ROI_PAC Documentation

Repeat Orbit Interferometry Package

Abridged version of Chapter 3 of a PhD thesis written by Sean Buckley CSR, UT Austin; edited by Paul Rosen and Patricia Persaud

May 10, 2000

Version 1.1

This document describes the algorithms and procedures implemented in ROI_PAC, the JPL/Caltech Repeat Orbit Interferometry Package. The overall flow of the processing stream is controlled by Perl scripts. Specific functions algorithmic functions are generally carried out by Fortran or C programs. Table 3.1 lists the major codes and their functions. These processing procedures form the basis of the data reduction for the bulk of geophysical applications that use InSAR.

Table 3.1. The general processing steps and the code elements

Processing Step	Code Element	Section
Process control	process.pl	3.0
SAR raw data conditioning.	make_raw.pl	3.1
SAR Image formation	roi	3.2
SAR Image registration	ampcor	3.3

Interferogram formation	resamp_roi	3.3
Baseline determination	baseest	3.4
Interferogram flattening	cecpxcc, rilooks	3.5
Correlation determination	Makecc; icu	3.6
Interferogram filtering	icu	3.7
Phase unwrapping and absolute phase determination	Icu; baseline	3.8
Deformation determination	diffnsim	3.9
Topography determination from unwrapped phase	inverse3d	3.10

3.0 PROCESS CONTROL

The algorithms comprising ROI_PAC are controlled by Perl Scripts. The major controlling script is process.pl. This script goes from raw data through geocoded deformation maps, in some cases, complete with model removed. The script is designed to allow partial processing of the data, so that if a particular part of the processing fails or is interrupted, the processing can be restarted where it left off (after the problem is fixed.). Process.pl calls many other perl scripts. All perl script usage descriptions are included in in-line “pod2man” documentation system. The calling sequence and the available intermediate processing stages are described in that documentation.

3.1 RAW DATA CONDITIONING

Radar interferometry requires the use of multiple radar data sets which have been processed consistently relative to each other. The need to control carefully the processing of the radar data requires the use of the so-called *raw signal data* or *Level 0 data.* Although the raw data is the most basic of the radar data levels, there are significant differences in the degree to which various processing facilities prepare the satellite downlink for distribution as raw signal data. This section discusses the issues related to preparing raw data for the image formation procedure. These issues include: data format, missing line correction, sampling window start time adjustment, data unpacking, and image formation parameter extraction.

3.1.1 Data Format

SAR raw data presently is delivered in the Committee on Earth Observation Satellites (CEOS) format. The CEOS-format SAR raw data nominally consists of a Volume Directory File, SAR Leader File, Raw Data File, and a Null Volume File. The Volume Directory File describes the arrangement of the data on the storage media. The SAR Leader File provides pertinent information about the specific SAR data set: Raw Data File size, spacecraft height and velocity, scene center latitude, longitude and time of acquisition, etc. The Raw Data File contains a header record plus the SAR raw data nominally stored as one data line per record. Each record consists of a prefix, raw data, and a suffix.

3.1.2 Missing Line Correction

As part of the preprocessing of the data at the archival facility, bad data lines may be identified and removed. Missing lines in the raw data are a problem for interferometry applications where two images must be registered to the sub-pixel level. To check for missing lines, the line counter may be extracted from the prefix information of each raw data record. Missing lines are identified by non-consecutive line counter values. To fix the raw data for a missing line, the previous data line simply is duplicated as a placeholder for the missing line.

3.1.3 Sampling Window Start Time Adjustment

Another data conditioning issue important to radar processing interferometry is the radar return sampling. The radar transmits a pulse of microwave energy which interacts with the surface, with some of the energy being scattered back toward the radar. The backscatter is sampled at the *range sampling frequency* starting at the *sampling window start time* (SWST). As the topographic height of the surface illuminated by the radar changes, the radar pulse total transit time changes. Occasionally, the SWST is adjusted to accommodate changes in the transit time. However, in the image formation stage, it is most convenient to work with a constant SWST. The procedure for fixing the SWST is to determine the minimum SWST for the raw data and zero-pad lines with SWST values greater than the minimum SWST.

3.1.4 Data Unpacking

The sampled radar return is conveniently stored as a complex number, with in-phase (real) and quadrature (imaginary) components. For each return sample, the in-phase and quadrature components are quantized integer values. For example, ERS SAR raw data is 5-bit quantized, i.e., each in-phase and quadrature component will have a value between 0 and 31.

Prior to image formation, the SAR raw data must be *unpacked*, a procedure whereby the distribution of values from each channel is adjusted to be zero-mean. To save hard disk storage space, unpacking typically is performed 'on-the-fly' as part of the data-read in the image formation software. It involves converting each SAR raw data sample component from a quantized integer to a real number and removing a bias from each channel. For ERS SAR raw data, the nominal in-phase and quadrature biases would be 15.5. It should be noted, deviations from the nominal bias values are common but small. The actual biases can be determined from the SAR raw data by computing in-phase and quadrature component means using every tenth sample, for example. Typically, this procedure will be performed by the archival facility with the deviations from the nominal bias values stored in the SAR Leader File.

3.1.5 Image Formation Parameter Extraction

As part of the image formation process, several input parameters are required. Some of the input parameters are radar system parameters: *pulse repetition frequency* (alternatively, *pulse repetition interval*), *sampling frequency*,

pulse length, *chirp slope*, and *wavelength* (alternatively, *carrier frequency*). The meaning of these radar system parameters and how they are used in the image formation process with be discussed in Section 3.2. For convenient reference, common ERS SAR system parameters are summarized in Table 3.1 (these can change depending on the observation).

Most, if not all the interferometric processing steps require parameters derived from the satellite orbit ephemeris. These parameters include the satellite body-fixed position and velocity, the height above the reference surface, and the earth radius. Significant issues have been raised as to the amount and quality of orbit information provided in the SAR Leader File. Some SAR archival facilities provide five or more state vectors spanning the time of the SAR acquisition, whereas others provide virtually no information. Throughout this dissertation, these issues were avoided by using ERS orbit ephemerides calculated at CSR.

Table 3.1. ERS SAR system parameters.

pulse repetition interval, t_{PRI}	595.27 μs
pulse repetition frequency, f_{PRF}	1679.9 Hz
sampling frequency, f_s	18.962 MHz
pulse length, τ_p	37.12 μs
chirp slope, K	418.91 GHz/s
wavelength, λ	5.6666 cm
carrier frequency, f_c	5.29 GHz

bandwidth, f_{BW}	15.55 MHz

An additional parameter required throughout the image formation process is the range to a given sample. The range may be determined from the number of samples relative to the first range sample, the *range sample spacing* (determined from the sampling frequency), and the *range to the first sample.* For the ERS SAR, the range to the first sample ρ_0 is computed as follows:

$$\rho_0 = \frac{c}{2} t_{SWST} \tag{3.1}$$

where

$$t_{SWST} = n_{SWST} t_n + 9 \; t_{PRI} - 6\text{x}10^{-6} \tag{3.2}$$

$$t_{PRI} = (n_{PRI} + 2) t_n \tag{3.3}$$

and c is the speed of light, t_{SWST} is the sampling window start time (in seconds), t_{PRI} is the pulse repetition interval (in seconds), n_{SWST} is the SWST count, n_{PRI} is the PRI count, and $t_n = 210.94\text{x}10^{-9}$ is the count duration (in seconds). The ERS SAR SWST count n_{SWST} and PRI count n_{PRI} are given in the prefix of each raw data record. The nominal value for the ERS SAR PRI count n_{PRI} is 2820, giving a pulse repetition interval t_{PRI} = 595.27 microseconds and a pulse repetition frequency f_{PRF} = 1679.9 Hz.

3.1.6 Concluding Remarks on Raw Data Conditioning

The remaining raw data conditioning and parameter extraction issues are developed best in the context of the image formation process itself and are discussed in the next section. These include the number of range samples to process (Section 3.2.3.2), the computation of the *doppler centroid coefficients* (Section 3.2.4), and the *azimuth compression patch size* and *valid azimuth samples per patch* (Section 3.2.6).

3.2 IMAGE FORMATION

There are two major categories of SAR image formation or focusing techniques: range-doppler processing and wavenumber domain processing [Bamler, 1992]. This section describes the implementation of a range-doppler processor for interferometry applications.

To begin, the transmitted signal or chirp pulse is defined. The fundamental quantity characterizing the image formation process is the received response from a point target on the surface. It is shown the received response to a point target, spread out in both range and azimuth, may be resolved with a matched filter. The matched filter first is implemented in a range compression procedure. The range-compressed point target response is developed and discussed. Because the range to the target changes as it is imaged in time, a range migration processing step must be added to the process chain. Finally, the target is fully resolved through implementation of a matched filter azimuth compression step.

3.2.1 Chirp Pulse Point Target Response

A radar pulse of unit amplitude and unit time duration centered at time $t = 0$ may be written as:

$$\text{rect}[t] = \begin{cases} 1, & |t| \le \frac{1}{2} \\ 0, & \text{otherwise} \end{cases} \tag{3.4}$$

Generalizing, a radar pulse of amplitude $A(t)$ and time duration τ centered at time $t = t_i$ may be written as:

$$A(t)\,\text{rect}\left[\frac{t - t_i}{\tau}\right] = \begin{cases} A(t), & t_i - \frac{\tau}{2} \le t \le t_i + \frac{\tau}{2} \\ 0, & \text{otherwise} \end{cases} \tag{3.5}$$

Following the above discussion, a *linear frequency-modulated (FM) pulse* $s_n(t)$ (also referred to as a *chirp pulse*) transmitted at time $t = -\tau_p/2$ and centered at time $t = 0$ may be written as:

$$s_n(t) = \exp\{j\psi\}\,\text{rect}\left[\frac{t}{\tau_p}\right] = \exp\left\{j2\pi\left[f_c t + \frac{K}{2}t^2\right]\right\}\text{rect}\left[\frac{t}{\tau_p}\right] \tag{3.6}$$

where ψ is the phase of the transmitted pulse, f_c is the carrier frequency, K is the chirp slope (in Hz/s), τ_p is the pulse duration, and an arbitrary amplitude factor has been neglected. When the chirp slope K is positive (negative), it is referred to as an up (down) chirp.

The linear FM pulse derives its name from the linear behavior of its instantaneous frequency with respect to time:

$$f = \frac{1}{2\pi}\frac{d\psi}{dt} = f_c + Kt, \qquad -\tau_p/2 \le t \le \tau_p/2 \tag{3.7}$$

It is noted, the linear FM pulse bandwidth f_{BW} can be closely approximated from the difference between the instantaneous frequency extrema:

$$f_{BW} = \left(f_c + K\tau_p/2\right) - \left(f_c - K\tau_p/2\right) = |K|\tau_p \tag{3.8}$$

This expression holds for a sufficiently large bandwidth time product $f_{BW}\tau_p$ and is relevant to the interpretation of the range-compressed point target response (Section 3.2.3.1) [Curlander and McDonough, 1991; Cook and Bernfeld, 1967].

To simplify the development of the chirp pulse point target response, the range between the radar and a given point target will be assumed fixed for a given pulse (but may change from pulse to pulse). *This is referred to as the start-stop approximation and is the reason why the two-dimensional image formation process can be split into the two one-dimensional frequency-domain processing steps in the range-doppler technique.* Accordingly, the time delay τ between radar transmission, illumination of a point target at a range ρ, and radar reception of a given pulse may be written as:

$$\tau(s) = 2\rho(s)/c \tag{3.9}$$

where the range to the point target (and hence the time delay) are a function of the along-track time s (referred to as *slow time*). Therefore, the point target response $r(s,t)$ may be written as:

$$r(s,t) = s_n(t-\tau) = \exp\left\{ j2\pi \left[f_c(t-\tau) + \frac{K}{2}(t-\tau)^2 \right] \right\} \text{rect}\left[\frac{t-\tau}{\tau_p} \right] \qquad (3.10)$$

Therefore, the point target response is a function both of slow time s and the time t associated with the range direction (referred to as *fast time*).

As part of the in-phase and quadrature sampling, the carrier frequency is removed from the radar return, a procedure referred to as *basebanding* or *heterodyning*. The point target response $r(s,t)$ after basebanding is:

$$r(s,t) = \exp\{-j4\pi\rho(s)/\lambda\} \exp\left\{ j\pi K\left[t-\tau(s)\right]^2 \right\} \text{rect}\left[\frac{t-\tau(s)}{\tau_p} \right] \qquad (3.11)$$

where $\lambda = c/f_c$ and λ is the wavelength.

During image formation, the sampled radar return is processed or filtered such as to identify the point targets and place them in the appropriate location in the output image. In other words, the point target response, distributed both in fast time and slow time, is compressed to the appropriate range and azimuth (fast time and slow time) location. The approach used to compress the sampled radar return is the matched filter, the subject of the next section.

3.2.2 Matched Filter

The derivation of the matched filter for use in radar was developed by Dr. D. O. North in a classified RCA Laboratories Report in 1943. Due to its significance to modern radar theory, the report was reprinted in the general literature in 1963 [North, 1963]. This section will provide a terse but adequate summary of the matched filter. For more details, the reader is referred to the review provided by Curlander and McDonough [1991].

From Section 3.2.1, it is obvious by inspection the model input Eqn. (3.6) and the model output Eqn. (3.10) represent a linear, time-invariant (LTI) system. Accordingly, the matched filter will be designed with LTI characteristics. For a one-dimensional LTI system, the system output $g(t)$ is the convolution of the system input $f(t)$ with the *system unit impulse response* $h(t)$:

$$g(t) = \int_{-\infty}^{\infty} f(t')h(t - t')dt' \tag{3.12}$$

where the unit impulse response is the response of the system to a unit impulse $f(t) = \delta(t - t')$:

$$\left[g(t)\right]_{f(t)=\delta(t-t')} = \int_{-\infty}^{\infty} \delta(t - t')h(t - t')dt' = h(t) \tag{3.13}$$

Therefore, an LTI system can be characterized completely by its unit impulse response.

The derivation of the matched filter can be stated as follows: given a transmitted radar signal $s_n(t)$ (Eqn. (3.6)) and the response associated with a point target at a range ρ (Eqn. (3.10)) corrupted by additive white noise, select a filter to identify the point target and assign it to the correct range. Consider first filtering of a pulse in range. The stop-start approximation, introduced in the previous section, allows the azimuth location or slow time dependence in Eqn. (3.10) to be ignored in this derivation. Accordingly, the matched filter response should be greatest at the time $t = \tau$ associated with the range ρ, significantly above the noise transmitted through the filter. This leads to the constraint of maximizing the filter signal-to-noise ratio (SNR) at time $t = \tau$. Since the matched filter is a LTI system, it can be characterized by its unit impulse response $h(t)$, which is found to be [Curlander and McDonough, 1991]:

$$h(t) = s_n^*(-t) \tag{3.14}$$

Application of this matched filter is referred to as *range compression.* The range-doppler processing algorithm applies another matched filter to compress the the radar return in the azimuth direction, called *azimuth compression*. Posing an additional complication, the slow time dependence of the range in Eqn. (3.11) results in a *range migration* of a point target response as the spacecraft moves along its trajectory. This is accommodated with a range migration correction implemented prior to azimuth compression.

3.2.3 Range Compression

3.2.3.1 Range-Compressed Point Target Response

Referring to the matched filter Eqn. (3.14) and basebanding the transmitted signal Eqn. (3.6), the *basebanded* range compression matched filter $h_r(t)$, also known as the *range reference function*, may be written as:

$$h_r(t) = \exp\left(-j\pi K t^2\right)\operatorname{rect}\left[\frac{t}{\tau_p}\right] \tag{3.15}$$

Substituting the filter input Eqn. (3.11) and the filter unit impulse response Eqn. (3.15) into the filter convolution Eqn. (3.12), the range compression matched filter output becomes:

$$g(s,t) = f_{BW}\operatorname{sinc}\left\{\pi f_{BW}\left[t - 2\rho(s)/c\right]\right\}\exp\left[-j4\pi\rho(s)/\lambda\right] \tag{3.16}$$

where the pulse bandwidth f_{BW} expression Eqn. (3.8) was utilized and $\operatorname{sinc}(x) \equiv \sin x/x$.

The range-compressed point target response modulus absolute value and power, centered on $t = \tau$, are shown in Figures 3.1 and 3.2, respectively. The peak between the nulls $f_{BW}t = \pm 1$ is referred to as the *main lobe* with the other peaks referred to as the *side lobes*.

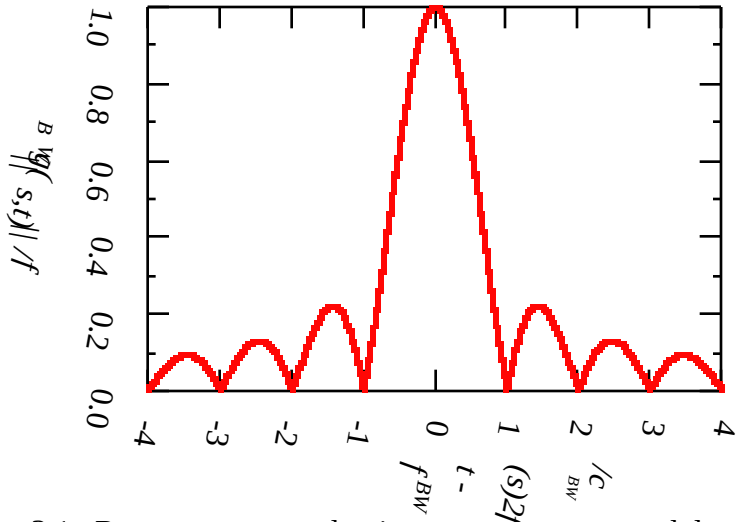


Figure 3.1. Range-compressed point target response modulus absolute value.

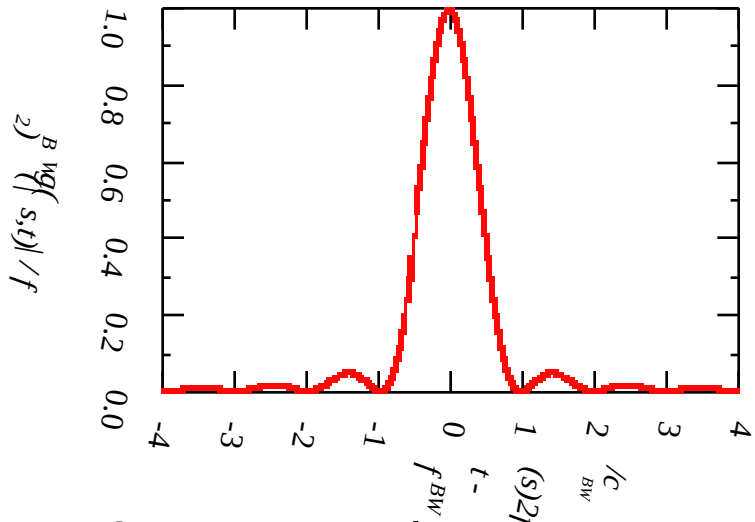


Figure 3.2. Range-compressed point target response output power.

The *slant range resolution* δr associated with the range-compressed point target response may be written as:

$$\delta r = c/2 f_{BW} = c/2|K|\tau_p \tag{3.17}$$

It can be seen from Fig. 3.2 and Eqn. (3.17) that increasing the pulse bandwidth f_{BW} narrows the main lobe of the range-compressed point target response, improving the slant range resolution. It is worth noting, one point of confusion for radar neophytes is the distinction between the slant range *resolution* δr in Eqn. (3.17) and the slant range *sample spacing* Δr:

$$\Delta r = \frac{c}{2 f_s} \tag{3.18}$$

where f_s is the sampling frequency.

In summary, because the radar matched filter is an LTI system, the range compression is the convolution of the raw data with the range reference function. This can be carried out efficiently as a complex multiplication in the (range) frequency domain. Therefore, the range compression is performed by taking the FT of the range reference function and the one-dimensional FT (in range) of the raw data, multiplying the complex-valued range reference function and raw data range transforms, and transforming the range-compressed data back to the time domain.

3.2.3.2 Chirp Extension

Chirp extension refers to the general procedure of extending the number of processed range samples beyond the limits of the raw range data extent in the near and far range. Because the signal return of a point target is spread over the time of the transmitted pulse, typically several hundred range pixels, the return of targets located outside the sampled data extent will be partially represented within the sampled data. So by processing "zero samples" within a chirp length of the data compressed targets are reconstructed. The additional pixels are of coarser resolution and contain more phase noise as compared with those pixels processed with the full pulse bandwidth.

Chirp extension, given as the number of additional range bins (pixels in the output image) in the near-range, uses what energy may be contained in the collected raw data samples to form the additional range bins. Specifying a large enough chirp extension in the near-range as well as a large enough number of total range bins to process assures one processes the data to the full extent possible in range.

3.2.4 Doppler History Considerations

Tracking the doppler shift history of a point target as it is illuminated by the radar provides the information necessary to resolve the azimuth location of the target. The doppler information is used both in the range migration and the azimuth compression processing steps. This section discusses the estimation of the doppler history as well as the reasons for processing radar data to the same doppler for interferometry applications.

3.2.4.1 Doppler Centroid and Doppler Rate Estimation

The doppler history of a point target varies both in slow time and in fast time. Typically, the doppler first is modeled as a linear function of slow time s:

$$f = f_{DC} + f_R s \tag{3.19}$$

where f is the instantaneous doppler frequency, f_{DC} is the *doppler centroid*, and f_R is the *doppler rate*. The doppler centroid and doppler rate either may be calculated from orbit and attitude data or estimated in an automated fashion from the data itself. Automatic estimation of the doppler centroid and doppler rate is referred to as *clutterlock* and *autofocus*, respectively.

Clutterlock may be performed on various levels of processed data (raw data, range-compressed data, range-compressed and azimuth-compressed data) and in either the time or frequency domain. For this research, clutterlock was performed with the correlation doppler estimator (CDE) proposed by Madsen

[1989]. The CDE is a time-domain estimator applied to the raw data. The CDE relies on the direct relation between the phase of the correlation function and the doppler centroid [Madsen et al., 1989].

The output for the CDE clutterlock algorithm is a set of data points of doppler centroid frequency versus slant range. In this work, the doppler centroid then is modeled as a quadratic function of slant range (fast time). A regression analysis may be performed on the CDE output to determine the doppler centroid quadratic coefficients. The doppler centroid quadratic coefficients (scaled by the PRF) may be expressed as a function of range bin in the raw data:

$$f_{DC} = \tilde{f}_{DC0} + \tilde{f}_{DC1} b + \tilde{f}_{DC2} b^2 \tag{3.20}$$

where

$$b = \mathrm{int}\left(\frac{\rho - \rho_0}{\Delta r}\right) \tag{3.21}$$

and b is the range bin (numbered from zero), ρ is the slant range from the antenna to range bin b, ρ_0 is the slant range to the zeroth range bin (i.e., the slant range to the first pixel), and $\mathrm{int}(\ )$ is the integer or truncation operator.

Alternatively, the doppler coefficients may be expressed as a function of slant range:

$$f_{DC} = f_{DC0} + f_{DC1}\rho + f_{DC2}\rho^2 \tag{3.22}$$

Equating Eqn. (3.20) and Eqn. (3.22) for three values of *b* (and the corresponding values of ρ), the bin-referenced doppler coefficients may be written as functions of the range-referenced doppler coefficients:

$$\tilde{f}_{DC0} = f_{DC0} + f_{DC1}\rho_0 + f_{DC2}\rho_0^2 \tag{3.23}$$

$$\tilde{f}_{DC1} = \left(f_{DC1} + 2f_{DC2}\rho_0\right)\Delta r \tag{3.24}$$

$$\tilde{f}_{DC2} = f_{DC2}\left(\Delta r\right)^2 \tag{3.25}$$

Similarly, the range-referenced doppler coefficients may be written as functions of the bin-referenced doppler coefficients:

$$f_{DC0} = \tilde{f}_{DC0} - \tilde{f}_{DC1}\left(\frac{\rho_0}{\Delta r}\right) + \tilde{f}_{DC2}\left(\frac{\rho_0}{\Delta r}\right)^2 \tag{3.26}$$

$$f_{DC1} = \frac{1}{\Delta r}\left[\tilde{f}_{DC1} - 2\tilde{f}_{DC2}\left(\frac{\rho_0}{\Delta r}\right)\right] \tag{3.27}$$

$$f_{DC2} = \frac{\tilde{f}_{DC2}}{(\Delta r)^2} \tag{3.28}$$

The relation between the doppler history, spacecraft orbit and spacecraft attitude are presented in Curlander and McDonough [1991]. In this work, the doppler rate f_R was calculated from the orbit information:

$$f_R = -2V^2/\lambda\rho \tag{3.29}$$

where V is an effective velocity related to the body-fixed spacecraft velocity. This is an approximation that ignores a range dependent acceleration term.

3.2.4.2 Processing to the Same Doppler

In general, the two raw data sets processed to form an interferogram will have different doppler histories. Two effects are apparent when forming an interferogram with images processed to different dopplers. First, the image geometries will be different. If the doppler histories are sufficiently close to one another, this can be adequately accommodated in the image registration process (see Section 3.3). Second, the interferogram formation will result in the different dopplers beating against each other. This amounts to an azimuth fringe rate which must be removed from the interferogram.

These complications may be avoided by processing the raw data sets to the same doppler. This results in one or both of the radar images being processed with doppler errors (deviations from the true doppler history). Doppler centroid errors lead to decreased SNR in the radar images [Li et al., 1985]. However, problems associated with decreased SNR may be reduced with proper weighting of the side lobe energy during azimuth compression (see Section 3.2.6). The approach taken in this work, and by most others in interferometry, is to process the raw data to the same doppler centroid.

Since the raw data sets generally are not co-registered, a relation must be derived to convert between doppler centroid coefficients taking into consideration a gross offset between the two images. Assuming one processes to the same doppler, the doppler centroid f_{DC} associated with a given range bin in both images may be written as:

$$f_{DC} = \left(\tilde{f}_{DC0}\right)_1 + \left(\tilde{f}_{DC1}\right)_1 b_1 + \left(\tilde{f}_{DC2}\right)_1 b_1^2 \tag{3.30}$$

$$f_{DC} = \left(\tilde{f}_{DC0}\right)_2 + \left(\tilde{f}_{DC1}\right)_2 (b_1 + \Delta b) + \left(\tilde{f}_{DC2}\right)_2 (b_1 + \Delta b)^2 \tag{3.31}$$

where b_1 is a given range bin in the first image, Δb is the pixel offset from b_1 to the corresponding range bin in the second image and the doppler centroid coefficient subscripts refer to the first or second image. Equating Eqn. (3.30) and Eqn. (3.31) for three arbitrary values of b_1 provides three independent equations which may be solved for the new doppler coefficients for one of the images.

For example, if one wishes to process both images to the first doppler centroid, the doppler coefficients to be used to process the second image would be:

$$\left(\tilde{f}_{DC0}\right)_2 = \left(\tilde{f}_{DC0}\right)_1 - \left(\tilde{f}_{DC1}\right)_1 \Delta b + \left(\tilde{f}_{DC2}\right)_1 (\Delta b)^2 \tag{3.32}$$

$$\left(\tilde{f}_{DC1}\right)_2 = \left(\tilde{f}_{DC1}\right)_1 - 2\left(\tilde{f}_{DC2}\right)_1 \Delta b \tag{3.33}$$

$$\left(\tilde{f}_{DC2}\right)_2 = \left(\tilde{f}_{DC2}\right)_1 \tag{3.34}$$

3.2.5 Range Migration Correction

After range compression, the point target response remains spread over many azimuth samples and requires the azimuth compression procedure to fully resolve the target. Adding complication to the azimuth compression process is the presence of the range-dependent term in the range-compressed point target response Eqn. (3.16). In other words, the range-compressed point target response is distributed over several range bins as the target is imaged in slow time. This is referred to range migration and must be accommodated prior to azimuth compression.

3.2.5.1 Range Migration

The phase ψ associated with the range-compressed point target response Eqn. (3.16) is:

$$\psi = -4\pi\rho(s)/\lambda \tag{3.35}$$

From Eqn. (3.35), the doppler centroid f_D and doppler rate f_R are:

$$f_D = \dot{\psi}/2\pi = -2\dot{\rho}(s)/\lambda \tag{3.36}$$

$$f_R = \ddot{\psi}/2\pi = -2\ddot{\rho}(s)/\lambda \tag{3.37}$$

The range to the target, which is a function of slow time s, may be written in terms of a Taylor series expansion about the slant range ρ_c (and its derivatives $\dot{\rho}_c$, $\ddot{\rho}_c$, ...) and slow time s_c when the target is in the center of the beam:

$$\rho(s) = \rho_c + \dot{\rho}_c(s - s_c) + \tfrac{1}{2}\ddot{\rho}_c(s - s_c)^2 \tag{3.38}$$

where the Taylor series has been truncated at second order. The derivatives of the range then may be expressed as:

$$\dot{\rho}(s) = \dot{\rho}_c + \ddot{\rho}_c(s - s_c) \tag{3.39}$$

$$\ddot{\rho}(s) = \ddot{\rho}_c \tag{3.40}$$

Utilizing Eqn. (3.36), Eqn. (3.37), Eqn. (3.39), and Eqn. (3.40), the range derivatives evaluated at the beam center may be written in terms of the doppler centroid f_{DC} and doppler rate f_R, resulting in the following expression for the range to the point target as a function of slow time s:

$$\rho(s - s_c, s_c, \rho_c) = \rho_c - \frac{\lambda f_{DC}}{2}(s - s_c) - \frac{\lambda f_R}{4}(s - s_c)^2 \tag{3.41}$$

It is noted the doppler centroid f_{DC} and doppler rate f_R are functions of s_c and ρ_c. In other words, the doppler centroid and doppler rate vary as the radar images point targets at different ranges through slow time.

Range migration is defined as the range to the point target *relative* to the beam center point target range as the radar images the point target through slow time. Therefore, the range migration $\Delta\rho$ may be written as:

$$\Delta\rho = \rho(s - s_c, s_c, \rho_c) - \rho_c = -\frac{\lambda f_{DC}}{2}(s - s_c) - \frac{\lambda f_R}{4}(s - s_c)^2 \tag{3.42}$$

The first term is referred to as *range walk* while the second term is referred to as *range curvature*. As indicated by Eqn. (3.42), more significant range migration occurs at longer wavelengths.

3.2.5.2 Stationarity of Range-Compressed Point Target Response

Fortunately, the weak dependence of the doppler centroid and doppler rate on s_c results in the range-compressed impulse response being approximately stationary in slow time s. Therefore, Eqn. (3.41) may be simplified:

$$\rho(s - s_c, s_c, \rho_c) \approx \rho(s - s_c, \rho_c) \equiv \tilde{\rho}(s, \rho_c) = \rho_c - \frac{\lambda}{2}\left(f_{DC}s + \frac{1}{2} f_R s^2\right) \tag{3.43}$$

where a change of variables from $s - s_c$ to s has been made.

Since the range-compressed impulse response is stationary in slow time, the range-compressed doppler spectrum can be evaluated by the *method of stationary phase*. The method of stationary phase as it applies to radar may be stated as follows [Cook and Bernfeld, 1967]: the major contribution to the FT integral being considered occur at the points $\hat{s}$ of stationary phase.

The phase $\Psi(s,\rho_c)$ of the complex-valued integrand of the doppler spectrum FT integral may be written as:

$$\Psi(s,\rho_c) = -\frac{4\pi}{\lambda}\,\tilde{\rho}(s,\rho_c) - 2\pi f s \tag{3.44}$$

where f is the doppler frequency. Applying the method of stationary phase:

$$\left[\dot{\Psi}(s,\rho_c)\right]_{s=\hat{s}} = 0 \tag{3.45}$$

results in *the locking relation between time and frequency*:

$$\hat{s} = (f - f_{DC})/f_R \tag{3.46}$$

Substituting the locking relation Eqn. (3.46) into Eqn. (3.43), the range associated with points of stationary phase may be written as:

$$\tilde{\rho}(\hat{s},\rho_c) = \rho_c - \frac{\lambda}{4f_R}\left(f^2 - f_{DC}^2\right) \tag{3.47}$$

In summary, the range-compressed point target response is spread over several range samples as the point target is imaged through slow time, a phenomenon referred to as range migration. The range migration correction is a procedure whereby the range associated with the point target response is adjusted to a constant value, the range ρ_c at beam center.

3.2.5.3 Interpolation of Range-Compressed Data

The range migration correction amounts to an interpolation of the range-compressed data. For small range walk, this interpolation can be done completely in the doppler frequency domain. On the other hand, large range walk complicates matters, requiring an extra processing step referred to as *azimuthal-range compression* or *secondary range compression* [Curlander and McDonough, 1991; Jin and Wu, 1984].

The range migration correction procedure may be split into two steps. First, the amount of range migration is computed in terms of an integer number and fractional number of range samples. Second, the interpolation is performed on the fractional number of range samples, taking into consideration the shift due to the integer number of range samples.

The interpolation scheme used in this work utilized a sinc-function interpolator. For a series of samples $s(i)$, the N-point sinc-function-interpolated value $s(x)$ at x would be:

$$s(x) = \sum_{i=-N/2+1}^{N/2} s(i)\,\mathrm{sinc}\left[\pi(x-i)\right] = \frac{\sin \pi x}{\pi} \sum_{i=-N/2+1}^{N/2} \frac{(-1)^i}{x-i}\, s(i) \tag{3.48}$$

For this work, an eight-point sinc-function interpolator was used. To save computations, the sinc-function evaluation in Eqn. (3.48) was performed as a look-up in an interpolation coefficient array generated during the processor initialization.

3.2.6 Azimuth Compression

Azimuth compression of the range-compressed, range-migrated data is analogous to range compression. The azimuth compression matched filter unit impulse response $h_a(s)$, also known as the *azimuth reference function*, may be written as:

$$h_a(s) = \exp\left\{ j2\pi \left(f_{DC} s - \frac{1}{2} f_R s^2 \right) \right\} \operatorname{rect}\left\{ \frac{s}{\tau_a} \right\} \tag{3.49}$$

where

$$\tau_a = \lambda \rho_c / 2 V_{st} \delta a \tag{3.50}$$

$$\delta a \geq L_a / 2 \tag{3.51}$$

and s is measured relative to s_c, τ_a is the synthetic aperture duration (also known as the SAR integration time), V_{st} is the spacecraft speed relative to the point target, h is the satellite height above the reference surface, r is the radius of the earth, L_a is the antenna size in the along-track direction, and δa is the *desired* azimuth resolution (theoretical limit is $L_a/2$).

In this situation, the azimuth resolution δa was treated as an input parameter. As can be seen from Eqn. (3.49) and Eqn. (3.50), specifying courser azimuth resolution reduces the required SAR integration time.

The modulus of the point target response after azimuth compression is:

$$|g(s)| = B_D \,\mathrm{sinc}(\pi B_D s) \tag{3.52}$$

where

$$B_D = |f_R|\tau_a \tag{3.53}$$

and B_D is the doppler bandwidth. The azimuth resolution δa may be written as:

$$\delta a = V_{st}\delta s = V_{st}B_D^{-1} = \frac{V_{st}}{|f_R|\tau_a} \tag{3.54}$$

where δs is the azimuth filter output main lobe half-power width. One interpretation of Eqn. (3.54) is finer azimuth resolution (from larger doppler bandwidth) is achieved with a longer synthetic aperture. For example, spotlight-mode SAR, as opposed to the fixed viewing geometry of stripmap-mode SAR, steers the radar footprint to illuminate a fixed location on the ground. The longer synthetic aperture results in dramatically finer azimuth resolution as compared with stripmap-mode SAR. Of course, this comes at the cost of reduced spatial coverage.

Azimuth compression is a convolution of the range-compressed, range-migrated data with the azimuth reference function. Similar to the range compression, the azimuth compression can be carried out efficiently in the

doppler frequency domain. This is done as a complex multiplication of the azimuth reference function doppler spectrum with the range-compressed, range-migrated data. Finally, the result is transformed back to the time domain.

However, azimuth compression has two additional complications not present in range compression. First, the azimuth reference varies as a function of the point target range ρ_c through the doppler centroid and the doppler rate. This requires the azimuth reference function doppler spectrum to be re-calculated as the azimuth compression progresses from near-range to far-range. Second, the SAR integration time also varies as a function of the point target range ρ_c. From Eqn. (3.50), SAR integration time increases as the range to the point target increases. This requires an increasing number of azimuth samples be used in the azimuth compression as one progresses from near-range to far-range.

3.2.7 Concluding Remarks on Image Formation

In summary, this section covered the basic operations of the range-doppler processor for interferometry applications. The range-doppler processor consists of three steps: range compression, range migration, and azimuth compression. Implementation of the range-doppler processor for long strips of data requires the data be split into patches. For each patch, there will be several azimuth lines at the beginning and end which will be resolved with less than the full doppler bandwidth. This is overcome by overlapping adjacent patches and writing as output only those azimuth lines in the patch which were processed with the full doppler bandwidth.

3.3 IMAGE REGISTRATION AND INTERFEROGRAM FORMATION

After the image formation process, the reference and slave images will be misregistered. The registration process requires knowledge of a set of range and azimuth offset measurements for each patch. The offset measurements then may be used to estimate a functional mapping to resample the slave image to the reference image.

To obtain a set of range and azimuth offset measurements, an amplitude correlation procedure was applied to patches or chips of the radar imagery. The reference image chip size was 64 x 64 pixels whereas the slave image chip size was 124 x 124 pixels (64 pixels plus a 30-pixel pull-in on either side). Each chip consists of real-valued amplitudes computed from the complex-valued single-look complex imagery. A normalized cross-correlation surface is computed between the reference image chip and a series of chips (the same size as the reference image chip) within the slave image chip. From this correlation surface, a peak correlation value is determined.

An edit criteria is used to keep chips with peak correlation SNR greater than a specified threshold and covariance less than a given threshold. For these chips, the reference and slave image chips are deramped and oversampled by a factor of two. The correlation procedure then is repeated. The resulting correlation surface itself is oversampled by a factor of two and the correlation peak determined. For the given chip, the real-valued range and azimuth offsets to the correlation peak are written as output. The user may cull through the registration offsets to discard any remaining outliers.

For this work, the slave image range offset $\Delta\rho$ and azimuth offset Δa were modeled as polynomial functions of the range/azimuth pixel location (x,y) in the reference image:

$$\begin{aligned}\Delta\rho = {} & C_1^\rho + C_2^\rho x + C_3^\rho y + C_4^\rho xy + C_5^\rho x^2 + C_6^\rho y^2 \\ & + C_7^\rho xy^2 + C_8^\rho x^2 y + C_9^\rho x^3 + C_{10}^\rho y^3\end{aligned} \tag{3.55}$$

$$\begin{aligned}\Delta a = {} & C_1^a + C_2^a x + C_3^a y + C_4^a xy + C_5^a x^2 + C_6^a y^2 \\ & + C_7^a xy^2 + C_8^a x^2 y + C_9^a x^3 + C_{10}^a y^3\end{aligned} \tag{3.56}$$

where C_i^ρ are the range offset fit coefficients and C_i^a are the azimuth offset fit coefficients. The range/azimuth fit coefficients are estimated through a single value decomposition procedure utilizing the offset measurements.

For a given pixel in the reference image, the corresponding slave image pixel location is determined by calculating the range and azimuth offsets from Eqn. (3.55) and Eqn. (3.56) and adding to the reference image pixel location. In general, the corresponding slave image pixel location will contain both integer and fractional parts. The resampled slave image complex pixel value is determined by *shifting* to the slave image location specified by the integer portion of the calculated location and *interpolating* based upon the fractional part.

The interpolation and interferogram formation procedures are complicated by two issues. First, the range and azimuth spectrum of a complex radar image are not centered on zero frequency. Therefore, the interpolator used to resample the slave image must have band-pass characteristics. Linear and quadratic

interpolators are unacceptable as they act as low-pass filters. For this work, a sinc-function interpolator was used [Rosen and Hensley, 1996a].

The second complication to the interferogram formation procedure is the wavenumber or *spectral shift* between two radar images. The difference in look angle between radar images results in a shift in the frequency of the ground reflectivity spectra. During inteferogram formation, the non-common portion of the spectra contributes noise to the interferogram. Fortunately, this source of noise may be alleviated by spectrally filtering each image prior to interferogram formation. This is done in range by shifting each image spectra toward the other by half the total frequency shift and truncating the non-overlapping portion [Gatelli et al., 1994].

In this work, the spectral filtering was implemented conveniently in the time domain. The spectral shift was executed with a complex modulation of each of the complex images, i.e., a phase ramp was removed. The truncation of each shifted image spectra was executed with a Hamming weighting function. The interferogram then was formed through pixel-by-pixel cross-multiplication of the complex images. The final step was to re-insert into the interferogram the phase ramp applied to each of the images. This preserves the relation between the interferometric phase and the range difference between the radar acquisitions.

At this stage, although the interferogram may contain the topographic or deformation signals of interest, the dominant signal in the interferogram will be due to the earth ellipsoid. Prior to removal of the earth ellipsoid signature through interferogram flattening, a precise estimate of the interferometric baseline must be determined.

3.4 BASELINE DETERMINATION

One of the more subtle issues in radar interferometry is the definition, computation, and implementation of the interferometric baseline. The degree to which the surface topography contributes to the interferometric phase depends upon the interferometric baseline. Accordingly, an error in the interferometric baseline will result in residual topographic fringes in a differential interferogram. In addition, baseline errors lead to errors in the height reconstruction process.

This section begins with the definition of the interferometric baseline and its representation in the sch coordinate system. Next, a baseline model and its relation to the available measurements is presented. The baseline estimation procedure is described. Finally, a summary of the various estimation measurements and parameters is provided.

3.4.1 Interferometric Baseline Definition

To define the interferometric baseline, consider a target imaged by two antennae:

$$\vec{T} = \vec{P}_1 + \rho_1 \hat{l}_1 = \vec{P}_2 + \rho_2 \hat{l}_2 \tag{3.57}$$

where $\vec{T}$ is the target position vector, $\vec{P}_i$ is the antenna i phase center position vector, ρ_i is the antenna i range to target, $\hat{l}_i$ is the unit look vector from antenna i to the target, and $i = 1,2$.

The interferometric baseline is defined as the difference between the antennae phase centers when a given target is imaged [Rosen and Hensley, 1996a]. From Eqn. (3.57), the interferometric baseline $\vec{b}$ may be written as:

$$\vec{b} \equiv \vec{P}_2 - \vec{P}_1 = \rho_1 \hat{l}_1 - \rho_2 \hat{l}_2 \tag{3.58}$$

Due to the presence of the unit look vectors from the antennae to the target in Eqn. (3.58), *the interferometric baseline is a function of time* (through the antennae positions) *and target position.*

3.4.2 Baseline in sch Coordinates

Appendix B contains a definition of the sch coordinate system. Figure 3.3 shows the radar imaging geometry in the sch coordinate system. The unit look vector $\hat{l}_i^{sch_i}$ from antenna *i* to the target in sch coordinates associated with antenna *i* may be written as [Rosen and Hensley, 1996a]:

$$\hat{l}_i^{sch_i} = \begin{bmatrix} l_i^{s_i} \\ l_i^{c_i} \\ l_i^{h_i} \end{bmatrix} = \begin{bmatrix} \cos\theta_i \sin\gamma_i \cos\beta + \sin\theta_i \sin\beta \\ -\cos\theta_i \sin\gamma_i \sin\beta + \sin\theta_i \cos\beta \\ -\cos\theta_i \cos\gamma_i \end{bmatrix} \tag{3.59}$$

where θ_i is the look angle associated with antenna *i*, γ_i is the pitch angle associated with antenna *i*, β is the azimuth angle, and a *left-looking* geometry has been assumed.

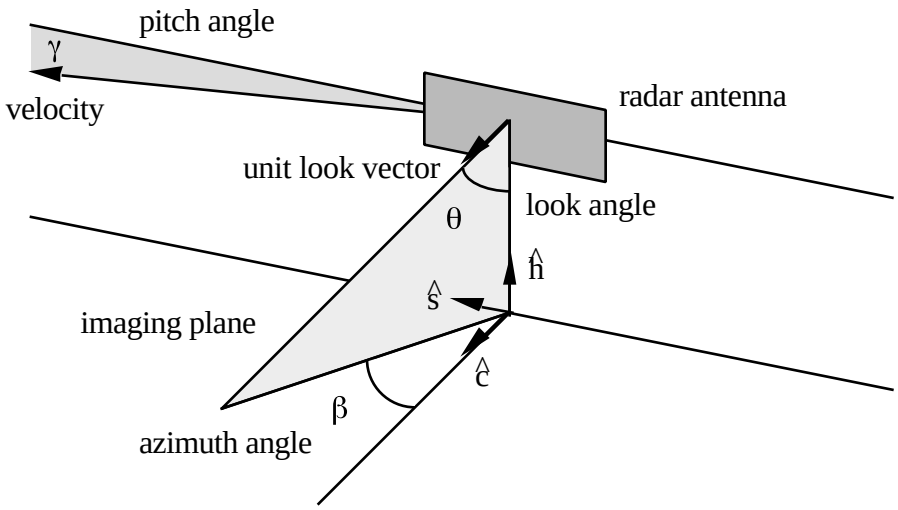


Figure 3.3. Radar imaging geometry in sch coordinate system.

Therefore, the unit look vectors for the antennae may be written as:

$$\hat{l}_1^{sch} = \begin{bmatrix} \sin\theta_1 \sin\beta \\ \sin\theta_1 \cos\beta \\ -\cos\theta_1 \end{bmatrix} \tag{3.60}$$

$$\hat{l}_2^{sch} = \begin{bmatrix} \cos\theta_2 \sin\gamma \cos\beta + \sin\theta_2 \sin\beta \\ -\cos\theta_2 \sin\gamma \sin\beta + \sin\theta_2 \cos\beta \\ -\cos\theta_2 \cos\gamma \end{bmatrix} \tag{3.61}$$

where the pitch angle γ_1 associated with antenna 1 was assumed zero, allowing the subscript to be dropped from the pitch angle γ_2 associated with antenna 2. To express the unit look vector associated with antenna 2 in the sch coordinates associated with antenna 1, a simple rotation may be applied:

$$\hat{l}_2^{sch_1} = M_{sch_2}^{sch_1}\hat{l}_2^{sch_2} = \begin{bmatrix} \cos\theta_2 \sin\gamma \cos(\beta-\eta) + \sin\theta_2 \sin(\beta-\eta) \\ -\cos\theta_2 \sin\gamma \sin(\beta-\eta) + \sin\theta_2 \cos(\beta-\eta) \\ -\cos\theta_2 \cos\gamma \end{bmatrix} \tag{3.62}$$

where

$$M_{sch_2}^{sch_1} = \begin{bmatrix} \cos\eta & -\sin\eta & 0 \\ \sin\eta & \cos\eta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.63}$$

and η is the divergence angle (in the *sc*-plane) from antenna track 1 to antenna track 2. Now assume the divergence angle η, the pitch angle , and the azimuth angle β are sufficiently small such that:

$$\tan\eta \approx \sin\eta \tag{3.64}$$

$$\tan\gamma \approx \sin\gamma \tag{3.65}$$

$$\cos\beta \approx 1 \tag{3.66}$$

$$\cos(\beta-\eta) \approx 1 \tag{3.67}$$

These are valid assumptions for the ERS SAR antennae. Finally, substituting Eqn. (3.60) and Eqn. (3.62) into Eqn. (3.58) and using the approximations given in Eqn. (3.64) through Eqn. (3.67), the baseline may be expressed in (antenna 1) sch coordinates:

$$\vec{b} = \begin{bmatrix} b_s \\ b_c \\ b_h \end{bmatrix} = \begin{bmatrix} b_c \tan\beta + \rho_2 \sin\theta_2 \tan\eta - \rho_2 \cos\theta_2 \tan\gamma \\ \rho_1 \sin\theta_1 \cos\beta - \rho_2 \sin\theta_2 \cos\beta \cos\eta \\ -\rho_1 \cos\theta_1 + \rho_2 \cos\theta_2 \cos\gamma \end{bmatrix} \tag{3.68}$$

The along-track component of the baseline in Eqn. (3.68) is referred as the target-dependent along-track offset Δs_T:

$$\Delta s_T = b_c \tan\beta + \rho_2 \sin\theta_2 \tan\eta - \rho_2 \cos\theta_2 \tan\gamma \tag{3.69}$$

3.4.3 Baseline Model Parameters

It is desirable to express the baseline as a function of a single along-track variable s which takes into consideration both the time-dependent and target-dependent portion of the baseline. For simplicity, the cross-track and vertical baseline components will be assumed to be linear functions of s.

By removing the target-dependent along-track offset Δs_T from the baseline given in Eqn. (3.68), one can utilize a modified baseline model:

$$\vec{b}' = \vec{b} - \begin{bmatrix} \Delta s_T \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} b_s \\ b_c \\ b_h \end{bmatrix} = \begin{bmatrix} \Delta s_0 + (k s_2 - s_1) - \Delta s_T \\ b_{c0} + \dot{b}_c (s_1 + \Delta s_T) \\ b_{h0} + \dot{b}_h (s_1 + \Delta s_T) \end{bmatrix} \tag{3.70}$$

where Δs_0 is the along-track constant offset, k is the along-track scale factor between the s_1 and s_2 coordinates, b_{c0} is the initial cross-track baseline, b_{h0} is the initial vertical baseline, $\dot{b}_c$ is the cross-track baseline rate of change, and $\dot{b}_h$ is

the vertical baseline rate of change. Introducing the along-track variable $s = s_1 + \Delta s_T$, this baseline model may be written concisely as:

$$\vec{b}'(s) = \begin{bmatrix} \Delta s_0 + (k s_2 - s) \\ b_{c0} + \dot{b}_c s \\ b_{h0} + \dot{b}_h s \end{bmatrix} \tag{3.71}$$

The target-dependent along-track offset Δs_T can be re-written in terms of the baseline rates of change:

$$\Delta s_T = b_c \tan\beta + \dot{b}_c \rho_2 \sin\theta_2 - \dot{b}_h \rho_2 \cos\theta_2 \tag{3.72}$$

where $\dot{b}_c = \tan\eta$ and $\dot{b}_h = \tan\gamma$.

3.4.4 Measurement and Model Parameter Relation

The measurements available for the baseline estimation are the range and azimuth offsets. The azimuth offset Δs_{offset} comes from the registration process whereas the range offset $\Delta\rho_{offset}$ comes from the registration process or the unwrapped phase.

Substituting $s_2 = s_1 + \Delta s_{offset}$ into the *s*-component of Eqn. (3.70), the relation between the azimuth offset measurement Δs_{offset} and the model parameters may be written as:

$$\Delta s_0 + k(s_1 + \Delta s_{offset}) - s_1 - \Delta s_T = 0 \tag{3.73}$$

where Δs_T is given by Eqn. (3.72). Similarly, the relation between the range offset measurement $\Delta\rho_{offset}$ and the model parameters may be written as:

$$\Delta\rho_{offset} + \Delta\rho_0 - \Delta\rho = 0 \tag{3.74}$$

where

$$\Delta\rho_{offset} = \left(\lambda/4\pi\right)\phi \qquad \text{(when using the unwrapped phase)} \tag{3.75}$$

$$\Delta\rho = \rho_1\left(\sqrt{\mu} - 1\right) \qquad \text{(}c.f.\text{ Eqn. (2.15))} \tag{3.76}$$

$$\mu = 1 - 2\left(\hat{l}_1 \cdot \vec{b}\right)\rho_1^{-1} + b^2\rho_1^{-2} \tag{3.77}$$

$\Delta\rho_0$ is the range constant offset and is the unwrapped, unflattened interferometric phase (*i.e.*, the unwrapped phase with the reference surface contribution reinserted). Table 3.2 summarizes which baseline parameters may be estimated from the range offsets, along-track offsets, and the unwrapped phase.

3.4.5 Baseline Parameter Estimation Procedure

To summarize, the parameter or state vector X may be written as:

$$X^T = \left[b_{c0} \quad b_{h0} \quad \dot{b}_c \quad \dot{b}_h \quad k \quad \Delta s_0 \quad \Delta\rho_0\right] \tag{3.78}$$

where the model parameters were defined in Eqn. (3.70).

Table 3.2. Baseline parameter estimation summary.

	range offsets	along-track offsets	unwrapped phase
along-track constant offset, Δs_0	no	yes	no
along-track scaling factor, k	no	yes	no
range constant offset, $\Delta\rho_0$	no	no	yes
initial cross-track baseline, b_{c0}	yes	no	yes
initial vertical baseline, b_{h0}	yes	no	yes
cross-track baseline rate of change, $\dot{b}_c$	yes	yes	yes
vertical baseline rate of change, $\dot{b}_h$	yes	yes	yes

The observation-state relation may be written as:

$$Y_i = G_i(X) + \varepsilon_i \qquad i = 1,...,l \tag{3.79}$$

where

$$Y_i = \begin{bmatrix} \left(\Delta\rho_{offset}\right)_i \\ \left(\Delta s_{offset}\right)_i \end{bmatrix} \qquad i = 1,...,l \tag{3.80}$$

$$G_i(X) = \begin{bmatrix} \Delta\rho - \Delta\rho_0 \\ \frac{1}{k}\left(s_1 + \Delta s_T - \Delta s_0\right) - s_1 \end{bmatrix} \qquad i = 1,...,l \tag{3.81}$$

and Y_i is the observation vector, $G_i(X)$ is the measurement model, ε_i is the measurement noise, and l is the number of offset measurements. $\Delta\rho$ is determined from Eqn. (3.76), Δs_T is determined from Eqn. (3.72), and s_1 is determined from the along-track spacing and the line number of offset measurement *i*.

It is noted this is an over-determined system, i.e., the number of observations is greater than the number of parameters to be estimated. Therefore, a set of constraints must be developed to select one estimate from the many possible solutions. The approach used here will be to minimize the sum of the square of the observation errors $J\left(X^*\right)$:

$$J\left(X^*\right) = \sum_{i=1}^{l} \varepsilon_i^{*T} \varepsilon_i^* = \sum_{i=1}^{l} \left[Y_i - G_i\left(X^*\right)\right]^T \left[Y_i - G_i\left(X^*\right)\right] \tag{3.82}$$

where X^* is initial guess for the parameter vector. This represents a nonlinear least squares problem which solves for an updated parameter vector. This procedure is iterated until the difference between updated and previous parameter vectors is less than a given threshold.

3.5 INTERFEROGRAM FLATTENING

The ellipsoid or reference surface contributes to the interferometric phase. Interferogram flattening is the linearization of the interferometric phase about the reference surface. That is to say, the interferometric phase due to the reference surface is removed (subtracted) from the interferogram.

The procedure for flattening the interferogram is summarized in Table 3.3. The equations necessary for flattening have been presented in previous sections and are included below for completeness.

Table 3.3. Interferogram flattening procedure.

Step	Equations
1. compute ρ_1, s_1, h	3.83 - 3.85
2. compute ρ_2 ($\Delta\rho_{ref} = 0$ first time around)	3.86
3. compute θ_{1ref}, θ_{2ref}, approximate b_c, b_s	3.87 - 3.89
4. compute b_c, b_h, b, $\hat{l}_{1ref}$ $\vec{b}$, approximate $\Delta\rho_{ref}$	3.90 - 3.94
5. repeat steps 2-4 for new $\Delta\rho_{ref}$	-
6. compute phase ϕ_{ref}, form complex number	3.95 - 3.96
7. conjugate multiplication	3.97
8. repeat steps 1-7 for all range bins in given azimuth line	-
9. repeat steps 1-8 for all azimuth lines	-

$$\rho_1 = \rho_0 + (i-1)\Delta r \qquad i = 1, \ldots, \text{number of range samples} \qquad (3.83)$$

$$s_1 = (j_a - 1)\Delta a \qquad j_a = 1, \ldots, \text{number of azimuth lines} \tag{3.84}$$

$$h = h_0 + \dot{h} s_1 + \ddot{h} s_1^2 \tag{3.85}$$

$$\rho_2 = \rho_1 + \Delta\rho_0 + \Delta\rho_{ref} \tag{3.86}$$

$$\cos\theta_{lref} = \frac{\rho_l^2 + (r+h)^2 - r^2}{2\rho_l (r+h)} \qquad l = 1,\ 2 \tag{3.87}$$

$$b_c \approx b_{c0} + \dot{b}_c s_1 \tag{3.88}$$

$$b_s = \Delta s_T = b_c \tan\beta + \dot{b}_c \rho_2 \sin\theta_2 - \dot{b}_h \rho_2 \cos\theta_2 \tag{3.89}$$

$$b_c = b_{c0} + \dot{b}_c (s_1 + \Delta s_T) \tag{3.90}$$

$$b_h = b_{h0} + \dot{b}_h (s_1 + \Delta s_T) \tag{3.91}$$

$$b = \sqrt{b_s^2 + b_c^2 + b_h^2} \tag{3.92}$$

$$\hat{l}_{1ref} \quad \vec{b} = b_s \sin\theta_{1ref} \sin\beta + b_c \sin\theta_{1ref} \cos\beta - b_h \cos\theta_{1ref} \tag{3.93}$$

$$\Delta\rho_{ref} = \rho_1 \left(\sqrt{1 - 2\left(\hat{l}_{1ref} \quad \vec{b}\right)\rho_1^{-1} + b^2 \rho_1^{-2}} - 1 \right) \tag{3.94}$$

$$\phi_{ref} = (4\pi/\lambda)\Delta\rho_{ref} \tag{3.95}$$

$$C_{\phi ref} = \cos\phi_{ref} + j \sin\phi_{ref} \tag{3.96}$$

$$C_{\phi\, flat} = C_\phi C_{\phi ref}^* \tag{3.97}$$

where ρ_0 is the range to the first range sample, Δr is the range sample spacing, Δa is the azimuth line spacing, h is the spacecraft height above the reference surface, r is the radius of the earth (assumes spherical earth), ϕ_{ref} is the interferometric phase due to the reference surface, $C_{\phi ref}$ is the complex interferogram due to the reference surface, C_ϕ is the complex interferogram, and $C_{\phi\, flat}$ is the complex flattened interferogram.

The approach taken in this work was to flatten using a baseline estimated from the registration offsets, unwrap the phase, estimate an improved baseline from the unwrapped phase, and flatten the original interferogram using the improved baseline estimate.

After flattening using the procedure described above, there still may be a residual tilt in the interferometric phase due to an erroneous baseline. This is particular true when using a baseline estimated from the registration offsets to flatten an interferogram. The residual tilt therefore must be removed empirically.

If the tilt is removed from an interferogram flattened using the baseline estimated from registration offsets, one may unwrap the phase, re-insert the tilt, unflatten, and estimate an improved baseline. If the tilt is removed from an interferogram flattened using the baseline estimated from the unwrapped phase, one must adjust the baseline parameters to take into consideration the applied phase ramp.

3.6 CORRELATION DETERMINATION

Whereas the phase of the complex-valued cross-correlation of two radar signals is related to the interferometric phase, the amplitude is related to the correlation coefficient. Noise sources lead to interferometric decorrelation (correlation coefficients less than the perfect correlation value of one). Noise can be considered as random fluctuations in the radar return. This leads to treatment of the reference signal S_1 and slave signal S_2 as random processes.

In general statistical terms, the (complex) cross-correlation or *coherence* γ of two complex random processes S_1 and S_2 may be expressed as [Ziemer and Tranter, 1995; Hagberg et al., 1995]:

$$\gamma \equiv \frac{\langle S_1 S_2^* \rangle}{\sqrt{\langle S_1 S_1^* \rangle \langle S_2 S_2^* \rangle}} = \frac{\mathrm{E}[S_1 S_2^*]}{\sqrt{\mathrm{E}[S_1 S_1^*]\mathrm{E}[S_2 S_2^*]}} \tag{3.98}$$

where E[X] and <X> are the expected value (statistical or ensemble average) and the time average, respectively, of the random process X. Eqn. (3.98) assumes S_1 and S_2 are *ergodic processes*, *i.e.*, statistical or ensemble averages of the random processes are equivalent to time averages of the random processes [Ziemer and Tranter, 1995]. In other words, the coherence may be determined from a spatial average of a single realization in time. It is noted, the *correlation coefficient* $|\gamma|$ simply is the amplitude of the complex coherence.

Computationally, the correlation coefficient at a given pixel is estimated as a weighted spatial average over a window of size N [Hagberg et al., 1995; Rosen and Hensley, 1996a]:

$$|\gamma| = \frac{|S_{12}|}{\sqrt{S_{11}}\sqrt{S_{22}}} \tag{3.99}$$

where

$$S_{12} = \sum_{k=0}^{N-1}\sum_{j=0}^{N-1} w_N(k,j) S_1(k,j) S_2^*(k,j) S_D(k,j) \tag{3.100}$$

$$S_{11} = \sum_{k=0}^{N-1}\sum_{j=0}^{N-1} w_N(k,j) S_1(k,j) S_1^*(k,j) \tag{3.101}$$

$$S_{22} = \sum_{k=0}^{N-1}\sum_{j=0}^{N-1} w_N(k,j) S_2(k,j) S_2^*(k,j) \tag{3.102}$$

and w_N is the scalar weighting array of window size N and S_D is a complex parameter used to detrend the interferometric phase in the window, removing a local phase slope which would otherwise bias the correlation estimate. The local phase slopes associated with a given pixel are calculated from the flattened interferometric phase using a similar windowing and weighting scheme.

In this work, the scalar weighting was implemented as a normalized Gaussian low pass filter:

$$w_N(k,j) = w'_N(k,j) \Big/ \sum_{k=0}^{N-1} \sum_{j=0}^{N-1} w'_N(k,j) \qquad j,k = 0,\ldots,N-1 \qquad (3.103)$$

where

$$w'_N(k,j) = \exp\left\{-\left[(k-M)^2 + (j-M)^2\right]/M\right\} \qquad (3.104)$$

$$M = \text{int}(N/2) \qquad (3.105)$$

There is a fundamental relationship between the correlation coefficient and the interferometric phase standard deviation. It has been show, as the number of looks taken to form the interferogram N_L increases, the interferometric phase standard deviation σ_ϕ asymptotically approaches the Cramer-Rao bound [Rodriguez and Martin, 1992; Rosen and Hensley, 1996a]:

$$\sigma_\phi = \frac{1}{\sqrt{2N_L}} \frac{\sqrt{1-|\gamma|^2}}{|\gamma|} \qquad (3.106)$$

From Eqn. (3.106), a smaller correlation coefficient leads to a larger interferometric phase standard deviation. Larger phase standard deviations complicate the phase unwrapping process and may lead to unwrapping errors. One approach to reducing phase unwrapping errors is to unwrap only the phase of those pixels above a given correlation threshold. Another approach is to filter the interferogram, the subject of the next section.

3.7 NONLINEAR ADAPTIVE INTERFEROGRAM FILTERING

This section describes a nonlinear spectral filter designed by Goldstein and Werner [1998] in two contexts. First, the reasoning behind the filter design is discussed. Second, the implementation of the filter algorithm is presented.

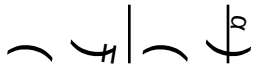
In a radar image, the observed surface backscatter is due to the coherent sum of individual scatterers within each resolution element. The local interferometric phase will be dominated primarily by a single fringe frequency. In other words, the local interferogram power spectrum will consist of narrow-band signal and broad-band noise.

Goldstein and Werner [1998] filter the local interferogram power spectrum to increase the signal to noise ratio at the cost of decreased spectral bandwidth associated with the range and azimuth coordinates. An adaptive filter must be used to take into consideration the non-stationary power spectrum, i.e., the spatially changing fringe frequency.

The power spectrum of a random signal may be estimated using the DFT. One such method utilizes *periodogram analysis*, where the power spectrum estimator is the so-called periodogram. The *periodogram* is the square of the magnitude of the two-dimensional DFT $\tilde{I}(u,v)$ of the interferogram $I(x,y)$ where (u,v) are the frequencies associated with the range and azimuth coordinates (x,y). In this case, the interferogram is split into patches and a single periodogram is computed for each patch. However, random variations in the interferogram phase result in a poor estimate of the power spectrum [Oppenheim and Schafer, 1989].

To smooth the power spectrum estimate, the method of periodogram averaging is utilized. In this method, the interferogram patch is split into overlapping subpatches and a periodogram is computed for each subpatch. The overall patch power spectrum $\hat{S}(u,v)$ is the weighted sum of the individual subpatch periodograms [Oppenheim and Schafer, 1989].

With the local power spectrum estimated, the nonlinear filter $Z(u,v)$ may be constructed:



(3.107)

where α is the filter weight. A filter weight of zero results in no filtering with larger filter values resulting in heavier filtering. Finally, the filtered interferogram $\hat{I}(x,y)$ is calculated:

$$\hat{I}(x,y) = \mathrm{FT}^{-1}\left\{\tilde{I}(u,v)Z(u,v)\right\} \tag{3.108}$$

where FT^{-1} represents the two-dimensional inverse Fourier transform. This procedure is repeated for all interferogram patches. To alleviate discontinuities at patch boundaries, the patches are overlapped and a triangular weighted sum of the overlapping patches is used [Goldstein and Werner, 1998].

3.8 PHASE UNWRAPPING AND ABSOLUTE PHASE DETERMINATION

In interferometry, it is the *absolute interferometric phase* which may be related directly to the topographic height and amount of deformation. However, the *measured interferometric phase* represents the absolute interferometric phase only modulo 2π. To go from the measured phase to the absolute phase, a two-step process is required: phase unwrapping and absolute phase determination. The relation between the measured interferometric phase $\phi_{measured}$, the *unwrapped interferometric phase* $\phi_{unwrapped}$, and the absolute interferometric phase $\phi_{absolute}$ may be summarized as follows:

$$\phi_{measured} = \mathrm{mod}(\phi_{absolute}, 2\pi) \tag{3.109}$$

$$\phi_{unwrapped} = (\phi_{measured})_{unw} = \phi_{absolute} + 2\pi n \tag{3.110}$$

$$\phi_{absolute} = \frac{4\pi}{\lambda}\delta\rho \tag{3.111}$$

where n is an integer, $\delta\rho$ represents the one-way path length difference due to topography and deformation, and $()_{unw}$ represents the phase unwrapping process.

3.8.1 Phase Unwrapping

There are two broad categories of radar interferometry phase unwrapping algorithms presently in use: residue-based algorithms and least-squares algorithms. All these algorithms differentiate the measured phase then integrate the resulting *phase gradient* to arrive at the unwrapped phase. The algorithms

differ in how true phase discontinuities (due to layover, noise, etc.) are accommodated. Residue-based algorithms limit the integration path by identifying and connecting phase discontinuity locations, called *residues.* Least squares algorithms integrate the phase gradient by applying various phase continuity or smoothness constraints [Zebker and Lu, 1997].

The phase unwrapping algorithm used in this work was a residue-based algorithm developed by Goldstein et al. [1988] referred to as the GZW algorithm (based upon the authors' last names). This algorithm may be split into three steps: residue identification, residue connection, and integration. As an added constraint, these steps are applied only to pixels above a given correlation threshold.

The GZW algorithm assumes the radar sampling was such that the unwrapped phase generally will not vary by more than one-half cycle from pixel to pixel. Locations where this assumption is violated are termed *residues.* The residue calculation may be summarized by considering the following general subset of measured phase values:

$$\begin{bmatrix} \phi(i,j) & \phi(i,j+1) \\ \phi(i+1,j) & \phi(i+1,j+1) \end{bmatrix} \tag{3.112}$$

The residue associated with these phase values and assigned to pixel (i,j), may be written as:

$$\text{residue}(i,j) = \frac{1}{2\pi}\left[\Delta\Phi_i(i,j) - \Delta\Phi_j(i,j)\right] \tag{3.113}$$

where

$$\Delta\Phi_i(i,j) \equiv \Phi_i(i,j+1) - \Phi_i(i,j) \tag{3.114}$$

$$\Delta\Phi_j(i,j) \equiv \Phi_j(i+1,j) - \Phi_j(i,j) \tag{3.115}$$

$$\Phi_i(i,j) \equiv \phi(i+1,j) - \phi(i,j) \tag{3.116}$$

$$\Phi_j(i,j) \equiv \phi(i,j+1) - \phi(i,j) \tag{3.117}$$

and Φ_i and Φ_j may need to be adjusted by a multiple of 2π to be consistent with the assumption that the phase change from one pixel to the next is less than one-half cycle. The residue computation given above results in only three possible values: *no residue* (0), a *positive residue* (+1), or a *negative residue* (-1). Appendix C provides an example of the phase unwrapping residue calculation.

Integration along a closed path containing a single residue results in a phase discontinuity when closing the integration path. For a positive/negative residue, this phase discontinuity is $+/-2\pi$. However, if an equal number of positive and negative residues (net neutral charge) are contained within the closed integration path, the phase remains continuous when closing the integration path. The goal of residue connection is to connect residues with *branch cuts* such as to arrive at an overall net neutral charge for the network of connections (referred to as the *tree*). In summary, neutral trees are used to eliminate full-cycle phase discontinuities when integrating along a closed path. However, it is noted the unwrapped phase will vary by more than a half-cycle when crossing any branch cut.

There are a plethora of ways to implement the residue connection process. The approach presented by Goldstein et al. [1988] and implemented in this work was to minimize the length of individual branch cuts in a given tree. This amounts to minimizing the total number of half-cycle discontinuities. The algorithm begins by searching the image for residues. Once a residue is found, a search is initiated to find the closest adjacent residue. The residues are connected with a branch cut. The algorithm proceeds to increase the tree length until the tree is neutralized. This procedure is repeated until all residues in the image are connected. The overall tree length is limited so as to keep large trees from isolating sections of the phase field [Zebker and Lu, 1997].

Once all the residues have been connected and all trees neutralized, the final step is to integrate to arrive at the unwrapped phase. The integration is initiated at a specified location or *seed*. The unwrapped phase is estimated for any pixel adjacent to the seed *which does not cross a branch cut* or already have an estimate. The unwrapped phase is calculated by adding or subtracting the necessary integer number of cycles to maintain the half-cycle phase difference assumption. This procedure is repeated for all pixels adjacent to the seed which were unwrapped. The unwrapped phase proceeds to grow outward from the seed.

In deformation studies, temporal decorrelation significantly complicates the unwrapping process. Areas of correlation may be isolated by vast areas of decorrelation. In this case, the unwrapping procedure may have to be repeated several times. This is done by placing a seed and unwrapping, placing another seed and unwrapping, etc. In general, each unwrapping result will be offset by an integer number of cycles relative to the others.

3.8.2 Absolute Phase Determination

After phase unwrapping, a final phase offset must be determined to relate the unwrapped phase to the absolute phase. There are several approaches to absolute phase determination. Two such approaches will be discussed here: the spectral approach and the ground control point approach.

The spectral approach presented by Madsen and Zebker [1992] splits the full range bandwidth, with center frequency f, into lower and upper bands, with center frequencies f_{lower} and f_{upper}. Two lower resolution interferograms are computed around the center frequencies f_{lower} and f_{upper} and a differential interferogram formed. The resulting fringe frequency of the differential interferogram is sufficiently low as to make phase unwrapping and absolute phase determination unnecessary (i.e., the measured, unwrapped, and absolute phases are one in the same). The differential phase scaled by the frequency ratio $f/\left(f_{upper} - f_{lower}\right)$ is equal to the full-resolution interferometric phase. By forming the difference between the scaled differential interferogram and the unwrapped phase of the full-resolution interferogram and averaging to reduce noise, the remaining phase offset may be determined [Madsen and Zebker, 1992; Madsen, Zebker and Martin, 1993].

The ground control point approach utilizes the unwrapped phase and surface height measurements at known locations to determine the phase offset. Specifically, a range constant offset is estimated as part of the baseline estimation process (Section 3.4).

3.9 DEFORMATION DETERMINATION FROM THE TWO-PASS TECHNIQUE

In differential radar interferometry, deformation is determined by isolating and removing the topographic contribution from an interferogram. In general, there are three ways to accomplish this topography removal: the two-pass, three-pass, or four-pass technique. The advantages and disadvantages of each of these techniques were discussed in Section 2.6.

In this work, a USGS DEM and the two-pass technique were implemented to remove the topographic contribution from the interferometric phase. The overriding reasons for utilizing the two-pass technique were: (1) USGS DEMs were readily available, and (2) current exorbitant costs of ERS radar data over North America limited the purchase of data for this project.

This section discusses the processing procedure for the two-pass technique. This includes interferogram simulation, simulation registration, and deformation determination.

3.9.1 Interferogram Simulation

Topography removal in the two-pass technique utilizes a DEM from a source other than an interferogram. In general, the DEM will have a sample spacing and alignment different than the interferogram of interest. Therefore, the DEM must be resampled and registered to the interferogram.

The resampling of the DEM is performed by mapping the DEM to the appropriate radar geometry through a geometric simulation of the imaging process. Referring to Fig. 3.3, the relation between the radar geometry and the DEM may be expressed as:

$$\hat{l} \cdot \hat{V} = \sin\beta \tag{3.118}$$

where

$$\hat{l} = \frac{\vec{P}(t) - \vec{P}_{DEM}}{\left|\vec{P}(t) - \vec{P}_{DEM}\right|} \tag{3.119}$$

$$\hat{V} = \frac{\vec{V}(t)}{\left|\vec{V}(t)\right|} \tag{3.120}$$

The simulation procedure is as follows: for every pixel in the DEM, iterate to find the time t_{DEM} when the spacecraft position and velocity satisfy Eqn. (3.118). The corresponding output radar geometry pixel (x, y) is:

$$x = \frac{\left|\vec{P}(t_{DEM}) - \vec{P}_{DEM}\right| - \rho_{ref}}{\Delta r} + 1 \tag{3.121}$$

$$y = \frac{\left|\vec{V}(t_{DEM})\right| t_{DEM} - a_{ref}}{\Delta a} + 1 \tag{3.122}$$

where x is the range pixel, y is the azimuth pixel, ρ_{ref} is the range to the first pixel in the output simulation, a_{ref} is the azimuth location of the first azimuth line in the output simulation, Δr is the range sample spacing, and Δa is the azimuth sample spacing.

There are two quantities assigned to each output pixel. The first quantity is the DEM height, adjusted for any biases between the DEM datum and the

output simulation datum (typically WGS-84). The second quantity is a simulated radar intensity I:

$$I = \frac{\cos^2 \theta_i}{\sqrt{1 - \cos^2 \theta_i}} \tag{3.123}$$

where

$$\cos\theta_i = -\hat{l} \cdot \hat{n} \tag{3.124}$$

and θ_i is the local incidence angle and $\hat{n}$ is the surface outward unit normal vector. The simulated radar intensity is used in the simulation registration process.

3.9.2 Simulation Registration

After the DEM has been simulated into radar coordinates, the radar simulation is coarsely registered with the interferogram. To finely register the simulation to the interferogram, several steps are required: simulation regridding, amplitude correlation, and simulation transformation.

It is noted the mapping from the input DEM to the output radar simulation is not one-to-one. In particular, input DEM pixels will be irregularly spaced in the output radar simulation. In addition, the DEM generally is of courser resolution than the radar simulation. This results in the output simulation being filled with holes. To alleviate these problems, the output simulation must be interpolated (regridded) to a uniform grid.

The regridding algorithm used in this work was a triangulation-based interpolator. The surface to be interpolated is split into a set of triangles according to the algorithm presented by Aikima [197X]. Each triangle is fit by a fifth-order polynomial with first-order and second-order partial derivatives continuous along triangle edges.

Once the simulation has been interpolated to a uniform grid, the registration process is identical to the procedure presented in Section 3.3. Amplitude correlation between the interferogram amplitude and the simulated radar intensity is used to determine a set of registration offsets. The offsets are used to estimate a set of affine transformation parameters. Finally, the affine transformation is applied to produce a regridded simulation finely registered to the interferogram.

3.9.3 Deformation Determination

With the regridded simulation finely registered to the interferogram, the topography now may be removed from the interferogram. The topography removal process is similar to the interferogram flattening procedure presented in Section 3.5 with two caveats. First, instead of the look angle being computed to the reference surface, it is computed to the topographic height z associated with a given pixel. In other words, the look angle computation in Eqn. (3.87) should be replaced with:

$$\cos\theta_{lref} = \frac{\rho_l^2 + (r+h)^2 - (r+z)^2}{2\rho_l(r+h)} \qquad l = 1,\ 2 \tag{3.125}$$

Second, if the baseline was estimated from the unwrapped phase (in areas where there was topography but no deformation), a range constant offset was estimated. The range constant offset $\Delta\rho_0$ is removed by replacing the reference surface interferometric phase computation in Eqn. (3.95) with:

$$\phi_{ref} = (4\pi/\lambda)\Delta\rho_{ref} + (4\pi/\lambda)\Delta\rho_0 \tag{3.126}$$

The final steps in the deformation determination process are to unwrap the differential interferometric phase and convert the phase values to deformation distances (i.e., scale the unwrapped phase by $\lambda/4\pi$). Ground truth measurements provide a means to verify the range constant offset was estimated properly.

3.10 TOPOGRAPHY GENERATION

After phase unwrapping and precision baseline estimation using the unwrapped phase, it is possible use the unwrapped phase to construct a topographic map. The equations are given in the class notes. More description to come.