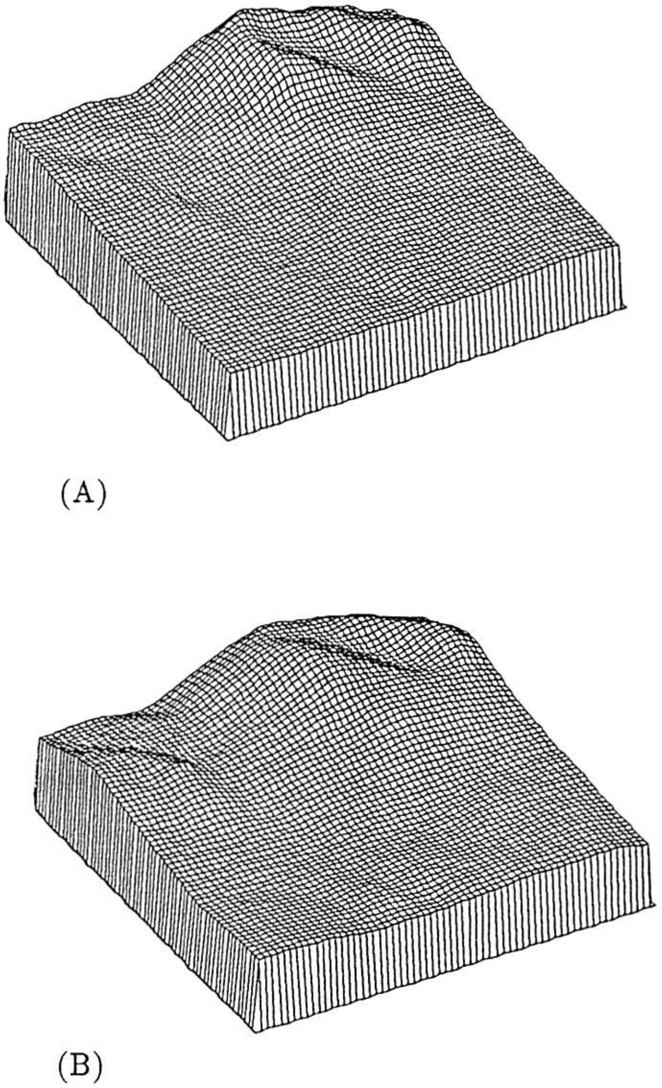
Figure 4.11: 64 by 64 sample overviews of the shape from shading results for the third SAR image in Figure 4.8. (A) DTM, (B) Surface reconstruction from SAR image and low resolution data.

for reference. Second, the standard deviation between the low resolution DTM and the high resolution DTM is provided as a basis of comparison. Third, the standard deviation of the surface reconstruction combining shape from shading and the low resolution DTM is given.

	SAR image		
	1	2	3
Standard deviation of high resolution DTM (meters)	84.3	84.0	80.0
Standard deviation of low resolution DTM (meters)	16.8	21.2	17.5
Standard deviation of surface reconstruction (meters)	14.4	17.6	14.1

Table 4.4: Surface height reconstruction error.

Table 4.4 shows that, in the experiments presented above, shape from shading provided a modest improvement over the RMS error of the low resolution surface data. Comparison of Figure 4.8 with 4.6 and the comparison provided by Figure 4.6 show that shape from shading provided intensity prediction capability and structural detail not available from the low resolution data.



(A)

(B)

Figure 4.12: Full resolution subarea view of the shape from shading results for the third SAR image in Figure 4.8. (A) DTM, (B) Surface reconstruction from SAR image and low resolution data.

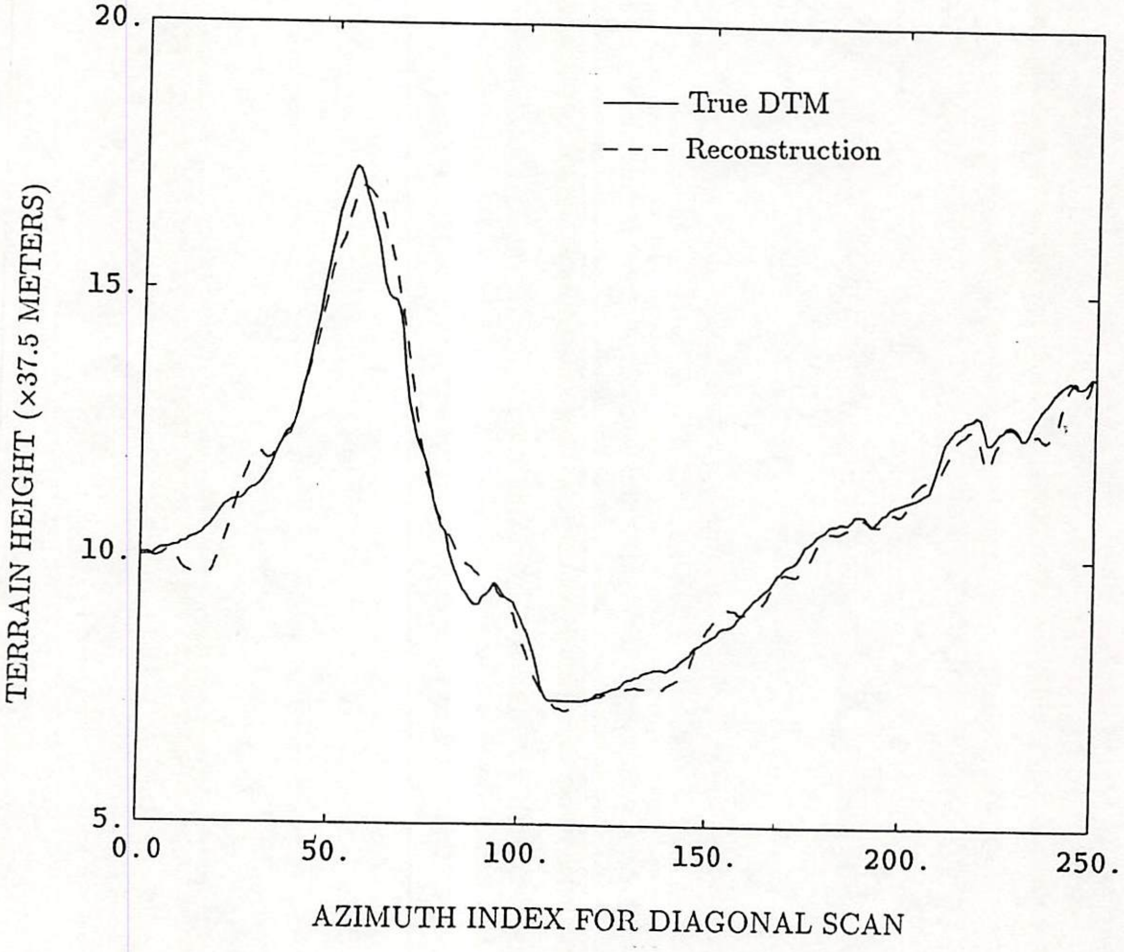


Figure 4.13: 1-D diagonal cuts across the DTM and surface reconstruction obtained from the first SAR image.

Chapter 5

A Computational Vision Approach to Image Matching

Scene matching is very important in both machine vision and remote sensing. We define the basic scene matching problem as follows: Given two images I_1 and I_2 of the same physical scene, find the correspondence between I_1 and I_2. *Correspondence* is defined as the coordinate transformation $\mathcal{T}$ that aligns the two images such that a similarity measure $\varrho(\mathcal{T}; I_1, I_2)$ is maximized. Classical "area correlation" fits into this framework with $\mathcal{T}$ strictly translation and ϱ the cross-correlation function between I_1 and I_2 as a function of their relative translation. A similarity measure (and some kind of continuity constraint) is needed also for symbolic matching of discrete features because of segmentation errors and the possibility of scene content changes from image to image. The problem of symbolic matching can therefore be considered as a generalization of this Chapter's concept where the similarity measure arises in a graph-matching proceedure.

$\mathcal{T}$ is needed for image registration which, in turn, is needed for any image processing or image analysis task using multiple images. If the sensor locations are known for both images then variations in $\mathcal{T}$ provide scene structure. This is known as stereopsis or stereogrammetry. If the location of one sensor is known relative to the scene then $\mathcal{T}$ determines the location of the second sensor relative to the scene. Navigation update is one application of that principle.

In this chapter a predictive approach is used to automatically register SAR images with DTMs, with other SAR images made from grossly different aspect angles, and with aerial photography.

5.1 Scene Matching as Prediction

Scene matching relies on some level of prediction; Otherwise there is nothing to match. Direct cross-correlation of image intensity relies on $I_1(\mathcal{T}(x,y))$ providing a good prediction of $I_2(x,y)$. If the images are made from similar sensors and shading differences are small then this is the most appropriate approach. If there is significant surface height variation, significant differences in illumination direction, and only weak albedo variation then cross-correlation of image intensity is not reliable.

Matching of edges and object boundaries relies on those features being predictable from one image to the next. This is the case when matching images of scenes with piecewise constant surface slopes, which map into piecewise constant image intensity. Then intensity discontinuities due to shading are relatively consistent between images. Boundary features are also appropriate for scenes dominated by abrupt albedo transitions that arise from abrupt transitions in the materials composing the surface. The actual value of the albedo may vary wildly as a function of wavelength, but the locations of albedo discontinuities often do not. Hence, boundaries are appropriate features for matching multi-spectral images of scenes dominated by transitions in surface materials. Of course, the correlation between albedo discontinuities for images close in wavelength, such as L-band versus X-band, should be much higher than for images from different parts of the spectrum, such as X-band versus infrared.

5.1.1 Matching Images to Scene Models

Feature-based approaches that match images to scene models have been considered for both SAR imagery [92,119] and imagery from visual and infrared sensors [16,131]. Unfortunately, when image intensity is not dominated by albedo discontinuities and slope discontinuities then segmentable features are not repeatable from image to image and some other approach is required.

Matching of images to scene models directly by image intensity prediction has a longer history and wider applicability. In the late 1940's a navigation system that correlates radar imagery with simulated imagery was developed for the United States Air Force and included into the MACE missile system [94]. The radar reference imagery was simulated by photographing a physical model of the terrain with albedo features painted in. The reference imagery was stored on film and correlated with the sensed radar imagery on board the missile using vacuum tube electronics. More recently, Horn and Bachman [72] registered Landsat images to DTMs by correlating real images with simulated images. The coordinate transformation $\mathcal{T}$ between the real image and simulated image was modeled by translation, rotation, and scaling. The registration error estimation algorithm selected values of those transformation parameters that maximized the normalized cross-correlation between the two images. This requires resampling one of the images into the new coordinate system defined by each candidate $\mathcal{T}_K$ and then calculating the correlation coefficient.

We have used a predictive approach to automatically register SAR images with DTMs, with other SAR images made from grossly different aspect angles, and with aerial photographs. We also demostrate automatic registration when the topography is not known a priori but reconstructed from the image using shape from shading techniques. Only scenes of continuous surfaces without detailed albedo variation are considered. Feature based scene matching typically fails on those types of scenes. We follow the approach taken by Horn and Bachman, extend it to SAR imagery, and demonstrate the application of surface reconstruction to match dissimilar images.

5.2 Registration Algorithm

The dominant unknowns between sensor coordinates and surface coordinates are modeled by 3-D translation, rotation and scaling. For orthographic projections from surface coordinates to image coordinates this corresponds to 2-D translation, rotation, and scaling between simulated and sensed images. Under these conditions, the transformation between two sensed images is the composition of two 2-D translation, rotation, and scaling operations; the result being a 2-D affine transformation.

For SAR images of everything but flat terrain, the projection from surface coordinates to image coordinates is not orthographic. Yet, as shown in Chapter 2, the projection from surface height above the slant plane to SAR image coordinates is approximately orthographic. Therefore, the incremental transformation required to register a real SAR image with a simulated SAR image from roughly the same orientation is closely approximated by simple 2-D translation, rotation, and scale. This transformation is given by

$$\mathcal{T}(r,y) \approx \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} s_r & 0 \\ o & s_y \end{pmatrix} \begin{pmatrix} r \\ y \end{pmatrix} + \begin{pmatrix} r_0 \\ y_0 \end{pmatrix} . \tag{5.1}$$

A similar principle holds for aerial photographs; The image is approximately an orthographic projection of the surface represented relative to a plane orthogonal to the line-of-sight (intead of a plane parallel to the line-of-sight as for SAR).

Matching is accomplished by selecting the parameters $(s_r, s_y, \varphi, r_0, y_0)$ from (5.1) which maximize the cross-correlation between the predicted and sensed images, i.e.

> Select $(s_r, s_y, \varphi, r_0, y_0)$ such that $\varrho\{I_2(r,y), I_1(\mathcal{T}(r,y))\}$ is maximized.

The most direct realization of this approach is an exhaustive search for the transformation parameters that maximize ϱ, an extremely computationally intensive procedure. Many different methods have been used to avoid the computational burden of exhaustive search, including multi-subarea correlation [46,105], multi-resolution hierarchical search [46,72,137], and Fourier transform methods [5,96].

We have decomposed above the five-dimensional search space into a 2-D search over translation followed by three 1-D searches over φ, s_r, and s_y, respectively. Significant error in any one of the parameters can cause errors in searching for the other parameters so that two or three iterations are sometimes required. To reduce computation the above five parameter search is first performed at reduced resolution and then repeated at higher resolutions—a multi-resolution hierarchical search. Finally, if rotational errors of more than about three degrees are present between the simulated SAR image and the sensed image then a new SAR image is simulated using the new slant plane orientation estimates and the entire search process is repeated. This minimizes shading errors and residual geometric distortions due to incorrect slant plane orientation. The overall algorithm structure is presented in Figure 5.1 and registration parameter search is outlined in Table 5.1.

- Start with initial parameter estimates $(\tilde{s}_r, \tilde{s}_y, \tilde{\varphi}, \tilde{r}_0, \tilde{y}_0)$.
- Search for translation.
    - For each candidate translation $(\tilde{\tilde{r}}_0, \tilde{\tilde{y}}_0)$ calculate $\varrho[I_1(r,y), I_2(\mathcal{T}(r,y))]$ for the parameter vector $(\tilde{s}_r, \tilde{s}_y, \tilde{\varphi}, \tilde{\tilde{r}}_0, \tilde{\tilde{y}}_0)$.
    - Interpolate and save the translation for maximum ϱ. Denote it as $(\tilde{r}_0, \tilde{y}_0)$.
    - Save the maximum value of ϱ for error screening.
- Search for rotation.
    - For each candidate rotation $\tilde{\tilde{\varphi}} = \tilde{\varphi} \pm k\Delta\varphi$ (with k selected to span the interval of uncertainty) calculate $\varrho[I_1(r,y), I_2(\mathcal{T}(r,y))]$ for the parameter vector $(\tilde{s}_r, \tilde{s}_y, \tilde{\tilde{\varphi}}, \tilde{r}_0, \tilde{y}_0)$.
    - Interpolate and save the rotation for maximum ϱ. Denote it as $\tilde{\varphi}$.
    - Save the maximum value of ϱ for error screening.
- Search for s_r (same as above but using $\tilde{\tilde{s}}_r$).
- Search for s_y (same as above but using $\tilde{\tilde{s}}_y$).
- Repeat at same resolution with $\sim \frac{1}{2}$ the search uncertainty for each parameter.
- Increase image resolution and repeat search.

Table 5.1: Registration parameter search outline.

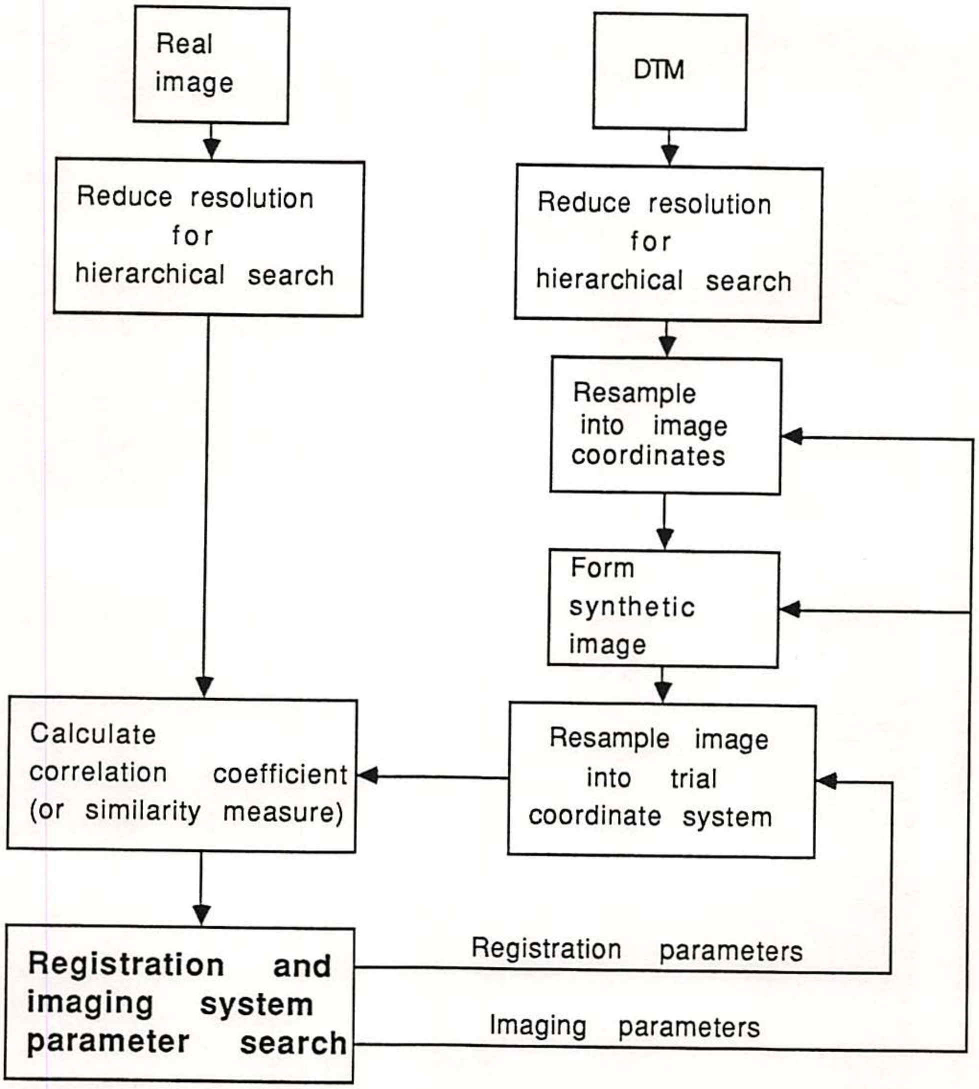


Figure 5.1: Algorithm structure for precise registration of images to DTMs.

A similar approach was used by Horn and Bachman [72] to register Landsat images to DTMs. This approach is now much more attractive, with commercially available image processing hardware designed for performing fast convolutions and fast coordinate transformations [155].

Step Size Selection

The maximum step size is selected for each search such that the entire region of uncertainty is covered without undersampling. The maximum step size for translational search is one pixel. In searching through rotation and scaling, one search step should not cause the outer portion of the image to move more than one pixel. We include a factor K to allow for slightly oversampled imagery, giving

$$\Delta\varphi = \tan^{-1}\left(\frac{2\sqrt{2}K}{\text{image size}}\right) \tag{5.2}$$

and

$$\Delta s_x = \Delta s_y = \frac{2K}{\text{image size}} \tag{5.3}$$

where K is between 1 and 2.

Error Detection and Recovery

Error detection is most easily accomplished by checking the results of multiple subarea matches for consistency as in [105]. An example of this is provided later. Although it is not possible to detect errors from a single match, it is possible to deduce a high error probability. Three indications have been considered; (1) The global maximum correlation is not at a local maximum, i.e. the maximum is on a border of the search region, (2) the global maximum correlation is not much larger than the second largest local maximum, and (3) the correlation peak is small. The first condition indicates a high likelihood that the true parameters are out of the search region. The second condition indicates a significant likelihood that noise or other errors have raised what should be a sidelobe to a level higher than the true mainlobe of the correlation function. That conclusion is supported by the following argument: The cross-correlation is modeled as a noisy version of the autocorrelation of I_1 after correction for the shape of the noise autocorrelation. If the signal autocorrelation is dominated by a peak and a sidelobe then there is a monotonic relationship between the cross-correlation peak to sidelobe ratio and probability that the cross-correlation mainlobe truly corresponds to the autocorrelation mainlobe. The third indicator, correlation peak value, provides an estimate of signal to noise ratio. All three reliability measures need to be reexamined for multiplicative noise.

Error recovery requires some means for providing alternate hypotheses and ranking them. Local maxima in the cross-correlation function provide good alternatives to the global maximum and consistency between multiple subareas provides a method (along with the actual cross-correlation value) for ranking the alternatives. This strategy was used to reject outliers in a multi-subarea correlation algorithm where local translational shifts of subareas distributed about the image are used to solve for an affine transformation model [46,105]. Our results indicate that the same approach is useful for providing error recovery in the five parameter search algorithm in this chapter. Instead of testing consistency between local translation estimates, consistency between local estimates of all five parameters $(s_r, s_y, \varphi, r_0, y_0)$ is tested (care must be taken to account for coupling of the local translation with φ, s_r, and s_y).

5.3 Registration of Images With DTMs

The SIR-B SAR images in the top row of Figure 4.8 were registered to a DTM of the same area using the algorithm described in Section 5.2. First, we started with uncertainties of about ±20 pixels in translation, ±10% in scale, and rotational errors of about 4 degrees. The starting uncertainty could be less than those values because of the spacecraft ephemeris data (or inertial navigation system data for airborne SAR) that is available, in principle. Nevertheless, it is important to be able to deal with higher uncertainties when navigation data is not conveniently obtained.

Figure 5.2a shows a plot of ϱ as a function of (r_0, y_0) with all other parameters fixed for the top left image in Figure 4.8 in combination with its simulated version. Figure 5.2b shows the same for the top right image in Figure 4.8. While the two images are formed with nearly orthogonal aspect angles relative to each other, both cross-correlation functions are more narrow in the range direction than in the azimuth direction. This is reasonable because, for constant albedo, the image intensity is a very strong function of the derivative of surface height in the range direction. The resulting image intensity tends to have a shorter correlation length in the range direction.

For each image, the registration parameter search converged well enough to provide sub-pixel accuracy across the entire image without the use of any error recovery mechanism. The DTM was then registered to the SAR image, with height measured relative to the slant plane. This precisely maps the surface topography onto the SAR image using the same resampling algorithm used for simulating the SAR image. It it also possible to perform the inverse mapping, i.e. map the SAR image onto the surface, using a similar algorithm. The accuracy was evaluated by flickering between the real SAR image and the synthetic image generated from the registered DTM on a COMTAL display. The visual effect of small errors in the registration parameters can be judged by slightly perturbing the registration parameters and then subjectively evaluating the jitter that appears when flickering between the two slighlty misregistered images. Translational misregistration of as little as 1/4 pixel are readily seen by the eye as jitter between the two images. For images of size 256 × 256, rotational errors of around 0.1 degrees were visually detectable, and coordinate scaling errors of about 0.2% were visually detectable. Since no such jitter was observed between automatically registered images, we conclude that the registration errors were less than the above values. Errors in surface height show up as shading changes and localized shifting of intensity features, as with stereo parallax measurement. Almost no shifting of local features were observed, with most of the intensity errors due to apparent albedo variations and random noise.

The aerial photograph in Figure 4.4 was also registered by the same method. In this case, albedo variations dominate the image intensity and the maximum correlation coefficient obtainable between the image and its simulation was $\varrho = 0.21$. This image was very difficult to match visually with the simulated image because of the albedo variation and, as a result, manual registration was

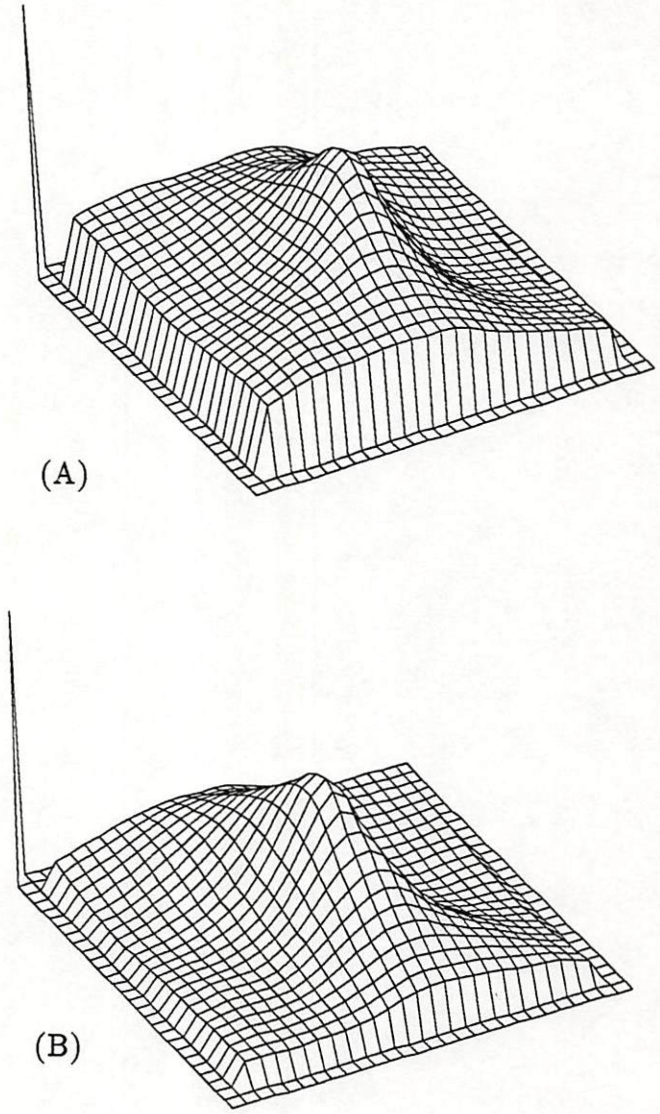


Figure 5.2: Correlation coefficient between real SAR image and synthesized image as a function of (r_0, y_0) with all other parameters fixed. (A) For the top left image in Figure 4.8. (B) For the top right image in Figure 4.8.

grossly in error. After the automatic registration was completed, the resulting rectification appeared to be within a pixel accuracy over the entire image except at mountain peaks, where uncompensated parallax errors appeared. This was evaluated by flickering between the two images on the same display and comparing subtle topography-induced intensity variations in the real aerial photograph with those of the simulated aerial photograph.

5.4 Registration of Dissimilar Images

Our primary motivation for performing the registrations described above was to compare shape from shading results with the DTM, estimate reflectance map parameters, and estimate albedo variation. Other applications could be supported by the same method. It is possible to register imagery from many sources by registering them all to a common database, e.g. the DTM.

For images made from very different geometries or from different sensors, both manual and automatic registration are difficult problems. If we have knowledge of surface structure then registration can be obtained using the method in Section 5.3. This is illustrated by attempting to correlate the top left image of Figure 4.8 directly with the top right image, a difficult problem due to the 81 degree aspect difference.

Figure 5.3a shows a plot of cross-correlation between the two SAR images as a function of (r_0, y_0) after correcting for rotation and scaling. The peak is flattened relative to Figure 5.2 with a peak correlation coefficient of only 0.38.

The registration algorithm latched onto a few distinct topographic features. The dominant topographic features have the largest stereo parallax so that, when they are registered, everything else is grossly misregistered. Alternately, the registration obtained by combining the two image–DTM registrations from Section 5.3 provides for a closer match on the more subtle features. There, most of the two scenes are in close alignment with the largest parallax errors occurring on large topographic features. It is also possible to correct for local foreshortening to get a very precise registration, although this was not implemented.

If we do not have a DTM, it is possible to construct one from the data as demonstrated in Chapter 4. This idea was tested by correlating the synthesized images in Figure 4.8 with the SAR images in the top row. That is, we attempted to register a SAR image to its prediction based on a surface *reconstruction.* The surface reconstruction was made from a real SAR image with a different aspect angle. In each case the resulting registration was visually correct to within one pixel overall but with some larger local registration errors appearing because of local errors in the surface reconstruction. These local errors were mostly in the downslope regions of the image used for reconstruction.

Figure 5.3b shows a plot of cross-correlation between the top left SAR image with it prediction from the top right image (the prediction is shown in the bottom left of Figure 4.8). The peak is slightly higher ($\varrho = 0.42$) and more prominent than that of Figure 5.3a.

We also tried registering the aerial photograph to the SAR image. As expected, direct correlation of the SAR image with the aerial photograph failed. The correlation coefficient was negative even for the best visually obtained offset. A synthesized image was then created from the surface reconstruction given the SAR image in Figure 4.8. The registration algorithm failed when applied to the entire frame. Registration was then attempted on three large randomly selected subareas. Two out of the three subareas were consistent with each other as determined by the multiple subarea correlation outlier rejection algorithm from [46]. The registration obtained by averaging the translation of the two consistent subareas resulted in a 3 pixel error when compared to visual estimates.

The experimental results from this Chapter are summarized in Table 5.2.

5.4.1 Local Matches

Although we have concentrated on registration of full image frames using one global transformation, local matches are needed for deriving stereo correspondences and removing high order misregistration. This has been done using both feature based techniques [110], and using intensity correlation [102,116,132]. Matching stereo pairs of SAR images presents the problem of "edge migration" due to shading changes between the two images [99]. The approach presented

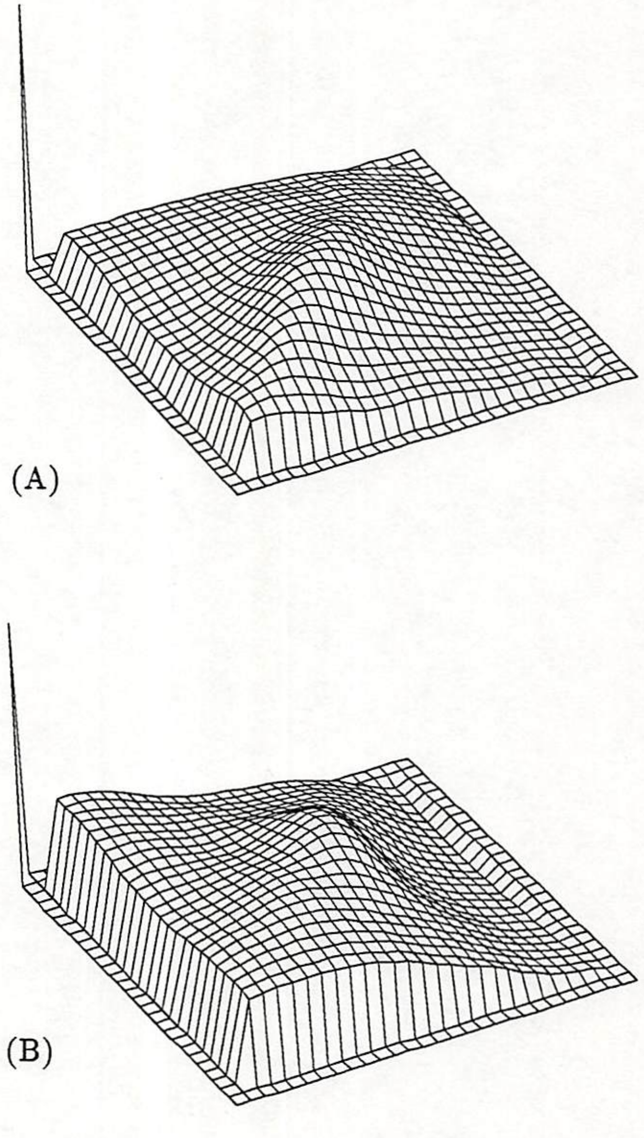


Figure 5.3: Cross-correlation of two SAR images of the same area but with near orthogonal aspect angles. (A) Plot of cross-correlation between the two SAR images after correcting for rotation and scaling, no correction for shading. (B) Cross-correlation after applying shading correction. Shading correction utilizes the surface reconstruction obtained by the shape from shading algorithm.

Images Correlated		Peak ϱ	Registration Results
SAR image 1	Prediction from DTM	.665	Global subpixel accuracy
SAR image 2	Prediction from DTM	.545	Global subpixel accuracy
SAR image 3	Prediction from DTM	.590	Global subpixel accuracy
SAR image 1	Prediction using SAR image 3	.422	$<$ 1 pixel average error Larger local misregistrations
SAR image 3	Prediction using SAR image 1	.418	$<$ 1 pixel average error Larger local misregistrations
SAR image 1	SAR image 3	.379	Two major peaks registered $\geq$ 4 pixel error elsewhere
Aerial Photo	Prediction from DTM	.214	$<$ 1 pixel error
Aerial Photo	SAR image 1	—	Unable to match: No positive correlation
Aerial Photo	Prediction from SAR image 1	.152	Single subarea: 11 pixel error Multiple subareas: $\approx$ 3 pixels error

Table 5.2: Summary of registration tests. The first three cases involved searching for translation, rotation, and scaling. All other cases only involved a translational search.

above, using single-image-based surface reconstruction to facilitate the matching may help to improve the accuracy of SAR stereo matching.

Chapter 6

Transformed-Gaussian Random Field Models

6.1 Introduction

The theory of Gaussian random fields and linear system are relatively mature and provide a good set of tools for model-based image processing. Yet image intensity seldom fits the models needed for applying those tools. Two simple methods have been applied for converting image intensity to Gaussian first order statistics. The first method compensates for a nonstationary image mean by subtracting a local intensity average from the image intensity [78],[95],[109]. This has also been extended to compensate for nonstationary variance [95]. The second method is to transform the intensity as close to Gaussian statistics as possible by passing it through an invertible point-nonlinearity [84]. We follow the second method, treating the issue of non-Gaussian statistics separately from the issue of nonstationarity. A combination of the two approaches appears to be appropriate in developing multi-resolution algorithms for image processing and image analysis.

This research provides models for spatial correlation which, together with the distribution model, are mathematically tractable and fit a variety of images. If the data is first transformed to Gaussian statistics with an invertible point nonlinearity then it can be modeled by extending the well-understood results from Gaussian random fields with linear spatial interaction. Lognormal random fields with multiplicative spatial interaction are a special case of the above transformed-Gaussian random fields which are of interest in radar and image processing. Other distributions such as the Weibull [56,143] and the K-distribution [81] have been considered for radar data, but do not lend themselves well to parameter estimation and synthesis. Transformed-Gaussian random fields are much more practical spatial models in this sense.

Several researchers have found that radar returns corrupted by speckle and also with speckle averaged out by noncoherent integration fit lognormal statistics [12,40,147,160]. This is based on examining histograms as well as formal goodness-of-fit tests such as the Kolmogorov-Smirnov (K-S) test [160]. Previously published results [12,40,47,147,160] suggest lognormal models for radar image intensity based on image statistics alone.

The use of lognormal models for radar and optical images has also been motivated by physical reasons. Lognormal models have been suggested for homomorphic filtering to separate multiplicative illumination and reflectivity components [153]. For imagery at visible and infrared wavelengths, spatial models which are linear in the logarithm of intensity have been suggested for use in filtering signal-dependent noise [89]. Radar returns can have very high dynamic range, up to 90 dB [113]. For this reason the logarithm of radar returns are often digitized and stored rather than the power itself. Even at visible wavelengths the logarithm of image intensity is often digitized rather than intensity itself due to dynamic range considerations. Another physical consideration is that image intensity is positive. An image can be thought of as an estimate of the power spectrum of aperture voltage and positivity of lognormal data is consistent with image intensity being a power signal.

Physiological models for the human visual system (HVS) provide further motivation for the use of lognormal stochastic models for imagery, when appropriate. The HVS response to the logarithm of image intensity is approximately linear [141,153]. With processing matched to the HVS receiver/observer characteristics it has been possible to improve visual quality over that of conventional techniques in both data compression [140] and image enhancement applications [153].

Many theoretical and phenomenological studies have been conducted investigating the first order density and average value of radar clutter under various conditions. Yet very little work has been published in modeling spatial correlation in radar imagery. Clearly, an independent and identically distributed (IID) model is too simplistic to be of much use.

We propose Gaussian autoregressive random field and noncausal Gaussian Markov random field (GMRF) models for the logarithm of radar image intensity in 2-D. This leads to the lognormal multiplicative AR (MAR) and lognormal multiplicative Markov random field (MMRF) models, discussed below, and provides a general but tractable statistical representation for lognormal radar imagery. Blanco [20] investigated the spatial correlation of 2-D radar clutter and mentioned the possibility of modeling the logarithm of clutter power as a Gaussian Markov process. However, no discrete formulation or method for synthesis were provided, and isotropic clutter was assumed. The isotropic assumption is restrictive for radar imagery. Suppose, for example, that the surface being illuminated is isotropic. Even then the resulting radar image is not, in general, isotropic due to the properties of image formation. This effect was illustrated in Figure 5.2. Lognormal time series models in one and two dimensions have been considered in the literature. Peebles [123] provided a method suitable for synthesizing correlated lognormal clutter given known correlations, however it was formulated for 1-D data only, no method was provided for parameter estimation, and no synthesis results were given. Kaufman et al. [89] suggested modeling the logarithm of image intensity as a unilateral AR process (the term "unilateral" is defined in Section 6.2). In that paper, techniques were developed for estimating the parameters of a unilateral AR process in the presence of noise for image restoration applications. Radar imagery, noncausal models, and model selection were not considered, and synthesis of textures visually similar to real images was not demonstrated. We have considered the problem of selecting random field models for radar imagery where model validity is judged by both the statistical goodness-of-fit of the model to empirical radar data and how closely the synthesized random fields visually resemble the original radar imagery.

Two lognormal random field models with multiplicative spatial interaction are discussed in Section 6.2. Markovianity, second order properties, and stability of lognormal random fields are then discussed. In Section 6.3, least squares (LS), and maximum likelihood (ML) estimates of model parameters are presented along with bounds on estimation error. In Section 6.4, a decision rule is formulated for selecting transformations to normality and model order for random fields which can be expressed as nonlinear transformations of Gaussian random fields. The objective is to decide, based upon image statistics alone, in what domain to treat the image as a Gaussian random field. Lognormal random fields are treated as a special case with the transformation to normality being logarithmic. Procedures for synthesizing unilateral and noncausal lognormal random fields are presented in Section 6.5. The approach is to estimate the random field parameters from empirical data, generate a Gaussian random field with appropriate parameters by passing white Gaussian pseudo-random numbers through a linear system and then exponentiating the Gaussian random field. In Section 6.6, the lognormal models discussed in this chapter are applied to radar image synthesis. Lognormal models were fit to some of the available SAR images and then synthetic textures were generated which closely resemble the original SAR images. Conclusions are presented in Section 6.7. Derivations of second order properties, and proofs of theorems on Markovianity can be found in [49].

6.2 Lognormal Multiplicative Random Field Models

Two models for 2-D lognormal random fields with multiplicative interaction are now examined. The first model is the unilateral multiplicative autoregressive random field (MAR), and the second model is the noncausal multiplicative Markov random field (MMRF). The MAR and MMRF models are natural extensions of Gaussian AR and GMRF models. Other 2-D multiplicative lognormal models follow as extensions of other Gaussian random field models such as

ARMA and ARIMA models [21,86]. We restrict our attention to the unilateral MAR and noncausal MMRF models, which are the most basic.

Case i: MAR. Let $v(s), \mathbf{s} \in \Omega$, be represented by the following white-noise-driven multiplicative system

$$v(\mathbf{s}) = \prod_{\mathbf{r} \in N} [v(\mathbf{s}+\mathbf{r})]^{\theta_{\mathbf{r}}} \nu(\mathbf{s}) \tag{6.1}$$

where $\Omega = \{0, 1, \ldots, M-1\} \times \{0, 1, \ldots, M-1\}$, N is the neighborset defining model support, $\nu(\mathbf{s})$ is a lognormal white noise process referred to as the driving process, and $\mathbf{s} = (m, n)$, a 2-D index.

The random field $v(\mathbf{s})$ is said to obey a lognormal MAR model if $w(\mathbf{s}) = \ln v(\mathbf{s})$ obeys the following Gaussian AR model with $\mu(\mathbf{s}) = \ln \nu(\mathbf{s})$

$$w(\mathbf{s}) = \sum_{\mathbf{r} \in N} \theta_{\mathbf{r}} w(\mathbf{s}+\mathbf{r}) + \mu(\mathbf{s}) \tag{6.2}$$

where $\mu(\mathbf{s})$ is zero mean white Gaussian noise. The nonzero mean case is addressed in [49]. The covariance of $\mu(\mathbf{s})$ is given by

$$C_\mu(\mathbf{r}) = \begin{cases} \sigma_\mu^2 & \mathbf{r} = (0,0) \\ 0 & \mathbf{r} \neq (0,0) \,. \end{cases} \tag{6.3}$$

For brevity, we only consider the MAR model in its unilateral form. The same basic definition and correlation structure applies for noncausal MAR models [28,85] but with the neighborhood of dependency including pixels in every direction. We have adopted the term unilateral instead of "causal", following Goodman and Ekstrom [58]. This is to avoid confusion with quarter-plane causality, where a pixel value depends only on its neighbors in one quadrant bounded by that pixel. The most general unilateral 2-D neighborset places dependency on past intensity values as the image is sequentially scanned, e.g. prior samples in a raster scanned image. This is referred to as nonsymmetric half-plane (NSHP) support. Quarter-plane causality is a special case of unilateral, or NSHP, model support.

The Gaussian AR model with unilateral dependency possesses the following unilateral Markov property;

$$p\left(w(\mathbf{s}) \mid w(\mathbf{s}_1)\ \forall \mathbf{s}_1 \in \Omega_{\mathbf{s},N}\right) = p\left(w(\mathbf{s}) \mid w(\mathbf{s}+\mathbf{r})\ \forall \mathbf{r} \in N\right) \tag{6.4}$$

where $\Omega_{\mathbf{s},N}$ is a definition of the "past" which depends on the structure of N as follows;

$$\begin{array}{ll} 1. & \mathbf{s} \notin \Omega_{\mathbf{s},N} \\ 2. & \mathbf{s}+\mathbf{r} \in \Omega_{\mathbf{s},N}\ \forall\, \mathbf{r} \in N \\ 3. & \mathbf{r} \in \Omega_{\mathbf{s},N} \rightarrow (\mathbf{r}+\mathbf{t}) \in \Omega_{\mathbf{s},N}\ \forall \mathbf{t} \in N, \text{ provided } \mathbf{r}+\mathbf{t} \neq \mathbf{s}\,. \end{array} \tag{6.5}$$

Using results from [49], the unilateral Markov property also holds for the log-normal MAR, $v(\mathbf{s})$.

In 1-D, any noncausal MAR model can be represented by an equivalent unilateral model. Thus, it suffices to consider only unilateral models in 1-D. In 2-D this is no longer true due to the lack of a general spectral factorization. Hence we also consider noncausal models and, in doing so, concentrate on the noncausal MMRF.

Case ii: MMRF. Suppose now that $w(\mathbf{s})$ obeys a noncausal Gaussian Markov random field (GMRF) model. Now $v(\mathbf{s})$ is represented by the multiplicative system in (6.1) and the underlying difference equation in (6.2) except that it is driven by a correlated noise with the following covariance [85]

$$C_\mu(\mathbf{r}) = \begin{cases} \sigma_\mu^2 & \mathbf{r} = (0,0) \\ -\theta_{\mathbf{r}}\sigma_\mu^2 & \mathbf{r} \in N \\ 0 & \text{otherwise}\ . \end{cases} \tag{6.6}$$

Then $v(\mathbf{s})$ is said to obey a noncausal MMRF model. We will refer to the noncausal MMRF simply as the MMRF.

The GMRF obeys the following Markov property with respect to noncausal neighborset N ;

$$p(w(\mathbf{s}) \mid w(\mathbf{s}_1),\ \forall \mathbf{s}_1 \neq \mathbf{s}) = p(w(\mathbf{s}) \mid w(\mathbf{s}+\mathbf{r}),\ \mathbf{r} \in N). \tag{6.7}$$

Here $\mathbf{s}$ represents the "future", $\{\mathbf{s}+\mathbf{r}\ \forall \mathbf{r} \in N\}$ denotes the "present", and all $\mathbf{s}_1 \notin \mathbf{s} \cup \{\mathbf{s}+\mathbf{r}, \mathbf{r} \in N\}$ denotes the "past". This is in contrast to the more familiar notion of a one-sided past, present, and future in (6.4). By Theorem 1, below, $v(\mathbf{s})$ also obeys the Markov property in (6.7).

The noncausal Gaussian AR random field and, therefore, the noncausal MAR also obey a noncausal Markov property but with respect to a larger set of neighbors than its model support.

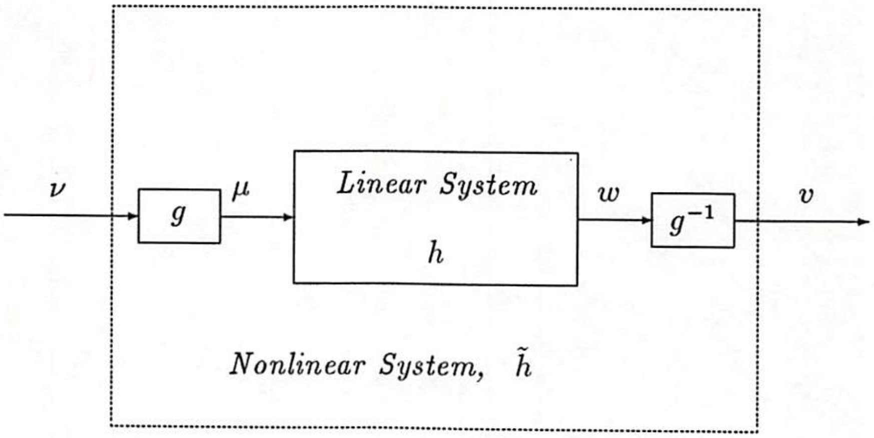


Figure 6.1: Nonlinear system obtained as nonlinear transformation of a linear system. Lognormal random field models are represented using $g(\nu) = ln(\nu)$, with ν lognormal.

6.2.1 Markovianity in Discrete Random Fields

The Markov property defined in (6.7) is sometimes referred to as local Markovianity. Global Markovianity is a more general definition used by Rosanov [136] and Woods [174]. For 2-D Gaussian random fields local Markovianity for all $\mathbf{s} \in \Omega$ implies global Markovianity [136]. This is also true of any random field representable as a Gaussian random field passed through an invertible point nonlinearity, and therefore lognormal random fields. Hence, we refer to (6.7) simply as the Markov property. This and other properties of discrete Markov random fields are stated as theorems in the following paragraph and proved in [49].

Consider random fields generated by the nonlinear system in Figure 6.1 with $v = g^{-1}(w)$, where g and g^{-1} are strictly invertible point nonlinearities. First, suppose that w is a Markov random field, not necessarily Gaussian.
Theorem 1 : w locally Markov implies v locally Markov.
Proof: See [49].

That is, an invertible point-transformation of a Markov random field is also a Markov random field. Theorem 1 does not assume any distribution for w or v. Now suppose that w is also Gaussian.

Theorem 2 : $v(\mathbf{s})$ locally Markov for all $\mathbf{s}$ and w is Gaussian implies v globally Markov .

Proof: See [49].

Therefore, in any transformed Gaussian random field, local Markovianity implies global Markovianity.

Now let $\mathcal{M}(v(\mathbf{s})|\gamma_{\mathbf{s}})$ be the conditional median of $v(\mathbf{s})$, where $\gamma_{\mathbf{s}} = \{v(\mathbf{s}_1), \forall \mathbf{s}_1 \neq \mathbf{s}\}$. The median is invariant under invertible point transformations and, for Gaussian random fields, the mean and median are equal. Hence, the conditional median is useful for deriving properties of transformed-Gaussian random fields from the well understood properties of Gaussian random fields.

Theorem 3 : If w is a GMRF then

$$\mathcal{M}(v(\mathbf{s})|\gamma_{\mathbf{s}}) = g^{-1}\left(\sum_{\mathbf{r}\in N} \theta_{\mathbf{r}} w(\mathbf{s}+\mathbf{r})\right). \tag{6.8}$$

Proof: See [49].

That is, the conditional median of v is obtained by transforming the conditional mean of w. If v is lognormal and w is Gaussian, then $g^{-1}(w) = e^{w}$, and $E(v|\gamma_{\mathbf{s}})/\mathcal{M}(v|\gamma_{\mathbf{s}}) = e^{\frac{1}{2}\sigma_{\mu}^2}$. This leads to the following two lemmas:

Lemma 1 : If v is lognormal and Markovian then

$$\mathcal{M}(v(\mathbf{s})|\gamma_{\mathbf{s}}) = \prod_{\mathbf{r}\in N} [v(\mathbf{s}+\mathbf{r})]^{\theta_{\mathbf{r}}}. \tag{6.9}$$

Lemma 2 : If v is lognormal and Markovian then

$$E(v(\mathbf{s})|\gamma_{\mathbf{s}}) = e^{\frac{1}{2}\sigma_{\mu}^2} \prod_{\mathbf{r}\in N} [v(\mathbf{s}+\mathbf{r})]^{\theta_{\mathbf{r}}}. \tag{6.10}$$

6.2.2 Second Order Properties for MAR and MMRF Models

The covariance and power spectrum of v is not easily obtained, in general. Fortunately, it often suffices to know only the covariance and power spectrum of w.

The covariance and power spectrum of w are discussed below and expressions for the autocorrelation, covariance, and variogram of v in terms of w are presented.

The MAR random field is represented by the multiplicative system in (6.1) driven by white lognormal noise, ν. Then w obeys a Gaussian AR model so that the power spectrum of the logarithm of the MAR random field is [87]

$$\mathcal{S}_w(\lambda) = \frac{\sigma_\mu^2}{|1 - \sum_{\mathbf{r} \in N} \theta_{\mathbf{r}} e^{\sqrt{-1}\ \lambda^T \mathbf{r}}|^2} \,. \tag{6.11}$$

The linear system which generates $w(\mathbf{s})$ is said to be bounded-input-bounded-output (BIBO) stable if bounded inputs guarantee bounded outputs. The necessary and sufficient condition for BIBO stability of w is [133]

$$\left\{1 - \sum_{(i,j) \in N} \theta_{(i,j)} Z_1^i Z_2^j \right\} \neq 0 \tag{6.12}$$

in the region defined by $\{(Z_1, Z_2) : \ |Z_1| \leq 1 \cap \ |Z_2| \leq 1\}$. It can easily be shown that the conditions for the lognormal random field $v(\mathbf{s})$ to be BIBO stable and positive are identical to the conditions for BIBO stability of $w(\mathbf{s})$. More precisely,

$$0 < \nu(\mathbf{s}) < \infty \quad \forall \mathbf{s} \in \Omega \quad \Longrightarrow \quad 0 < v(\mathbf{r}) < \infty \quad \forall \mathbf{r} \in \Omega \tag{6.13}$$

if and only if

$$|\, \mu(\mathbf{s}) \,| < \infty \quad \forall \mathbf{s} \in \Omega \Longrightarrow |\, w(\mathbf{r}) \,| < \infty \ \ \forall \mathbf{r} \in \Omega \,.$$

Here $\nu(\mathbf{s})$ and $v(\mathbf{r})$ are restricted to positive values since $0 < e^w < \infty$ for all $|\, w \,| < \infty$. Similar stability results hold for more general transformed Gaussian random fields when appropriate bounds are placed on ν and v.

The MMRF is represented as the output of the multiplicative system in (6.1) except that it is driven by a correlated noise. For MMRF models the covariance of the logarithm of the driving process is

$$C_\mu(\mathbf{r}) = \begin{cases} \sigma_\mu^2 & \mathbf{r} = (0,0) \\ -\theta_{\mathbf{r}} \sigma_\mu^2 & \mathbf{r} \in N \\ 0 & \text{otherwise} \,. \end{cases} \tag{6.14}$$

For the power spectrum $\mathcal{S}_w(\lambda)$ to be real requires $\theta_{\mathbf{r}} = \theta_{-\mathbf{r}}$ [18]. Therefore, we define a new neighborset N' for Markov random fields, which contains the half of N corresponding to the unique half of the parameter set. Specifically,

$$\mathbf{r} \in N' \Rightarrow -\mathbf{r} \notin N'; \;\; N = \{\mathbf{r} : \mathbf{r} \in N'\} \cup \{\mathbf{r} : -\mathbf{r} \in N'\}. \tag{6.15}$$

The resulting power spectrum for the logarithm of a MMRF is [85]

$$\mathcal{S}_w(\lambda) = \frac{\sigma_\mu^2}{1 - 2\sum_{\mathbf{r} \in N'} \theta_{\mathbf{r}} \cos(\frac{2\pi}{M}\lambda^T \mathbf{r})} \,. \tag{6.16}$$

Again, the conditions for stability of $v(\mathbf{s})$ are identical to the conditions for stability of $w(\mathbf{s})$. A sufficient condition for stability of $w(\mathbf{s})$ is that [85]

$$\left\{ 1 - 2 \sum_{\mathbf{r} \in N'} \theta_{\mathbf{r}} \cos\left(\frac{2\pi}{M}\lambda^T \mathbf{r}\right) \right\} > 0 \tag{6.17}$$

for all $\lambda \in \Omega$.

In [49] it is shown that the autocorrelation of v is proportional to the exponentiated covariance of w, i.e.

$$R_v(\mathbf{s}) = m_v^2 e^{C_w(\mathbf{s})} \,. \tag{6.18}$$

By Fourier transforming the Taylor series expansion of C_w we obtain a generalization of the Middleton expansion [1] for the power spectrum of v

$$\mathcal{S}_v(\lambda) = m_v^2 \sum_{k=0}^{\infty} \frac{1}{k!} B_k(\lambda) \tag{6.19}$$

where

$$B_k(\lambda) = \begin{cases} \mathcal{S}_w'(\lambda) \overset{k}{*} \mathcal{S}_w'(\lambda) & k \neq 0 \\ \\ \delta(\lambda) & k = 0 \end{cases} \tag{6.20}$$

and $\overset{k}{*}$ denotes k-fold convolution, and $\mathcal{S}_w'(\lambda) = FT\,[C_w(\mathbf{s})]$.

The asymptotic properties of $\mathcal{S}_v$ are summarized here. If $\sigma_w^2 \ll 1$ then v has approximately the same spectrum as w except for a DC term which ensures positivity of v. When $\sigma_w^2 \gg 1$ then the spectrum of v approaches a Gaussian shape with root-mean-square (RMS) bandwidth σ_w times that of $\mathcal{S}_w$, provided

$\mathcal{S}_w$ is well-behaved. For high contrast radar images this provides some theoretical justification for the practice of assuming a Gaussian shape for the power spectrum of image intensity. Thus, the relationship between the power spectra of w and v varies between adding a DC offset to broadening into a Gaussian shape with RMS bandwidth proportional to σ_w.

The covariance and variogram of v are derived as functions of C_w in [49]. The covariance of v is

$$C_v(\mathbf{s}) = m_v^2[e^{C_w(\mathbf{s})} - 1] \tag{6.21}$$

and the resulting normalized variogram is

$$\mathcal{V}_v(\mathbf{s}) = 2\left[1 - \frac{e^{C_w(\mathbf{s})} - 1}{e^{C_w(0)} - 1}\right] \tag{6.22}$$

The shape of $R_v, \mathcal{V}_v$, and $\mathcal{S}_v$ are highly dependent on σ_w, while the shape of R_w, $\mathcal{V}_w$, and $\mathcal{S}_w$ are independent of σ_w. Numerical techniques for calculating R_v, $\mathcal{V}_v$, and $\mathcal{S}_v$ are discussed in [49].

6.3 Parameter Estimation

For compactness, define the image intensity vector $\mathbf{v}$ as the lexicographic ordered or row-scanned vector

$$\begin{aligned} \mathbf{v} \;=\; & [v(0,0), v(0,1), \ldots, v(0, M-1), \\ & v(1,0), \ldots, v(1, M-1), \ldots, v(M-1, M-1)]\,. \end{aligned} \tag{6.23}$$

We define the parameter vector as

$$\underline{\theta} = col\,[\theta_{\mathbf{r}}, \mathbf{r} \in N]\ . \tag{6.24}$$

The covariance matrix of $\mathbf{v}$ is the $M^2 \times M^2$ matrix

$$\mathbf{C_V} = [C_v(\mathbf{s} - \mathbf{r}); \mathbf{s}, \mathbf{r} \in \Omega], \tag{6.25}$$

where Ω is defined in (6.1). Then, the difference equation for $\mathbf{w} = \ln(\mathbf{v})$ becomes

$$\mathbf{H}(\underline{\theta})\mathbf{w} = \mu + m_w\mathbf{1} \tag{6.26}$$

where a nonzero mean m_w is now included, the vector $\mathbf{1}$ is a $M^2 \times 1$ vector of $1's$, and $\mathrm{H}(\underline{\theta})$ is a $M^2 \times M^2$ matrix. Finally, to simplify notation let

$$\rho = \sigma_\mu^2 \,. \tag{6.27}$$

6.3.1 Least-Squares Estimates

For MAR models, the least-squares (LS) parameter estimates based on $\mathbf{w}$ are [85]

$$\hat{\underline{\theta}} = \left[\sum_{\mathbf{s}\in\Omega} \mathbf{z}(\mathbf{s})\mathbf{z}^T(\mathbf{s})\right]^{-1} \left[\sum_{\mathbf{s}\in\Omega} \mathbf{z}(\mathbf{s})\left(w(\mathbf{s}) - \hat{m}_w\right)\right] \tag{6.28}$$

with

$$\hat{\rho} = \frac{1}{M^2}\sum_{\mathbf{s}\in\Omega}\left(w(\mathbf{s}) - \hat{m}_w - \hat{\underline{\theta}}^T\mathbf{z}(\mathbf{s})\right)^2 \tag{6.29}$$

and

$$\hat{m}_w = \frac{1}{M^2}\sum_{\mathbf{s}\in\Omega} w(\mathbf{s}) \tag{6.30}$$

where

$$\mathbf{z}(\mathbf{s}) = col\left[w(\mathbf{s}+\mathbf{r}) - \hat{m}_w, \mathbf{r}\in N\right] \,. \tag{6.31}$$

It turns out that, for the unilateral lognormal MAR random field, the LS estimates above are identical to the maximum-likelihood (ML) estimates of $\underline{\theta}$ given lognormal observations v. Teekens [156] has shown that the joint-complete sufficient statistics for the parameters of a 1-D MAR process are given by the LS parameter estimates for the underlying Gaussian AR process. This is also true for unilateral 2-D MAR random fields.

For the MMRF and other noncausal models, the LS and ML estimates are not equal. The following LS estimates for MMRF parameters are consistent, though not efficient [85]:

$$\hat{\underline{\theta}} = \left[\sum_{\Omega} \mathbf{q}(\mathbf{s})\mathbf{q}^T(\mathbf{s})\right]^{-1} \left[\sum_{\Omega} \mathbf{q}(\mathbf{s})\left(w(\mathbf{s}) - \hat{m}_w\right)\right] \tag{6.32}$$

with

$$\hat{\rho} = \frac{1}{M^2}\sum_{\Omega}\left(w(\mathbf{s}) - \hat{m}_w - \hat{\underline{\theta}}^T\mathbf{q}(\mathbf{s})\right)^2 \tag{6.33}$$

where

$$\mathbf{q}(\mathbf{s}) = col\left[w(\mathbf{s}+\mathbf{r}) + w(\mathbf{s}-\mathbf{r}) - 2\hat{m}_w, \mathbf{r} \in N'\right] \tag{6.34}$$

and $\hat{m}_w$ is the sample mean as before.

6.3.2 Maximum-Likelihood Estimates

The joint probability density function for a lognormal random field $\mathbf{v}$ is

$$p(\mathbf{v}) = \left[(2\pi)^{\frac{M^2}{2}} |\mathbf{C_w}|^{\frac{1}{2}} \prod_{\Omega} v(\mathbf{s})\right]^{-1} \exp\left[-\frac{1}{2}(\ln(\mathbf{v}) - m_w\mathbf{1})^T \mathbf{C_w}^{-1} (\ln(\mathbf{v}) - m_w\mathbf{1})\right] . \tag{6.35}$$

To find the ML estimates we need the likelihood function for $\mathbf{v}$ conditioned on the model parameters. For unilateral MAR random fields, the determinant of the conditional covariance reduces to $|\mathbf{C}_\mu| = \rho^{M^2}$, hence the joint log-likelihood of the observations given the parameters reduces to

$$\begin{aligned} \ln p(\mathbf{v}|\theta, \rho, m_w) &= -\tfrac{M^2}{2} \ln 2\pi\rho - \tfrac{1}{2\rho} \textstyle\sum_\Omega \left(\ln(v(\mathbf{s})) - m_w - \underline{\theta}^T \mathbf{z}(\mathbf{s})\right)^2 \\ &\quad - \textstyle\sum_\Omega \ln v(\mathbf{s}) \end{aligned} \tag{6.36}$$

ignoring boundary conditions. By maximizing the above equation with respect to $\underline{\theta}$, ρ, and m_w, the ML estimates can be shown to be identical to the LS estimates in unilateral MAR models, and identical to the ML estimates for the parameters of the underlying Gaussian AR random field, $\mathbf{w}$.

The same is not true, however, for the noncausal MMRF model. The conditional log-likelihood function in this case is

$$\begin{aligned} \ln p(\mathbf{v}|\, \underline{\theta}, \rho, m_w) &= \tfrac{1}{2} \textstyle\int \ln\left(1 - 2\underline{\theta}^T \phi_\lambda\right) d\lambda - \tfrac{M^2}{2} \ln 2\pi\rho \\ &\quad - \tfrac{1}{2} (\ln(\mathbf{v}) - m_w\mathbf{1})^T \mathbf{H}(\theta) (\ln(\mathbf{v}) - m_w\mathbf{1}) \\ &\quad - \textstyle\sum_\Omega \ln v(\mathbf{s}) \end{aligned} \tag{6.37}$$

where

$$\phi_\lambda = col\left[\cos\left(\frac{2\pi}{M}\lambda^T\mathbf{r}\right), \mathbf{r} \in N'\right] . \tag{6.38}$$

Note that the log-likelihood is non-quadratic in $\underline{\theta}$, unlike for the unilateral model. Equation (6.37) is the same as the log-likelihood of the underlying GMRF except for the last term, which is not a function of any unknown parameters. Hence, the MLE for the parameters of $\mathbf{v}$ is obtained by maximizing the same criterion function as for $\mathbf{w}$. The ML estimates for the mean and variance terms of the GMRF are [18]

$$\tilde{m}_w = \hat{m}_w, \tag{6.39}$$

and

$$\tilde{\rho} = \frac{1}{M^2} \sum_{\Omega} \left(w(\mathbf{s}) - \tilde{m}_w - \tilde{\underline{\theta}}^T \mathbf{q}(\mathbf{s}) \right) \left(w(\mathbf{s}) - \tilde{m}_w \right) \tag{6.40}$$

$$= \left[\hat{C}_w(0) - 2 \sum_{N'} \tilde{\theta}_{\mathbf{r}} \hat{C}_w(\mathbf{r}) \right] \tag{6.41}$$

where $\hat{C}_w(\mathbf{r})$ are the sample correlations given by

$$\hat{C}_w(\mathbf{r}) = 1/M^2 \sum_{\Omega} \left(w(\mathbf{s}) - \tilde{m}_w \right) \left(w(\mathbf{s}+\mathbf{r}) - \tilde{m}_w \right) . \tag{6.42}$$

By substituting for $\tilde{m}_w$, and $\tilde{\rho}$ and dropping the terms which are independent of $\underline{\theta}$, the MLE of $\underline{\theta}$ is the value which maximizes

$$-\frac{M^2}{2} \ln \left(\hat{C}_w(0) - 2 \sum_{N'} \tilde{\theta}_{\mathbf{r}} \hat{C}_w(\mathbf{r}) \right) - \frac{1}{2} \int \ln \left(1 - 2 \tilde{\underline{\theta}}^T \phi_\lambda \right) d\lambda . \tag{6.43}$$

Applying recent results for Gaussian random fields [145], the above solution always exists, and is unique, asymptotically consistent and efficient.

6.3.3 Estimation Error

The ML parameter estimates are asymptotically efficient, i.e. their covariance for large sample sizes approaches the Cramér-Rao lower bound (CRLB). For unilateral models, the LS estimates are equivalent to the ML estimates and are therefore asymptotically efficient. For noncausal models, the LS and ML estimates are not equivalent and the LS estimates are consistent but not asymptotically efficient [85]. The CRLB is given below for the parameters of the unilateral MAR random field and the noncausal MMRF. The variance of the LS estimates for MMRF parameters is also presented.

The covariance matrix of any unbiased estimate of $\underline{\theta}$ is bounded below by the inverse of the Fisher information matrix,

$$\mathbf{C}_{\underline{\theta}} \geq \mathbf{J}^{-1}(\underline{\theta}) \tag{6.44}$$

where the elements of the Fisher information matrix are

$$J_{\mathbf{rs}}(\underline{\theta}) = -E\left\{\frac{\partial}{\partial\theta_{\mathbf{r}}}\frac{\partial}{\partial\theta_{\mathbf{s}}}\ln p_v(\mathbf{v}|\,\underline{\theta})\right\}, \tag{6.45}$$

and the inequality $\mathbf{A} \geq \mathbf{B}$ for matrices denotes $\mathbf{A} - \mathbf{B}$ is positive semi-definite. Evaluation of (6.44) with the likelihood function in (6.36) gives the following error bounds for the parameters of the unilateral MAR random field

$$\begin{array}{lcl} \mathbf{C}_{\hat{\underline{\theta}}} & \geq & \frac{1}{M^2}\rho\mathbf{P}^{-1} \\ \mathrm{var}\,(\hat{\rho}) & \geq & \frac{2}{M^2}\rho^2 \\ \mathrm{var}\,(\hat{m}_w) & \geq & \frac{1}{M^2}\rho\,(1-\sum_{r\in N}\theta_{\mathbf{r}})^{-2} \end{array} \tag{6.46}$$

where $\mathbf{P} = E[\mathbf{z}(\mathbf{s})\mathbf{z}^T(\mathbf{s})]$ and $\mathrm{var}(\cdot)$ denotes error variance. For the MMRF we evaluate (6.44) with the likelihood function in (6.37), giving the following inequalities

$$\begin{array}{lcl} \mathbf{C}_{\tilde{\underline{\theta}}} & \geq & \mathbf{J}^{-1}(\underline{\theta}) \\ \mathrm{var}\,(\tilde{\rho}) & \geq & \frac{2}{M^2}\rho^2 \\ \mathrm{var}\,(\tilde{m}_w) & \geq & \frac{1}{M^2}\rho\left(1-2\sum_{\mathbf{r}\in N'}\theta_{\mathbf{r}}\right)^{-1} \end{array} \tag{6.47}$$

where

$$J_{\mathbf{rs}}(\underline{\theta}) = 2\int\frac{\cos\left(\frac{2\pi}{M}\mathbf{r}\lambda^T\right)\cos\left(\frac{2\pi}{M}\mathbf{s}\lambda^T\right)}{\left[1-2\sum_{\mathbf{t}\in N'}\theta_{\mathbf{t}}\cos\left(\frac{2\pi}{M}\mathbf{t}\lambda^T\right)\right]^2}d\lambda\ . \tag{6.48}$$

If we approximate the block-Toeplitz matrix $\mathbf{C}_w$ by a block-circulant matrix or, equivalently, assume a toroidal lattice, then the integral above is replaced by a summation over $\lambda \in \Omega$.

The covariance of the LS estimate for the parameters of the MMRF is [85]

$$\mathbf{C}_{\hat{\underline{\theta}}} = \frac{1}{M^2}\left[\rho\mathbf{Q}^{-1} + 2\rho^2\mathbf{Q}^{-2} - \frac{\rho}{M^2}\mathbf{Q}^{-1}\sum\sum\nolimits_{(\mathbf{s}-\mathbf{r})\in N}\hat{\theta}_{\mathbf{s}-\mathbf{r}}\mathbf{T}_{\mathbf{r},\mathbf{s}}\mathbf{Q}^{-1}\right] \tag{6.49}$$

where

$$\begin{array}{lcl} \mathbf{Q} & = & E\left[\mathbf{q}(\mathbf{s})\mathbf{q}^T(\mathbf{s})\right] \\ \mathbf{T}_{\mathbf{r},\mathbf{s}} & = & E\left[\mathbf{q}(\mathbf{r})\mathbf{q}^T(\mathbf{s})\right] \end{array} \tag{6.50}$$

with $\mathbf{q}(\mathbf{s})$ as defined in (6.32).

6.4 Decision Rules for Model Selection

This approach for model selection follows that of Kashyap [84]. The basic idea is to select one of a finite set of model structures $\{C_k\}$ which best fits the observed image intensity $\mathbf{v}$. Each model structure includes a strictly monotonic transformation to normality $g_k(\cdot)$ and neighborset N_k. This is in contrast to traditional hypothesis testing approaches such as the K-S goodness-of-fit test where one particular model is tested against all alternatives. The traditional goodness-of-fit tests are valid only for IID data and are very difficult to generalize, even for very simple models of spatial interaction [166] while the approach developed below takes into account the spatial correlation of the data.

We follow a special case of the Bayesian approach for model selection. With the general Bayesian approach, the model structure C_k is selected such that

$$P\left(C_k|\mathbf{v}\right) > P\left(C_j|\mathbf{v}\right) \tag{6.51}$$

for all $C_j \neq C_k$. From Bayes rule

$$P\left(C_k|\mathbf{v}\right) = p\left(\mathbf{v}|C_k\right) P\left(C_k\right) / p\left(\mathbf{v}\right). \tag{6.52}$$

We drop $p(\mathbf{v})$ since it is independent of model class and therefore irrelevant. We also drop $P\left(C_k\right)$ since there is no reason to believe that the classes are not equally likely. Then it remains to solve for $p\left(\mathbf{v}|C_k\right)$. Regardless of any distribution assumptions

$$p\left(\mathbf{v}|C_k, \underline{\theta}_k, \rho_k\right) = p\left(\mathbf{w}_k|C_k, \underline{\theta}_k, \rho_k\right) |\partial \mathbf{w}_k / \partial \mathbf{v}|, \tag{6.53}$$

where ρ_k is defined in (6.27), $|\partial \mathbf{w}_k / \partial \mathbf{v}|$ is the Jacobian of the transformation between $\mathbf{w}$ and $\mathbf{v}$, and the matrix $[\partial \mathbf{w}_k / \partial \mathbf{v}]$ is diagonal. Therefore

$$p\left(\mathbf{v}|C_k, \underline{\theta}_k, \rho_k\right) = p\left(\mathbf{w}_k|C_k, \underline{\theta}_k, \rho_k\right) \prod_{\mathbf{s}\in\Omega} g_k^{'}(\mathbf{s}) \tag{6.54}$$

where

$$g_k^{'}(s) = \frac{dg_k(\alpha)}{d\alpha}|_{\alpha=v(\mathbf{s})} \quad . \tag{6.55}$$

Now suppose that $\mathbf{w}_k$ is Gaussian. Under this hypothesis, Kashyap and Chellappa [85] have derived an approximate expression for $p\left(\mathbf{w}|C_k\right)$ by asymptotically integrating over $\underline{\theta}_k$ and ρ_k with suitable priors. We obtain the test statistic

T_k by using this approximation to $\ln p(\mathbf{w}_k | C_k)$, including the Jacobian term $g_k^{'}(s)$, and dropping constant terms giving

$$T_k(\mathbf{v}) = \sum_{\mathbf{s} \in \Omega} \ln g_k^{'} + \int \ln \frac{1}{\tilde{\rho}_k} \mathcal{S}_{w_k} \left(\lambda, \underline{\tilde{\theta}}_k \right) d\lambda - \frac{M^2}{2} \ln \tilde{\rho}_k - \frac{n_k}{2} \ln M^2 \tag{6.56}$$

where $\tilde{\rho}$ and $\underline{\tilde{\theta}}_k$ are suitable parameter estimates obtained by fitting model C_k to the data, and n_k is the number of unique entries in $\underline{\theta}_k$. We then select model structure C_k which maximizes $T_k(\mathbf{v})$.

Under the hypothesis that $\mathbf{w}_k$ obeys a unilateral Gaussian AR model, the integral term in (6.56) vanishes. Therefore the test statistic becomes [47]

$$T_k(\mathbf{v}) = \sum_{\mathbf{s} \in \Omega} \ln g_k^{'} - \frac{M^2}{2} \ln \tilde{\rho}_k - \frac{n_k}{2} \ln(M^2). \tag{6.57}$$

The decision rule is to select the model which maximizes T_k in (6.57).

Under the hypothesis that $\mathbf{w}_k$ obeys a noncausal GMRF model, the test statistic in (6.57) reduces to

$$T_k(\mathbf{v}) = \sum_{\mathbf{s} \in \Omega} \ln g_k^{'} - \frac{M^2}{2} \ln \tilde{\rho}_k - \frac{n_k}{2} \ln(M^2) + \frac{1}{2} \int \ln \left(1 - 2\underline{\tilde{\theta}}_k^T \phi_{k\lambda} \right) d\lambda. \tag{6.58}$$

We then select the model class which maximizes (6.58). For all but very small sample sizes, the final term in (6.58) with the integral replaced by summation over $\lambda \in \Omega$ provides a reasonable approximation to the integral.

Example 1:

Suppose we want to compare the fit of two unilateral AR models with the same number of parameters n but different intensity transformations g_1 and g_2. Let hypothesis C_k denote that $w_k(\mathbf{s}) = g_k[v(\mathbf{s})]$ obeys a Gaussian AR model of order n . The resulting decision rule is

$$\left(\frac{M^2}{2} \right) \ln \frac{\tilde{\rho}_1}{\tilde{\rho}_2} \underset{C_1}{\overset{C_2}{\gtrless}} \sum_{\mathbf{s} \in \Omega} \ln g_1^{'} - \sum_{\mathbf{s} \in \Omega} \ln g_2^{'} \quad . \tag{6.59}$$

Example 2:

Suppose that we want to select either $C_1 =$ Gaussian AR model of order n, or $C_2 =$ lognormal MAR model of order n, both unilateral models. Then $g_1(v) = v$, and $g_2(v) = \ln v$. The decision rule from equation (6.59) becomes

$$\left(\frac{M^2}{2} \right) \ln \frac{\tilde{\rho}_1}{\tilde{\rho}_2} \underset{C_1}{\overset{C_2}{\gtrless}} \sum_{\mathbf{s} \in \Omega} \ln v(\mathbf{s}) \quad . \tag{6.60}$$

Note that, for the unilateral models considered in the two examples, we select the transformation resulting in the lowest residual variance after accounting for the Jacobian of g_k.

The usefulness of these decision rules in determining the appropriate model for real SAR images is illustrated by the experimental results in Section 6.6.

6.5 Synthesis of Lognormal Random Fields

Synthesis of lognormal data is accomplished by first synthesizing Gaussian data and then exponentiating it. This approach has been used for IID lognormal data [3], 1-D lognormal data with arbitrary autocorrelation [123], and unilateral MAR random fields [89,156]. A model based procedure for synthesis of 2-D lognormal random fields based on fitting multiplicative lognormal models to empirical data is given below.

Lognormal Random Field Synthesis Procedure:

(1) Take the logarithm of the sample data.

(2) Fit the appropriate Gaussian random field model to the logarithm using the methods in Section 6.3.

(3) Synthesize the Gaussian random field $\mathbf{w}$ using the methods described below.

(4) Exponentiate $\mathbf{w}$.

A nonzero mean is easily included by first synthesizing a zero-mean Gaussian random field and then adding m_w as represented in (6.26).

The methods for synthesizing unilateral AR and noncausal GMRF models in step (3) are outlined below. Further details can be found in [26].

6.5.1 Synthesis of Unilateral Gaussian Random Fields

Due to the causality of the model and the whiteness of $\mu(\mathbf{s})$, a unilateral AR random field can be generated directly with convolution on a finite lattice given a set of white-Gaussian pseudo-random numbers for $\mu(\mathbf{s})$ in (6.2). However, care must be taken so that boundary values obey the model. It is usually necessary to save boundary samples from the original data or to synthesize several start-up rows and columns which are discarded later.

6.5.2 Synthesis of Noncausal Gaussian Random Fields

Synthesis of noncausal random fields cannot be accomplished by direct convolution. In general, synthesizing a noncausal GMRF or AR random field requires inverting the $M^2 \times M^2$ matrix $\mathbf{H}(\underline{\theta})$. Since M can be quit large, this is impractical. Instead, the noise driven linear system h is implemented by Fourier techniques. Suppose that $\mathbf{w}$ obeys a GMRF or Gaussian AR model defined on a toroidal lattice. Then the covariance matrix of $\mathbf{w}$ is block-circulant and therefore exactly diagonalized by the DFT [28]. We can then directly calculate the eigenvalues of $\mathbf{C_w}$ as the power spectrum of $\mathbf{w}$ evaluated at discrete frequency values $\lambda \in \Omega$. In practice, approximating $\mathbf{H}(\underline{\theta})$ by a block-circulant matrix, i.e. assuming a toroidal lattice, affects only boundary values and is therefore a good approximation for all but very small data sets [97]. Another simplification with a toroidal lattice is that the integral is replaced by a summation over $\lambda \in \Omega$ in the likelihood equations (6.37) and (6.43), and the decision rule (6.58).

The procedure for synthesis of a noncausal GMRF on a toroidal lattice is then:

(1) Generate white Gaussian pseudo-random data $\mathcal{U}(\lambda)$.

(2) Calculate the power spectrum $\mathcal{S}_w(\lambda, \underline{\theta})$ for $\lambda \in \Omega$.

(3) Get the Fourier domain realization of w

$$\mathcal{W}(\lambda) = \mathcal{U}(\lambda)\sqrt{\mathcal{S}_w(\lambda, \underline{\theta})}.$$

(4) Inverse DFT $\mathcal{W}(\lambda, \underline{\theta})$ to get $w(\mathbf{s})$.

Experimental results for synthesis of lognormal random fields are given in Section 6.6.

6.6 Application to Radar Image Synthesis

Coherent radar imagery is corrupted by speckle, which is often modeled as multiplicative noise. Ignore thermal noise for now and model the observed square-law detected image I as

$$I(\mathbf{s}) = v(\mathbf{s})n(\mathbf{s})\ , \qquad (6.61)$$

where $v(\mathbf{s})$ is the noncoherent image proportional to the ensemble average radar backscatter for the object at the physical location corresponding to $\mathbf{s}$. Variation in $n(\mathbf{s})$ as a function of $\mathbf{s}$ is referred to as speckle, while pass-to-pass variation of n for a fixed position on the object is referred to as scintillation. For fully-developed speckle, $n(\mathbf{s})$ is Gamma distributed and independent of $v(\mathbf{s})$. Presence of thermal noise and sidelobe noise requires a three-parameter lognormal distribution.

Suppose that v is in fact lognormal, and n is Gamma distributed. It has been shown that, for sufficiently large σ_w, observed power I is approximately lognormal also [103]. Therefore, if the lognormal noncoherent image has sufficient contrast, then the speckled image is lognormal. For high contrast images we can then synthesize the speckled image $I(\mathbf{s})$ directly as a lognormal random field. For low contrast images or if pass-to-pass variations in the same scene are to be simulated then it is more appropriate to synthesize v and n separately and then multiply them to get I.

6.6.1 Decision Rule Results

SEASAT SAR imagery was tested for normality and lognormality. The data included imagery from (1) rural and suburban areas of the San Fernando Valley, California, (2) hills, mountains, and flat land in the Dominican Republic, (3) hilly land in Kentucky, (4) sea clutter during a hurricane. Two sets of data were created. The first was standard SEASAT imagery consisting of the intensity average of four looks [130]. The second set was at reduced resolution to average out speckle. Resolution reduction was performed on intensity rather than the logarithm of intensity in order to avoid biasing the outcome towards the lognormal case. Using the decision rule in (6.57) the fit of the nearest-neighbor unilateral AR models with $N = \{(0,-1),(-1,0)\}$ on twelve 64 by 64 pixel SAR images were compared for the Gaussian vs. lognormal cases. The results are summarized in Table 6.1. The lognormal AR model was preferred over the Gaussian AR model for each of the twelve SAR images tested. Eight images at visible wavelengths were also tested. Among them all but two preferred the Gaussian AR model over the lognormal model for intensity.

IMAGE TYPE	NUMBER TESTED	LOGNORMAL	GAUSSIAN
4-Look SAR	6	6	0
Spatially Averaged SAR	6	6	0
Brodatz Textures	5	1	4
Real World Photographs	3	1	2

Table 6.1: Summary of model selection tests; causal AR model with nearest-neighbor support, lognormal versus Gaussian.

The same set of SAR images was also tested for normality vs. lognormality using the K-S goodness-of-fit test. Since the K-S test is valid only for IID data, the images were undersampled by selecting every fourth pixel in both the vertical and horizontal directions, reducing the 64 by 64 images to 16 by 16.

The lognormal IID model was accepted at the 95% significance level for nine of the twelve SAR images, while the Gaussian IID model was accepted for only three of the twelve images. For one image, both the Gaussian and lognormal models were rejected by the K-S test, while the decision rule from (6.57) selected the lognormal model. A second difficulty which can arise is that, with smaller sample size, the power of the K-S test is reduced resulting in cases where neither model is rejected. The decision rule developed in this chapter yields a systematic decision in both the above cases while the K-S test does not.

6.6.2 Synthesis Results

Two model-based texture synthesis experiments are demonstrated here. First, a set of five different synthetic textures, shown in the left column of Figure 6.2 were generated. These were generated using noncausal lognormal MMRF models with arbitrarily selected parameters and neighborsets, and with each texture using the same set of pseudo-random numbers, $\{\mu(\mathbf{s})\}$. The parameter values are given in Table 6.2. Least-squares estimates of the parameters of each texture were then used in generating new textures with the same visual properties, shown in the right column of Figure 6.2. Use of different random number sequences does not change the visual properties of the textures but creates a new realization with similar, though not identical, features. This has been documented in [27].

In the second experiment, four textures taken from the reduced resolution SEASAT SAR data set were fitted with noncausal MMRF models and unilateral MAR models using the methods in Section 6.5. The results are shown in Figure 6.3. There the first column shows the original SAR images, the second column shows synthetic textures obtained by fitting noncausal MMRF models, the third column shows synthetic textures obtained by fitting unilateral MAR models. In each case, twelve parameters were estimated from the data in addition to the sample mean and driving process variance. Note in Figure 6.3 the striking similarity between the synthetic textures and the original SAR images. Further, note that four distinctly different original SAR images are shown. Inspection of the textures on a high quality display show somewhat better visual quality for the noncausal models. Also note that the Markov and AR models best represent the small features in textures, e.g. ranging from about 1 to 15 pixels in extent.

The synthesis results are of interest for two reasons. First, they further validate the lognormal MAR and MMRF models for radar imagery and radar clutter. Second, synthesis is a useful application in its own right for radar Monte Carlo simulations [123], image data compression , and simulation of radar images from cartographic data [69].

Variogram Plots

Variogram plots are used in image processing and time series analysis as an indication of the adequacy of a particular class of stochastic models. We have calculated the theoretical variograms given noncausal MMRF parameters estimated from radar images. These are plotted against empirical variograms obtained from the sample covariance of the original radar image and also the corresponding synthetic texture. In Figure 6.4 such a plot is shown for the third

Image	Parameter	True Value	Estimated Value
1	$\theta_{1,0}$	0.1600	0.1643
	$\theta_{0,1}$	0.1000	0.1091
	$\theta_{3,0}$	0.1200	0.1334
	$\theta_{0,3}$	−0.1400	−0.1423
	m_x	128.00	128.034
	ρ	300.00	293.481
2	$\theta_{1,0}$	0.1700	0.1829
	$\theta_{0,1}$	0.1700	0.1805
	$\theta_{3,0}$	−0.1500	−0.1564
	$\theta_{0,3}$	−0.1500	−0.1313
	m_x	128.00	128.025
	ρ	300.00	296.500
3	$\theta_{1,0}$	0.3000	0.3040
	$\theta_{0,1}$	0.1900	0.1943
	m_x	128.00	128.157
	ρ	300.00	293.665
4	$\theta_{1,0}$	0.4900	0.4925
	$\theta_{0,1}$	0.0020	0.0042
	m_x	128.00	128.169
	ρ	300.00	298.996
5	$\theta_{1,0}$	0.4000	0.4114
	$\theta_{0,1}$	−0.0940	−0.0840
	m_x	128.00	128.038
	ρ	300.00	295.560

Table 6.2: Model parameters and their least-squares estimates for the synthetic textures in Figure 6.2.

image from the top of Figure 6.3. Note the close fit of all three curves in each scan direction. Figure6.4.A shows that the model has captured the damped oscillation of the original image in the vertical direction, while also conforming to the much more highly damped behavior of the original in the horizontal direction shown in Figure 6.4.B. Similar results were obtained for other textures.

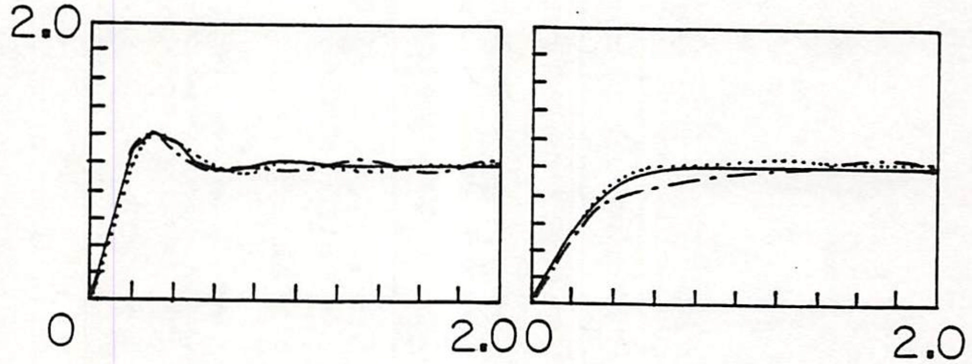
6.7 Conclusions

Lognormal random fields can be expressed as transformations on Gaussian random fields. Consequently, many theoretical results for lognormal random fields are obtained as extensions of the well-understood properties of Gaussian random fields. For example, using the formulations in this chapter, parameter estimation techniques and stability criteria for a lognormal random field are the same as for the underlying Gaussian random field.

The notion of a conditional mean and median are equivalent for Gaussian random fields but not for lognormal random fields. The conditional median is easily obtained for transformed Gaussian random fields by transforming the conditional median of the underlying Gaussian random field.

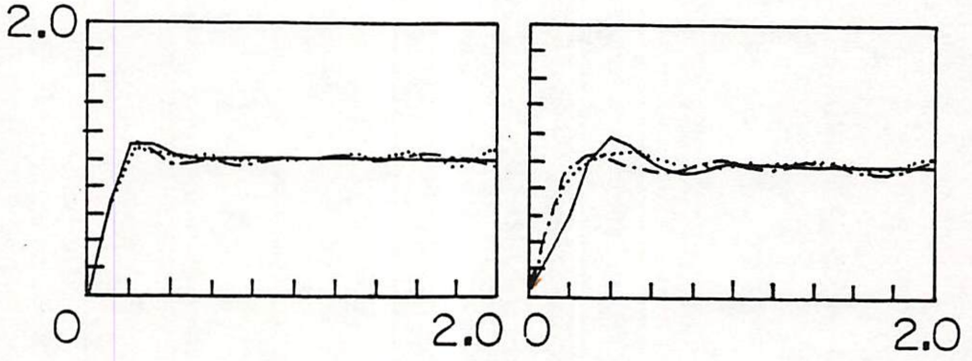
The decision rule developed in this chapter chooses from a set of competing models that which maximizes a test statistic derived from the observations $\mathbf{v}$. Suppose we hold the model order fixed. Then a transformation g_k is selected which results in the lowest variance for the transformed and whitened data after accounting for the Jacobian of g_k and the Jacobian of the whitening transformation. Experimental results using this decision rule showed lognormal models to be preferred over Gaussian models for the available radar imagery. The K-S test also showed the lognormal distribution to provide a good fit to available radar imagery.

The synthesis procedure for lognormal random fields is to first synthesize the underlying Gaussian random field and then exponentiate it. MAR and MMRF models were fit to sections of four distinctly different radar images and the resulting synthetic textures closely resembled the original radar images. The



A. Vertical Scan

B. Horizontal Scan



C. Down-Right Scan

D. Down-Left Scan

——— Theoretical Variogram Given MMRF Model Parameters

-—-— Empirical Variogram of Radar Image

········· Empirical Variogram of Synthetic MMRF

Figure 6.4: Theoretical and empirical variograms for lognormal MMRF and SAR image.

synthesis results serve to demonstrate that the lognormal MAR and MMRF models are practical and useful models for radar imagery.

Chapter 7

Conclusions and Topics for Future Research

7.1 Conclusions

A computational vision model for SAR imagery was presented and applied to surface topography estimation and image registration. SAR image based shape from shading algorithms were developed for surface topography estimation. The slant plane parameterization of surface height and the use of Fourier basis functions in the shape from shading formulation provide a conceptually simple formulation and allow a computationally efficient implementation. Use of auxiliary low resolution surface data as a constraint in shape from shading was introduced and successfully applied to SAR imagery. The low resolution data was also used for estimating reflectance map parameters. The utilization of image-derived topographic information for image registration and very precise alignment of SAR imagery with DTMs was demonstrated. A transformed-Gaussian random field model formulation was introduced. This model formulation was applied to texture synthesis and automatic model identification for SAR imagery and should be considered for other image processing and analysis applications.

7.2 Topics for Future Research

Several promising topics for future study have become apparent during the course of this research.

The SAR shape from shading algorithm may be useful for constructing topographic maps of the surface of Venus from SAR imagery and altimetry data collected during the Magellan mission. Several practical issues must be addressed first. The altimetry data must be registered with the SAR imagery. It is necessary to achieve registration accuracy comparable to the SAR image resolution, which is much finer than the altimetry resolution. While this is possible, it requires a very large integration area, perhaps extending thousands of SAR image pixels in range and azimuth. The orthographic projection approximation may not be valid over the entire area of interest and, therefore, some sort of compensation may be required in the processing. A radar cross-section model suitable for the conditions on Venus should be considered. With a suitable RCS model, the estimates of albedo and terrain surface roughness, obtained as a byproduct of shape from shading, may be accurate enough to be useful in other geological investigations. Future improvements in terrain surface estimation techniques, considered below, may also be useful for Magellan.

Additional research is needed for handling layover. The possibility of layover introduces both a spatial ambiguity and a radiometric ambiguity not considered in this thesis. The radiometric ambiguity is essentially the same as that which occurs in visual imagery at points of the surface that are normal to the incident illumination ($\alpha_i = 0$) [19]. For example, consider the partial sphere surfaces in Chapter 3. If the illumination was from directly overhead, it would be impossible to tell from image intensity alone whether the partial sphere extended above the plane or below the plane. In visual imagery (assuming the geometry of Figure 2.2) the geometric ambiguity occurring where the surface is normal to the illumination is an ambiguity in the z direction only—it is not observable in the image. For SAR imagery, a zero crossing of α_i means a change of the slant range ordering of features as ground range increases. It may be possible to locate and analyze layover conditions by detecting likely zero crossings of α_i from image intensity.

If shadows can be reliably detected then they provide information on surface slope at the shadow entry points and the relative height between shadow entry and exit points along lines of constant cross-range. Unfortunately, shadow boundaries are difficult to reliably identify in noisy imagery. Often there is only a small difference in mean intensity between a shadow and the dark but non-shadowed downslopes of a mountain. The uncertainty introduced by additive thermal and sidelobe noise tends to swamp those subtle intensity differences. Similarly, it is difficult to distinguish between shadows and areas of low backscatter due to low albedo or, for example, the specular reflection typical of paved roads. A predictive approach to shadow detection may reduce this difficulty, as discussed in Chapter 3.

Because vision is a highly underdetermined problem, fusion of information from different images and from different cues within a single image is becoming more important as the utilization of individual information sources matures. For example, it has been suggested that shape from shading complements stereogrammetry and radar interferometry [39,99]. We have already discussed the notion that stereo provides either low resolution or sparse surface structure information while shading provides either high resolution structure or fills the gaps between stereo matches.

An additional role for shading is to screen false stereo matches. It is often difficult to obtain reliable stereo matches, especially in the presence of speckle noise. One method for screening false matches is to require the image shading predicted by the stereoscopically-derived surface reconstruction to approximately fit the shading in the observed image. Further, the variance of the observed image intensity should exceed the variance of the stereo-predicted shading component by at least some threshold (the threshold can be predicted by speckle characteristics). Thus, the fusion of stereo and shading may enhance the reliability of automatically derived stereo matches.

Similarly, the precision of stereo matches can be improved by accounting for shading differences that occur between SAR stereo image pairs. Fullerton, et al. [54] reported a technique for approximate shading correction of opposite side SAR stereo image-pairs, where the illumination geometry difference causes dark areas in one image to be bright in the other. A simple correction procedure that maps high intensities into low intensities in one of the images was effective in that situation. Two additional correction methods, intensity prediction and intensity compensation, are discussed here. The intensity prediction method extracts high frequency surface information using shape from shading techniques and then forms a predicted image, as in Chapter 4. Stereo matching of, say, the first SAR image with its prediction given the second SAR image allows the computation of a residual parallax error. This provides a residual surface that, when added to the shape from shading results, approximates the desired surface estimate. The intensity compensation method starts with a stereo-derived surface reconstruction and then predicts a local shading ratio between the two images. Given the surface reconstruction, the aspect difference between the two images, knowledge of the parallax errors, and the reflectance map it is relatively easy to compute the shading compensation. If the additive noise terms are low, the correction is insensitive to albedo variations. After shading compensation the stereo matching procedure is repeated with, hopefully, greater achievable accuracy.

Photometric (or radiometric) stereo can also be extended to SAR imagery and applied simultaneously with traditional geometric stereo. The difficulty in applying radiometric stereo to SAR imagery is that it requires a precise registration between two or more images made with different illumination angles. When the illumination geometry affects the image coordinate system, as with SAR imagery, this registration requires knowledge of surface topography, the unknown that we seek. A similar requirement arises in the intensity prediction and correction methods for combining stereo and shading information. Perhaps the local registrations provided by the initial stereo correspondences will be sufficiently accurate to allow radiometric stereo, geometric stereo, and monocular shading cues to bootstrap each other. New methods for extracting stereo depthmaps from visual images [8,150], if re-evaluated for SAR imagery, may provide better methods for utilizing both radiometric and geometric information.

The shape from shading results presented in this thesis covered situations where stereo matching faces limitations: The image intensity is dominated by the shading of relatively smooth topographic features with minimal albedo variation. The image registration results dealt with the same cases. Matching of boundary features typically fails in scenes not dominated by discontinuities in surface slope and albedo so that an intensity prediction approach was used instead. More general scenes, containing a mixture of smooth and abrupt variations in albedo, surface slopes, and the reflectance map are often encountered in practice. This leads to one of the toughest machine vision problems, separation of the albedo and surface structure components under general conditions.

It has been shown that images from multiple spectral bands [41], from multiple frequency radar imagery [15], and multiple polarization radar imagery [163] can be helpful in that endeavor. More research is needed in the estimation and segmentation of albedo and reflectance map variations. Models are needed which are general enough to account for the effects of surface roughness, the dielectric properties of the surface, polarization diversity, and frequency diversity yet are tractable enough to be useful for image analysis. The stochastic approaches to vision problems considered by Marroquin [111], Geman and Graffigne [55], Derin and Elliott [37], and Chellappa [29] may help to provide a framework for algorithm development given a suitable model.

It appears that the full power of numerical methods for solving differential equations has not been fully utilized in shape from shading and other image analysis applications. Simchony and Chellappa [146] have recently made progress in this area that is potentially useful for SAR imagery. Fourier transform methods were applied to directly solve linearized versions of the type of differential equations encountered in shape from shading [146] and to enforce integrability. For nonlinear reflectance maps a few iterations of the direct algorithm are required. With the direct approach, it is possible to approximate the surface slopes consistent with a given intensity function, apply smoothness constraints, enforce fairly general boundary conditions, enforce integrability, and fuse shading information with stereo information in a unified algorithm. The direct shape from shading approach should be considered, in the future, for SAR imagery

Estimation theoretic approaches have been successful in other image analysis applications involving noise, underdetermined parameters, or both [25,29,111,181].

In the past, shape from shading has not been approached from an estimation theoretic standpoint. The experiments in Chapter 4 seem to indicate that improvements could be obtained by explicitly considering observation noise and randomly occurring modeling errors.

The stochastic models presented in Chapter 6 are useful for applications beyond the demonstrated texture synthesis. They are useful in providing parametric spectral estimates needed for data compression and speckle reduction filtering. It has been suggested that significant progress is still possible in SAR image data compression if speckle noise is considered as an integral part of the data compression problem [6]. This requires a statistical model for the speckle-free component of the SAR image.

The transformed-Gaussian Markov random field models help provide a tractable approach for SAR texture classification and segmentation through extensions of earlier work [25,29,55]. The detection of objects on a diffuse clutter background may similarly benefit from the random field models—Previous models considered for radar clutter have not adequately accounted for spatial correlation. Transformed-Gaussian models suggest simple extensions of constant false-alarm rate approaches developed for Gaussian image data [157]. Edge detection in SAR imagery is difficult due to speckle noise. Improvements over past edge detectors for SAR imagery may be attainable by re-examining noise tolerant edge detection approaches [181] in the context of a statistical model for the speckle-free component of SAR imagery.

In this thesis, general problem formulations and model formulations were considered and then special cases were examined. A modular approach was followed where concepts from the literature, previously considered for visual imagery were re-examined for SAR imagery using appropriate SAR image models. The same process should be useful for developing future SAR image analysis techniques and, perhaps, will be crucial for successful large scale multi-sensor fusion. Again, similar paradigms may be useful for processing data from many different sensors but the specific image models are likely to be very different and the knowledge base required for scene analysis may be very different. A thorough understanding of the similarities as well as the differences between each modality is needed for designing an appropriate suite of sensors and for timely development of effective processing systems.

References

[1] N. Abramson, "Bandwidth and Spectra of Phase- and Frequency- Modulated Waves", *IEEE Trans. Commun. Sys.*, Vol. C-11, pp. 407–414, December 1963.

[2] N. Ahuja and A. Rosenfeld, "Mosaic Models for Textures", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. PAMI-3, pp. 225–235, January 1981.

[3] J. Aitchison and J. A. C. Brown, *The Lognormal Distribution*, Cambridge Univ. Press, 1957.

[4] P. E. Anuta, "Digital Registration of Multispectral Video Imagery", *SPIE J.*, Vol. 7, pp. 168–175, September 1969.

[5] P. E. Anuta, "Spatial Registration of Multispectral and Multitemporal Digital Imagery Using FFT Techniques", *IEEE Trans. Geoscience Electronics*, Vol. GE-8, No. 4, pp. 353–368, October 1970.

[6] H. H. Arsenault, "Information Extraction from Images Degraded by Speckle", *International Geoscience and Remote Sensing Symposium*, Ann Arbor, MI, pp. 1317–1320, May 1987.

[7] E. Bahar and S. Chakrabarti, "Full-Wave Theory Applied to Computer Aided Graphics for Objects", *IEEE Computer Graphics and Applications*, Vol. 7, pp. 46–60, July 1987.

[8] S. T. Barnard, "A Stochastic Approach to Stereo Vision", *AIAA Fifth Int. Conf. Artificial Intell.*, Philadelphia, PA, pp. 676–680, August 11–15, 1986.

[9] D. E. Barrick, "Rough Surface Scattering Based on the Specular Point Theory", *IEEE Trans. Antennas Propagat.*, Vol. AP-16, pp. 838–850, July 1968.

[10] D. E. Barrick, "Relationship Between Slope Probability Density Function and the Physical Optics Integral in Rough Surface Scattering", *Proc. IEEE*, Vol. 56, pp. 1728–1729, October 1968.

[11] H. G. Barrow and J. M. Tenenbaum, "Computational Vision", *Proc. IEEE*, Vol. 69, pp. 572–595, May 1981.

[12] D. K. Barton, Ed., *Radars–Volume V, Radar Clutter*, Artech House, Mass., 1975.

[13] P. Beckmann and A. Spizzichino, *The Scattering of Electromagnetic Waves from Rough Surfaces*, Pergamon Press, 1963.

[14] P. Beckmann, "Scattering by Composite Rough Surfaces", *Proc. IEEE*, Vol. 53, pp. 1012–1015, August 1965.

[15] A.R. Benton and R.W. Newton, "The Utility of Multi-Frequency Radar in the Coastal Zone Environment", *IEEE Geosci. and Remote Sens. Soc. Newsletter*, Vol. 9, pp. 4–8, June 1985.

[16] J. E. Berry, "Three-dimensional Autonomous Scene Matching", *Proc. SPIE*, Vol. 238, pp. 138–145, August 1980.

[17] J. E. Besag and P. A. P. Moran, "On the Estimation and Testing of Spatial Interaction in Gaussian Lattice Processes", *Biometrika*, Vol. 62, pp. 555–562, 1975.

[18] J. E. Besag, "Spatial Interaction and the Statistical Analysis of Lattice Systems", *J.R. Statist, Soc. B*, Vol. 36, pp. 192–236, 1974.

[19] A. Blake, "On the Geometric Information Obtainable from Simultaneous Observation of Stereo Contour and Shading", In *Report CSR 205-86*, Department of Computer Science, University of Edinburgh, 1986.

[20] M. A. Blanco, "On Clutter Modeling and the Spectra of Two-Dimensional Gauss-Markov Random Signals", *IEEE Trans Aerosp. Elect. Syst.*, Vol. AES-18, pp. 228–234, March 1982.

[21] G. E. P. Box and G. M. Jenkins, *Time Series Analysis*, Holden-Day, Oakland, California, 1976.

[22] M. J. Brooks and B. K. P. Horn, "Shape and Source From Shading", *Proc. Int. Joint Conf. on Artificial Intell.*, Los Angeles, pp. 932–936, August 1985.

[23] A. R. Bruss, "The Image Irradiance Equation: Its Solution and Application", PhD Thesis, MIT, June 1981.

[24] D. B. Campbell, et al., "Venus: Identification of Banded Terrain in the Mountains of Ishtar Terra", *Science*, Vol. 221, pp. 644–647, August 1983.

[25] S. Chatterjee and R. Chellappa, "Model Based Texture Segmentation", *20th Asilomar Conf. on Signals, Systems and Computers*, Monterey, CA, November 1986.

[26] R. Chellappa and R. L. Kashyap, "Synthetic Generation and Estimation in Random Field Models of Images", *IEEE Conference on Pattern Recognition and Image Processing*, Dallas, TX, 1981.

[27] R. Chellappa, "Fitting Markov Random Field Models to Images", Tech. Report, TR-994, Computer Science Ctr., Univ. of Maryland, College Park, MD, January 1981.

[28] R. Chellappa and R. L. Kashyap, "Digital Image Restoration Using Spatial Interaction Models", *IEEE Trans. Acoust., Speech, Signal Proc.*, Vol. ASSP-30, pp. 461–472, June 1982.

[29] R. Chellappa, "Model Based Approaches for Some Image Understanding Problems", *Proc. SPIE*, Vol. 758, pp. 39–49, January 1987.

[30] D. S. Chen and J. P. Allebach, "Analysis of Error in Reconstruction of Two-Dimensional Signals from Irregularly Spaced Samples", *IEEE Trans. Acoust., Speech, Signal Processing*, Vol. ASSP-35, pp. 173–180, February 1987.

[31] R. E. Clapp, "A Theoretical and Experimental Study of Radar Ground Return", MIT Radiation Laboratory, Tech. Rept. 1024, 1946.

[32] R. L. Cosgriff, W. H. Peake, and R. C. Taylor, *Terrain Scattering Properties for Sensor System Design*, Tech. Report, Engineering Experiment Station, Ohio State University, Columbus, 1960.

[33] C. Cox and W. Munk, "Statistics of the Sea Surface Derived From Sun Glitter", J. Marine Research, Vol. 13, pp. 198–227, November 1954.

[34] C. Cox and W. Munk, "Measurement of the Roughness of the Sea Surface from Photographs of the Sun's Glitter", J. Opt. Soc. Amer., Vol. 44, pp. 838–850, November 1954.

[35] J. C. Curlander, "A Post-Processing System for Automated Rectification and Registration of Spaceborne SAR Imagery", *Int. J. Remote Sensing* to Appear, 1987.

[36] E. Cutting, J.H. Kwok, and S.N. Mohan, "The Venus Radar Mapper (VRM) Mission", *AIAA 22nd Aerospace Sciences Meeting*, Reno, Nevada, January 1984.

[37] 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, pp. 39–55, 1987.

[38] H. Derin, et al., "Unsupervised Segmentation of Speckled Images", *International Geoscience and Remote Sensing Symposium*, Ann Arbor, MI, pp. 1349–1356, May 1987.

[39] G. Domik, "Methods and Application of Surface Shape Reconstruction from Multiple SAR Images", *International Geoscience and Remote Sensing Symposium*, Ann Arbor, MI, pp. 1551–1555, May 1987.

[40] T.H. Einstein, "Effect of Frequency Averaging on Estimation of Clutter Statistics Used in Setting CFAR Detection Thresholds", MIT Lincoln Lab. TT-60, AD A131947, November 1982.

[41] P. T. Eliason, et al., "Extraction of Topgraphic and Spectral Albedo Information from Multispectral Images", *Photogrammetric Eng. and Remote Sensing*, Vol. 47, pp. 1571–1579, 1981.

[42] A. E. Ennos, "Speckle Interferometry", in *Laser Speckle and Related Phenomena*, J. C. Dainty, Ed., Berlin: Springer-Verlag, 1975.

[43] J. V. Evans and G. H. Pettengill, "The Scattering Behavior of the Moon at Wavelengths 3.6, 68, and 784 Centimeters", *J. Geophysical Res.*, Vol. 68, 1963.

[44] D. Fischer, et al., "On Fourier-Toeplitz Methods for Separable Elliptic Problems", *Math. of Computation*, Vol. 28, No. 26, pp. 349–368, April 1974.

[45] P. G. Ford and G. H. Pettengill, "Venus: Global Surface Emissivity", *Science*, Vol. 220, pp. 1379–1381, June 1983.

[46] R. T. Frankot, "SAR Image Registration by Multi- Resolution Correlation", *SPIE Proceedings*, Vol. 432, pp. 195–203, August 1983.

[47] R.T. Frankot and R. Chellappa, "A Decision Rule for the Choice of Gaussian or Lognormal Models for Images", *IEEE Conf. Comput. Vision and Patt. Recog.*, San Francisco, CA, 1985.

[48] R. T. Frankot and R. Chellappa, "An Improved Algorithm for the Shape from Shading Problem", *Conference on Systems and Signal Processing*, Indian Institute of Science, Bangalore, India, December 11–13, 1986.

[49] R. T. Frankot and R. Chellappa, "SAR Image Synthesis Using Lognormal Random Field Models", *IEEE Trans. Geosci. Remote Sensing*, Vol. GE-25, pp. 195–207, March 1987.

[50] R. T. Frankot and R. Chellappa, "Application of a Shape from Shading Technique to Synthetic Aperture Radar Imagery", *International Geoscience and Remote Sensing Symposium*, Ann Arbor, MI, pp. 1323–1329, May 1987.

[51] R. T. Frankot and R. Chellappa, "A Method for Enforcing Integrability in Shape from Shading Algorithms", *International Conference on Comp. Vision*, London, England, June, 1987. Also accepted for publication in *IEEE Trans. Pattern Analysis and Machine Intelligence.*

[52] V.S. Frost, et al, "A model For Radar Images and its Application to Adaptive Digital Filtering", *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. PAMI-4, pp. 157–166, March 1982.

[53] V. S. Frost, K. S. Shanmugan, and J. C. Holtzman, "The Influence of Sensor and Flight Parameters on Texture in Radar Images", *IEEE Trans. Geosci. and Remote Sensing*, Vol. GE-22, pp. 440–448, Sept. 1984.

[54] J. K. Fullerton, F. Leberl, R. E. Marque, "Opposite-Side SAR Image Processing for Stereo Viewing", *Photogrammetric Engineering and Remote Sensing*, Vol. 52, pp. 1487–1498, September 1986.

[55] S. Geman and C. Graffigne, "Markov Random Field Models and their Applications to Computer Vision", *Proceeedings Int. Congress of Mathematicians 1986*, Ed. A.M. Gleason, Am. Math. Soc., Providence, 1987.

[56] G.B. Goldstein, "False-Alarm Regulation in Lognormal and Weibull Clutter", *IEEE Trans. Aerosp. Elect. Syst.*, Vol. AES-9, pp. 84–92, January 1973.

[57] H. Goldstein, "Frequency Dependence of the Properties of Sea Echo", *Phys. Rev.*, Vol. 70, pp. 938–946, December 1946.

[58] D. M. Goodman and M. P. Ekstrom, "Multidimensional Spectral Factorization and Unilateral Autoregressive Models", *IEEE Trans. Automat. Contr.*, Vol. AC-25, pp. 258–262, April 1980.

[59] J. W. Goodman, *Fourier Optics*, McGraw-Hill, San Francisco, 1968.

[60] J. W. Goodman, *Statistical Optics*, John Wiley & Sons, New York, 1986.

[61] L. C. Graham, "Synthetic Interferometer Radar for Topographic Mapping", *Proc. IEEE*, Vol. 64, pp. 763–768, June 1974.

[62] W. E. L. Grimson, *From Images to Surfaces*, MIT Press, Cambridge, Mass., 1981.

[63] T. Hagfors, "Backscattering from an Undulating Surface with Applications to Radar Returns from the Moon", *J. Geophys. Res.*, Vol. 69, pp. 3779–3784, 1964.

[64] R.R. Hansen, Jr., and R. Chellappa, "2-D Spectrum Estimation for Imperfectly Observed Lattice Data", *Proc. International Conf. on Acoustics, Speech, and Signal Processing*, Dallas, TX, pp. 1605–1608, April 1987.

[65] R.M. Haralick and S. Wang, "Relative Elevation Determination from LANDSAT Imagery", *IEEE COMPCON*, Arlington, VA, pp. 284–292, Fall 1983.

[66] R.M. Haralick, J.B. Campbell, S. Wang, "Automatic Inference of Elevation and Drainage Models from a Satellite Image", *Proc. IEEE*, Vol. 73, pp. 1040–1053, June 1985.

[67] R. O. Harger, "Sea Surface Height Estimator Using Synthetic Aperture Radar Complex Imagery", *IEEE J. Oceanic Eng.*, Vol. OE-8, No. 2, pp. 71–78, April 1983.

[68] P. S. Heckbert, "Survey of Texture Mapping", *IEEE Computer Graphics and Applications*, Vol. 6, pp. 56–67, November 1986.

[69] J. C. Holtzman et al., "Radar Image Simulation ", *IEEE Trans. Geosci. Remote Sens.*, Vol. GE-16, pp. 296–303, October 1978.

[70] B.K.P. Horn, "Obtaining Shape From Shading Information", in: P.H. Winston, Ed., *The Psychology of Machine Vision*, McGraw-Hill, NY, 1975, pp. 115–155.

[71] B.K.P. Horn, "Understanding Image Intensities", *Artificial Intelligence*, Vol. 8, pp. 201–231, April 1977.

[72] B.K.P. Horn and B.L. Bachman, "Using Synthetic Images to Register Real Images With Surface Models", *Comm. ACM*, Vol. 21, pp. 914–924, November 1978.

[73] B.K.P. Horn "Hill-Shading and the Reflectance Map", *DARPA Image Understanding Workshop*, Menlo Park, CA, pp. 79–120, April 1979.

[74] B.K.P. Horn and M.J. Brooks, "The Variational Approach to Shape From Shading", *Comput. Vision Graphics Image Process.*, Vol. 33, pp. 174–208, February 1986.

[75] B.K.P. Horn and B.G. Schunk, "Determining Optical Flow", *Artificial Intelligence*, Vol. 17, pp. 185–203, August 1981.

[76] B.K.P. Horn and R.W. Sjoberg, "Calculating the Reflectance Map", *Applied Optics*, Vol. 18, pp. 1770–1779, June 1979.

[77] *Magellan Project Radar System Analysis Methods and Performance Estimates*, Hughes Aircraft Company, DRD No. SE011, February 1986.

[78] B. R. Hunt and T. M. Cannon, "Nonstationary Assumptions for Gaussian Models of Images", *IEEE Trans. Systems, Man and Cybernetics*, pp. 876–882, December 1976.

[79] K. Ikeuchi and B.K.P. Horn, "Numerical Shape from Shading and Occluding Boundaries", *Artificial Intelligence*, Vol. 17, pp. 141–184, August 1981.

[80] V. Isham, "An Introduction to Spatial Point Processes and Markov Random Fields", *International Statistical Review*, Vol. 49, pp. 21–43, 1981.

[81] E. Jakeman and P. N. Pusey, "A Model for Non-Rayleigh Sea Echo", *IEEE Trans. Antennas and Propag.*, Vol. AP-24, pp. 806–814, November 1976.

[82] J. K. Jao, "Amplitude Distribution of Composite Radar Clutter and the K-Distribution", *IEEE Trans. Antennas and Propag.*, Vol. AP-32, pp. 1049–1062, October 1984.

[83] H. Jinchi, T. Simchony, and R. Chellappa, "Stochastic Relaxation for MAP Restoration of Gray Level Images with Multiplicative Noise", *Proc. International Conf. on Acoustics, Speech, and Signal Processing*, Dallas, TX, pp. 1236–1239, April 1987.

[84] R.L. Kashyap, "A Bayesian Comparison of Different Classes of Dynamic Models Using Empirical Data", *IEEE Trans. Automat. Contr.*, Vol. AC-22, pp. 715–727, October 1977.

[85] R.L. Kashyap and R. Chellappa, "Estimation and Choice of Neighbors in Spatial-Interaction Models of Images", *IEEE Trans. Inform. Theory*, Vol. IT-29, pp. 60–72, January 1983.

[86] R.L. Kashyap and P.M. Lapsa, " Synthesis and Estimation Using Long-Correlation Models", *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. PAMI-6, pp. 800–809, November 1984.

[87] R. L. Kashyap and A. R. Rao, *Dynamic Stochastic Models From Empirical Data*, Academic Press, 1976.

[88] M. Katzin, "On the Mechanisms of Radar Sea Clutter", *Proc. IRE*, Vol. 45, pp. 44–54, Jan. 1957.

[89] H. Kaufman, J.W. Woods, S. Dravida, and A.M. Tekalp, " Estimation and Identification of Two-Dimensional Images", *IEEE Trans. Automat. Contr.*, Vol. AC-28, pp. 745–756, July 1983.

[90] V. H. Kaupp, "Radar Image Simulation", D.Engr. Thesis, University of Kansas, 1979.

[91] W. Keydel, "Application and Experimental Verification of an Empirical Backscattering Cross Section Model for the Earth's Surface", *IEEE Trans. Geosci. Remote Sensing*, Vol. GE-20, pp. 67–71, January 1982.

[92] G. K. Kiremidjian, "Registration of a Synthetic Aperture Radar Reconnaissance Image with a Map Reference Database", *Proc. SPIE*, Vol. 281, pp. 110–115, 1981.

[93] G. K. Kiremidjian, "Issues in Image Registration", *Proc. SPIE*, Vol. 758, pp. 80–87, January 1987.

[94] R. F. Koch and D. C. Evans, "ATRAN Terrain Sensing Guidance—the Grand-daddy System", *Proc. SPIE*, Vol. 238, pp. 2–9, August 1980.

[95] D. T. Kuan, et al., "Adaptive Noise Smoothing Filter for Images with Signal-Dependent Noise", *IEEE Trans. Pattern Anal. and Machine Intell.*, Vol. PAMI-7, pp. 165–177, March 1985.

[96] C. D. Kuglin et al., "Map Matching Techniques for Terminal Guidance using Fourier Phase Information", *Proc. SPIE*, Vol. 186, May 1979.

[97] H. Kunsch, "Approximations to the Maximum Likelihood Equations for Some Gaussian Random Fields", *Scandinavian Journal of Statistics*, Sept. 1983.

[98] F. Leberl, "Accuracy Analysis of Stereo Side-Looking Radar", *Photogrammetric Engineering and Remote Sensing*, Vol. 45, pp. 1083–1096, August 1979.

[99] F. Leberl, et al., "Multiple Incident Angle SIR-B Experiment Over Argentina: Stereo-Radargrammetric Analysis", *IEEE Trans. Geosci. Remote Sensing*, Vol. GE-24, pp. 482–491, July 1986.

[100] C. H. Lee and A. Rosenfeld, "Improved Methods of Estimating Shape from Shading Using Light Source Coordinate System", *Artificial Intelligence*, Vol. 26, pp. 125–143, 1985.

[101] D. Lee, "A Provably Convergent Algorithm for Shape from Shading", *DARPA Image Understanding Workshop*, Miami Beach, Florida, December 1985.

[102] M. D. Levine, et al., "Computer Determination of Depth Maps", *Computer Graphics and Image Processing*, Vol. 2, pp. 131–150, October 1973.

[103] D. J. Lewinski, "Nonstationary Probabilistic Target and Clutter Scattering Models ", *IEEE Trans. Antennas and Propag.*, Vol. AP-31, pp. 490–498, May 1983.

[104] F. Li and R. Goldstein, "Studies of Multi-baseline Interferometric Synthetic Aperture Radars", *International Geoscience and Remote Sensing Symposium*, Ann Arbor, MI, pp. 1545–1550, May 1987.

[105] T. K. Lo and G. Gerson, "Guidance Position Update by Multiple Subarea Correlation", *Proc. SPIE*, Vol. 186, May 1979.

[106] M. W. Long, *Radar Reflectivity of Land and Sea*, Artech House, Mass., 1983.

[107] B. Mandelbrot, *Fractals: Form, Chance, and Dimension*

[108] P.A. Maragos, R.W. Schafer, and P.M. Mersereau, "Two-Dimensional Linear Prediction and its Application to Adaptive Predictive Coding of Images", *IEEE Trans. Acoust., Speech, Signal Proc.*, Vol. ASSP-32, pp. 1213–1229, December 1984.

[109] A. Margolit, I. S. Reed, and R. M. Gagliardi, "Adaptive Optical Target Detection using Correlated Images", *IEEE Trans. Aerospace and Electronic Systems*, Vol. AES-21, pp. 394–405, May 1985.

[110] D. Marr, *Vision: A Computational Investigation in the Human Representation and Processing of Visual Information*, W. H. Freeman, San Francisco, 1981.

[111] J. Marroquin, S. Mitter, and T. Poggio, "Probabilistic Solution of Ill-Posed Problems in Computational Vision", *Journal of the American Statistical Assoc.*, Vol. 82, pp. 76–89 , March 1987.

[112] H. Masursky, et al., "Pioneer Venus Results: Geology From Images and Altimetry", *J. Geophys. Res.*, Vol. 85, pp. 8232–8260, December 1980.

[113] R. H. Mitchel and S. Marder, "Synthetic Aperture Radar Image Quality Considerations", *Optical Engineering*, Vol. 21, pp. 48–55, January 1982.

[114] J. W. Modestino, et al., "Stochastic Image Models Generated by Random Tesselations of the Plane", *Computer Graphics and Image Processing*, Vol. 12, pp. 74–98, 1980.

[115] P. Moon and D. E. Spencer, *The Photic Field*, MIT Press, Mass., 1981.

[116] K. M. Mori, et al., "An Iterative Prediction and Correction Method for Automatic Stereocomparison", *Computer Graphics and Image Processing*, Vol. 2, pp. 386–405, July 1973.

[117] H. Mostafavi and F.W. Smith, "Image Correlation With Geometric Distortion. Part II: Effect on Local Accuracy ", *IEEE Trans. Aerosp. Electron. Syst.*, Vol. AES-14, pp. 494–500, May 1978.

[118] M. Nakagami, "The m Distribution - A General Formulation of the Intensity Distribution of Rapid Fading", in *Statistical Methods in Radio Wave Propagation*, W.C. Hoffman, Ed., New York: Pergamon, 1960.

[119] L. M. Novak, "Correlation Algorithms for Radar Map Matching" *IEEE Trans. Aerosp. Electron. Syst.*, Vol. AES-14, pp. 641–648, July 1978.

[120] R. J. Papa, et al., "The Need for an Expanded Definition of Glistening Surface", Rome Air Development Center, Tech. Rept. RACD-TR-82-271, 1982.

[121] R. J. Papa, et al., "An Analysis of Physical Optics Models for Rough Surface Scattering", Rome Air Development Center, Tech. Rept. RACD-TR-84-195, 1984.

[122] H. B. Papo and E. Gelbman, "Digital Terrain Models for Slopes and Curvatures", *Photogramm. Engr. and Rem. Sens.*, Vol. 50, pp. 695–701, June 1984.

[123] P.Z. Peebles, "The Generation of Correlated Lognormal Clutter for Radar Simulations", *IEEE Trans. Aerosp. Elect. Syst.*, Vol. AES-7, pp. 1215–1217, November 1971.

[124] A.P. Pentland, *The Visual Inference of Shape: Computation From Local Features*, Ph.D. Dissertation, Department of Psychology, MIT, Cambridge, MA, 1982.

[125] A.P. Pentland, "Fractal-Based Description of Natural Scenes", *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. PAMI-6, pp. 661–674, November 1984.

[126] A. P. Pentland, "Shading into Texture", *Artificial Intelligence*, Vol. 29, pp. 147–170, 1987.

[127] G. H. Pettengill, et al., "Pioneer Venus Results: Altimetry and Surface Properties", *J. Geophys. Res.*, Vol. 85, pp. 8261–8270, December 1980.

[128] G. H. Pettengill, P. G. Ford, and S. Nozette, "Venus: Global Surface Radar Reflectivity", *Science*, Vol. 217, pp. 640–642, August 1982.

[129] B. T. Phong, "Illumination for Computer Generated Pictures", *Comm. ACM*, pp. 311–317, June 1975.

[130] S.H. Pravdo et al., *SEASAT Synthetic Aperture Radar Users Manual*, JPL Publication 82-90, March 1983.

[131] K. Price and D. R. Reddy, "Matching Segments of Images", *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. PAMI-1, pp. 110–116, 1979.

[132] H. K. Ramapriyan et al., "Automated Matching of Pairs of SIR-B Images for Elevation Mapping", *IEEE Trans. Geosci. Remote Sensing*, Vol. GE-24, pp. 462–472, July 1986.

[133] R. R. Read, et al., "Two-Dimensional Recursive Filtering", in *Picture Processing and Digital Filtering*, T.S. Huang, Ed., Springer-Verlag, 1979.

[134] T. Reindfleisch, "Photometric Method for Lunar Topography", *Photogrammetric Eng.*, Vol. 32, pp. 262–276, 1966.

[135] B. D. Ripley, "Modeling Spatial Patterns", and discussion, *Journal of the Royal Statistical Society*, pp. 172–212, 1977.

[136] Y. A. Rosanov, "On Gaussian Fields With Given Conditional Distributions", *Theory of Probability and its Applications*, Vol. 12, pp. 381–391, 1967.

[137] A. Rosenfeld and G. J. VanderBrug, "Coarse-Fine Template Matching", *IEEE Trans. Systems, Man and Cybernet.*, Vol. SMC-7, pp. 104–107, 1977.

[138] G. T. Ruck, Ed., *Radar Cross-Section Handbook*, Plenum Press, New York, 1970.

[139] W. Rudin, *Principles of Mathematical Analysis*, New York: McGraw-Hill, Third Edition, 1976.

[140] J. L. Mannos and D. J. Sakrison, "The Effects of a Visual Fidelity Criterion on the Encoding of Images", *IEEE Trans. Inform. Theory*, Vol. IT-20, pp. 525–536, July 1974.

[141] D. J. Sakrison, "On the Role of the Observer and a Distortion Measure in Image Transmission", *Trans. Commun. Syst.*, Vol. COM-25, pp. 1251–1267, November 1977.

[142] M. S. Scivier, M. A. Fiddy, and R. E. Burge, "Estimating SAR Phase from Complex SAR Imagery", *J. Phys. D: Appl. Phys.*, Vol. 19, No. 3, pp. 357–362, March 1986.

[143] M. Sekine, et al, "Suppression of Ground and Weather Clutter", *IEE Proc.*, Vol. 128, Pt. F, pp. 175–178, June 1981.

[144] B. J. Schachter, A. Rosenfeld, and L. S. Davis, *IEEE Trans. Systems, Man, and Cybernetics*, Vol. SMC-8, pp. 216–224, September 1978.

[145] G. Sharma and R. Chellappa, "A Model-Based Approach for Estimation of Two-Dimensional Maximum Entropy Power Spectra", *IEEE Trans. Inform. Theory*, Vol. IT-31, pp. 90–99, January 1985.

[146] T. Simchony and R. Chellappa, "Direct Analytic Methods for Solving Poisson Equations in Computer Vision Problems", to appear *IEEE Computer Society Workshop on Computer Vision*, Miami Beach, FL, December 1987.

[147] M. I. Skolnik, Ed., *Radar Handbook*, McGraw-Hill, New York, 1970.

[148] F. W. Smith and J. A. Malin, "Models for Radar Scatterer Density in Terrain Images", *IEEE Trans. Aerospace and Electronic Sysytems*, Vol. AES-22, pp. 642–647, September 1986.

[149] G. B. Smith, "The Relationship Between Image Irradiance and Surface Orientation", *IEEE Conf. Comput. Vision Pattern Recog.*, Washington D.C., June 1983.

[150] G. B. Smith, "Stereo Integral Equation", *AIAA Fifth Int. Conf. Artificial Intell.*, Philadelphia, PA, pp. 689–694, August 11–15, 1986.

[151] D. L. Snyder, *Random Point Processes*, Wiley-Interscience, New York, 1975.

[152] L. M. Spetner and I. Katz, "Two Statistical Models for Radar Terrain Return", *IEEE Trans. Antennas and Propagation*, Vol. AP–8, pp. 242–246, 1960.

[153] T.G. Stockham, "Image Processing in the Context of a Visual Model", *Proc. IEEE*, Vol. 60, pp. 828–842, July 1972.

[154] T. M. Strat, "A Numerical Method for Shape from Shading from a Single Image", M.S. Thesis, Dept. of Electrical Engineering and Computer Science, MIT, 1979.

[155] M. Sugai, et al., "VLSI Processor for Image Processing", *Proc. IEEE*, Vol. 75, pp. 1160–1166, September 1987.

[156] R. Teekens, *Prediction Methods in Multiplicative Models*, Rotterdam Univ. Press, 1972.

[157] C. W. Therrien, et al., "Statistical Model-Based Algorithms for Image Analysis" *Proc. IEEE*, Vol. 74, pp. 532–551, April 1986.

[158] K. E. Torrance and E. M. Sparrow, "Theory for Off-Specular Reflection from Roughened Surfaces", *J. Opt. Soc. Am.*, Vol. 57, pp. 1105–1114, September 1967.

[159] H. J. Trussell and M. Civinlar, "The Feasible Solution in Signal Restoration", *IEEE Trans. Acoust., Speech, Signal Processing*, Vol. ASSP-32, No. 2, pp. 201–212, April 1984.

[160] F. T. Ulaby , "Vegetation Clutter Model", *IEEE Trans. Antennas and Propag.*, Vol. AP-28, pp. 538–545, July 1980.

[161] F.T. Ulaby, R.Y. Li, and K.S. Shanmugan, " Crop Classification Using Airborne Radar and Landsat Data", *IEEE Trans. Geosci. Remote Sens.*, Vol. GE-20, pp. 42–51, January 1982.

[162] F. T. Ulaby, R. K. Moore, A. K. Fung, *Microwave Remote Sensing, Active and Passive*, Addison Wesley, Mass, 1982.

[163] F. T. Ulaby, et al., "Relating Polarization Phase Difference of SAR Signals to Scene Properties", *IEEE Trans. Geosci. Remote Sensing*, Vol. GE-25, pp. 83–92, January 1987.

[164] J. van Diggelen, "A Photometric Investigation of the Slopes and Heights of the Ranges of Hills in the Maria of the Moon", *Bulletin of the Astronom. Inst. of the Netherlands*, Vol. 11, July 1951.

[165] M. von Rohr, Ed., *Geometrical Investigation of the Formation of Images in Optical Instruments*, English Translation by R. Kanthack, His Majesties Stationary Office, London, 1920.

[166] M.S. Weiss, "Modification of the Kolmogorov-Smirnov Statistic for Use With Correlated Data", *J. Amer. Stat. Assoc.*, Vol. 73, pp. 872–875, December 1978.

[167] R. L. Wildey, "Topography from Single Radar Images", *Science*, Vol. 224, pp. 153–156, April 1984.

[168] R. L. Wildey, "Radarclinometry for the Venus Radar Mapper", *Photogrammetric Engineering and Remote Sensing*, pp. 41–50, January 1986.

[169] R. L. Wildey, "Radarclinometry", *Earth, Moon, and Planets*, Vol. 36, pp. 217–247, 1986.

[170] A.P. Witkin, "Recovering Surface Shape and Orientation from Texture", *Artificial Intelligence*, Vol. 17, pp. 17–46, August 1981.

[171] R.J. Woodham, "Photometric Method for Determining Surface Orientation from Multiple Images", *Optical Engineering*, Vol. 19, pp. 139–144, January 1980.

[172] R. J. Woodham, "Using Digital Terrain Data to Model Image Formation in Remote Sensing", *Proc. SPIE*, Vol. 238, pp. 361–369, August 1980.

[173] R.J. Woodham, "A Computational Approach to Remote Sensing", *IEEE Comput. Vision and Pattern Recog. Conference*, June 1985, San Francisco, pp. 2–12.

[174] J.W. Woods, "Two-Dimensional Discrete Markovian Random Fields", *IEEE Trans. Inform. Theory*, Vol. 18, pp. 232–240, March 1972.

[175] J. W. Woods, "Two-Dimensional Kalman Filtering", in *Two-Dimensional Digital Signal Processing*, Vol. 1, T. S. Huang, Ed., Springer-Verlag, 1981.

[176] S.T. Wu and A.K. Fung, "A Noncoherent Model for Microwave Emissions and Backscattering from the Sea Surface", *J. Geophysical Res.*, Vol. 77, pp. 5917–5929, October 1972.

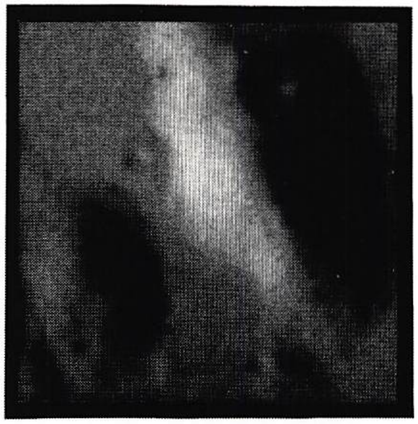
[177] C. Wu, et al., *An Introduction to the Interim Digital SAR Processor and the Characteristics of the Associated Seasat SAR Imagery*, JPL Publication 81-26, April 1981.

[178] D. C. Youla and H. Webb, "Image Restoration by the Method of Convex Projections: Part 1-Theory", *IEEE Trans. Med. Imaging*, Vol. MI-1, pp. 81–94, October 1982.

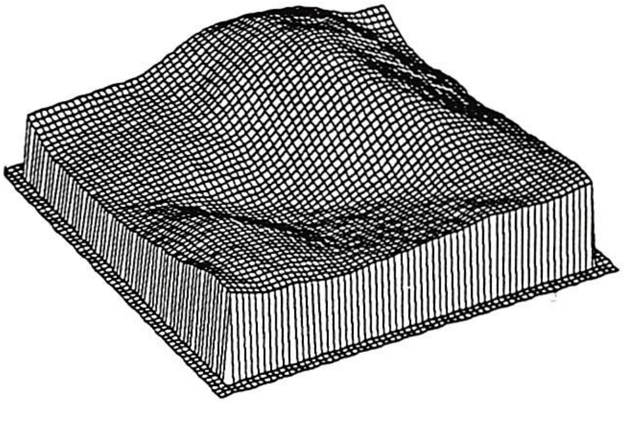
[179] S. H. Yu, "Implementation of Shape-From-Shading Algorithms", USC Intelligent Systems Group Report 104, pp. 85–92, October 1983.

[180] H. A. Zebker and R. M. Goldstein, "Topographic Mapping from Interferometric Synthetic Aperture Radar Observations", *International Geoscience and Remote Sensing Symposium*, Amherst, MA, 1985.

[181] Yi-Tong Zhou and R. Chellappa, "Linear Feature Extraction Based on an AR Model Edge Detector, *Proc. International Conf. on Acoustics, Speech, and Signal Processing*, Dallas, TX, pp. 555–558, April 1987.



(A)



(B)

Figure 3.3: Shape from shading results for real imagery of the moon. (A) Shows the observed image, and (B) is the estimated surface.

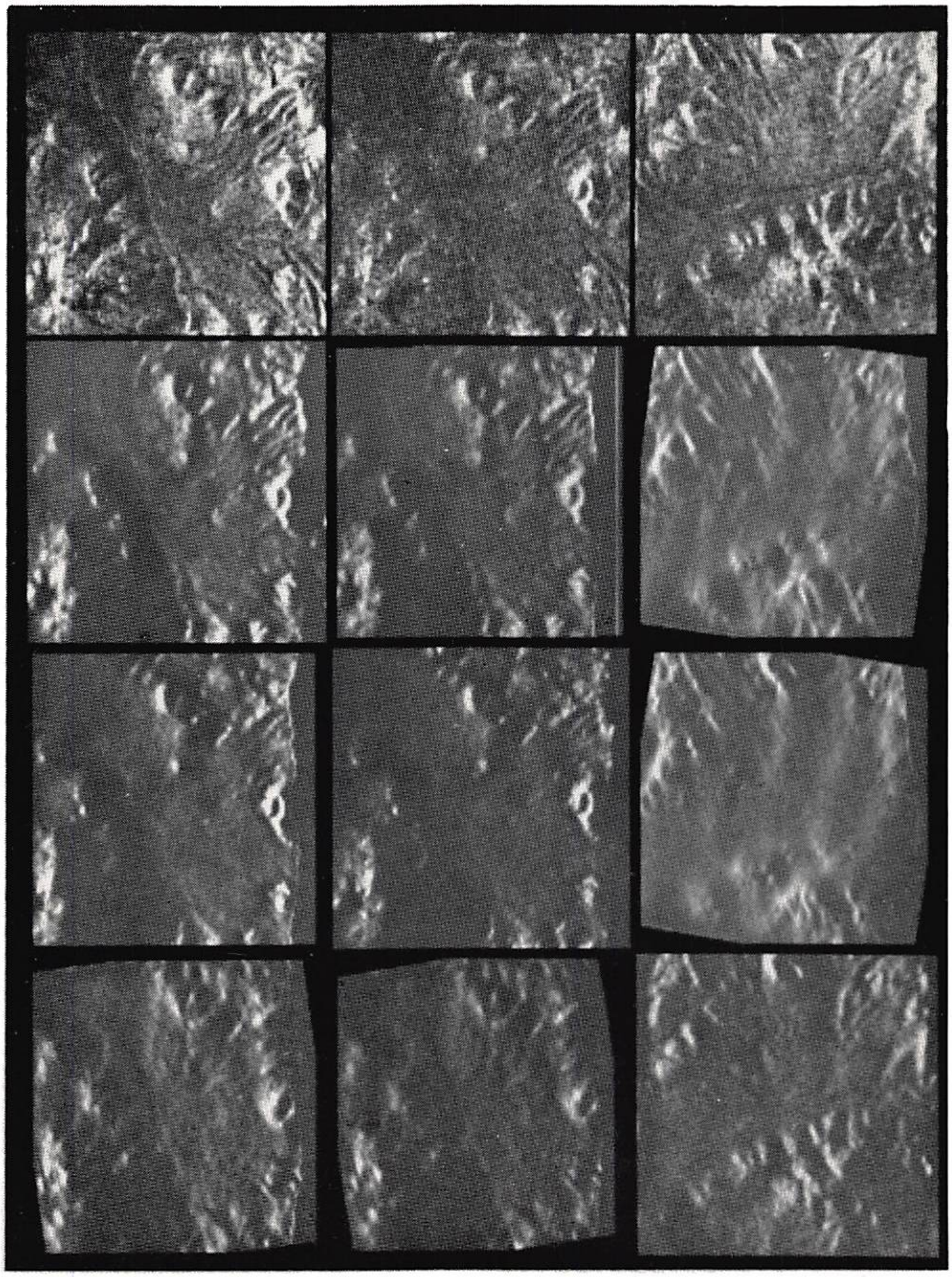


Figure 4.8: Three original SAR images and and their intensity predictions given shape from shading results.

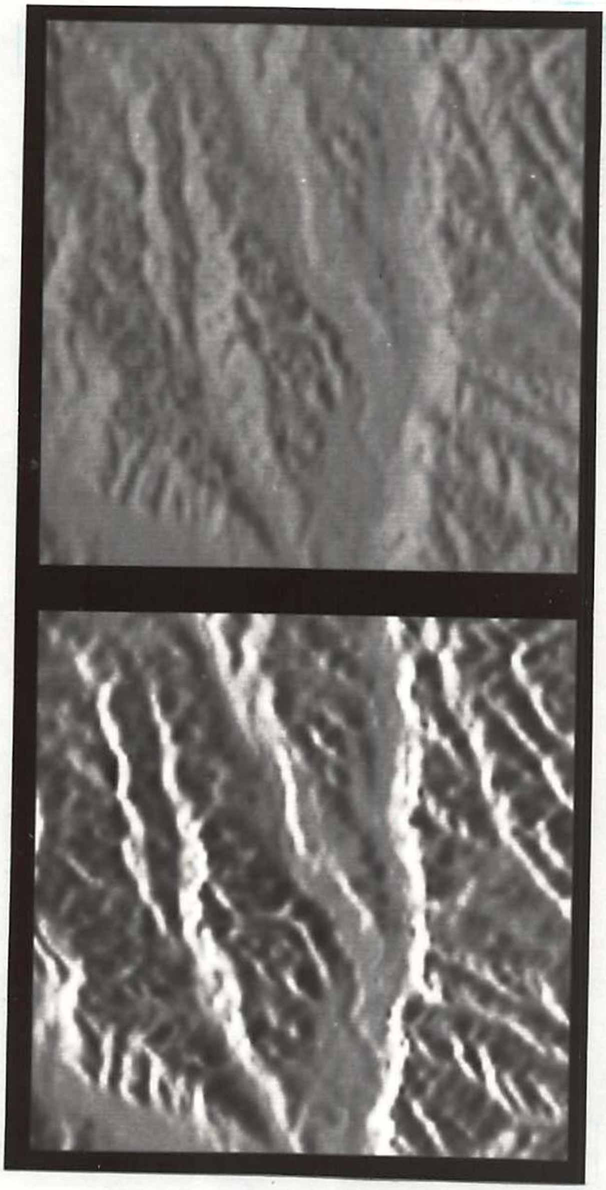


Figure 2.4: Top: Simulated aerial photograph, Bottom: Simulated SAR image. The highly directional reflectivity and the foreshortening of mountain upslopes is apparent in the SAR image.

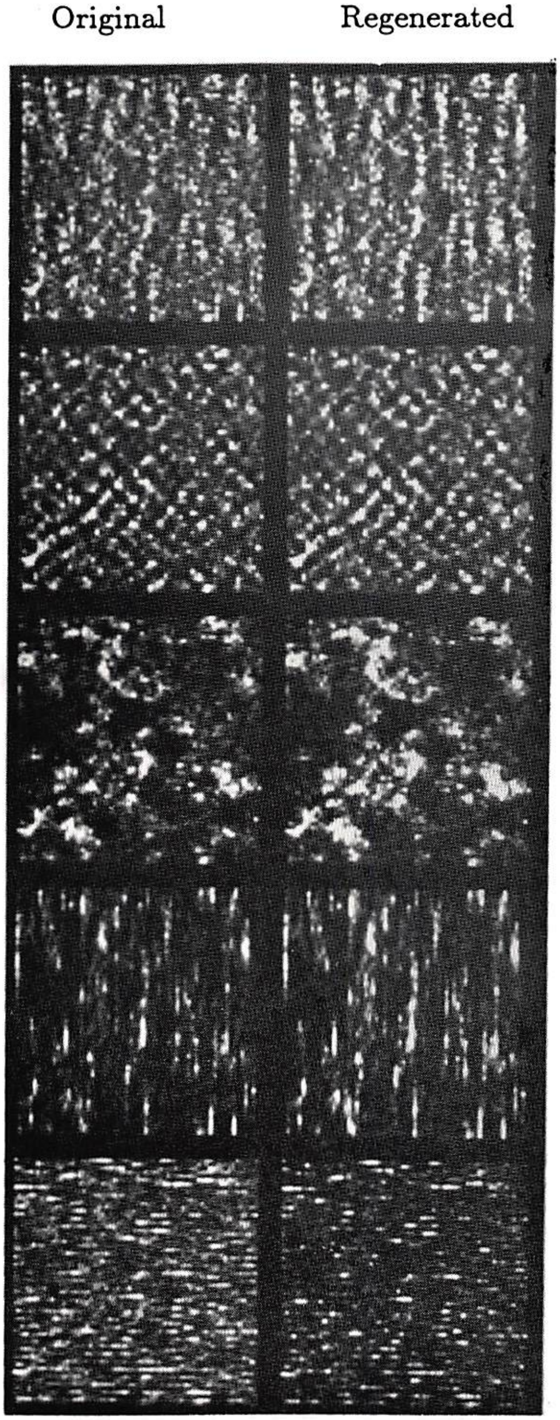


Figure 6.2: Purely synthetic textures obeying lognormal MMRF and regenerations using LS parameter estimates.

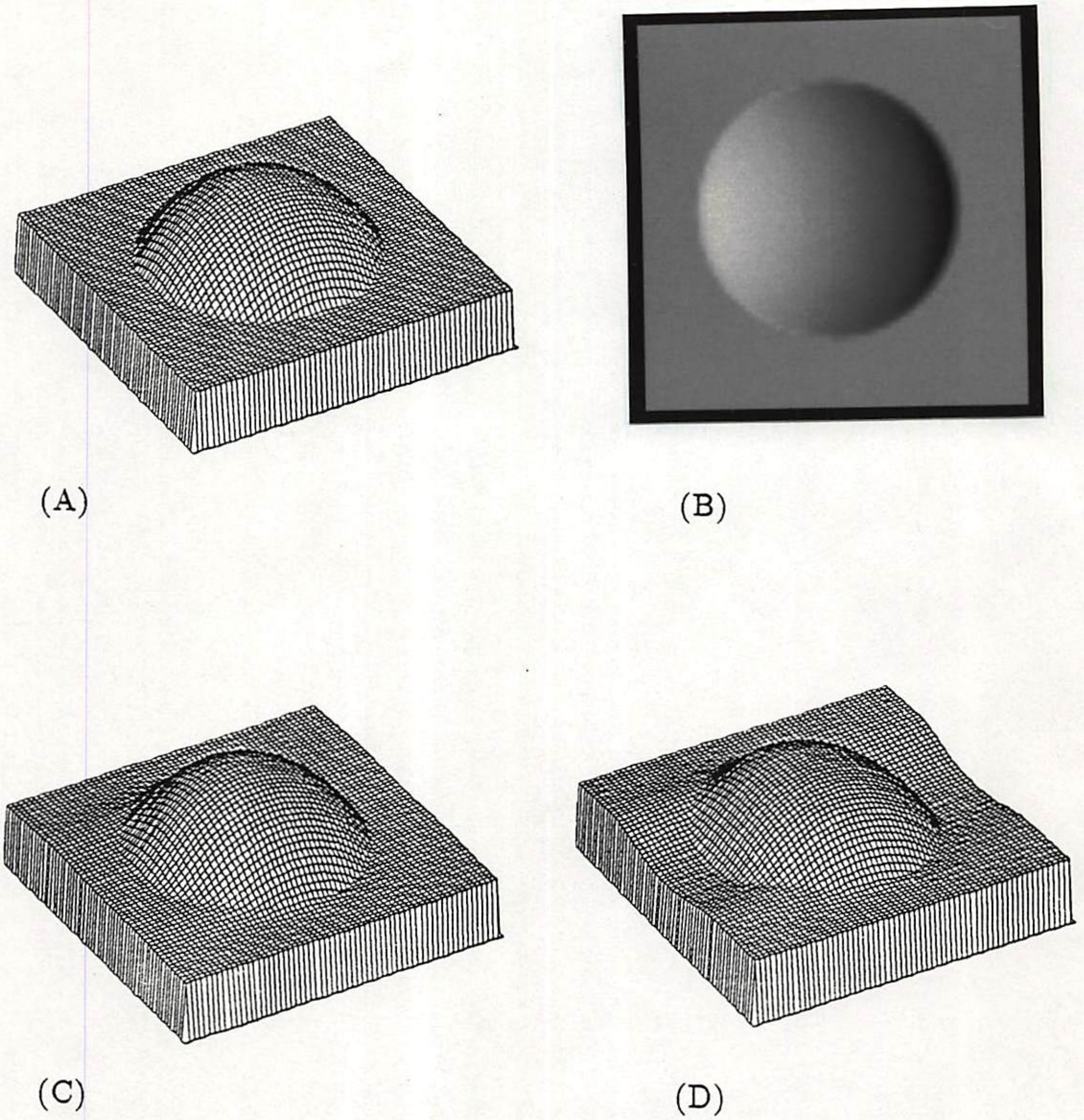
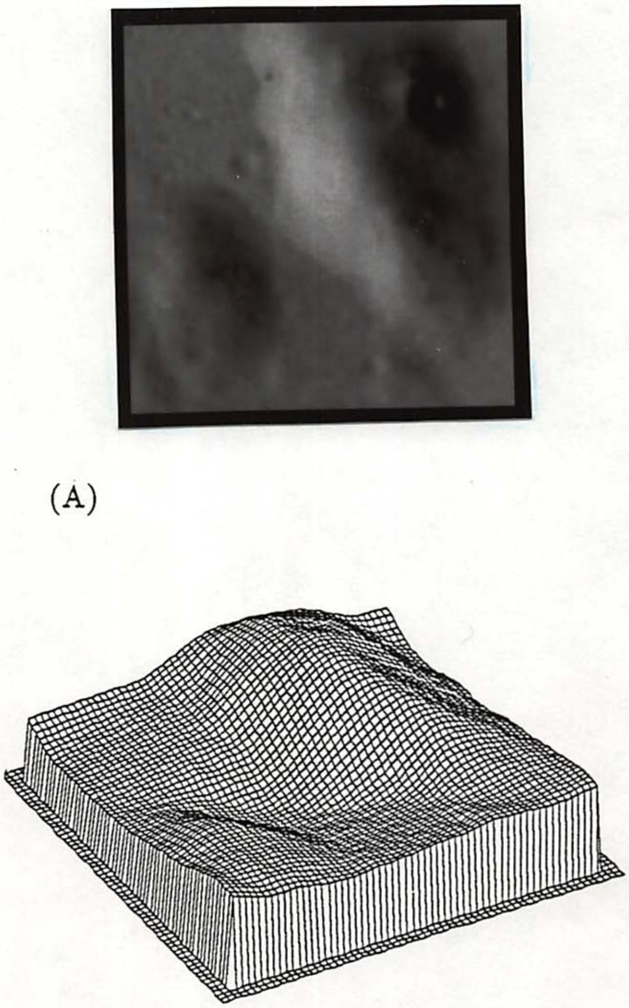


Figure 3.2: Shape from shading results using simulated sphere image. (A) Shows the true surface shape, (B) shows an image simulated from that surface, (C) is the surface estimated by assuming that all surface slopes around the border of the image are zero, (D) is the surface estimated with unknown boundary slopes.



(A)

(B)

Figure 3.3: Shape from shading results for real imagery of the moon. (A) Shows the observed image, and (B) is the estimated surface.

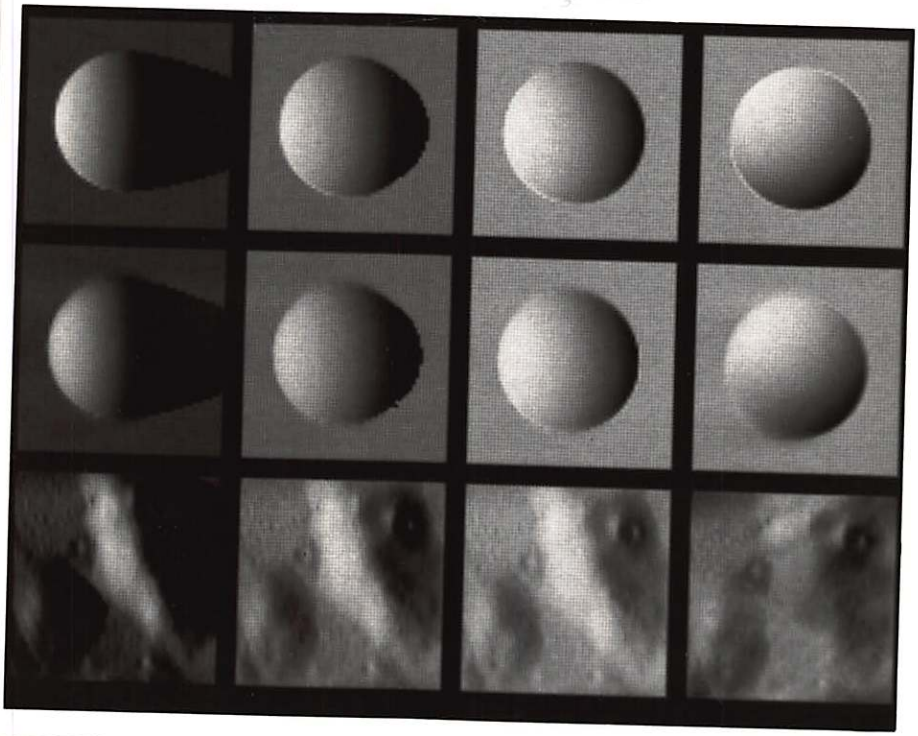


Figure 3.7: Simulated images for various imaging geometries given shape from shading results. The first row shows the images simulated from the true partial sphere surface of Figure 3.2. The second row uses the reconstructed surface given only the image intensity in Figure 3.2, and the third row uses a surface estimated from the moon image of Figure 3.3.

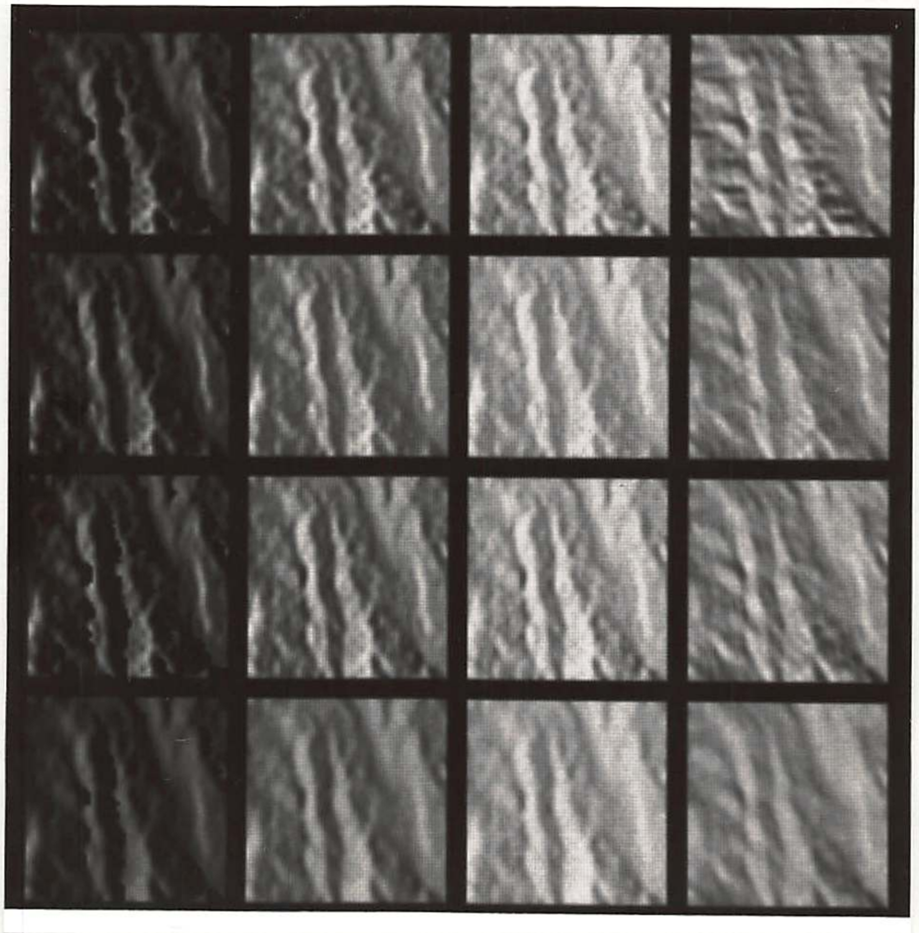


Figure 3.8: Simulated images given true DTM compared with images predicted by shape from shading results. The first (top) row shows images simulated from the DTM for various illumination directions. The second row contains predicted images for those same illumination geometries but using the surface estimate given the first image from row one. The third, fourth, and fifth rows use the surfaces estimated given the second, third, and fourth images respectively from row one.

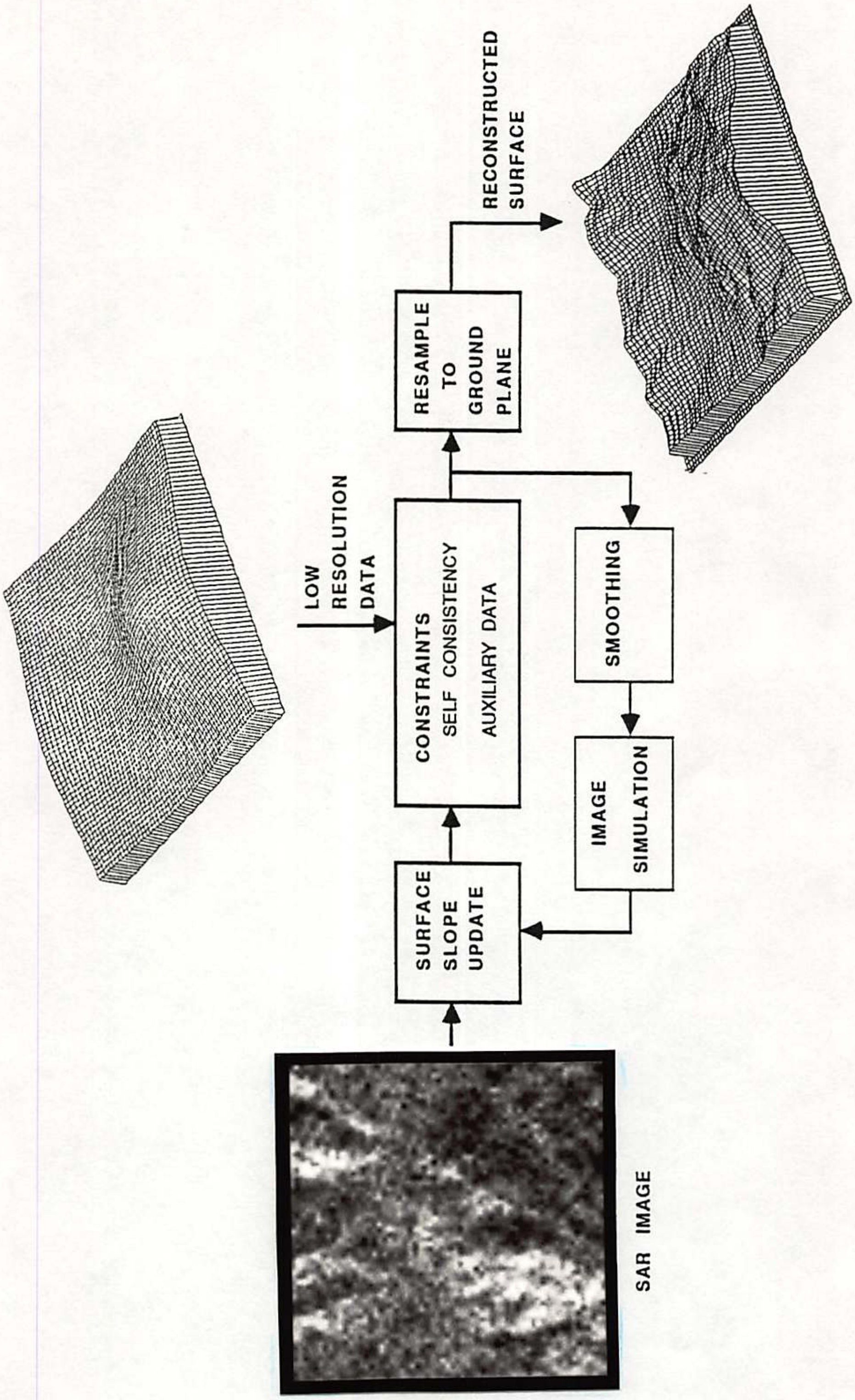
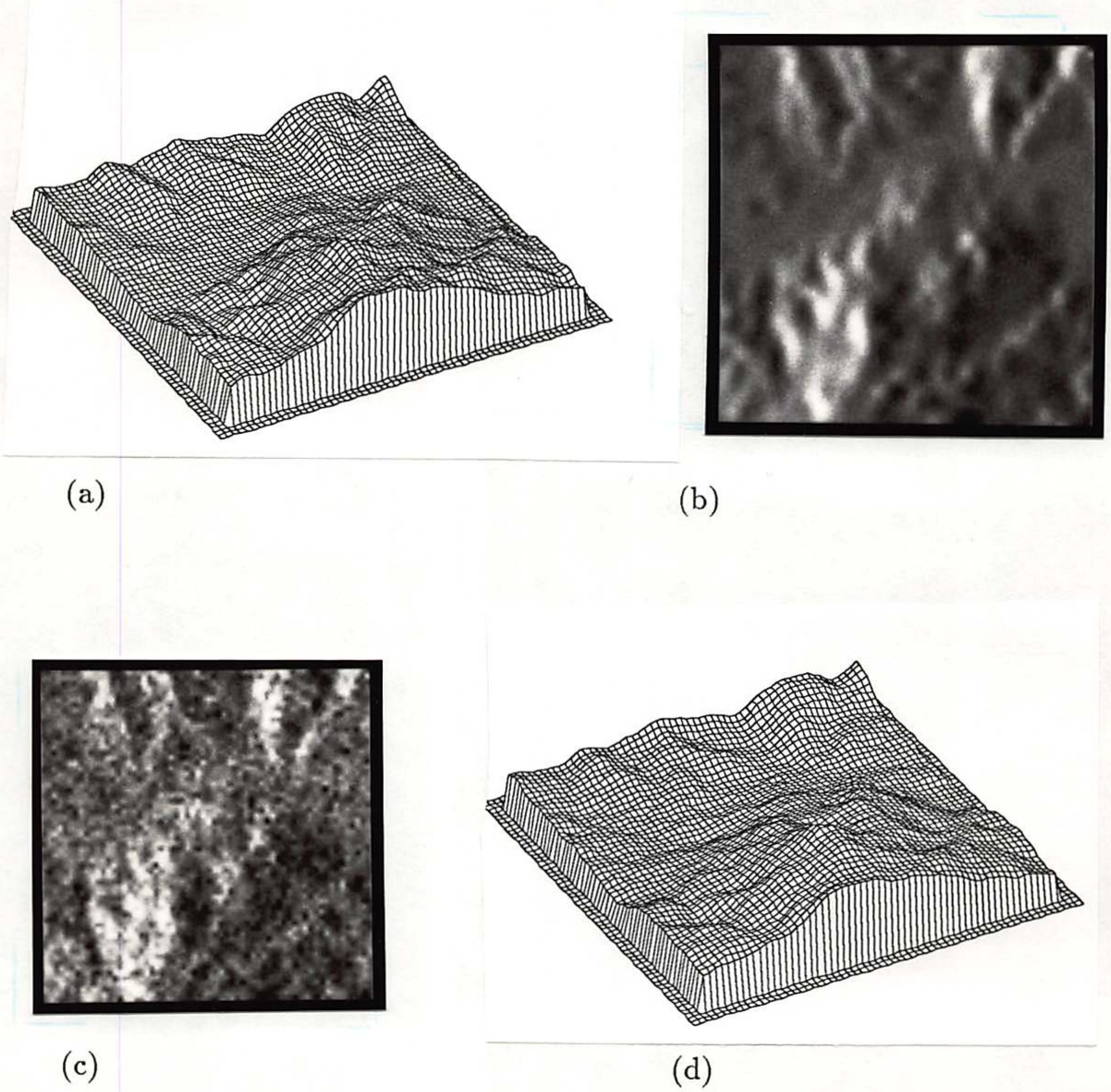


Figure 4.1: Block diagram of SAR shape from shading approach.



(a) (b)

(c) (d)

Figure 4.2: (a) Original DTM, (b) simulated SAR image, no noise, (c) simulated SAR image with speckle and sidelobe noise after 16 look noncoherent averaging, (d) surface reconstructed from noisy image.

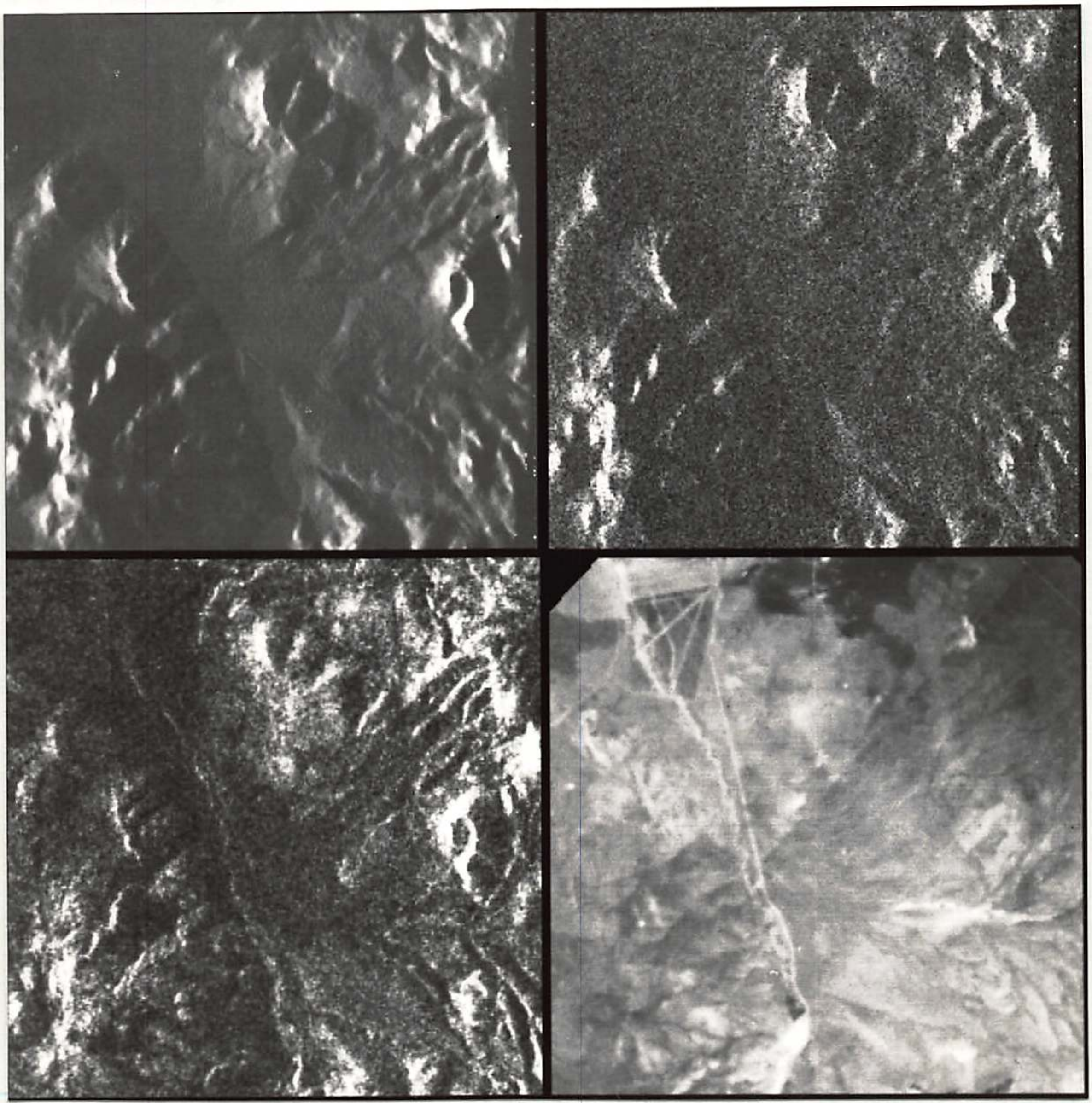


Figure 4.4: Comparison between simulated SAR image, real SAR image, and an aerial photograph. Top-left: simulated noise-free SAR image given a DTM (includes additive noise bias term). Top-right: Simulated SAR image including 28-look speckle pattern. Bottom-left: Real SIR-B SAR image. Bottom-right: Aerial photograph.

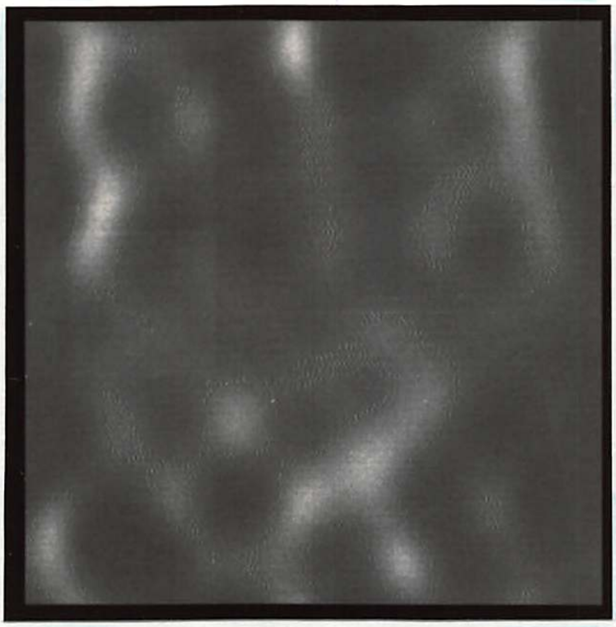
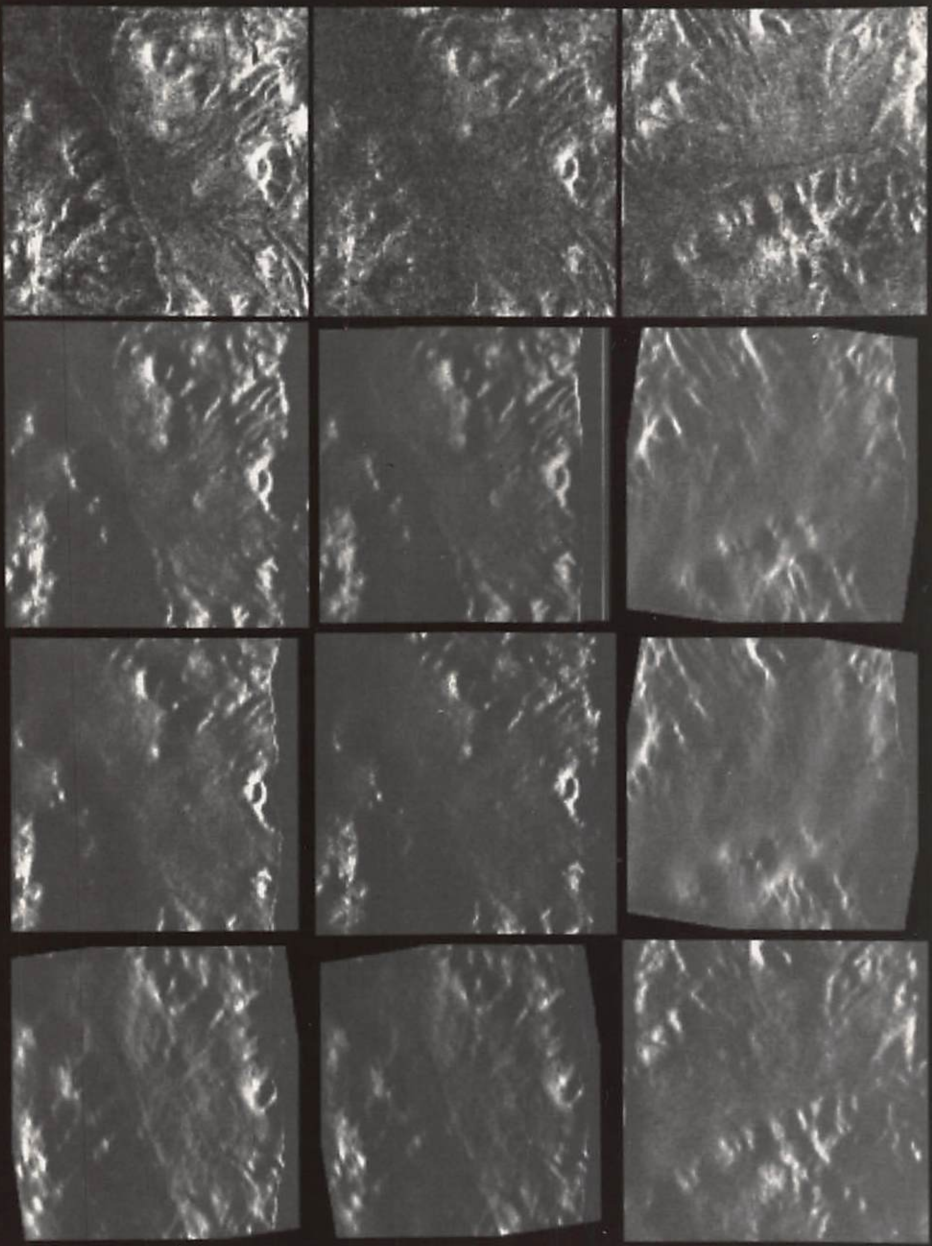


Figure 4.7: SAR image simulated from the same low resolution DTM used to aid surface reconstruction. This illustrates how little low frequency surface information is available in image intensity.



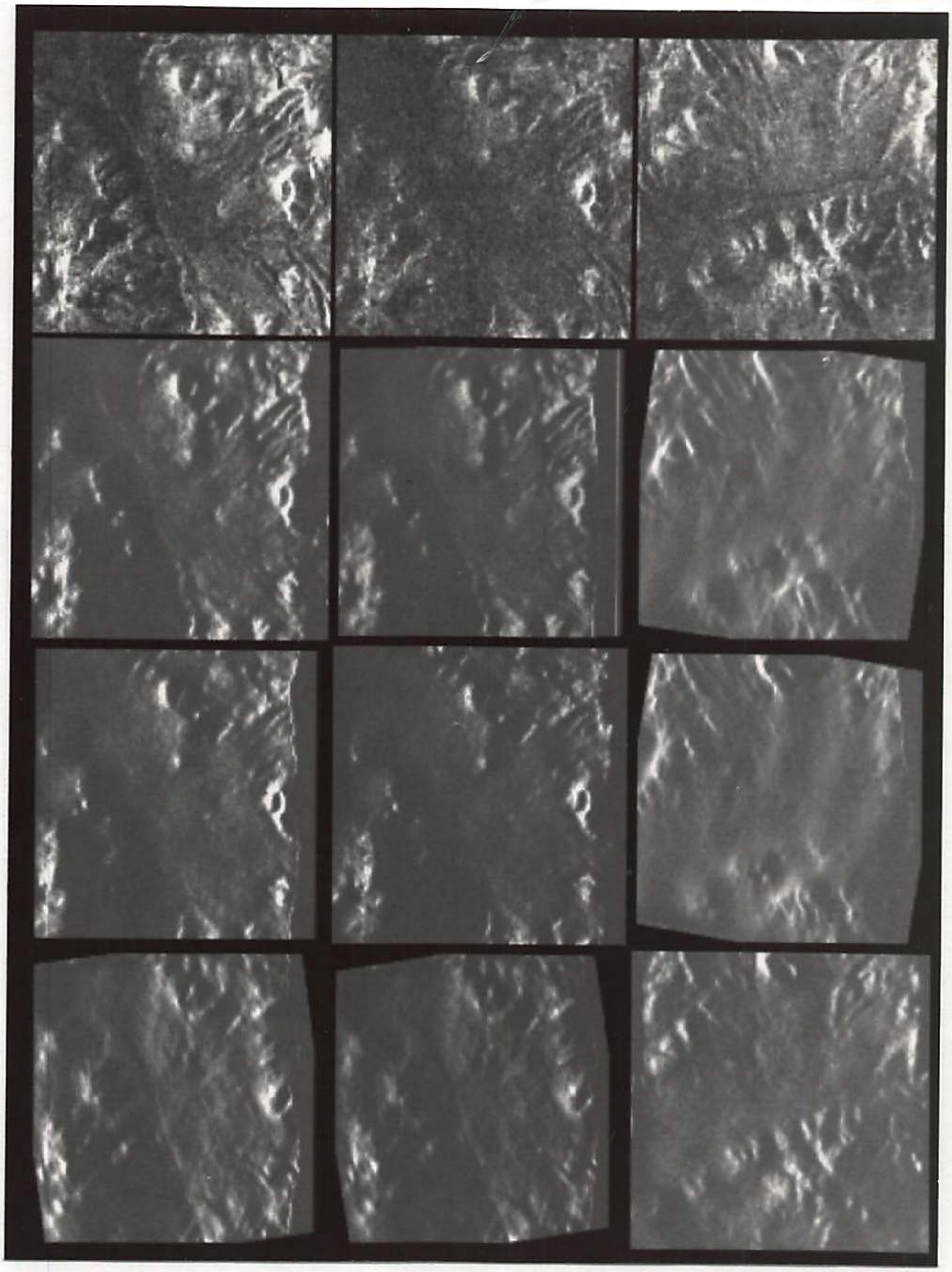


Figure 4.8: Three original SAR images and and their intensity predictions given shape from shading results.

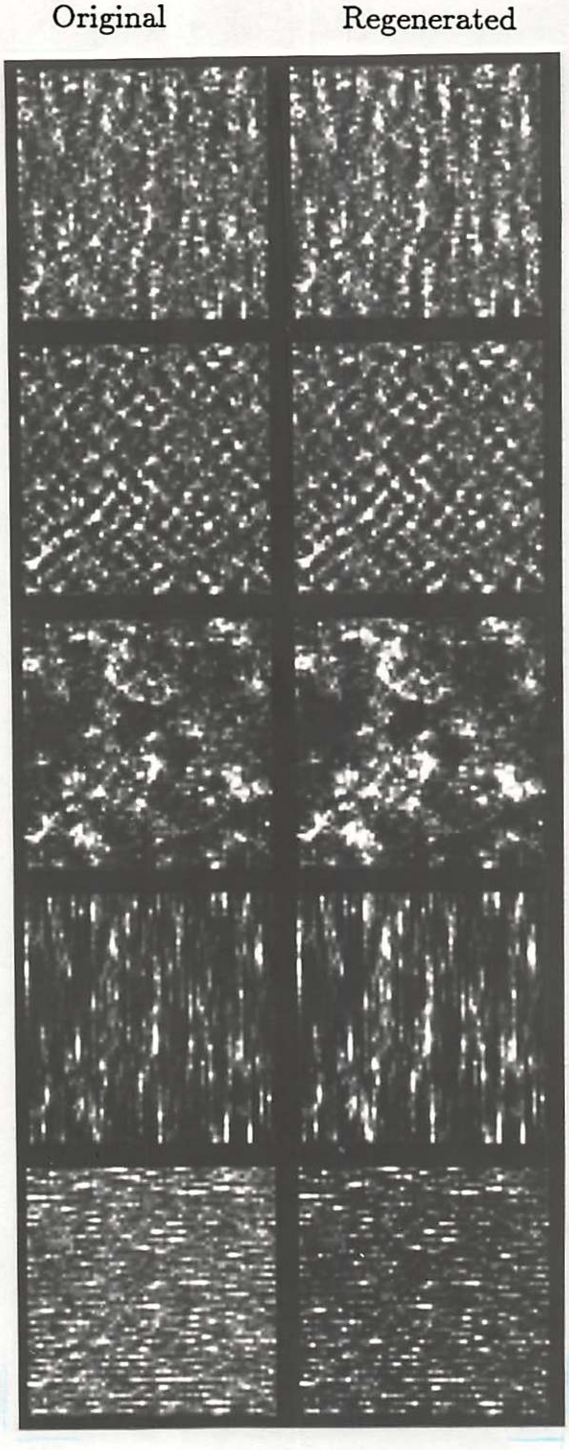


Figure 6.2: Purely synthetic textures obeying lognormal MMRF and regenerations using LS parameter estimates.

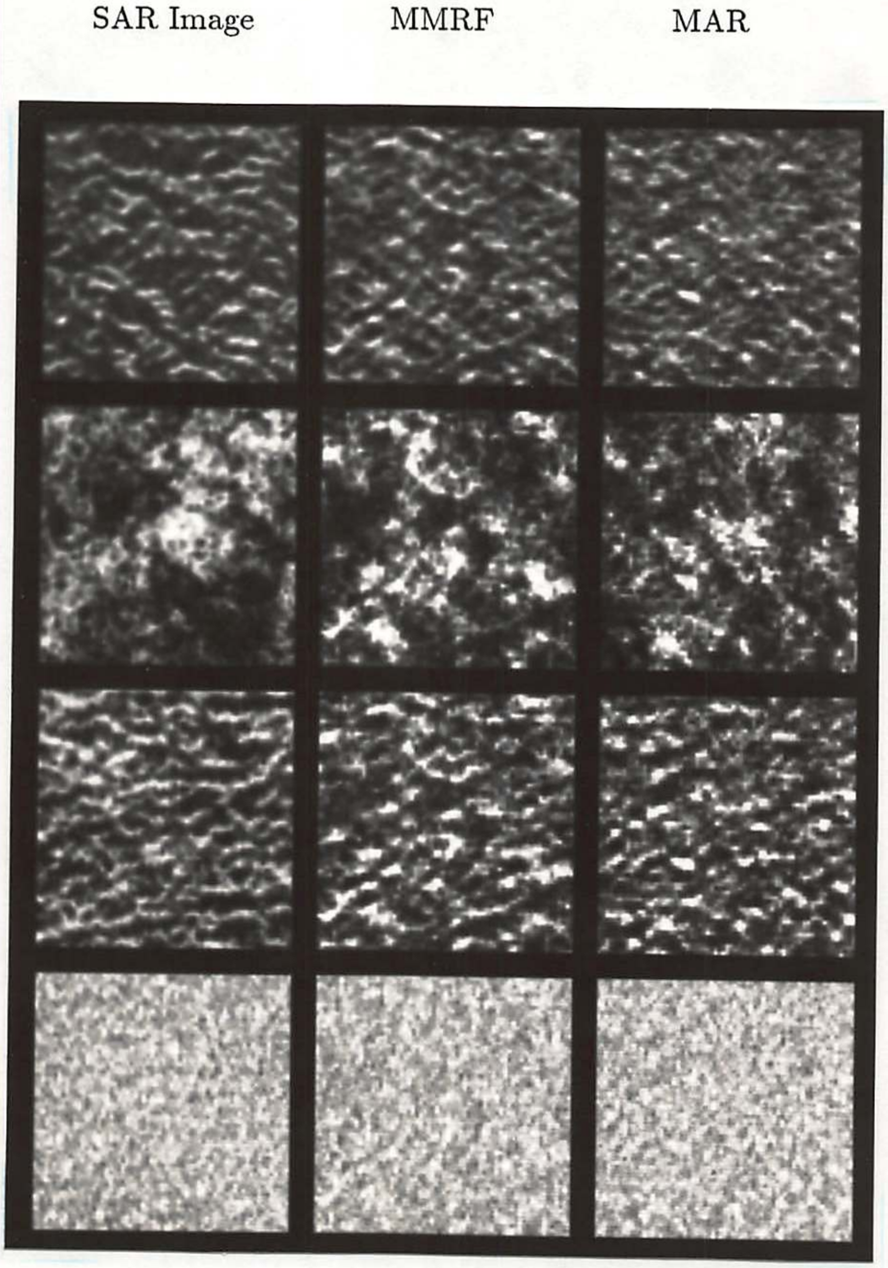


Figure 6.3: Radar images and lognormal synthetic textures. Images are 64×64 pixels.