# Topographic Reconstruction from Radar Imagery 

by

Joseph R. Matarese

Submitted to the Department of Earth, Atmospheric, and Planetary Science in partial fulfillment of the requirements<br>for the degree of<br>Master of Science

at the

## MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1988
(C) MASSACHUSETTS INSTITUTE OF TECHNOLOGY 1988

Signature of Author $\qquad$

Certified by

M. Nafi Toksöz Director, Earth Resources Laboratory

Thesis Supervisor

Accepted by
William F. Brace Chairman, Department Committee

## DISCLAIMER OF QUALITY

Due to the condition of the original material, there are unavoidable fiaws in this reproduction. We have made every effort possible to provide you with the best copy available. If you are dissatisfied with this product and find it unusable, please contact Document Services as soon as possible.

Thank you.

The grayscale figure pages within this thesis are photo copies and lack the broader tonality of the originals. The scanned images, however, represent the best availble copies.

To my mother and grandparents, who taught me the value of outlook and perspective without forcing any on me.

VAX, MicroVAX, 11/750, and 11/780 are trademarks of Digital Equipment Corporation. UNIX is a trademark of AT\&T Bell Laboratories.

# Topographic Reconstruction from Radar Imagery <br> by 

Joseph R. Matarese

Submitted to the Department of Earth, Atmospheric, and Planetary Sciences on May 23, 1988 in partial fulfillment of the requirements for the degree of Master of Science in Earth, Atmospheric, and Planetary Sciances


#### Abstract

Many geologically interesting features (e.g. fault zones) coincide with areas of high terrain relief. The rugged topography causes distortion of radar imagery for these regions. Although radar foreshortening can enhance large-scale geologic features like ridge offsets or scarps, this effect may hide detailed structures that provide important insight. To obtain this "encrypted" information so that it may be registered with other geologic and geophysical data, one must minimize the adverse influence of topography on the radar image.

This thesis proposes migration, an active area of research in seismic exploration for imaging irregular boundaries from backscattered seismic waves, as a fundamental step in the topographic correction of radar images. Based upon the migration technique, an automatic method for producing digital elevation data from radar imagery is presented.


Thesis Supervisor: M. Nafi Toksöz
Professor of Geophysics
Director, MIT Earth Resources Laboratory

## Acknowledgments

First I would like to thank my advisor, Naf Toksöz, who had the foresight to grab onto imaging radar as a useful tool for geophysical and geologic exploration. Nafis sense of perspective in pioneering the earthward applications of radar provided me with direction in this new, and sometimes chaotic, field.

I greatly appreciate the guidance of Michael Prange, who introduced me to the fundamentals of radar imaging and who encouraged me constantly through the various stages of this research.

Thanks to various people in EAPS for their support and enthusiasm during my stay at MIT: in particular, Erik Mollo-Christensen and Ken Morey, who taught me basic image processing during my initial MIT years and who even paid me for it; Tim Grove, my undergraduate advisor, who buffered me from the MIT bureaucracy; and also Sean Solomon and Stan Hart who made my graduate stay possible by providing me with a crucial TA'ship.

I am indebted also WHOI remote sensing specialists Hans Graber, Jules Jaffe, and Kathie Kelly who succeed courageously at encapsulating a wealth of oceanographic remote sensing knowledge into a semester-long class.

I would like to thank NASA and JPL for funding the radar imaging project (JPL 956943 under NASA Prime) and to especially thank JPL's Diane Evans, Mike Kobrick, and John Curlander for enabling the results of this thesis to be realized.

My gratitude extends to the people of ERL, who have created a working environment that I'd rather not give up.

And, finally, my thanks goes to Ken Rugg, Bob Chase, George Franks, and John McCrea my friends and housemates, who paid most of the rent. Now it's my turn to relax and have a homebrew!

## Contents

1 Introduction ..... 7
1.1 Objectives ..... 10
1.2 Outline ..... 11
2 Radar Imaging ..... 13
2.1 The SAR System ..... 13
2.2 Geometry ..... 14
2.2.1 Orbital Configuration ..... 14
2.2.2 Slant Range vs. Ground Range ..... 17
2.2.3 Look Angle vs. Incidence Angle ..... 19
2.2.4 Relief Displacement ..... 19
2.2.5 Resolution ..... 21
2.3 Radiometry ..... 25
2.3.1 The Radar Equation ..... 25
2.3.2 Instrument Parameters ..... 26
2.3.3 Range and Scattering Area ..... 27
2.3.4 The Normalized Radar Cross Section ..... 28
2.4 Multiple Incidence Angle SAR ..... 29
2.5 Synthetic Imaging ..... 29
2.5.1 Input Data ..... 31
2.5.2 Procedure ..... 33
2.5.3 Comparison ..... 35
3 Radar Inversion ..... 37
3.1 Radar Migration ..... 39
3.1.1 A Simple Seismic Analogy ..... 40
3.1.2 Additional Constraints ..... 42
3.2 Two-dimensional (Raster) Migration ..... 46
3.2.1 Receiver Array Orientation ..... 48
3.2.2 Histogram Equalization ..... 48
3.2.3 Backprojection ..... 50
3.2.4 The Variance Field ..... 50
3.2.5 Surface Determination ..... 50
3.2.6 Results ..... 51
3.2.7 Additional Comments ..... 57
3.3 Hierarchical Boundary Matching ..... 58
3.3.1 Chronology ..... 58
3.3.2 The Algorithm ..... 62
3.3.3 Automated Migration ..... 64
3.3.4 Filtering ..... 66
4 Implementation and Results ..... 80
5 Related Work ..... 100
5.1 Radarclinometry ..... 100
5.2 Interferometry ..... 102
5.3 Stereomapping ..... 105
6 Conclusions and Future Work ..... 106
A Edge-detection ..... 108
B Surface Fitting ..... 112

## Chapter 1

## Introduction

In order to establish a background for this thesis, before presenting its outline and objectives, I will first provide a brief discussion of radar imaging applications - their history and current status with respect to remote sensing in general and geoscience as a whole. My introduction serves not only to motivate this specific work by citing the benefits of terrestrial radar imagery, but also to suggest the importance of the goal to which this research contributes: the development of a broad geoscience database.

## Chronology

The pioneering interest in radar as a geologic tool emerged after World War II with the declassification of radar-derived data; and by the late 1960's, "radar geology" had become an active field. The Plan Position Indicator (PPI), constructed for military use during the war, presented the first radar data. PPI images, though poor in resolution by today's standards, allowed the development of two image analysis schemes: the compilation of radar image interpretation keys (Hoffman, 1954), and the identification of geologic structural and textural features from radar imagery (Feder, 1959). With the advent of side-looking airborne radar (SLAR) instrumentation and synthetic aperture radars (SAR) having high resolving capabilities, image coverage increased enormously and led to many geological applications. The results of these efforts are detailed in various bibliographical publications (Bryan, 1973; Bryan, 1979; Moore et al., 1974; Sabins, 1976).

Since the early 1970's, NASA has taken a leading role in the production and distribution
of SAR imagery. After a brief hiatus, during the initial years of the Landsat program, NASA refocused its energy toward microwave imaging and proposed Seasat (Pravdo et al., 1983), the first spaceborne SAR, in 1973. The Seasat sensor, which was configured for oceanographic experimentation, flew for just over 3 months in 1978, before suffering an equipment failure. However, in its short life span, Seasat successfully provided synoptic radar imagery of both oceanic and continental surfaces. Using Seasat SAR spare parts, a new radar was constructed and flown aboard the space shuttle in 1981 (Shuttle Imaging Radar - Phase A or SIR-A) and in 1984 (SIR-B).

The first SIR missions (Cimino and Elachi, 1982; Ford et al., 1986) were geared toward continental observations and were designed as "proof-of-concept" experiments to determine the importance of various instrumentation parameters for imaging. The SIR-A SAR sensor had entirely fixed properties, including wavelength, polarization, and look angle; its primary objective was to pick up where airborne SAR left off, furthering our understanding of land radar "signatures". SIR-B employed a mechanically steerable antenna, which allowed for multiple look angle (MLA) imaging, a technique that was also performed with airborne SAR. However, the relatively steady spaceborne platform, i.e. the shuttle, made accurate surface reconstruction from MLA imagery feasible. The Challenger accident has delayed the SIR-C mission (SIRC Science Plan, 1986), which plans to incorporate, variable wavelength, polarization, and look angle capabilities. The SIR-C SAR will eventually be modified and flown on the Earth Orbiting Satellite (EOS) (EOS SAR Science Plan, 1988), scheduled for launch in 1994.

## Radar Geoscience

As the organization of this chapter might lead one to believe, microwave remote sensing has progressed in an applications-driven manner. Radar observation was borne out of military and space programs and thus funding has been prejudiced toward development of equipment and toward experimentation, rather than toward analysis. While saving a detailed description of the radar imaging process for Chapter 2, I find it necessary to briefly explain what the state-of-the-art microwave instrumentation affords the geoscience community - especially in comparison with optical systems.

The value of microwave remote sensing is derived from distinct differences between SAR
and optical imaging geometry, and more importantly, radiometry. Microwave penetration of the atmosphere gives radar imaging a fundamental advantage over optical systems. With the deployment of spaceborne platforms, radar is an all-weather tool, having applications not only for cloud-covered regions of this planet, but also for other celestial bodies with dense atmospheres, e.g. Venus.

The reason for this atmospheric transparency stems from a four or five order-of-magnitude difference in frequency between microwave and visible/infrared radiation. While optical antennae record energy on the scale of one micron, making them sensitive to the molecular structure of surface scatterers, SAR receivers measure electromagnetic energy of centimeter-scale wavelength, owing to the physical nature of scatterers, e.g. roughness and moisture content.

SAR's unique geometry also provides for topological observation. Unlike optical imaging devices, which are passive systems - the satellite focusses its "camera" straight down at the earth's surface and records scattered sunlight, radar imaging requires that microwave energy be transmitted toward the surface, where it is "backscattered" by terrain elements and recorded at the receiving antenna. Thus, one refers to radar as an active imaging system. Landforms and tectonic fabrics have distinctive radar signatures; lineament mapping is perhaps the most widespread application of SAR imagery.

This particular thesis topic evolved from an ongoing effort at the Earth Resources Laboratory to study some of the continental applications of radar imaging, namely mapping active fault zones. As with most other geologic applications of radar, the initial purpose (Prange, 1985) was simple: determine the radar signatures of various fault zone features over a wellknown region and use these findings to improve mapping in remote terrain. This corresponded to a qualitative classification scheme complete with a "training" area - around the San Andreas/Garlock Fault junction, northwest of Los Angeles, California, and an exploratory location - along the North Anatolian Fault Zone (NAFZ) in Turkey. Manual lineament mapping was chiefly used to define topographic features.

Because of the labor-intensive nature of lineament mapping and classification of image textures, we began to experiment with automated, and therefore more quantitative, interpretation techniques. This lead to the implementation of computer algorithms for edge-detection and lineament formation (Toksöz et al., 1986; Guler et al., 1986) to assist radar geologists in structural
mapping. Another goal of quantitative analysis was multispectral evaluation of remote sensing data, i.e. the comparison of surficial scattering at different frequency bands. The Landsat Multi Spectral Scanner (MSS) and Thematic Mapper (TM) produced this data set for optical wavelengths. We desired to combine SAR imagery with Landsat; however, the distortion of radar imagery, owing to topography, precluded attempts at accurately registering SAR images with corresponding Landsat data.

This exposes the problem of merging radar imagery with other geographically organized information, as well as optical imagery.

## The Geoscience Database

As optical and thermal imagery (e.g. Landsat, SPOT) and radar imagery (Seasat, SIR-A \& B) become increasingly familiar to researchers, and as applications for remotely sensed images become more diverse, techniques for quantitative analysis of the data will come under great demand. The U.S. government (USGS, NASA), as well as major oil companies, including ARCO, Chevron, and Mobil are currently developing a computational framework through which this analysis can proceed: the geoscience database.

One important objective of geoscience database management systems is the registration of satellite imagery, ground truth, and subsurface information to a common geographic coordinate system. This provides prospective researchers with a medium through which to quickly and efficiently compare all necessary and available data for a desired locale. For example, geologists mapping a remote region might wish to augment ground truth data (geologic maps, paleontologic data, seismics) with satellite imagery. Or an oil exploration company, planning new offshore drill sites, would want bathymetry, 3-D seismic data, and ocean-floor geochemistry.

### 1.1 Objectives

In the case of radar imagery, registration of image data to a geographic reference grid is no simple matter, especially in the absence of digital terrain elevation models (DTEM). Using USGS DTEM data, which equates to a digitized "topo map", terrain distortions may be modeled by generating synthetic SAR images. When a favorable match between simulated and real data is obtained, knowledge of the forward model can be applied to unwarp the real image. In the
absence of DTEM data, one must attempt to infer shape, specifically slope, from the radar image brightness values. Variation of radar return with parameters other than local slope, and lack of a reference value severely limit the accuracy of this method.

However, given multiple look angle radar data and the active imaging nature of SAR, a migration technique may be used to obtain geometrically restored SAR images, producing DTEM data as a by-product; this paves the way for radiometric restoration (i.e. finding "correct" brightness values, normalized for the topographic effect) of the radar image, as well as for registration with other geoscience data. To itemize:

- generate digital terrain data (in DTEM form) from MLA radar imagery;
- use the derived DTEM to correct the SAR images for topography;
- correlate SAR and DTEM data to radiometrically correct imagery; and,
- register restored SAR data with geoscience database.


### 1.2 Outline

Chapter 3 contains a description of the radar migration technique, which is the crux of this thesis. The applicability of migration for topographic correction of SAR imagery is developed through analogy to seismic migration. Key differences between SAR and seismic migration, in their nature and implementation, are noted. The bulk of this chapter devotes attention to the actual radar migration procedure, which is described as a series of "modular" steps.

Those unfamiliar with radar imaging should read through Chapter 2, which serves as somewhat of a tutorial on SAR. The second through fourth sections of this chapter describe various aspects of the forward model, i.e. imaging, that are crucial to understanding the inversion. This chapter ends with a discussion of other interesting SAR features, not currently essential to radar migration.

Chapters 4 and 5 present trials and conclusions from the inversion. Migration is applied to synthetic SAR data first, and its accuracy is assessed. Results are then obtained for real radar images over known topography, so that SAR-derived DTEM accuracy may be determined, and over a remote regions where no (unclassified?) terrain elevation data currently exists.

This thesis completes with a discussion of related inversion techniques. The relative merits of each are identified, with emphasis on potential collaboration to improve accuracy.

## Chapter 2

## Radar Imaging

### 2.1 The SAR System

SAR employs a more complex set of instrumentation than does a passive optical imager, and as a result it introduces additional geometric and radiometric considerations; however, the main theme of SAR - the manner in which it produces high-resolution two-dimensional images - is beautifully straightforward. This chapter mentions the more pertinent aspects of SAR; for a more detailed "tutorial", see (Tomiyasu, 1978).

The SAR system consists of three basic components: a pulsed transmitter, an antenna, and a phase-coherent receiver. Since both transmitted and received microwave energy pass through the antenna, a circulator is required to coordinate timing. The circulator prompts the transmitter to generate a pulsed signal at a specific rate, called the pulse repetition frequency (PRF). The term interpulse period (IPP) refers to the gap between successive pulses and may be related simply to the PRF:

$$
\begin{equation*}
I P P=1 / P R F+\tau \tag{2.1}
\end{equation*}
$$

where $\tau$ is the transmitted pulse width. During the IPP, the backscattered return is picked up by the antenna and directed by the circulator to the receiver, which records the signal.

Since the definitive measure of an imaging system's capability is resolution and since active systems suffer notoriously in this respect (think of how many shots and geophones must go into the average 3-D seismic survey), SAR utilizes some "trickery". The safest answer to the question, "How does one improve resolution?", is the reply, "Build a bigger antenna!" In the
case of seismic surveying, where both surficial and vertical techniques (VSP) exist, one may cost-effectively design a large imaging array. For remote sensing purposes, large antennae make for heavy and expensive (as well as offensive) payloads, and the logistics of flying a receiver array comprising several satellites is harrowing. The SAR geometry allows that pulse traveltime sampling determine range resolution (which is nothing new) and that the antenna size determine azimuthal resolution in a directly (rather than inversely) proportional manner! In other words, SAR attains finer resolutions with a smaller antenna. For the uninitiated, this important feature will be detailed in a later section on resolution, providing a "bridge" between discussions of SAR geometry and radiometry.

### 2.2 Geometry

To provide the setting for this explanation of SAR, I first present the imaging configuration along with its jargon.

SAR platforms have flown on several aircraft, on a satellite (Seasat), and twice (so far) on the Space Shuttle. The latter deployments (Seasat and SIR-A/B) fall under the category of spaceborne imaging radar. Because these missions provided the data for this work, I will bias the discussion toward them (away from airborne radars).

### 2.2.1 Orbital Configuration

Spaceborne SAR platforms circle the Earth along a path that is determined by many parameters including the orbit's altitude, inclination, and periodicity. The trace of this "flight" path along the Earth's surface is termed the nadir or ground track (see Figure 2-1); the platform velocity vector v provides the direction of azimuth or along-track, mentioned earlier. Naturally the axis of altitude lies perpendicular to the surface, and completing the orthogonal coordinate system is the ground range or cross-track axis.

The radar platform is positioned on the spacecraft so that the antenna array faces in a direction $\hat{\mathbf{n}}$ perpendicular to the velocity vector and toward a portion of the surface known as the footprint (see Figure 2-2). The footprint defines the surface area illuminated by the radar at a given instant, and its dimensions (in azimuth $W_{a}$ and in range $W_{g}$ ) are determined by the


Figure 2-1: Spaceborne SAR imaging geometry.


Figure 2-2: SAR beam footprint.
following relations:

$$
\begin{gather*}
W_{a}=\frac{\lambda R}{D_{a}} \equiv \beta_{a} R  \tag{2.2}\\
W_{g}=\frac{\lambda R}{D_{r} \cos i} \equiv \beta_{r} R / \cos i \tag{2.3}
\end{gather*}
$$

where
$\lambda$ is the radar wavelength ( 23.5 cm for $\mathrm{SIR}-\mathrm{A} / \mathrm{B}$ ),
$R$ is the slant range from antenna to target (footprint center),
$D_{a}$ and $D_{r}$ are the horizontal and vertical dimensions, respectively, of the antenna,
$i$ is the incidence angle (at which the radar signal impinges upon the target), and
$\beta_{a}$ and $\beta_{r}$ are the azimuth and range antenna beamwidths.
As the footprint moves (with spacecraft velocity) across the surface, it traces out the image swath.

### 2.2.2 Slant Range vs. Ground Range

At present, synthetic aperture radar image products come in three successive grades, "raw" images, ground range images, and geocoded images. Each grade represents conversion of the image dimensions to a slightly more user-friendly coordinate system. Geocoded images attempt to show the microwave backscatter in coordinates of latitude and longitude along a geoidal Earth surface; over regions of shallow relief, this grade is desirable for purposes of registration. Ground range images try simply to map the backscatter intensities back onto the image swath. Both of these products require some knowledge of spacecraft ephemeris in the processing. The raw image, however, records the return signal amplitude in the slant range image plane.

The slant range image plane is defined by the spacecraft velocity vector and the antenna beam axis (see Figure 2-3), but this coordinate system is somewhat misleading. Slant range is not a straightforward direction but rather a measure of the one-way travel time of the radar signal as it is backscattered from the surface, i.e.

$$
\begin{equation*}
R=c t / 2 \tag{2.4}
\end{equation*}
$$



Figure 2-3: The slant range image plane vs. ground range.
where $t$ is the two-way travel time, and $c$ the velocity of light.
To summarize the distinction between slant and ground ranges: slant range is defined as a function of time while ground range is defined as the distance from nadir to target along a spheroidal Earth surface.

### 2.2.3 Look Angle vs. Incidence Angle

While two distinct "ranges" exist in SAR lingo, four different angular measures of the antenna orientation and the direction of radar signal propagation appear. These are look angle $\theta$, incidence angle $i$, depression angle $\alpha$, and local incidence angle $\theta_{\text {local }}$. Incidence angle, as noted earlier, measures the angle between the spheroidal surface normal at the target and the direction of microwave signal propagation. The local incidence angle is a slight, yet important variation, which takes into account the actual surface normal (a function of topography) at the target. Look angle measures the orientation of the antenna beam axis with respect to vertical at the antenna/spacecraft, and the depression angle is its complement.

Owing to the great distances involved in the imaging configuration, the Earth's curvature may not be neglected, and the look angle will differ from the incidence angle by several degrees. In fact, the incidence angle may change by several tenths of a degree over the ground range dimension of the image.

These angles are summarized graphically in Figure 2-4.

### 2.2.4 Relief Displacement

Geometrical distortions owing to topography may be explained through the concept of relief displacement. Simply put, mountains and ridges that strike parallel to the flight path appear to lean toward the radar in a direction exactly perpendicular to the platform velocity vector. For example, if one erects a (highly reflective) flagpole a given geographical location (latitude and longitude) and then images the pole on a passing orbit, the receiver will register a return from the top of the pole before it obtains the return from the pole's base.

Relief displacement manifests itself on the SAR image in two ways. It modulates ground resolution and, as a result, affects the return signal power. These topics will be discussed in the next several subsections.


Figure 2-4: Various imaging angles defined.

### 2.2.5 Resolution

This section begins by addressing the question, "How did SAR get its name?", because the idea of a synthetic aperture grew out of the azimuth resolution problem.

Conventional radars employed a physically large antenna, which comprised an array of small, identical radiating elements. This array effectively focussed the radiation pattern, narrowing the beamwidth in both range and azimuth. The resolution of this system was given by the dimensions of the beam footprint (see earlier equations). The range resolution, however, could be improved by time sampling (or "range gating") the received signal; the azimuth resolution could only be increased by expanding the array's azimuth dimension.

To eliminate this problem, researchers developed a method for synthesizing a large antenna by recording a target's "history" as the physical antenna moves along the orbit. This target history, in effect, represents the returned signal that would be recorded by a larger physical antenna having an azimuthal length,

$$
\begin{equation*}
L_{e f f}=v T_{D} \tag{2.5}
\end{equation*}
$$

where $v$ is the platform velocity, and $T_{D}$ the dwell time (time interval in which the target falls within the footprint of the antenna).

Earlier, in the discussion of radar imaging geometry, azimuth resolution was given by the width of the antenna beam footprint (Equation 2.2); however, one must modify this for SAR applications. With conventional radar, antenna beamwidth only effects reception (i.e. scattering area is illuminated once, and any backscattered return within the beamwidth of the antenna is recorded). However, with SAR, angular "selectivity" applies to transmission as well (the target is illuminated by a succession of bursts; see Figure 2-5), and the phase difference over the two-way path determines the antenna pattern. The effective azimuthal beamwidth for SAR is approximately

$$
\begin{equation*}
\beta_{a, e f f}=\lambda / 2 L_{e f f}, \tag{2.6}
\end{equation*}
$$

and the azimuthal resolution becomes

$$
\begin{equation*}
\delta_{a}=\beta_{a, e f f} R=\frac{\lambda R}{2 L_{e f f}} \tag{2.7}
\end{equation*}
$$



Figure 2-5: SAR target imaging geometry.

In the final step, one reconsiders Figure 2-5to equate the synthetic aperture length, $L_{e f f}$ to that of the physical antenna, $D_{a}$. From the geometry,

$$
\begin{equation*}
L_{e f f}=R \lambda / D_{a} . \tag{2.8}
\end{equation*}
$$

By combining the last two equations, one solves for the azimuthal resolution,

$$
\begin{equation*}
\delta_{a}=\frac{\lambda R}{2} \frac{D_{a}}{R \lambda}=\frac{D_{a}}{2} . \tag{2.9}
\end{equation*}
$$

This well-known result, which is derived more vigorously elsewhere (Cutrona, 1970), shows that azimuthal resolution for SAR is independent of range and of the synthetic aperture length. More profoundly, smaller low-gain physical antennae produce finer resolution with one obvious constraint: very high transmitter power is required for target detection.

Range resolution is derived from time sampling the returned signal with respect to a the transmitted pulse width. This range gating technique provides a range resolution,

$$
\begin{equation*}
\delta_{r}=c \tau_{e f f} / 2 \tag{2.10}
\end{equation*}
$$

where $\tau_{\text {eff }}$ is the effective transmitted pulse width of the radar.
Signal processing techniques, generally labeled pulse compression, increase this resolution by decreasing $\tau_{e f f}$. Commonly, the SAR system designer accomplishes this through linear frequency modulation; if the transmitted pulse is linearly frequency-modulated, or "chirped",

$$
\begin{equation*}
\tau_{e f f}=\frac{2.8}{\alpha \tau} \tag{2.11}
\end{equation*}
$$

where $\alpha$ is twice the FM rate (Cutrona, 1970). Other techniques, including non-linear FM and phase-coding may be employed.

Currently operational spaceborne SARs resolve to tens of meters both in azimuth and in range; however, rugged terrain can severely effect the actual ground resolution. To illustrate this, one may consider the ground resolution for inclined planar surfaces. Figure 2-6 shows two cases: (a) a surface sloping away from the radar and (b) a surface sloping toward the radar.

Relative to a flat (uninclined) surface, slopes facing away from the SAR exhibit one of two resolution distortions - elongation, or in more severe cases, shadowing. Ground resolution is given by the relation:

$$
\begin{equation*}
\delta_{g}=\delta_{r} / \cos \theta_{l o c a l}, \tag{2.12}
\end{equation*}
$$

(a)

(b)


Figure 2-6: The effect of surface slope on ground resolution.
where $\theta_{\text {local }}$ is the local incidence angle. Elongation, or "stretching" of the imaged surface occurs whenever $\theta_{\text {local }}$ exceeds the incidence angle (for a flat surface) but remains less than $90^{\circ}$ (left hand side of Figure 2-6a). For these slopes, the ground range resolution approaches $\delta_{r}$. Shadowing occurs if the local incidence angle exceeds $90^{\circ}$ (right hand side of Figure 2-6a). These slopes are occluded and produce no return.

Similarly, two types of distortions, one more severe than the other, exist for slopes facing the radar. Foreshortening decreases ground resolution by increasing the scattering area that is sampled within one range resolution cell (Figure 2-6b - right). As the local incidence angle approaches zero, the incline becomes parallel to the microwave fronts, and the whole section of slope images at one travel time (Figure 2-6b - middle). As $\theta_{\text {local }}$ goes negative, the surface becomes overlayed - backscatter from the top of the slope arrives before scatter from the bottom. Layover also provides an opportunity for multiple scattering (Figure 2-6b - left), or "double-bouncing". Resolution in this region loses meaning.

Figure 2-7 presents an actual SAR image of Mt. Shasta to demonstrate these distortions. The flight path of the SAR platform is shown, confirming the poor ground resolution (foreshortening) on the slope facing toward the radar and the increased resolution (elongation) on the slope facing away.

### 2.3 Radiometry

In closing the preceding section with a discussion on the geometry of resolution, I stopped short of addressing a closely related topic - the effect of geometry on radar backscatter. Before reaching this point, and tying this chapter together, I will provide an outline of SAR radiometry.

### 2.3.1 The Radar Equation

The radar equation provides a simple description of the factors influencing radar return. The relationship may be posed in two ways: a sequential factorization of the elements involved in transmitting a pulse, scattering it at the imaged surface, and receiving the returned signal; or as a multiplication of radar instrumentation characteristics and those of the scattering surface.

The first form (Skolnik, 1970) gives the received signal power $P_{r}$ as

$$
\begin{equation*}
P_{r}=\frac{P_{t} G_{t}}{4 \pi R^{2}} \cdot \frac{\sigma}{4 \pi R^{2}} \cdot A_{r} \tag{2.13}
\end{equation*}
$$

The radar transmitter radiates an amount of power $P_{t}$ through a SAR antenna having a gain $G_{t}$. The denominator of the first factor accounts for spherical spreading of the transmission with range $R$ (in meters). Therefore, the first term as a whole defines the power density at a distance (or range) $R$ from the SAR antenna. The second factor identifies the amount of power per square meter that is both scattered by the surface and returned toward the radar. Again, a spherical spreading factor is required in the denominator. The numerator involves an allencompassing variable, having units of square meters, which sums up the scattering properties of the surface; this is referred to as the radar target cross section $\sigma$. Finally, a portion of the returned power $P_{r}$ is intercepted by the SAR antenna and is passed along to the receiver. The effective aperture of the antenna is given by the term $A_{r}$ in square meters and is related to the antenna gain $G_{t}$ as

$$
\begin{equation*}
A_{r}=G_{t} \lambda^{2} / 4 \pi, \tag{2.14}
\end{equation*}
$$

where $\lambda$ is the electromagnetic wavelength.
The second manner in which to factor the radar equation,

$$
\begin{equation*}
P_{r}=(\text { instrument factor }) \cdot(\text { scatterer factor }), \tag{2.15}
\end{equation*}
$$

allows one to separate the controllable parameters from those which vary according to the imaged surface properties.

### 2.3.2 Instrument Parameters

The primary SAR instrument parameters are the transmitted power $P_{t}$, the antenna gain $G_{t}$, and the radar wavelength $\lambda$ as described above. These terms combine to form the instrument factor of the simple radar equation.

An important characteristic of the SAR imaging which does not enter into the simple radar equation is noise. Noise comes in two forms - additive and multiplicative. Noise produced by the electrical system components (thermal noise) and by voltage accumulation in the antenna sidelobes (Frankot and Chellappa, 1987) combines to bias the return power. Multiplicative
noise, often referred to as clutter or speckle, affects more strongly the total noise power. However, since speckle originates from coherent interference (owing to rough surface scattering) in the return from a resolution cell, it may looked upon as a feature of the surface rather than a measurement uncertainty (George, 1980). These noise sources may be incoherently smoothed by increasing the pulse repetition frequency so that a given resolution cell is sampled (or "looked" at) more than once. Naturally, this also beats down the electrical (thermal) noise, which adds random power to the system (Blake, 1970). Thermal noise power $P_{n}$ is given by the equation

$$
\begin{equation*}
P_{n}=k T_{s} B \tag{2.16}
\end{equation*}
$$

where $k$ is Boltzmann's constant, $T_{s}$ the system noise temperature (derived empirically through testing electrical system components), and $B$ the system bandwidth ( $B=1 / \tau_{\text {eff }}$ ).

The signal-to-noise ratio of a SAR system is simply the ratio $P_{r} / P_{n}$. To achieve suitable signal-to-noise in SAR imagery, one desires a bandwidth as large as possible to keep noise to a minimum; but certainly more freedom exists in obtaining a strong return signal power $P_{r}$. A radar system may be designed with a maximum (far) range in mind such that at the far range the signal-to-noise ratio falls below a predetermined threshold.

### 2.3.3 Range and Scattering Area

Since the instrumentation parameters are known for a particular mission, one is solely concerned with how the return power is modulated by the surface properties within the image swath. The macroscopic reflectivity of the surface is controlled by its shape (i.e. topography). The microscopic reflectivity arises from a number of factors which are collectively named the normalized radar backscatter cross section $\sigma^{\circ}$; the normalized cross section is equal to $\sigma$ in the radar equation divided by the area of the scattering target $A_{t}$.

Topography warps the SAR image by controlling target range and scattering area. Though technically a geometric effect, the distortions characterize themselves radiometrically, i.e. show up as variations in backscatter on a typical SAR image. The key relationship is that within an image, ground resolution is equivalent to the scattering area, which in turn is proportional to the return power:

$$
P_{r} \propto \sigma
$$

$$
\begin{align*}
& \propto \sigma^{\circ} A_{t} \\
& \propto \sigma^{\circ} \delta_{g} \delta_{a} . \tag{2.17}
\end{align*}
$$

This characteristic of radar is actually a feature of any active imaging system (e.g., seismics), but it sets radar apart from most other remote sensing systems. To emphasize this point, one may contrast SAR with optical imagery, where surface resolution depends on viewing angle. In a Landsat image, or even a basic photograph, object resolution increases as one moves from viewing the surface obliquely toward viewing normal to the surface. The opposite holds true for radar. Of course, the actual angular resolution of the optical camera is fixed, just as the slant range resolution of SAR is fixed.

To summarize, slopes facing toward the radar have poorer resolution than, and therefore appear "brighter" than slopes facing away from the radar.

### 2.3.4 The Normalized Radar Cross Section

What happens to microwave energy when it is scattered by a surface is summarized by the normalized radar cross section $\sigma^{\circ}$. Geologists, agriculturalists, and biologists have devoted much effort to determining the components of this variable, and the most conclusive result suggests that these components are many and that $\sigma^{\circ}$ has a large range of values. Although scattering scenarios have been found to fit theoretical models (e.g. densely vegetated terrain approximates an ideal volume scatterer, and fine-grained sedimentary cover provides a good specular scatterer), almost all surface types combine several scattering properties.

Scattering may be divided into two elements: volume scattering and surface scattering. Surface scattering depends entirely on the roughness properties of the boundary layer and may be subdivided into two end-members - diffuse and specular scatter corresponding respectively to very rough and very smooth surfaces. Volume scattering often depends on subsurface properties like soil moisture, which controls the dielectric constant of medium and determines microwave absorption, and inhomogeneity of buried sediments or debris. Volume scattering may also originate from thick forest canopies. Generally, penetration depth distinguishes volume from surface scattering; whenever the penetration depth of the boundary layer exceeds zero, one expects some volume scattering component.

Further complicating the fact that many different scattering types exist, is the observation
that widely varying types may occur in the same region, (e.g. a given image swath). It is hardly a coincidence that the most interesting geological regions have the most complex scattering mechanisms. One can afford to make few assumptions about the normalized cross section without some a priori knowledge of the imaged region (i.e. ground truth).

### 2.4 Multiple Incidence Angle SAR

The data set used for this work is produced by special style of imaging called multiple incidence angle SAR. The objective of this approach is to image a given ground swath at different antenna orientations so that backscatter from the surface can be analyzed as a function of incidence angle. To accomplish this, a given region of the Earth's surface provides the objective image swath for several successive platform orbits. On each orbit the radar antenna is tilted to direct the beam toward the desired swath (see Figure 2-8).

The degree of topographic distortion varies with incidence angle. At low incidence angles, the antenna is pointed near-vertically to the Earth's surface. Elevation is roughly in the direction of wave propagation so that relief displacement is strong. At high incidence angles, the antenna is aimed obliquely to the surface and relief displacement is weak. Resolution is always better at higher incidence angles, but this results in poorer signal to noise with respect to low incidence angles. At very high angles, terrain shadowing may occur, while conversely at low angles layover tends to mar resolution. Interpreters of SAR imagery of rugged topography desire a medium angle ( $\sim 45^{\circ}$ ) as a suitable trade-off between signal-to-noise and resolution, avoiding the effects of shadowing and layover.

### 2.5 Synthetic Imaging

In keeping with the theory that one can't invert what cannot be properly modeled, the ground work for this thesis involved creating synthetic radar image products and comparing them to the real thing (i.e. SIR-B and Seasat). Specifically, I attempted to qualitatively identify features of the real images that were suitably reproduced in the synthetics, and to eliminate from consideration those features not amenable to modelling.


Figure 2-7: Multiple incidence angle SAR geometry.

### 2.5.1 Input Data

To generate the synthetics, I chose a site near the San Andreas/Garlock Fault Zone junction north of Los Angeles, California. Figure 2-9 details the area, which includes Pyramid Lake, with Interstate 5 on its northeastern side and the San Gabriel Fault Zone to the southwest. This location was imaged during the SIR-B mission (scene AM-098.20-017), owing to its proximity to the major fault zones. The rugged terrain ranges from $574 m$ downstream from the lake to $2437 m$ in the range to the southwest according to digital topography purchased from the USGS.

Digital terrain elevation model (DTEM) data and normalized radar cross section $\sigma^{\circ}$ information provide the two "ground truth" sets required for a complete simulation; in this case, however, I chose to perform only a partial simulation without $\sigma^{\circ}$ data. The DTEM used here has a grid resolution of 30 m and corresponds to portions of two 15 minute quadrangles, 'Liebre Mt." and "Black Mt.", as mapped by the USGS.

I excluded $\sigma^{\circ}$ data for two reasons, which are somewhat related. First, the normalized radar cross section is, as previously mentioned, an abstract property of the surface and cannot be easily ascertained. In certain cases (Saunders et al., 1980), (Frankot and Chellappa, 1987), a region may be classified by one or more categories of scattering properties. For example, one would create a map of the region and divide the surface cover into several classes for which the backscatter properties are known. "Thick forest canopy" or "sparsely vegetated dry soil" are examples of terrains for which backscatter models have been experimentally determined. However, the Pyramid lake site exhibits a wide variety of undefined scattering regimes, corresponding to different geological structures and vegetation covers.

Secondly, I wanted to determine how well topography alone influences the radiometric properties of the image. Knowing this allows one to decide between a purely geometrical inversion or a combination geometric/radiometric inversion. If topography dominates all other surface parameters in determining the backscatter, one would make use of amplitude data in the inversion. If vegetation, cultural features (buildings, roads), and other surface properties dominate topography, then amplitude data would be useful only in determining boundaries rather than predicting the surface shape. I will discuss this further in Chapter 3.


Figure 2-8: Pyramid Lake, CA test site for SAR data simulation. Blank areas in the lower left and lower right corners fall outside the boundary of digital elevation (DTEM) data. Topographic contours are shown in the background; notable surface features appear on the overlay. The entire region is imaged by SIR-B scene AM-098.20-017.


Figure 2-8: Pyramid Lake, CA test site for SAR data simulation. Blank areas in the lower left and lower right corners fall outside the boundary of digital elevation (DTEM) data. Topographic contours are shown in the background; notable surface features appear on the overlay. The entire region is imaged by SIR-B scene AM-098.20-017.

### 2.5.2 Procedure

The algorithm for synthesizing a radar image from topography comprises two main parts: range location and amplitude determination. For each geographic location (or grid point) in the DTEM, the distance to the shuttle orbit and time of imaging are calculated (using ephemeris data). This provides the range and azimuth coordinates of the ground target in the synthetic image. The amplitude of the return is then computed for this range and azimuth. If the normalized radar cross section, $\sigma^{\circ}$ is set to 1 , then the return power is a function only of scattering area, which may be computed from the incidence angle and local slope (of the surface) as shown in Figure 2-10:

$$
\begin{align*}
P_{r} & =C A  \tag{2.18}\\
& =C L_{a} L_{g}  \tag{2.19}\\
& =C \frac{L_{g r i d}^{2}}{\cos \phi_{a} \cos \phi_{g}}, \tag{2.20}
\end{align*}
$$

where
$C$ equals the product of the instrument parameters (see previous discussion),
$A$ is the scattering area of the target,
$L_{a}$ is the azimuth dimension of the target, and
$L_{g}$ is the range dimension of the target.
The latter two parameters are rewritten in terms of $\phi_{a}$ and $\phi_{g}$, the local slope in azimuth and range respectively, and the grid spacing $L_{\text {grid }}(30 \mathrm{~m})$.

For most surface geometries, the target's range dimension is actually controlled by the radar range resolution $\delta_{r}$, and by the local incidence angle $\theta_{\text {local }}-\phi_{g}$, rather than by the local slope:

$$
\begin{equation*}
L_{g}=\delta_{r} / \sin \left(\theta_{\text {local }}-\phi_{g}\right) . \tag{2.21}
\end{equation*}
$$

In fact, the local slope in the range direction only acts to limit the return when the denominator of the preceding equation becomes too small, i.e. when the wavefront is parallel to the local surface. Because of this, $L_{g}$ is calculated both ways, and the smallest value is chosen.


Figure 2-9: Scattering area geometry. See text for explanation.

This last point leads to another important consideration: for very steep slopes facing toward the radar, several grid points of the same azimuth may also have approximately the same range; they fall into the same resolution cell. In these cases the return powers corresponding to each grid point are added and their sum assigned to the cell.

Finally, the return power at each cell location in the synthetic image is converted to a $D N$ ( Digital Number $=\sqrt{P_{r}}$ ), in accordance with real SAR image processing, and a synthetic slant range radar image results. The real and synthetic images for the Pyramid Lake site are shown in Figure 2-11.

In the creation of the synthetic images shown, no allowance for spherical spreading of the microwave energy is made. This simply keeps the DN bandwidth to a minimum and provides for more display contrast.

### 2.5.3 Comparison

Aside from the absence of noise in the synthetic example, one notices that the topography does strongly influence the return, but surface cover has an effect that cannot be easily separated. One can visually correlate many features, even those of fine resolution, between the two images. The ridges at the top and bottom of the images are well defined. However, the apparently diffuse scattering nature of the rough terrain produces a stronger return than predicted. The normalized radar cross section is greater than unity for this region. Conversely, the lake surface and valley (left center) have smoother surface cover and tend to scatter specularly; the return from these regions is less than predicted ( $\sigma^{\circ}<1$ ). It is also less, in general, than the return from the backsides of the ridges!

This forward modelling helps to direct efforts for inversion. The inability to closely predict backscatter using only topography suggests that simple shape-from-shading algorithms employing one image (incidence angle) will give suspect results for most terrain (exceptions exist (Frankot and Chellappa, 1987)). This leaves at least two possible paths which will be explored in the remainder of this thesis: the use of multiple incidence angle imagery to improve shape-from-shading results, and the use of MLA imagery with geometrical inversion (boundary matching).


Figure 2-10: Comparison of synthetic (a) and SIR-B (b) SAR data for the Pyramid Lake test site.

## Chapter 3

## Radar Inversion

Existing radargrammetric techniques take many forms: shape-from-shading, interferometry, stereo, and DN-matching; these are detailed in Chapter 5. Shape-from-shading is perhaps the simplest technique, deriving surface elevation information from the image DN, which correspond to backscattered power. Interferometry determines surface location by comparing backscattered phase information across an antenna baseline. Stereo-radargrammetry is an interactive method by which an interpreter visually matches features between a stereo pair of MLA images. DN-matching attempts to automatically emulate stereo-radargrammetry, performing two-dimensional cross-correlation between MLA images at coarse-to-fine resolution.

In studying these methods, one notes that only shape-from-shading and interferometry take advantage of the active nature of radar imaging. The other techniques are simply applications of photogrammetry to radar. Shape-from-shading uses range information along with a priori knowledge of the surface microwave reflectance to determine the surface shape from a single radar image view. Interferometry uses range and phase information to locate the surface. Stereo-radargrammetric algorithms match conjugate points between MLA images based on DN statistics. Only after the set of conjugate points pertaining to the imaged surface is produced, range information from the imagery is used to construct from these points the surface.

Stereo methods have proven successful, however, because regardless of the type of imaging, passive or active, visible surface features provide a wealth of clues about the surface shape. The human visual system doesn't need an interferometer to tell it the location and shape of an object!

The key in comparing the various techniques is to analyze the assumptions that each makes, and shape-from-shading arguably makes the most tenuous - that surface reflectance is known a priori. Surface reflectance depends on several parameters - roughness, moisture, and type of scattering (e.g. volume, surface, multiple). Not only do many types of surfaces exist, each with its own characteristic combination of these variables, but these surfaces coexist at a fine scale. Theoretically, there are regions of the Earth and of other planets where the scattering properties may fit a well known empirical law, and shape-from-shading can be applied with accurate results. In general, however, shape does not exclusively determine shading, and shading cannot exclusively determine shape.

If one accepts that shape-from-shading won't work, one also accepts, in effect, that topography cannot be conceived from one image; at least two images (preferably more) are needed. The fundamental requirement of any algorithm which uses MLA imagery to determine topography is that a given surface feature can be identified in each image. This seems obvious for stereo-mapping, but it applies to interferometry as well.

Interferometric systems simply compute the difference in phase, as received across the antenna baseline, of the radiation arriving from a specific point on the surface. The surface point is "imaged" at each of the receiving antennae. The longer the baseline between the two antennae, the more accurate the angular (and therefore, surface) resolution. For spaceborne applications, owing to the long slant range to target, the baseline length requirement suffers; launching and deploying such a system proves difficult. To circumvent this problem, Zebker and Goldstein (1985) , and later Li and Goldstein (1987) , described a method for using repeat orbits of Seasat to simulate an interferometer. (The details are discussed in Chapter 5.) First, they chose a surface site and assembled the images from each of the orbital passes. Then they "used a statistical correlation technique to determine the range and azimuth offsets between any two images." After completing this, they computed for each surface point the phase information necessary for the interferometry. This last step is redundant, however, because once the conjugate image points are determined, one is left with a simple stereo problem; the interferometry is only as good as the correlation permits. To use their words, "misregistration of the two images relative to each other in the coherent combination process can lead to severe degradations in the height measurement accuracies."

So, one needs to somehow identify surface features in each of the images, and then use these conjugate points to locate the surface. In attempting to assume as little as possible about the scattering properties of the surface, one must ask the following question, "What are the most reliable surface markers in (radar) imagery?" The answer at which I will arrive is "edges".

Before explaining my answer, I must first define what it is I call an edge. (A more general definition is given in Appendix A.) An edge (as viewed in SAR images) is the image representation of a surface boundary between two distinct scattering regimes. A shoreline provides an obvious example, where the smooth, specularly reflecting waters of a lake meet rough-scattering, vegetated land. Other examples include a fault exposure where different geologic units abut, or a boundary between field and forest. The contrast in scattering properties across the boundary manifests itself as an abrupt change in the amount power scattered back to the receiver. The result, as depicted in a radar image, is a "sharp" change in DN or pixel intensity. By "sharp", I specifically mean a local maximum in the DN gradient with respect to the direction normal to the boundary. Referring back to the shoreline example, the image of the shoreline would show a darker area (the water) abutting a lighter area (the land).

The argument for using edges to locate the surface arises from a comparison between edge information and actual DN information, which is normally used for radar inversion. Automated DN-matching algorithms operate on the assumption that radar-bright regions in one image of the MLA data set correspond to radar-bright areas in a conjugate image. This is usually the case for parallel, same-side imaging configurations, but this certainly does not apply to opposite-side or convergent observations. Secondly, DN-matching employs two-dimension cross-correlation between images to determine conjugate points, but this correlation is susceptible to the same topographic image distortions that it intends to remove. Some rationale for the use of edges rather than DNs is presented in this chapter. The final judgement can only come when the end results (DTEMs) of each method are evaluated; I offer this for future work.

### 3.1 Radar Migration

In defining radar migration, the work in this chapter plays to the discussion above. Since the primary development of this thesis is the application of migration, in the sense of seismic inversion, to radargrammetry, I first justify radar migration by analogy to seismic migration. Next,

I explore the important differences between the two applications and consider the assumptions that must be satisfied to successfully implement radar migration. "Raster migration" provides the context for this examination. Finally, I present an algorithm (Hierarchical Boundary Matching) for generating topography, based on the assumption that surface boundaries produce the most easily identifiable radar terrain signatures.

### 3.1.1 A Simple Seismic Analogy

The essence of migration, as presented here, is to project the wavefield backward from a receiver to recover the loci and associated amplitudes of all possible scattering points for that receiver (French, 1974; Schneider, 1978; Robinson, 1983; Stolt and Benson, 1986; Keho, 1986). In seismic applications, this "back-projection" is performed for many different receiver locations, and the geometrical shape of the subsurface scatterer is resolved by simply summing the amplitudes where the loci intersect.

Figure 3-1 diagrams the configuration/procedure. In performing the imaging, a linear array of receivers (depicted as triangles) is placed along the earth's surface. Generally, a source is buried at either end of the array and exploded, scattering seismic waves off of the subsurface and on to the receivers, where ground displacement is recorded with respect to travel time (time from blast). In this particular set-up, each receiver has a collocated source; the wavefield generated from the explosion of this, and only this, source is recorded at the receiver.

The initial wavefront generated from each source explosion travels a distance, shown by the vertical displacement of the circles below the surface, before interacting with the subsurface interface. The seismogram, or wavefield record, generated at the receiver is plotted above the surface, and scaled as half-travel-time (assuming unit velocity, this is equivalent to the depth scale). The amplitude peaks on each of the traces correspond to the half-travel-time of the reflected wavefront.

One can determine the position of the subsurface simply by fitting a line (labeled the "imaged subsurface") to the travel time data; however, this imaged interface does not correspond exactly to the actual position of the subsurface. To get the actual position of the subsurface, one may perform an elementary Kirchoff migration - backproject the wavefield from each receiver and stack the amplitudes to resolve the interface. The diagram depicts this as drawing circular arc


Figure 3-1: Schematic of seismic migration. See text for details.
centered at the source/receiver and having a radius corresponding to the half-travel-time of the wavefront. If enough collocated source/receivers are employed and the respective wavefields are stacked (i.e. summed), the maximum amplitudes should line up along the position of the subsurface.

### 3.1.2 Additional Constraints

At first glance, radar image restoration represents a simple form of migration, i.e., migration in a homogeneous medium with the Earth's surface as the scatterer; but complications to this straightforward method result from the sparse configuration of sources and receivers provided in radar imaging. For the Shuttle Imaging Radar-B (SIR-B) mission, designed specifically for multiple look angle imaging, one could expect up to four MLA images for a given ground swath. To make radar migration more closely analogous to seismic migration, the number of required shuttle passes per location would need to increase by at least an order of magnitude.

To demonstrate the difficulties of applying a simple Kirchoff migration to the spaceborne radar inversion problem, one may design an imaging/inversion example. Starting with a synthetic topographic profile, I simulate the imaging process by assuming several spacecraft passes in a direction perpendicular to the profile, so that the profile is imaged as one cross-track raster in each of the MLA images. The method for computing the brightness information for this raster is discussed in Chapter 2. Figure 3-2 shows the sample topography and the slant range images generated for several different look angles.

Once the images have been constructed, one performs the inversion by backprojecting each of the illuminations and stacking the results. Figure 3-3 shows the backprojections for four same-side look angles. In stacking the results, it is readily apparent that the surface is poorly resolved; this exposes another important difference between seismic and radar imaging - seismic imaging (and thus migration) depends greatly on specularly scattered waves, while radar imaging depends on backscatter (diffuse scattering). One could potentially employ a greater number of look angles in the imaging and subsequent migration to enhance the surface reconstruction. Figure 3-4 shows stack using a fuller set of look angles, but the results are still relatively poor.

To compensate for the small number of MLA images, on must more carefully scrutinize the amplitude information; that is, migrated amplitudes from the MLA imagery cannot simply be


Figure 3-2: (a) Sample topographic profile imaged at (b) $30^{\circ}$ and (c) $60^{\circ}$ illumination from the left, and (d) $30^{\circ}$ and (e) $60^{\circ}$ from the right. Panels (b) through (e) are slant range rasters with range along the horizontal. For panels (b) and (c), range increases from the left; range increases from the right for panels (d) and (e).


Figure 3-3: Slant range rasters for (a) $30^{\circ}$, (b) $40^{\circ}$, (c) $50^{\circ}$, and (d) $60^{\circ}$ look angles. Result (e) of backprojecting and stacking the data to resolve the surface. Panel (e) has ground range along the horizontal and elevation along the vertical.


Figure 3-4: Migration stack generated from images having look angles varying from $60^{\circ}$ left to $60^{\circ}$ right in $10^{\circ}$ increments.
summed, but instead must be compared somehow in order to determine the surface shape.
Two methods of performing this amplitude "scrutinization" are discussed in the remainder of this chapter. The first, which I worked on for approximately a year, is the two-dimensional migration problem illustrated previously with an added assumption about surface scattering (to be explained). The section devoted to this method was presented at IGARSS (International Geoscience and Remote Sensing Symposium) ' 87 in Ann Arbor, MI and is abstracted from those proceedings (Matarese et al., 1987). The second technique, hierarchical boundary matching (HBM), is based upon my belief that edge information is superior to amplitude information for use in the inversion. This refinement was prompted after my stay at Jet Propulsion Laboratory, Pasadena, CA, during the summer of 1987, where I gained a better working knowledge of the spaceborne radar imaging system.

### 3.2 Two-dimensional (Raster) Migration

Migration, like radar-stereo-mapping, takes advantage of the two-dimensional nature of the topographic distortions in radar, i.e. each cross-track raster in the radar image is a recording of a slice of topography, or topographic profile, perpendicular to the orbital (or flight) path of the spacecraft. In producing a MLA image set of a given ground swath, several nearly-parallel spacecraft overpasses are used (as described in Chapter 2). This configuration allows for the imaging of each topographic profile by one cross-track raster in each of the MLA images. If the azimuth resolution for each of the images is equivalent, the images may be "aligned" in azimuth (along-track direction) so that the only apparent distortions in the imagery are in the range direction. For images with good signal-to-noise this alignment is used for stereo viewing. (One may just put two of the MLA images side-by-side and cross one's eyes.) Figure 3-5 shows a cartoon illustration of the alignment procedure.

To make the radar inversion amenable to raster migration, alignment of the MLA imagery is critical; however, owing to the uncertainty in shuttle ephemeris, use of orbital parameters alone cannot adequately accomplish this task. A method for establishing this alignment through the use of ground control points (GCPs) to refine the orbit information is summarized by (Leberl et al., 1986b). This ephemeris correction, or a similar one, is required as a preprocessing step for the migration.


Figure 3-5: Azimuth alignment of MLA imagery.

After aligning the MLA imagery, the migration process begins. The steps are listed here with discussion to follow.

1. Corresponding (or conjugate) cross-track rasters are taken from each of the MLA images.
2. For each raster, the pixel DNs within the raster are normalized to span a common range of grey-levels.
3. Each of the conjugate rasters is backprojected into the object space.
4. The variance of the backprojected pixel values at each location in the object space is computed.
5. An optimization routine combines the variance data with surface shape constraints to determine the topography.

### 3.2.1 Receiver Array Orientation

The first step in the process, gathering conjugate cross-track rasters (which I will call simply "rasters" from here on) from the MLA images, is straightforward once the images have been properly aligned in the pre-processing. In gathering the conjugate rasters, one establishes the two-dimensional migration problem. The seismic analog to each of the rasters is a seismogram; each raster measures return power versus travel time just as each seismic trace records displacement versus travel time. The set of conjugate rasters is analogous to a linear array of geophones. One important difference, however, is the scattering surface resolution perpendicular to the receiver line. The topographic profile corresponding to the MLA rasters has an approximate width of $25 m$ (the azimuthal resolution of the radar), while the seismic resolution of a subsurface is much worse. (To apply radar antenna terminology, seismic sources have large beamwidths.)

### 3.2.2 Histogram Equalization

Histogram equalization describes the normalization of the DN in the MLA rasters to a common range of grey-levels ( 0 to 255 ). Since the brightness data has already been processed in generating the image product, this amounts to a re-normalization. Usually, this equalization is
carried out in the form of a linear stretch, whereby the maximum and minimum DN for each raster are determined, and new DN are calculated as follows:

$$
\begin{equation*}
D N_{n o r m}=\frac{\left(D N-D N_{\min }\right) \times 255}{D N_{\max }-D N_{\min }} . \tag{3.1}
\end{equation*}
$$

This portion of the procedure, though it seems harmless, attempts to derive surface information from the DN amplitudes. The normalization to a common range of DN serves to remove the differences in mean brightness between each of the rasters; one expects images having small look angles to be uniformly brighter than images having large look angles. (Remember that brightness decreases with increasing range, owing to spherical spreading, and that range is inversely proportional to the cosine of the look angle.) The normalization also attempts to identify a brightness value with each surface location, i.e. a resolution cell along the topographic profile will produce the same normalized brightness on each of the MLA rasters.

The backscatter assumption which histogram equalization attempts to take advantage of more precisely involves brightness relationships between locations along the topographic profile. If point $A$ images brighter than point $B$ at one look angle, then $A$ must return more power than $B$ at all other same-side look angles. Naturally, this will break down wherever layover effects exist, so one should use an MLA image set with a minimum look angle of, say, $30^{\circ}$. I'll refer to this as the assumption of look-angle independent relative return.

Since a linear stretch normalization proceeds slightly beyond this simple assumption, one can design a "scaling" procedure for performing the equalization more rigorously. The linear stretch also assumes that one knows the relationship between brightness and look angle to be linear. Unfortunately, this relationship is not only nonlinear, but it is unknown for most terrestrial surfaces. A better approach is to perform a ranking of the DN in each raster, assigning normalized DN on a percentile basis, i.e. if a given DN is brighter than $70 \%$ of the DN in the raster, compute its value as $0.7 \times 255$.

One difficulty relating to migration that these normalization techniques cannot address is that each of the corresponding MLA rasters does not exactly image the same portion of the topographic profile. Surface locations that are imaged in one MLA raster and not in the others will provide backscatter data that corrupt the normalization procedure. One must therefore attempt to remove this data, before normalization is attempted, so that each MLA raster corresponds to the same section of topography. Of course, this amounts to a restatement of the
problem one wishes to solve, and as a result, one can only hope that these corruptions are not significant.

### 3.2.3 Backprojection

The third step begins the actual migration process, using the refined orbital parameters of each pass to project conjugate rasters into the object space. The object space comprises a grid of elevation and ground range coordinates which act as a "window" containing the surface profile. The dimensions of the object space come from a priori knowledge of the imaged area - orbital parameters of the spacecraft, swath width, and approximate maximum and minimum terrain elevations. In effect, the object space is the smallest window which one expects to encompass the topographic profile; smaller windows lead to reduced computation.

As illustrated earlier in the seismic case, migration involves projecting a raster, which displays return power as a function of travel time, into the object space window. Specifically, for each pixel in the raster, its one-way travel time is used to place its DN onto that portion of a circular arc which intersects the object space. This backprojection is performed for each look angle so that each grid point in the object space window has a DN value from each of the MLA rasters.

### 3.2.4 The Variance Field

After the conjugate DN values have been assembled for each grid location in the object space window, one computes the sample variance; this provides the variance field. If the assumption about surface scattering properties holds, one expects points along the surface to have a low DN variance, although other positions in the object space, not coinciding with the surface profile, may have associated low DN variances also. The variance field only indicates that locations having low DN variance are more likely surface points than locations having high DN variance.

### 3.2.5 Surface Determination

Identifying the topographic profile of the scatterer from the variance field involves fitting a surface to the low DN variance values. An important constraint to this surface fitting routine applies a priori knowledge of the terrain slope characteristics. Since the angle of repose for
most geologic structures rarely exceeds $30^{\circ}$ from the horizontal, one must find a low DN variance profile meeting this criterion. In reality, some small scale surface features, e.g. cliffs or scarps, violate this rule, so this slope constraint will act somewhat like a high pass filter on the topography. The specific method I chose to accomplish this surface fitting was to apply a stochastic optimization technique which I discuss in Appendix B.

### 3.2.6 Results

To illustrate the raster migration procedure, I procured a digital terrain elevation model (DTEM) from the United States Geological Survey. The DTEM, as well as the procedure for generating synthetic SAR images, is described in Section 2.5. I then chose a topographic profile, a horizontal raster, from the image to test the migration algorithm.

In this example, the migration is performed using two MLA images, having look angles of $33.8^{\circ}$ and $50.1^{\circ}$ respectively. (Why I chose those angles I can no longer remember.) Figure 36 shows these MLA images. Using the raster from each MLA image which corresponds to the topographic profile, performing the backprojection, and computing the variance field, one arrives at a representation of the surface. Figure 3-7 shows the backprojected wavefields for each of the looks and the variance field (Figure 3-8) computed from the backprojections. After applying the surface fitting routine, I generated elevation data for the topographic profile which nearly matched the original elevations; this is displayed in Figure 3-9.

The artifacts in the results are caused by incorrect surface location. The program predicts the surface based upon low variance values; if these values appear at points in the object space that don't coincide with the surface, the program can potentially produce erroneous results. Low variance is a result of "mismatching" surface points based upon their associated DN values. As illustrated in Figure 3-10, one can imagine a profile containing two points, $A$ and $B$. If $A$ and $B$ have identical scattering properties and produce identical normalized DN values on the MLA rasters, then the migration may mismatch the backprojected DN value corresponding to $A$ in one raster to the backprojected DN value corresponding to $B$ in the other raster. The variance at points $A B^{\prime}$ and $A B^{\prime \prime}$, as well as at points $A$ and $B$, is zero, and all points are considered equally likely to lie on the profile. I will return to this problem later.


Figure 3-6: Synthetic SAR images generated from the Pyramid Lake DTEM, using bore angles of (a) $33.8^{\circ}$ and (b) $50.1^{\circ}$. Vertical axis denotes the flight direction; horizontal axis is ground range.


Figure 3-7: The object space window containing backprojected rasters is shown for the (a) $33.8^{\circ}$ and (b) $50.1^{\circ}$ looks. Elevation is in the vertical direction; ground range is in the horizontal direction. The variance field representation of the topographic profile is given on the next page.


Figure 3-8: The variance field expresses the likelihood of a point lying on the surface. This likelihood increases with decreasing pixel intensity, so that the surface appears as a dark curve.


Figure 3-9: A comparison of the surface locations calculated by migration (data points) and the digital topography (solid line). Data spacing is 30 m ground range and 6 m elevation (chosen arbitrarily). Data points at far left suffer from incomplete image data. Note the artifact at 4400 m ground range.


Figure 3-10: The mismatch problem for raster migration. See text for details.

### 3.2.7 Additional Comments

Unfortunately, this example only shows that given a topographic profile, and synthetic data generated from that profile which satisfies the scattering brightness assumption, one can invert for the surface; this method has yet to be tested on real data. The reason I am unable to apply raster migration to real data owes to the orbital geometry of the spaceborne SAR - no two MLA orbits are parallel enough to produce conjugate rasters of surface topographic profiles. The migration of spaceborne MLA images cannot be made two-dimensional. This method may still have potential for aircraft SAR image inversion, where parallel flight lines may be achieved.

Before I explain how to handle the problem of non-parallel orbits, I want to leave a few hints on alternative routes for raster migration (in case one chooses to apply it to airborne data). The fragile step in this method involves the histogram equalization; does the assumption about relative return I made earlier actually hold?

Noise certainly influences the answer. Images with low SNR will disrupt histogram equalization to some extent. It may be possible to use a hierarchical approach to reduce noise effects; one could start with a coarse resolution migration on smoothed raster data and determine the topography at a large scale. Then one could iterate to finer resolutions.

More importantly, however, a systematic error may occur when a particular surface harbors scattering properties which violate the relative brightness assumption. Conceivably, scattering surfaces exist for which return power may decrease over a range of decreasing look angles. A less extreme, and quite probable situation, might be a transposition of backscatter efficiency between two scattering regimes with a change in look angle. For example, let areas $A$ and $B$ neighbor each other. On image one, area $A$ might appear brighter than $B$, but on the other image $B$ might "outshine" $A$; this scenario probably isn't restricted to opposite looks (Leberl, 1979), though I know of no examples from the literature.

Other methods that may be used in place of the histogram equalization step include crosscorrelation and edge-detection. Cross-correlation between pairs of MLA rasters may be applied to take advantage of local relationships in surface scattering properties. Given two corresponding MLA image rasters $x$ and $y$, one first notes that each is simply a time series of return power.

The normalized cross-correlation between the the rasters is computed

$$
\begin{equation*}
c\left(\tau_{1}, \tau_{2}\right)=\frac{\sum_{t=-n / 2}^{n / 2} x\left(\tau_{1}+t\right) y\left(\tau_{2}+t\right)}{\left[\sum_{t=-n / 2}^{n / 2} x\left(\tau_{1}+t\right)^{2}\right]^{1 / 2}\left[\sum_{t=-n / 2}^{n / 2} y\left(\tau_{2}+t\right)^{2}\right]^{1 / 2}}, \tag{3.2}
\end{equation*}
$$

where
$\tau_{1}$ and $\tau_{2}$ are lags corresponding to rasters $x$ and $y$ respectively, $n$ is the number of pixels (surrounding $\tau_{1}$ and $\tau_{2}$ ) in each raster to be correlated, and $-1 \leq c\left(\tau, t_{0}\right) \leq 1$,

For same-side look angles, one calculates the lags for which $c$ is maximized, and then uses these lags to migrate for the location of the corresponding surface position. This computation is the basis for automated stereomapping algorithms, which are discussed in Chapter 5.

The MLA rasters may also be edge-detected, and the resulting "edge maps" migrated. The surface can then be resolved where these backprojected edges intersect in the object space. It is precisely this concept which I pursue in the next round.

### 3.3 Hierarchical Boundary Matching

Hierarchical Boundary Matching (HBM) incorporates three major concepts gleaned from my work with raster migration:

- the assumption that boundaries are reliable surface markers;
- the use of migration-style matching to reconstruct the location of these boundaries based upon edge information in the MLA imagery; and
- a hierarchical iteration to locate coarse-resolution boundaries first and then successively finer-resolution features.


### 3.3.1 Chronology

The driving force toward the development of the HBM method was my desire to tailor the radargrammetric inversion to spaceborne imaging SAR. Spaceborne platforms, such as Seasat
and the Shuttle Imaging Radars, have proven successful for large-scale mapping of of terrestrial and oceanographic surfaces. Though both programs were aborted, owing to equipment failures (to put it mildly in the case of SIR-B), they provided tantalizing imagery of the Earth's surface and prompted the design of a sophisticated orbital SAR to occupy a spot on the future EOS platform. Additionally, results from the Magellan mission may someday spawn an EOS-type orbital platform for Venus. Any useful radargrammetric techniques must allow for the imaging geometry of satellites.

This places raster migration at a dead end, since the laws of orbital mechanics prohibit exactly parallel overflights. Relief displacement in SAR imagery operates in a direction perpendicular to the ground track of the radar platform. This is an inherent feature of SAR design, and it cannot be manipulated in post-processing (Leberl, 1979).

One remaining alternative involves the backprojection of MLA image data from the nonparallel sub-tracks into a three-dimensional object space. The object space grows from a window, in the raster migration scenario, to a volume, but the same backprojection principles apply. The user supplies a range of elevations to encompass the topographic surface, which now has latitude and longitude dimensions instead of the simple ground range dimension of a topographic profile. The algorithm then attempts to resolve the surface by matching the conjugate features of the MLA imagery in the three-dimensional object space.

As one might expect, this process has the potential for computational nightmarishness, lending importance to any pre-processing step which can reduce the amount of data provided in the MLA imagery. Histogram equalization accomplishes this type of filtering in the preparation of the raster DN values for raster migration; however, this normalization becomes more difficult to apply when backprojecting image data into a three-dimensional object space. Figure 3-11 illustrates this problem. The ground track (azimuth) directions of two MLA radar passes are given by vectors $P_{1}$ and $P_{2}$. Both passes image the target $A$. However, $P_{1}$ images $A$ as part of surface profile $S_{1}$, which lies perpendicular to its nadir. Likewise, $P_{2}$ images $A$ as part of $S_{2}$. Because of the angle (or skew) between the ground tracks of the two orbits, the surface profiles $S_{1}$ and $S_{2}$ intersect. One cannot reduce the inversion to two dimensions, as in raster migration, because MLA rasters corresponding to a single profile cannot exist. In the raster migration case, the histogram equalization succeeds because all of the return power in each of the MLA


Figure 3-11: The difficulty in performing histogram equalization to prepare MLA image data for migration into the object space volume. See text for details.
rasters originates from the same topographic profile; for spaceborne MLA images, this is no longer the case.

Of course, one can attempt to equalize each MLA image in its entirety, but two problems arise. Surface regions imaged on one pass may not have been imaged on the others. This introduces many DN values in each image that do not correspond to those of the other images; these "exclusive" values disrupt the normalization. Secondly, the illumination vector for each MLA image is different; surface slopes "seen" in one image are not the same as those "seen" in the others. This problem can be corrected for in the post-processing of the imagery because the illumination direction is somewhat controllable by way of squinting (SIR-C Science Plan, 1986). Histogram equalization of the MLA radar imagery and subsequent 2-D cross-correlation form the basis of DN-matching automated stereo techniques.

Edge-detection, or the identification of surface boundary features in the MLA imagery, offers another solution to this problem by reducing spatial information to binary point data, i.e. a given location in the image either does or does not occur along a boundary between distinct scattering regimes. Automated edge-detection turns an image into an edge map, a binary representation of all of the abrupt DN changes in the image. (The edge-detection algorithm used in this thesis was developed by Canny (Canny, 1983) and is summarized in Appendix A.) Therefore, as histogram equalization prepares MLA rasters for migration into the object space window, edge-detection prepares MLA images for backprojection into the object space volume.

The hierarchy of resolutions used in HBM also helps to speed the inversion. As I will show presently, merely creating an edge map (by means of the edge-detector) of a full-resolution SAR image still leaves a glut of information for the boundary matching algorithm to process. To expedite computation, the algorithm first matches coarse-resolution edges between the MLA images to locate their corresponding surface boundary. After generating digital elevation data for these large-scale boundary features, the program moves to finer-resolution edge information to extract additional surface boundary locations/elevations, filling in the DTEM with each iteration.

### 3.3.2 The Algorithm

HBM is marked by five major sub-processes which combine to derive surface information from multiple look angle imagery.

Hierarchy - Creation of multiple resolution data sets.

Edge-detection - Demarcation of the image edges.
Pre-filtering - Selection of edge information to migrate. (Start of coarse-to-fine resolution processing loop.)

Automated Migration - Identification and matching of conjugate edges to locate the position of associated surface boundaries.

Post-filtering - Application of slope criterion to surface location data to remove anomalous information.

I broadly divide the method into preparative (steps one and two) and iterative (remaining steps) domains. The creation of the multiple resolution data sets from the MLA images and the subsequent imaging of this multiple resolution data simply transforms the data from MLA images to MLA multiple resolution edge maps for use in the pre-filtering. The last three steps equate to the general "pre-filter, backproject, and post-filter" approach of raster migration. This portion of the algorithm resolves the surface boundaries from the MLA edge maps, iterating from coarse to fine resolution. The flow of the algorithm through each of these steps is diagramed in Figure 3-12.

For the remainder of this section, I will assume a same-side MLA configuration for the SAR data. Indeed, all of the MLA radar images collected thus far have conformed to this geometry. Difficulty in performing the required orbital maneuvers has precluded opposite-look imaging efforts, although future missions may allow for this possibility. I will also assume that one's choice of look angles must lie in the range $30^{\circ}$ to $60^{\circ}$. Between look angles, one can assume that MLA imagery will remain clear of large-scale shadowing and layover effects owing to rugged topography with steep slopes (see Chapter 2).


Figure 3-12: Flowchart of the Hierarchical Boundary Matching (HBM) algorithm.

### 3.3.3 Automated Migration

Instead of proceeding through the algorithm in a step-by-step manner, I'll start at the core of the method, the automated backprojection routine. For the purposes of this work, I modified a program originally developed at JPL (Curlander, 1984) for projecting slant range SAR images onto the Earth's surface. This routine uses the orbital ephemeris of the radar platform in conjunction with existing digital terrain data for the image swath to convert the slant range SAR image to a ground range (latitude/longitude) representation. In the absence of DTEM data, the terrain surface is assumed to be at sea level, and image distortions appear wherever the assumption fails. This process is described in Chapter 2.

To illustrate how this backprojection routine is incorporated into the migration, one may consider the following example:

A corner reflector is placed in the middle of a (virtually) specularly scattering terrain having minimal relief.

The local relief varies between a minimum elevation $h_{\min }$ and a maximum elevation $h_{\max }$, but is not known exactly for the location of the corner reflector.

MLA imagery of the region shows the strong return of the reflector, surrounded by weak, uniform return from specularly scattering surface.

Given the slant range coordinates (azimuth, range) of the reflector in image $A$ and the orbital ephemeris for image $A$, one can calculate the ground position as a function of elevation for the reflector, based upon the range of possible surface elevations, $h_{\min }$ through $h_{\max }$. One repeats this step for the other images.

If the ephemeris and imaging parameters are accurate, the calculations of ground position versus elevation for the MLA images should produce one common result which identifies the location of the reflector on the surface. Figure 3-13 illustrates this procedure, assuming three MLA images having parallel orbital paths (for simplicity of drawing).


Figure 3-13: Location of a corner reflector (surface point) using migration. Scene (a) shows the topographic surface containing the corner reflector at $A$; the slant ranges $-r_{1}, r_{2}$, and $r_{3}$ - from target to spacecraft are recorded during imaging. This information, along with a priori constraints on target elevation $h$ and on geographic position $x, y$ are later used to invert (b) for target position. In the inversion, $P_{i}(x, y, \mathfrak{k})$ represent arcs of constant radius $r_{i}$ from the associated shuttle pass (these are arcs - not spheres - owing to the Doppler constraint). The position of the corner reflector is resolved where the $P_{t}$ intersect.

## Resolution

One also notes that the resolution, in elevation and ground range, at which the reflector is located improves with increasing look angle disparity, $\left|\theta_{2}-\theta_{1}\right|$. The theoretical formulation of surface resolution error as a function of the radar pulse resolution for two look angles is given by (Leberl, 1979).

$$
\begin{align*}
& \sigma_{h}=\sigma_{r} \frac{\left(\sin ^{2} \theta_{2}+\sin ^{2} \theta_{1}\right)^{1 / 2}}{\sin \left(\theta_{2}-\theta_{1}\right)}  \tag{3.3}\\
& \sigma_{g}=\sigma_{r} \frac{\left(\cos ^{2} \theta_{2}+\cos ^{2} \theta_{1}\right)^{1 / 2}}{\sin \left(\theta_{2}-\theta_{1}\right)} . \tag{3.4}
\end{align*}
$$

where
$\sigma_{r}$ equals half of the radar pulse (or slant range) resolution, assuming this resolution error is normally distributed.

### 3.3.4 Filtering

I have described HBM as an automated process that resolves the surface by migrating MLA image edge maps in much the same way that a human being can stereo-interpret an MLA image pair; however, while a human interpreter can easily identify spatial features - a pond, for example - in an MLA image pair, the computer only "sees" the edges corresponding to the pond's boundaries. The algorithm has no way of knowing a priori that these edges mark a pond shoreline; it must derive this information from other criteria. These criteria form the basis of the filtering routines, which attempt to ensure that HBM properly reconstructs the surface.

Before describing the theory and practice of filtering, however, I must first explain the two preparative steps in HBM - creating multiple resolution images and edge-detecting.

## Multiple Resolution Images

Creating a multiple resolution image set simply equates to pixel averaging. Starting with a 512 by 512 portion of a full resolution MLA image (pixel size $=(12.5 m)^{2}$ ), one can produce a 256 by 256 image of the same area by averaging 4 -pixel neighborhoods from the original image into single pixels of size $(25 m)^{2}$. Likewise, one can continue to produce coarser resolution images. An example of this pixel averaging is presented in Figure 3-14.


Figure 3-14: Creation of a multiple resolution image set. The original 512 by 512 image (a) is reduced to 256 by 256 (b), 128 by 128 (c), 64 by 64 (d), 32 by 32 (e), and 16 by 16 (f). With each reduction, the resolution of the image coarsens by a factor of 2 in both the range and azimuth directions.

## Edge-detection

After the multiple resolution image set is generated for each of the look angles, the imagery is edge-detected using the Canny algorithm (with $\sigma=1.0$ ) resulting in a multiple resolution edge map set. Implementation and sample results of the Canny edge detector are described in Appendix A. In this application, edge-detection is performed in the range direction only, so that the edge maps reflect abrupt changes in DN with respect to changes in range. In effect, the edge-detection is only sensitive to surface boundary orientations with a component in the azimuth direction. Another notable feature is that the separation between adjacent, azimuthally aligned edges can be as small as, but never less than, one image pixel. Localization by the edge-detector prevents edges from occupying neighboring range pixels. These last two points are illustrated in Figure 3-15. Figure 3-16 displays each of the images from Figure 3-14 after undergoing edge-detection. The "grey" edges mark abrupt increases in DN with increasing range while "white" edges mark abrupt decreases in DN with increasing range.

I edge-detected exclusively in the range direction because the MLA images have nearly parallel orbit paths. One can imagine a surface boundary running perpendicular to the orbit path (in the cross-track direction). This boundary would produce an edge oriented in the range direction on each of the MLA images. If these edges are migrated to locate the position of the associated surface boundary, the applicable range resolution $\sigma_{r}$ is proportional to the length of the edge rather than half of the pulse resolution (as mentioned earlier). Therefore, relative to azimuthally aligned edges, these range-oriented edges do not allow good resolution of their associated surface boundaries.

The effect of edge-detecting the images at multiple resolutions is to identify edge features corresponding to those resolutions; i.e. the edges in the coarse-resolution edge map correspond to large-scale boundary features on the surface, while the edges in the fine-resolution edge map correspond to the smallest-scale surface boundaries resolvable by the radar.

Looking at this fine-resolution edge map, one would not be surprised that in matching conjugate edge information between the several look angles, an automated algorithm will produce spurious results. Even a human stereo-interpreter would find the identification of conjugate features in a stereo-pair of edge maps to be an arduous task.


Figure 3-15: Alignment and localization of edges identified by the detection algorithm.


Figure 3-16: Edge maps corresponding to the multiple resolution imagery. The " 16 by 16 " coarse-resolution map is enlarged and shown at upper right.

## The Mismatching Problem

As alluded to earlier, in describing raster migration, mismatching presents the key obstacle toward achieving accurate surface location. In order to justify the pre- and post-filtering steps, I'll describe how the mismatch problem applies to HBM.

Simply stated, mismatching leads to the the calculation of incorrect surface positions. To explain this, I return to the raster migration scenario employing two edge-detected MLA images. As illustrated in Figure 3-17, a topographic profile is imaged by a cross-track raster in each of the two MLA images. The propagation (raster) orientations for each of the MLA images are labeled $p_{1}$ and $p_{2}$. In the forward model, point $A$, which

When the inversion is performed (Figure 3-17(d)), migrating the two rasters into the object space window, the surface is resolved wherever the backprojected edge data for each raster intersect. This intersection occurs at points $A^{\prime}$ and $B^{\prime}$ which correspond, respectively, to points $A$ and $B$ on the topographic profile. However, the backprojected edges also intersect at points $C^{\prime}$ and $D^{\prime}$, which do not correspond to actual surface positions. $C^{\prime}$ and $D^{\prime}$ are, therefore, incorrect locations which resulted from mismatching edge information. Specifically, the edge in raster $p_{2}$ corresponding to point $A$ intersects the edge in raster $p_{1}$ associated with $B$; since the algorithm identifies the surface based upon these edge intersections, it obediently (but wrongly) places the surface at point $C^{\prime}$. The false location $D^{\prime}$ is computed similarly.

## Pre-filtering

The HBM method relies upon two techniques for preventing the occurrence of mismatching. Obviously, the use of more than two MLA images helps significantly in that these additional images act to verify the surface locations determined in matching the initial pair. However, the only rigorous method for mismatch suppression is to enforce a minimum distance between the surface boundary locations to be determined. This concept is illustrated in Figure 3-18. If the geographic distance (in the cross-track direction) between $A$ and $B$ exceeds a certain value, $x$, then no mismatching can occur.

Since one has no control over the location of scattering boundaries on the surface, one must instead filter the edge information, which corresponds to these boundaries, as it appears in the MLA imagery. In the current example, the edges produced by $A$ and $B$ are separated by at


Figure 3-17: Imaging of a topographic profile (a) containing surface boundaries at positions $A$ and $B$. The orientation of two MLA image azimuths are assumed parallel so that the profile is imaged by one cross-track raster at each angle of incidence (panels (b) and (c)). Panel (d) shows the corresponding inversion.


Figure 3-18: Prevention of the mismatch problem. If the geographic separation between points $A$ and $B$ exceeds a certain value, $x$, no mismatching can occur. Panel (a) presents the topographic profile; panels (b) and (c) are MLA rasters generated from the topographic profile. The raster locations of edges corresponding to the surface boundary points $A$ and $B$ are shown. Panel (d) shows the result of migrating the rasters to identify the surface positions for $A$ and $B$. The circular arcs of the previous illustration are approximated by lines (since the distance from spacecraft to target is great).
least $x \sin \theta_{i}$ on each of the MLA rasters, where $\theta_{i}$ is the look angle for raster $i$. Therefore, one can migrate the rasters to uniquely determine the positions of $A$ and $B$. However, if another surface boundary $C$, producing an edge in the MLA imagery, lies between $A$ and $B$, such that both the distance between $C$ and $A$ and the distance between $C$ and $B$ are now less than $x$, then mismatching will prevent any of the points, $A, B$, or $C$, from being uniquely determined in the migration (Figure 3-19).

One can compute, for the purpose of pre-filtering, the minimum boundary separation distance, $x$, from a priori knowledge of the elevation spread for the ground region. For example, if one knows that elevations range from $h_{\min }$ to $h_{\max }$ for the swath imaged by the MLA data, one can safely assume that no mismatching will occur for a minimum boundary separation,

$$
\begin{equation*}
x>h\left(\frac{1}{\tan \theta_{2}}-\frac{1}{\tan \theta_{1}}\right)=h \frac{\sin \left(\theta_{1}-\theta_{2}\right)}{\sin \theta_{1} \sin \theta_{2}}, \tag{3.5}
\end{equation*}
$$

where
$h$ is the difference between a priori maximum and minimum elevations, $h_{\max }-h_{\min }$, and $\theta_{1}$ and $\theta_{2}$ are two of the MLA look angles.
(Note that this equation assumes that the radar is in the far field, i.e. arcs of constant range translate to lines. I will adhere to this approximation in the discussion and illustrations to follow.) Inspecting this equation, one notes that as $\theta_{1}$ approaches $\theta_{2}, x$ is minimized; but this is balanced by resolution considerations. Obviously, if the two look angles are equal, i.e. $\theta_{1}=\theta_{2}$, a degeneracy occurs - we have no point of intersection between the backprojected edges.

For example, given two MLA images with look angles $30^{\circ}$ and $60^{\text {circ }}$, one may use the above findings to calculate the mismatch-avoidance criterion. The minimum slant range $s_{30}$ between the edges in the $30^{\circ}$ raster should be

$$
\begin{equation*}
s_{30}=x \sin 30^{\circ}=h \frac{\sin 30^{\circ} \sin \left(60^{\circ}-30^{\circ}\right)}{\sin 60^{\circ} \sin 30^{\circ}}=h / \sqrt{3} . \tag{3.6}
\end{equation*}
$$

Similarly, for the $60^{\circ}$ raster, $s_{60}=h$.
To provide perspective, remember that the separation $s$ between adjacent edges, as detected by the Canny algorithm, may be as small as one pixel. For a full resolution, edge-detected image, the range dimension of one pixel is $12.5 m$ ! From Equation 3.5, unless the $h$ is also suitably


Figure 3-19: Addition of surface boundary point $C$. Panel (a) shows the location of $C$ on the topographic profile. Panels (b) and (c) show the appearance of the edge corresponding to $C$ in each of the MLA rasters. Panel (d) shows the result of migrating these edges. Note the occurrence of incorrect surface positions $C A$ and $C B$ owing to mismatching.
small (if it were, we would not need to determine the topography), mismatching will present a problem.

The pre-filtering procedure addresses this issue by using the multiple resolution edge maps, where the edge pixel separation (in meters) is greater. As stated earlier, the coarse-resolution edge maps show the large-scale boundary features in the image. In pre-filtering, one removes all edge information except that which identifies these boundaries. For example, say that in order to avoid mismatching, one must achieve an edge separation of 400 m . If the edge spacing in the full-resolution ( 512 by 512 ) edge map is $12.5 m$, one must appeal to the coarse-resolution ( 16 by 16) edge map which satisfies the 400 m spacing requirement.

My pre-filtering algorithm works by pinpointing at fine-resolution the coarse-resolution boundary. First, one determines the resolution which satisfies the separation requirement. This edge map is then used to "mask" the next-finer-resolution edge map to locate the boundaries more precisely. Iterative masking is performed until the position of the boundary on the full-resolution edge map is determined. Figure 3-20 illustrates this procedure. This pre-filtering produces edge maps, for each of the MLA images, which are ready for automated migration.

Before heading on to the subject of post-filtering, I make one final note about look angles. If one were given more than two MLA images having look angles $\theta_{1}$ through $\theta_{n}$, one would determine the surface resolution capability from the pair of look angles, $\theta_{i}$ and $\theta_{j}$, having the greatest disparity. However, the spacing requirement $s$ would be determined from the pair having the least disparity; this reduces the amount of necessary pre-filtering.

## Post-filtering

After migrating the first set of edge maps produced by way of pre-filtering, one has a reliable source of surface location data for the large scale boundary features. The next step is to begin "filling in" the data by deriving additional surface information for the finer-resolution features. This involves pre-filtering the edge map data again, but this time one starts with the next-to-coarsest-resolution edge map, rather than the coarsest one used in the first pass. One then migrates this new edge information, generating surface locations which must be combined with the existing data.

Post-filtering dictates the manner in which the new data is amalgamated. Because this data


Figure 3-20: The pre-filtering routine.
has the potential for mismatching (the edge maps from which it was derived do not satisfy the spacing requirement), it must be compared with the existing data to identify inconsistencies. This is accomplished by applying a slope criterion. For each new surface location, one can compute both the planimetric distance $x$ to each existing surface position and the corresponding elevation difference $y$. If $\phi=\tan ^{-1}(x / y)$ is greater than a certain angle (slope), the new location is thrown out (Figure 3-21); otherwise it is added into the database. Regional angle of repose serves as the basis for the slope criterion; in my experiments I have used an angle of $30^{\circ}$.

This "pre-filter, migrate, post-filter" loop may be repeated until all possible surface information has been extracted from the edge maps, however this proved to be beyond the computational limits of my CPU. The results that I was able to achieve are given next.


Figure 3-21: Post-filtering based upon the slope criterion. If $A$ and $B$ are known surface locations, then the post-filtering constraint allows only those positions (denoted by circles) which do not cause the surface gradient to exceed a reasonable angle of repose (in this case $\approx 45^{\circ}$.) Other computed surface points (denoted by $x$ 's) not meeting this requirement are ignored.

## Chapter 4

## Implementation and Results

For better or worse, I performed all of the coding and testing of this algorithm on the Earth Resources Laboratory's DEC VAXen, including a MicroVAX, an 11/780, and an 11/750. All of the routines, including the modified migration routine (originally written in FORTRAN), are written in the $C$ language for greater compatibility with the UNIX operating system used at the Earth Resources Lab. Assuredly, this environment is not optimum for handling imagery without the aid of an array processor, however it certainly promotes austere and efficient programming! A major breakthrough came in February (1988) when I figured out a way to reduce the computation time for my test data down to seven hours from roughly 100. (Up to that point, debugging had been a real horror story.)

As stated in the previous chapter, the algorithm comprises five steps, which equate to six distinct routines - one each for creating multiple resolution images, edge-detection, pre-filtering, and post-filtering; two for the migration. Pixel averaging, as well as general image processing, employed software developed by Peter Ford (Ford, 1985) of the Center for Space Research. Edge-detection was performed using an in-house implementation of the Canny algorithm; I had originally written this code for the identification of linear fault features in the SAR imagery. I also accept responsibility (or blame) for the filtering routines. By far, migration is the ratedetermining step for HBM. Its implementation is two-pronged: the modified JPL algorithm is used to calculate the possible surface positions (on a latitude/longitude/elevation grid) for each of the MLA edge maps, and then another routine searches the grid (object space volume) for edge intersections.

To test HBM, I chose a site on the northwest side of Mt. Shasta, CA, containing a basalt flow. Owing to the contrast in scattering properties between the basalt flow surface and the surrounding terrain, the boundaries of the flow are imaged clearly by MLA imagery. The basalt flow is also well defined on $15^{\prime}$ USGS topo maps (Hotlum and Juniper Flat quadrangles). The basalt flow spans approximately 7 km in length (NW to SE) and 2.5 km in width ( SW to NE ).

Between October 27 and 30, 1984, the Shasta region (including the basalt flow) was imaged by SIR-B as described in the following table:

| Data Take (Scene \#) | Azimuth <br> Angle | Center Incidence <br> Angle |
| :--- | ---: | ---: |
| AL-039.4 (019) | $118.0^{\circ}$ | $60.1^{\circ}$ |
| AL-055.4 (035) | $118.8^{\circ}$ | $53.5^{\circ}$ |
| AL-087.4 (022) | $120.4^{\circ}$ | $29.5^{\circ}$ |

Azimuth angles indicate the direction of the orbital ground track with respect to true north; one should note that these azimuths are nearly, but not precisely, parallel. Both the ground range imagery and accompanying listings of the orbital ephemeris are presented in Figures 4-1-4-6. In each image, a box surrounds the basalt flow area.

In the MLA imagery, one can clearly see Mt. Shasta, which rises to 4300 m in elevation, and the basalt flow, situated to the left of the mountain on each of the images. The images are all oriented with respect to the shuttle pass so that azimuth is along the horizontal and range is along the vertical. Owing to differences in the imaging parameters between data takes, the swath width varies. The SAR illuminated the images from the top, i.e. the platform passed to the northeast of the mountain. The image distortions caused by the extreme relief make the mountain appear to lean toward the radar and also affect the positioning of the basalt flow with respect to the mountain. The terrain distortions are especially apparent on data take AL-087.4, which has the smallest look angle.

In order to extract elevation data for the area surrounding the basalt flow, I localized a 512 by 512 region from each of the MLA slant range images. These full resolution subimages (Figure 4-1), along with the corrected shuttle ephemeris and an a priori stipulation that calculated elevations fall between sea level and 4800 m , provided the input for HBM

I use the phrase "corrected ephemeris" because ground control was required to improve


Figure 4-1: Ground range image for data take AL-039.4, scene 019.
－－ロッ 1


Figure 4－2：Imaging parameters for data take AL－039．4，scene 019.

 Kobrick VAX Gaussian Stretch ' User TGF/Box 30 .

Figure 4-3: Ground range image for data take AL-055.4, scene 035.


Figure 4－4：Imaging parameters for data take AL－055．4，scene 035.


Figure 4-5: Ground range image for data take AL-087.4, scene 022.


REMARKS: SLANT PANGE AND RECTIFIED IMAGE TAPEESUR SDFTWABE VERSIOM NMI.: N/A

Figure 4-6: Imaging parameters for data take AL-087.4, scene 022.


Figure 4-7: Full resolution slant range images centered about the basalt flow. All images are 512 by 512 pixels. (a) AL-039.4. (b) AL-055.4. (c) AL-087.4.
the ephemeris to the accuracy of a resolution cell. The sources of ephemeris error include inaccuracies in spacecraft tracking and orbital modelling, as well as perturbations caused by shuttle crew movement. In improving the ephemeris data for the region surrounding the basalt flow, however, these errors reduced to a simple translation of the shuttle position. For this example, I used five ground control points (GCP) to perform the correction. Since the ultimate goal of automated migration is to avoid using GCP at all, I did not analyze their effect on ephemeris improvement; this information is provided elsewhere (Leberl et al., 1986b). More accurate spacecraft tracking may soon eliminate the need for ground control.

Having corrected the orbital ephemeris, one can compute the expected resolution error in elevation and ground range using Equation 3.3. For $\sigma_{r} \approx 15 m, \theta_{1}=29.5^{\circ}$, and $\theta_{2}=60.1^{\circ}, \sigma_{h}$ and $\sigma_{g}$ are both $\approx 15 \mathrm{~m}$.

To compute the edge separation requirement $s, \mathrm{I}$ used $\theta_{1}=53.5^{\circ}, \theta_{2}=60.1^{\circ}$, and $h=$ 4800 m . From Equation $3.5, s=792 \mathrm{~m}$. Since the full-resolution image pixels have a range width of 12.5 m , this would require pixel averaging the 512 by 512 MLA imagery down to 8 by 8 images having $(800 m)^{2}$ pixels. However, these 8 by 8 images did not contain enough information to accurately edge detect, so I settled for 16 by 16 images with ( 400 m$)^{2}$ pixels and hoped to avoid mismatches. Figures 4-8-4-10 show the multiple resolution image sets generated for each of the MLA images. The multiple resolution imagery underwent edge-detection to produce the edge maps shown in Figures 4-11-4-13.

Using the 16 by 16 edge map as a basis, I pre-filtered the edge map data to identify only those edges associated with the large-scale surface boundaries. This filtering resulted in the edge maps shown in Figure 4-2. One notes that the edges correspond to the northeast and southwest boundaries of the flow.

After pre-filtering, I migrated the edge maps to locate the surface boundaries shown in Figure 4-3. For each data point shown in the figure, the elevation has been computed. When plotted on a properly scaled latitude/longitude grid, one can compare the shape of the flow as migrated with the shape as defined on the USGS topo map.

Finally, I filled in intraflow boundaries by migrating higher-resolution edge maps and amalgamating them with the previous data by way of post-filtering. These contoured results are presented in Figure 4-4.


Figure 4-8: Multiple resolution imagery of basalt flow for data take AL-039.4.


Figure 4-9: Multiple resolution imagery of basalt flow for data take AL-055.4.


Figure 4-10: Multiple resolution imagery of basalt flow for data take AL-087.4.


Figure 4-11: Multiple resolution edge maps of basalt flow for data take AL-039.4.


Figure 4-12: Multiple resolution edge maps of basalt flow for data take AL-055.4.


Figure 4-13: Multiple resolution edge maps of basalt flow for data take AL-087.4.


Figure 4-14: Edge maps marking the northeast and southwest boundaries of the basalt flow. (a) AL-039.4. (b) AL-055.4. (c) AL-087.4.


Figure 4-15: Comparison of migrated (top) and USGS-mapped (bottom) flow boundaries.


Figure 4-16: Contoured elevation data, produced by HBM (top) and taken from a USGS topographical map (bottom), for the basalt flow.

Owing to the unavailability of DTEM data (on order from JPL) at the time of this writing, I did not compute the rms elevation "error" of my data; however, I believe that such calculation is still premature. Qualitative comparison of the contoured data with actual USGS (Figure 4-5) topography clearly reveals the two primary results for elevation analysis.

- Given an object space volume spanning from sea level to 4800 m , HBM accurately locates the surface in the 1000 m to 1700 m range. The general slope of the basalt flow (i.e. sloping down toward the northwest) is also correctly determined.
- At the 100 m contour interval, distinct differences exist between the USGS and HBM contours.

Possible explanations for the second observation are detailed in Chapter 6.

## Chapter 5

## Related Work

The purpose of this chapter is to briefly describe and compare several other methods for generating topography data from radar images; they are:

- radarclinometry,
- interferometry, and
- stereo mapping.

The important questions one must ask in assessing these radargrammetric techniques include:

- What input data does each require?
- What assumptions does each make?
- How do the procedures work?
- What kind of results do they generate?


### 5.1 Radarclinometry

Radarclinometry (Frankot and Chellappa, 1987; Wildey, 1986), or "shape-from-shading", is perhaps the simplest approach in that it requires the least data; however, for this very same reason it is also the most difficult for which to formulate assumptions.

Shape-from-shading is a popular technique in the field of computational vision (Horn, 1986), for recovering the shape of objects having known reflectance properties. However, a reflectance
map is a two-dimensional representation of a three dimensional surface and is not sufficient to fully define the surface. To address this problem, one can obtain additional reflectance images of the surface for different illuminations and/or observations (stereo), or one can constrain the surface shape by way of continuity assumptions and boundary conditions.

The reflectance map, $R(p, q)$ (using the notation of Horn) can be generated for any given surface whose reflectance distribution is known. For example, a matte surface which reflects all incident radiation equally in all directions obeys Lambert's law of reflection,

$$
\begin{equation*}
L=\frac{1}{\pi} E \cos i \tag{5.1}
\end{equation*}
$$

where
$L$ is the scene radiance (reflectance),
$E$ is the illuminating source's radiance, and
$i \geq 0$ is the local incidence angle.
In order to relate this to the surface, $z(x, y)$, one may first define the unit surface normal,

$$
\begin{equation*}
\hat{\mathbf{n}}=\frac{(-p,-q, 1)^{T}}{\sqrt{1+p^{2}+q^{2}}} \tag{5.2}
\end{equation*}
$$

where

$$
\begin{aligned}
& p=\partial z / \partial x, \text { and } \\
& q=\partial z / \partial y .
\end{aligned}
$$

For a distant source, the unit illumination vector, $\hat{\mathbf{s}}$, is given by

$$
\begin{equation*}
\hat{\mathbf{s}}=\frac{\left(-p_{s},-q_{s}, 1\right)^{T}}{\sqrt{1+p_{s}^{2}+q_{s}^{2}}} \tag{5.3}
\end{equation*}
$$

where $\left(p_{s}, q_{s}\right)$ is the gradient of the incoming rays with respect to the surface coordinate system. Taking the dot product of the two unit vectors one derives an equation for the local incidence angle, which for Lambertian surfaces is related to the reflectance,

$$
\begin{equation*}
R(p, q)=C \cos i=C \hat{\mathbf{n}} \cdot \hat{\mathbf{s}}=\frac{1+p_{s} p+q_{s} q}{\sqrt{1+p^{2}+q^{2}} \sqrt{1+p_{s}^{2}+q_{s}^{2}}} \tag{5.4}
\end{equation*}
$$

where $C=E / \pi$.
Radarclinometry applies shape-from-shading to radar. For purposes of radargrammetry, one knows the illumination and observer locations (they coincide) just as before, and one assumes the surface reflectance properties based upon radar scattering theory and past experimentation.

Combining Equation 5.4 with surface constraints, one can use a radar image to solve for $z(x, y)$. For example, consider a topographic profile across a series of parallel ridges in a direction perpendicular to the ridge axes; one can choose this direction as the x -axis. Imaging (with SAR) the surface along this profile, one obtains a cross-track raster which equates to a reflectance profile ( $R$ no longer varies with $q$ ). This reflectance map is then converted to local incidence angle for each slant range pixel. For any two neighboring points on the surface starting with the ground control, one can compute the average $i$. The separation $d$ between the points is given by

$$
\begin{equation*}
d=r / \sin i, \tag{5.5}
\end{equation*}
$$

and using the look angle $\theta$, elevation $z=d \sin \theta$ and ground separation $x=d \cos \theta$ between the points are determined.

If reflectance is a function of $q$, as well as $p$, one must appeal to other forms of surface constraint (Wildey, 1986).

If asked to derive topography from only one SAR image, radarclinometry provides the means, however microwave reflectance properties of the surface, though possibly more uniform than those of other EM wavelengths, are still not well-determined for most terrestrial and planetary surfaces. Methods also exist for handling noise speckle in the radar images (Frankot and Chellappa, 1987); speckle poses yet another problem for shape-from-shading. However, a combination of radarclinometry and MLA inversion methods may prove highly synergistic.

### 5.2 Interferometry

In a sense, SAR interferometric methods bridge the gap between single image and MLA image methods. Interferometry involves the use of a second antenna, whether mounted on the same radar platform as the first or simulated by a near-repeat orbital pass. The former technique first appeared in the work of Graham (Graham, 1974), while the latter was attempted with the Seasat SAR (Zebker and Goldstein, 1985). Figure 5-1 diagrams the two methods.


Figure 5-1: Interferometric methods: (a) dual-antennae airborne; (b) Seasat repeat orbit. See text for description of parameters.

Interferometry measures the phase difference across the antenna pair of radiation backscattered by the surface. As shown in the figure, the two antennae are separated by a baseline $B$. For a given point $A$ on the surface, the path difference

$$
\begin{equation*}
\delta r=r_{2}-r_{1}=B \sin \theta, \tag{5.6}
\end{equation*}
$$

where $\theta$ is the look angle. The interference pattern of this configuration projects into space such that it's measured phase at point $A$ is

$$
\begin{equation*}
\phi=2 \pi \delta r / \lambda, \tag{5.7}
\end{equation*}
$$

where $\lambda$ is the radar wavelength. Since the vertical distance $h$ between the platform and point $A$ is given by

$$
\begin{equation*}
h=r \cos \theta, \tag{5.8}
\end{equation*}
$$

this can be rewritten

$$
\begin{equation*}
h=r \sqrt{1-\frac{\phi^{2} \lambda^{2}}{(2 \pi B)^{2}}} . \tag{5.9}
\end{equation*}
$$

In the case of Seasat, measurement of the interferometric phase required co-registration of the images produced on each of the two overpasses. Only after conjugate pixels, corresponding to a point on the surface, were determined, could the phase (and therefore the elevation) of the surface point be calculated.

I mentioned earlier that this seems redundant because given these conjugate pixels, one can determine the surface position stereoscopically; actually it exemplifies a synergy between interferometry and stereo. For nearly equal look angles, one achieves very poor resolution of the surface using stereo. The more precise nature of interferometric phase measurement, however, allows this method to excel. If one can determine conjugate points to a pixel accuracy using stereo, one can refine the surface location using interferometry. A similar "bootstrapping" method is used in global positioning (Hatch, 1982).

Zebker and Goldstein (1985) reported rms elevation accuracies of $6 m$ for airborne SAR and $5 m$ to $100 m$ for Seasat. Li and Goldstein (1987) attempted to identify the component of this error owing to phase measurement inaccuracies, and they theorized that an optimal baseline may exist. Unfortunately, no digital terrain data with accompanying rms elevation error estimates has not been easily generated, and this method requires more extensive consideration.

### 5.3 Stereomapping

The work presented in this thesis addresses one aspect of stereo-radargrammetric automation - namely, the use of boundary information to resolve the surface; another group of automated algorithms (Strong and Ramapriyan, ; Ramapriyan et al., 1986; Thomas et al., 1987) exist to attempt stereomapping from a different standpoint - DN matching. I presented arguments for the use of HBM rather than these methods, but again, a synergism exists between the two.

Automated DN matching implements a type of two-dimensional cross-correlation between MLA imagery. This correlation appears in the form of a match criterion; for example, Thomas et al (1987) use

$$
\begin{equation*}
F(k, l)=\frac{\sum_{i, j}|A(i, j)-B(i-k, j-l)|}{\left(\sum_{i, j} A^{2}(i, j)+\sum_{i, j} B^{2}(i-k, j-l)\right)}, \tag{5.10}
\end{equation*}
$$

where
$i$ and $j$ are range and azimuth pixel coordinates in the image reference window, and
$k$ and $l$ are (relative) range and azimuth pixel disparities between the reference window locations in each image.

The algorithm establishes a match ( $k, l$ ) for which $F$ is minimized, and uses this result to stereoscopically compute the surface position. To facilitate computation, a hierarchical approach is applied to this method also; i.e. matching is performed for coarse resolution (smoothed imagery) first and then performed for finer resolutions.

Although I know of no rms surface location errors that have been calculated for the results of these algorithms, Strong (personal communication) indicates, at least for Hierarchical Warp Stereo (HWS), that the computed elevation data appear "smoothed" relative to actual topography. He suggests that this effect is caused by the method's resolution limitations - for small reference window dimensions at high resolution, the DN statistics become insufficient to evaluate the match accurately. Apparently, either a boundary matching algorithm (HBM) or interferometry may aid in resolving to higher resolution.

## Chapter 6

## Conclusions and Future Work

The primary conclusion of this work, as shown by the success of HBM, is that edge-detection based, automated matching is well-suited for use in radargrammetry. Though the algorithm must (and shall) be applied to additional data in order to better determine its accuracy, surface location results calculated from the Mt. Shasta imagery agree qualitatively with USGS topography.

However, the HBM results, when compared with USGS terrain data, appear to exceed theoretical error expectations (Equation 3.3). Possible explanations include the following.

1. HBM limitations: mismatching errors which survive filtering.
2. Data limitations: (as described by Leberl (Leberl et al., 1986a))
(a) effects of "specular point migration" and migration of edges [owing] to differences in illumination;
(b) changes in backscatter [owing] to differences in incidence angles;
(c) variations in orbit azimuths and resulting edge migration; and
(d) effects of noise.
3. USGS topography inaccuracies: relatively small ( $\approx 5 m$ ) interpolation errors.

Quantifying these limitations necessitates three future tasks:

- comparison of HBM and stereomapping techniques on various data sets to evaluate HBM mismatching errors;
- comparison of mismatch-free HBM results with USGS DTEM to determine relative accuracy; and
- evaluation of HBM results with regard to geodetic satellite surveying.

Unfortunately, the Mount Shasta region affords one of the few locations for which accurate error analysis can be performed. A limited set of MLA imagery exists at present; this, coupled with the lack of accurate DTEM data for areas having rugged topography, precludes rigorous evaluation of HBM over various terranes. However, HBM results for Mount Shasta should be wholly applicable elsewhere, justifying further investigation.

## Appendix A

## Edge-detection

One of the primary features of a "typical" image is the edge. An edge is simply defined to be step in brightness values across a boundary as shown in Figure A-1. However, owing to the imperfect nature of imaging devices, including instrument noise, and "smooth' reflectance properties of particular surfaces, edges must be alternatively defined as local maxima in the brightness gradient at a given point in the image.

The task here is to "mark" or enhance these edges, and numerous discrete methods have been suggested for accomplishing this. An image strip, or brightness profile, is perhaps the easiest example to begin with. To locate the edges in this strip, one may convolve it (Figure A-2) with a differencing operator described by a template. The strip resulting from this convolution will have maxima corresponding to downsteps in the direction of convolution and minima corresponding to upsteps. Another method involves applying a second-derivative filter to the brightness strip and marking the edges at zero-crossings in the output.

Since images are generally two-dimensional, however, one must appeal to more complex operators for detecting and localizing edges. Marr \& Hildreth (1980) developed a popular method which first applied a rotationally symmetric Gaussian filter, followed by a Laplacian operator. Canny (Canny, 1983) later designed an operator, based upon the Gaussian, which held the optimum trade-off between edge detection and localization in the presence of noise. This rotationally symmetric operator featured orthogonal separability - i.e. it could be formed by successive convolution along columns and rows of an image array - and it contained an adjustable parameter $\sigma$ for detection of edges at different spatial frequencies. After convolving




Figure A-1: Convolutional edge-detection. An image strip (top) is convolved by a " $1,0,-1$ " template (middle), transforming the $D N$ boundary into a $D N$ gradient maximum or edge (bottom).
the image with this operator, one can directionally-differentiate the image to identify edges at various orientations. Figure A-3 presents a simple application of the Canny detector.

Because of the speckle noise inherent in SAR imagery, Canny's operator has been applied in a variety of radar image inversion problems. Matarese (Toksöz et al., 1986) first applied it to SIR-A/B images of active fault zone lineaments in eastern Turkey and southern California. This prompted its use in the Hierarchical Boundary Method for automated radargrammetry described in this report. Recently, scientists at JPL have successfully incorporated it into a motion-detection algorithm for studying ice flow movement in the arctic with airborne SAR.

Most research in edge-detection has come from the machine vision community; Horn provides an excellent summary of edge-detection and its applications in his book Robot Vision.

(a)

(b)

(d)

Figure A-2: A boundary between two regions of constant brightness in images (a) without noise and (b) with noise. Edge maps, (c) and (d), correspond to images (a) and (b), respectively. Edge map (d) resulted from the application of Canny's edge detector.

## Appendix B

## Surface Fitting

In my Chapter 3 discussion of the mismatch problem, I described the minimum parallax criterion and explained how it necessitated the pre-filtering of edge map data in HBM. Specifically, the range spacing between neighboring edges must exceed minimum allowable parallax, as computed from the multiple look angles and from a priori knowledge of maximum and minimum surface elevations. Pre-filtering of the edge map data, based upon the use of multiple resolution images, allowed this criterion to be met, resulting in the generation of accurate (mismatch-free) surface position data for the first HBM iteration. Subsequently, higher resolution edge map data not meeting the parallax criterion could be migrated during successive iterations and the resulting surface locations "fitted" to the initial data through post-filtering.

Surface fitting provides an alternative method to the hierarchical approach mentioned above; it instead ignores the mismatch problem until after the migration, sorting the correct surface locations from the incorrect ones as a stochastic, post-filtering process. Before abandoning raster migration for HBM, I made use of surface fitting for converting variance field data to a terrain profile. Though my implementation is far from rigorous, it presents an interesting example of probabilistic modeling, and I include it for completeness.

## The Stochastic Model

The variance field translates to a probability representation of the surface. As shown in Figure 38 , the low DN variance values correspond to likely positions of the surface while high DN
variances correspond to unlikely surface locations. Low variances do not insure that the surface passes through the associated grid point because of possible mismatching. Likewise, owing to noisy image data or sampling artifacts, high variances might actually appear at grid locations coincident with the surface. Qualitatively, however, the surface suggested by the variance field may be readily apparent (Figure 3-8).

Surface fitting, through an optimization technique, converts the variance field representation of the surface to an actual terrain profile. The object space is taken as a statistical model of the surface, specifically a one-dimensional Markov random field along the horizontal (ground range) axis of the object space window. The vertical, or elevation, axis supplies the possible states for the surface position. The goal is find a "smooth" surface passing through most-likely states. Surface fitting actually combines two probability fields, an exterior field given by the variance data and an interior field given by the smoothness of the surface to be determined.

The interior field applies a probabilistic slope criterion to the surface, i.e. the slope between neighboring points on the surface, as represented in the object space grid, should be minimized (topographic gradients rarely exceed $30^{\circ}$ ). Figure B-1 shows how the interior field is represented on the object space grid. In treating the object space window as a Markov chain, one assumes that the probability $P$ of the surface elevation $e$ at ground range $n$ equaling $e_{0}$, given the surface elevations at all other ground ranges, is the same as $P$ given the surface elevations at neighboring ground ranges $n-1$ and $n+1$.

$$
\begin{equation*}
P\left(e(n)=e_{0} \mid e(m), m \neq n\right)=P\left(e(n)=e_{0} \mid e(n-1), e(n+1)\right) . \tag{B.1}
\end{equation*}
$$

Therefore, a smooth surface, as defined here is one which minimizes the euclidean distance between the surface position at neighboring ground range locations.

This type of statistical representation compares with the Ising model (Kindermann and Snell, 1980) for predicting the alignment of a matrix of charged particles in the presence of a magnetic field. Each particle has two possible states, positive ( + ) and negative (-) spin. The interior field represents the influence of neighboring particle spins upon each other. The exterior field describes the magnetic field applied to the particles. A comparison of the Ising model with surface fitting is given in the following table.


Figure B-1: The interior field, as determined by a surface slope criterion. Rows $m$ through $m+4$ represent possible elevation states for the surface passing through columns $n$ and $n+1$.

|  | Ising Model | Surface Fitting |
| :--- | :--- | :--- |
| possible states | $2(+/-$ spin) | $n$ (elevations) |
| interior field | spin of neighbors | elevation of neighbors |
| exterior field | magnetic field <br> (uniform for <br> all positions) | variance field <br> (unique for each <br> ground range) |
|  |  |  |

The procedure which combines the field information to locate the surface profile is a modified Metropolis (Metropolis et al., 1953) algorithm like that used in simulated annealing techniques (Kirkpatrick et al., 1983). Simulated annealing treats the statistical model as a high "temperature" random field which, when cooled, crystallizes into a final state, determined by the global energy minimum. In the case of surface fitting, this final state is the terrain profile, which attempts to minimize the interior and exterior field energies by locating a smooth surface coincident with low DN variances.

Surface fitting requires two temperatures $T_{e}$ and $T_{i}$, which control the influence of the exterior and interior fields, respectively, on the annealing process. The exterior field temperature modulates the influence of the variance field likelihoods on the surface position. High $T_{e}$ dictates that the DN variances have little effect on the surface location, while low $T_{e}$ favors positioning of the surface at grid points having low DN variances. The interior field temperature applies similarly to the slope constraint: high $T_{i}$ mitigates the influence of the elevation at one ground range location on the elevations at neighboring ground range locations while low $T_{i}$ strengthens this local surface constraint.

Given these temperature controls the algorithm proceeds as follows, with object space col$u m n$ denoting the ground range position and row denoting the elevation.

1. Choose $T_{i}$.
2. Initialize $T_{e}$.
3. Initialize a random surface state.
4. While lowering $T_{e}$,
(a) Randomly choose a column.
(b) Randomly choose a new row for that column.
(c) If the new row has a lower DN variance than does the old row and if the new elevation (associated with that row) is nearer the elevations of the surface in neighboring columns,
move the surface position for that column to the new row.
(d) else,
move the surface position for that column to the new row, based on a computed probability $P$.

For each potential movement of the surface position from an old row (elevation) to a new row, one calculates an "energy" change, $\delta E$, which determines the probability of accepting the move. Since surface fitting consists of two criteria, the interior and exterior field, one has correspondingly two types of energy changes, $\delta E_{e}$ and $\delta E_{i}$. The change in exterior field energy is calculated as the difference in DN variance between the old and new rows. The change in interior field energy is calculated as the difference in the sum of euclidean distance to the surface position in neighboring columns.

$$
\begin{equation*}
\delta E_{i}(n)=\left(D^{\prime}(n-1)+D^{\prime}(n+1)\right)-(D(n-1)+D(n+1)) \tag{B.2}
\end{equation*}
$$

where,
$n$ is the column (ground range position) in question.
$D$ is the distance between the current surface elevation in column $n$ and the current surface elevation in either column $n-1$ or $n+1$.
$D^{\prime}$ is the distance between the new surface elevation in column $n$ and the current surface elevation in either column $n-1$ or $n+1$.

For $\delta E_{e} \leq 0$ and $\delta E_{i} \leq 0$, the move is accepted automatically. If either (or both) $\delta E$ are greater than zero, then one generates two random numbers, $r_{e}$ and $r_{i}$, between 0 and 1 . If

$$
\begin{equation*}
r_{e} \leq \exp \left(-\delta E_{e} / T_{e}\right) \tag{B.3}
\end{equation*}
$$

and

$$
\begin{equation*}
r_{i} \leq \exp \left(-\delta E_{i} / T_{i}\right) \tag{B.4}
\end{equation*}
$$

accept the move; otherwise, reject the move, leaving unchanged the surface elevation for the column.

Choosing the initial field temperatures and determining the rate at which to cool $T_{i}$ tend to frustrate this application. In Ising's model, the interaction between neighboring elements and the interaction between an element and the exterior field differ only in magnitude. In the case of surface fitting, the variance field and slope constraint are not easily related; one can only minimize some combination of their respective field energies. Deciding on the combination is a subjective matter. With this point in mind, I arrived at the surface fitting result shown in Chapter 3 (Figure 3-9) only after extensive experimentation with initial temperature values and with cooling rates. Certainly, a more rigorous study is warranted if this technique is to be used in the future.

## References

Blake, L., Prediction of Radar Range, chapter 2, McGraw-Hill, Inc., 1970.
Bryan, M., Radar Remote Sensing for Geosciences, Technical Report, ERIM, University of Michigan, Ann Arbor, 1973.

Bryan, M., Bibliography of Geologic Studies Using Imaging Radar, JPL Publication 79-53, NASA Jet Propulsion Lab, California Institute of Technology, Pasadena, 1979.

Canny, J., Finding Edges and Lines in Images, Artificial Intelligence Laboratory Tech. Rep. 720, Massachusetts Institute of Technology, 1983, 145 pages.

Cimino, J. and C. Elachi, Shuttle Imaging Radar-A (SIR-A) Experiment, JPL Publication 8277, NASA JPL, California Institute of Technology, Pasadena, 1982.

Curlander, J., Utilization of Spaceborne SAR Data for Mapping, IEEE Transactions on Geoscience and Remote Sensing, 22, 106-112, March 1984.

Cutrona, J., Synthetic Aperture Radar, chapter 23, McGraw-Hill, Inc., 1970.
Earth Observing System SAR, NASA Publication, 1988.
Feder, A., Radar as a Terrain-Analysing Device, Geol. Soc. Amer. Bull., 70, 1804-1805, 1959.
Ford, P., A GIPS Tutorial, Center for Space Research, Massachusetts Institute of Technology, 1985.

Ford, J., J. Cimino, B. Holt, and M. Ruzek, Shuttle Imaging Radar Views the Earth From Challenger: The SIR-B Experiment, JPL Publication 86-10, NASA JPL, California Institute of Technology, Pasadena, 1986.

Frankot, R. and R. Chellappa, Application of a Shape From Shading Technique to Synthetic Aperture Radar Imagery, in Proc. International Geoscience and Remote Sensing Symposium, "Remote Sensing: Understanding the Earth as a System", pp. 1323-1329, Univ. Michigan, Ann Arbor, May 1987.

French, W., Two-dimensional and Three-dimensional Migration of Model-Experiment Reflection Profiles, Geophysics, , 265-277, 1974.

George, N., Speckle, SPIE, 249, 124-140, 1980.

Graham, L., Synthetic Interferometer Radar for Topographic Mapping, Proc. of the IEEE, 62, 763-768, June 1974.

Guler, S., G. Garcia, L. Gulen, and M. Toksöz, The Detection of Geological Fault Lines in Radar Images, in Proc. NATO Advanced Study Institute, "Pattern Recognition: Theory and Applications", SPA-Balmoral Belgium, June 1986.

Hatch, R., The Synergism of GPS Code and Carrier Measurements, in Proc. Srd International Geodetic Symposium, pp. 1213-1231, February 1982.

Hoffman, P., Interpretation of Radar Scope Photography, Photogrammetric Eng., 20, 406-411, 1954.

Horn, B., Robot Vision, MIT Press and McGraw-Hill Book Company, 1986, 509 pages.
Keho, T., The Vertical Seismic Profile: Imaging Heterogeneous Media, PhD thesis, Massachusetts Institute of Technology, 1986.

Kindermann, R. and J. Snell, Markov Random Fields and their Applications, pp. 1-23, vol. 1, American Mathematical Society, Providence, RI, 1980.

Kirkpatrick, S., C. Gelatt, and M. Vecchi, Optimization by Simulated Annealing, Science, 220, 671-680, May 1983.

Leberl, F., Accuracy Aspects of Stereo Side-Looking Radar, JPL Publication 79-17, NASA JPL, California Institute of Technology, Pasadena, 1979, 58 pages.

Leberl, F., G. Domik, J. Raggam, J. Cimino, and M. Kobrick, Multiple Incidence Angle SIRB Experiment Over Argentina: Stereo-Radargrammetric Analysis, IEEE Transactions on Geoscience and Remote Sensing, 24, 482-491, July 1986a.

Leberl, F., G. Domik, J. Raggam, and M. Kobrick, Radar Stereomapping Techniques and Application to SIR-B Images of Mt. Shasta, IEEE Transactions on Geoscience and Remote Sensing, 24, 473-480, July 1986b.

Li, F. and R. Goldstein, Studies of Multi-baseline Spaceborne Interferometric Synthetic Aperture Radars, in Proc. International Geoscience and Remote Sensing Symposium, "Remote Sensing: Understanding the Earth as a System", pp. 1545-1550, Univ. Michigan, Ann Arbor, May 1987.

Marr, D. and E. Hildreth, Theory of Edge Detection, in Proc. Royal Society of London B, pp. 187-217, 1980.

Matarese, J., M. Prange, and M. Toksöz, Topographic Correction of Radar Images, in Proc. International Geoscience and Remote Sensing Symposium, "Remote Sensing: Understanding the Earth as a System", pp. 1537-1540, Univ. Michigan, Ann Arbor, May 1987.

Metropolis, N., A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, J. Chem. Phys., 1953.
Moore, R., L. Dellwig, S. Parashar, N. Hardy, and M. Soule, Applications of Imaging Radars, A Bibliography, Remote Sensing Lab Report 265-2, University of Kansas, Lawrence, 1974.

Prange, M., An Investigation of the Use of Shuttle Imaging Radar - A (SIR-A) Images for Fault Detection, February 1985, Dept. of Earth, Atmospheric, and Planetary Sciences, M. I. T.

Pravdo, S., B. Huneycutt, B. Holt, and D. Held, Seasat Synthetic-Aperture Radar Data User's Manual, JPL Publication 82-90, NASA JPL, California Institute of Technology, Pasadena, 1983.

Ramapriyan, H., H. Strong, Y. Hung, and C. Murray, Jr., Automated Matching of Pairs of SIRB Images for Elevation Mapping, IEEE Transactions on Geoscience and Remote Sensing, 24, 462-471, July 1986.

Robinson, E., Migration of Geophysical Data, International Human Resources Development Corporation, Boston, 1983, 208 pages.

Sabins, F., Geologic Applications of Remote Sensing, in Remote Sensing of the Environment, pp. 508-579, Addison-Wesley Publishing Co., Reading, MA, 1976.

Saunders, R., J. Holtzman, and C. Elachi, Simulation of Orbital Radar Images, in Radar Geology: An Assessment, pp. 45-63, NASA JPL, California Institute of Technology, Pasadena, 1980.

Schneider, W., Integral Formulation for Migration in Two and Three Dimensions, Geophysics, , 49-76, 1978.

Shuttle Imaging Radar-C Science Plan, JPL Publication 86-29, NASA JPL, California Institute of Technology, Pasadena, 1986.

Skolnik, M., An Introduction to Radar, chapter 1, McGraw-Hill, Inc., 1970.
Stolt, R. and A. Benson, Seismic Migration - Theory and Practice, in Handbook of Geophysical Exploration, edited by K. Helbig and S. Treitel, 1986, 382 pages.

Strong, J. and H. Ramapriyan, Massively Parallel Correlation Techniques to Determine Local Differences in Pairs of Images, copy of manuscript.

Thomas, J., V. Kaupp, W. Waite, and H. MacDonald, Considerations for Optimum Radar Stereo, in Proc. International Geoscience and Remote Sensing Symposium, ${ }^{\alpha}$ Remote Sensing: Understanding the Earth as a System", pp. 1531-1536, Univ. Michigan, Ann Arbor, May 1987.

Toksöz, M., L. Gulen, M. Prange, J. Matarese, G. Pettengill, and P. Ford, Delineation of Fault Zones Using Imaging Radar, in Proc. SIR-B Science Conference, JPL, Pasadena, CA, April 1986.

Tomiyasu, K., Tutorial Review of Synthetic-Aperture Radar (SAR) with Applications to Imaging of the Ocean Surface, Proc. of the IEEE, 66, 563-583, May 1978.

Wildey, R., Radarclinometry for the Venus Radar Mapper, Photogrammetric Eng. and Remote Sensing, 52, 41-50, 1986.

Zebker, H. and R. Goldstein, Topographic Mapping from Interferometric Synthetic Aperture Radar Observations, in Proc. International Geoscience and Remote Sensing Symposium, "Remote Sensing Instrumentation: Technology for Science and Applications", pp. 113-117, Univ. Massachusetts, Amherst, October 1985.

