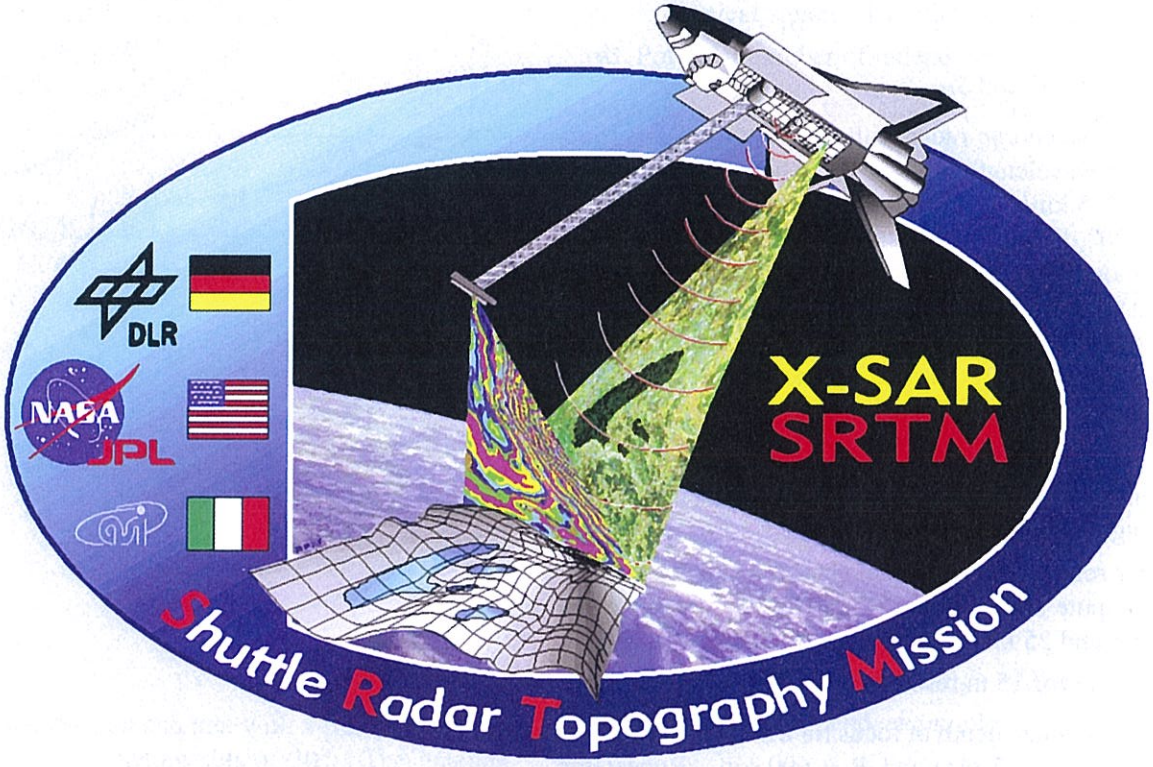
CHAPTER

15

Interferometric Synthetic-Aperture Radar



SRTM mission logo

CONTENTS

Overview

A radar echo contains both ***amplitude*** and ***phase*** information. Most imaging radar applications use algorithms that utilize only the amplitude to extract geophysical and biophysical information about the imaged scene. In some cases, the phase difference between the hh- and vv-polarized scattering amplitudes is used as well. In this chapter we consider how to exploit the ***phase difference*** between two radar echoes, associated with the same image pixel, but measured by two different systems along different directions. Measurement of such a phase difference is known as ***interferometry***, which is a well-developed technique in optics.

We can apply many of the same approaches used in optical interferometry to the modeling and analysis of radar signals. ***Radar interferometry*** allows us to measure the surface elevation of individual image pixels, which leads to the generation of ***topographic maps*** with height information. Radar interferometry also is used to measure surface motion. Because most of the applications of radar interferometry rely on the availability of high-resolution images of the terrain, the interferometric data are generally collected and processed using SAR techniques. Hence, such a radar is referred to as an ***interferometric SAR***, or InSAR or IFSAR for short.

In this chapter we introduce InSAR concepts and models, show how the interferometric radar data are used to extract useful geophysical information, and illustrate the InSAR technique with multiple examples.

15-1 Brief History of Radar Interferometry

Radar interferometry was used in Earth-based observations of Venus (Rogers and Ingalls, 1969), as a means to separate ambiguous range-Doppler echoes reflected simultaneously from Venus's northern and southern hemispheres. Subsequently, the technique was used to derive height information about the surfaces of the Moon (Zisk, 1972a, b) and Venus (Rumsey et al., 1974). Following these early planetary explorations, radar interferometric methods were adapted to measure the topography of the Earth's surface, from both aircraft and satellite platforms. Observations of interference fringes modulated by Earth surface topography were recorded by an aircraft interferometric radar and reported by Graham (1974). However, even though the recorded interference patterns clearly demonstrated that topography creates a predictable phase signal when observed interferometrically, they were difficult to analyze, and it proved impractical to extract topography information from the phase signatures.

In the early 1980s, Zebker and Goldstein (1986) implemented an airborne interferometric radar system on a NASA CV990 aircraft outfitted with a digital data-recording system. They recorded two radar channels, one from each of two separately-located radar antennas, and then processed the data using a JPL digital correlator system. They demonstrated that the phase difference between the two signals contained sufficient information so as to reconstruct the topography of the imaged radar scenes. Moreover, it was possible to correct the measured data for ever-present deviations of the aircraft flight path, which caused artifacts in the radar images. It was concluded that the digitally correlated radar interferometric technique can provide digital elevation information with an accuracy on the order of a few meters, which was far superior to the accuracy that was possible with the radar-stereo technique in use at the time.

The Goldstein-Zebker team extended the capabilities of radar interferometry by exploring the possibility of forming ***interferograms*** using radar antennas displaced *along* the flight path (rather than *across* the flight path), corresponding to temporal rather than spatial separation of the images. With this technique, they were able to map the velocities of ocean waves and currents with good accuracies (Goldstein and Zebker, 1987).

Through further work, JPL extended the use of radar interferometry to spaceborne platforms by interferometrically processing data acquired a decade earlier (1978) by the SEASAT SAR. The two major advances here were the ability to detect deformations at cm scale (Gabriel et al., 1988) and the ability to reconstruct continuous functions of radar phase, so-called ***phase***

unwrapping (Goldstein et al., 1988). The processed images detected surface deformations (between different dates) as small as a few centimeters and produced unambiguous maps of ground elevation, but despite these and other similar demonstrations, it was difficult to extend the capability of interferometric radar to mapping surface topography or deformation on a large scale, primarily due to the lack of a sufficiently large archive of interferometric radar data with supportable *in-situ* measurements.

The status of radar interferometry changed drastically after the launch of the ERS-1 satellite by the European Space Agency (ESA) in 1991. The satellite's C-band SAR acquired vast amounts of data on a global scale. At the French Center for Space Studies (CNES), Massonnet et al. (1993) developed an innovative method to isolate the interferometric deformation signature from radar interferograms of the surface, making it possible to readily view the surface displacements of geophysical processes. This development imparted a galvanizing effect on the radar community, and served to establish the credibility of radar interferometry. It also spurred the introduction of new methods and algorithms by numerous investigators (e.g., Zebker and Villasenor, 1992; Just and Bamler, 1994; Goldstein, 1995; Massonnet and Feigl, 1998; Zebker et al., 1997; Rosen et al., 2000) and led to two InSAR satellite missions: the ***Shuttle Radar Topography Mission*** (SRTM) in 2000 and the ***TanDEM-X*** mission in 2010.

The launch of the Envisat satellite by ESA in 2002 extended the continuous record of worldwide observations begun 10 years earlier with the ERS satellites. Other satellites, such as the Canadian Radarsat series, also added to the growing archive of available data. As data coverage became more and more routine, the InSAR technique matured into a standard analytical tool for the geophysics community (Peltzer et al., 2001; Bürgmann et al., 2002; Fielding et al., 2005; Pritchard and Simons, 2004). The technique further advanced by coupling InSAR analysis more closely with crustal deformation modeling (Amelung et al., 2000; Jonsson et al., 2002), by adding new capabilities such as vector deformation mapping (Bechor and Zebker, 2006), and also by moving into new application areas such as hydrological applications (Amelung et al., 1999; Hoffmann et al., 2003).

Since 2000, the focus of technological development in the radar interferometry arena has shifted towards the production of time series of interferograms so as to reconstruct surface-deformation histories, rather than snapshots at a single time of deformation. A number of new methods, including ***stacking*** (Lyons and Sandwell, 2003), ***persistent scattering*** (Ferretti et al., 2000), and ***small baseline subset analysis*** (Berardino et al., 2002), have become commonplace. The persistent-scattering approach was developed by Ferretti and colleagues (Ferretti et al., 2000, 2001, 2004), who applied it mainly to the analysis of urban structures with well-defined scattering centers. Its applicability to natural terrains was pioneered by Hooper et al. (2007), who showed that many surfaces or vegetation covers in fact exhibit scattering that does not scintillate in phase over many months or years. Improved persistent-scattering detection algorithms based on maximum likelihood methods (Shanker and Zebker, 2007) have led to more robust implementations exploiting these longer-lived pixels.

15-2 2-D versus 3-D Measurements

When we measure the returns of multiple pulses and then process them together to realize high-resolution capability in both the range and azimuth directions, we generate a radar image that displays the 2-D backscatter distribution of the imaged scene. Each pixel in the image has a time delay associated with its distance R from the radar, and a corresponding two-way ***phase delay*** given by

$$\phi = -\frac{4\pi}{\lambda} R, \tag{15.1}$$

where λ is the EM wavelength. The measured phase delay may include a noise contribution, but for simplicity we ignore it for the present time.

15-2.1 Interferometric Phase

► InSAR methods are based on combining radar returns from two different antennas, displaced from each other in space or time. ◄

If the antennas are displaced in space, we denote the spatial separation as the radar ***spatial baseline***. The two radar images of the scene differ due to the ***parallax*** of the two observations, generating a distortion of the one image compared with the other. The distortion is related to the surface topography.

The scenario shown in Fig. 15-1 depicts a SAR imaging a scene with topographic variations. The radar return signal from a particular range bin includes contributions from all equirange points on the surface corresponding to that range bin. Except for the few situations in which the radar shadow can be used to estimate the height of an object or feature, it is difficult to extract height information from single radar images, because the radar image is fundamentally two-dimensional in nature, whereas the scene is three-dimensional. Many surface points at different heights can return echoes simultaneously.

When two antennas are used to generate two separate images, a point P on the terrain surface, as depicted in Fig. 15-2, is at the intersection of two circles, one of radius R_1—with its center at antenna A_1—and another of radius R_2. If we know the positions of the two antennas, and if we have accurate measurements of R_1 and R_2, we can determine the 3-D coordinates of P unambiguously.

Prior to the advent of radar interferometry, the traditional method that was used to extract topographic information from pairs of radar images of the same scene relied on the radar-stereo technique described in Section 14-11. The stereo technique does not require or use phase information, but it has limited sensitivity to topography and it is inherently susceptible to the presence of speckle in the radar images. Furthermore, the two ranges R_1 and R_2 to a specific feature in the scene have to be significantly different in length in order for the stereo technique to provide a useful estimate of height.

The two receiver configurations commonly used in radar interferometry are displayed in Fig. 15-3:

(1) ***Cross-track with two independent radars***, each with its own transmit/receive antenna [Fig. 15-3(a)].

(2) ***Cross-track with common transmit antenna***, but separate receive antennas [Fig. 15-3(b)]. The transmit antenna also serves as the receive antenna for one of the radars and an auxiliary antenna is used as the receive antenna for the second radar.

(3) ***Along-track*** with two independent radars separated in time [Fig. 15-3(c)].

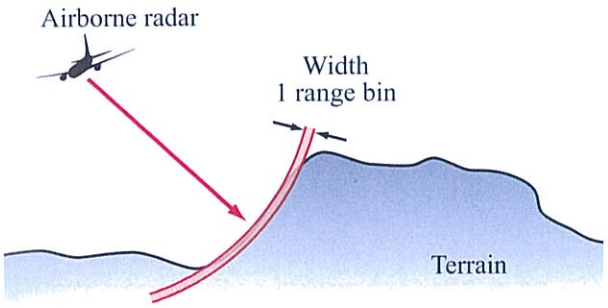


Figure 15-1: Radar returns from all points within any range bin are received at once. Thus conventional radar images cannot properly depict topographic information.

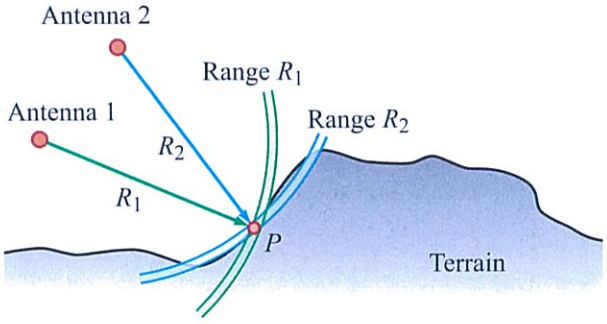
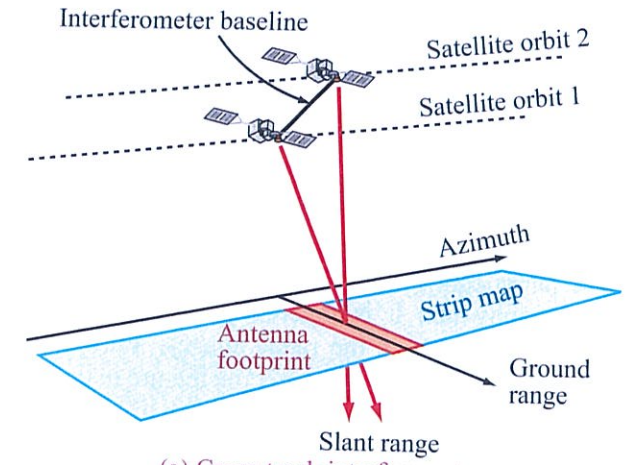
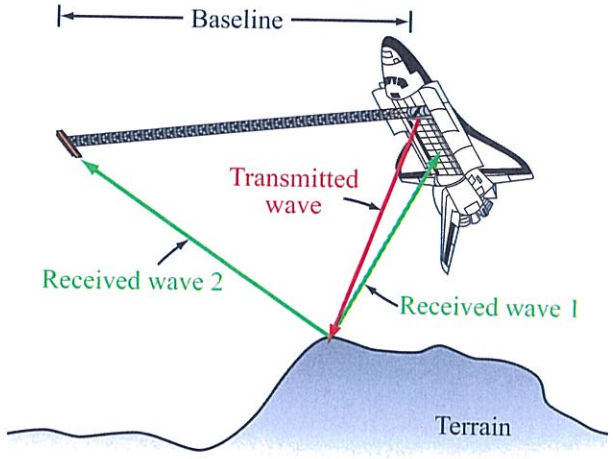


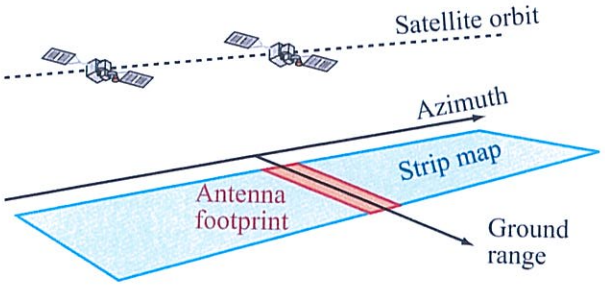
Figure 15-2: Distances to a point P from Antenna 1 and Antenna 2 are R_1 and R_2, respectively. The two arcs of these radii cross at the true position of the point in three-dimensional space.



(a) Cross-track interferometer with two independent radars



(b) Cross-track interferometer with common-transmitter radars



(c) Along-track interferometer

Figure 15-3: Cross-track interferometric configurations in (a) and (b) and along-track configuration in (c).

The two antennas may be be separated from one another in the cross-track direction, as shown in Fig. 15-3(a) and (b), which means that the two antennas comprising the interferometer are flown on separate, parallel tracks. Such an interferometer is called a ***cross-track interferometer***. Alternatively, the two antennas can be flown along the same track [Fig. 15-3(c)], with one displaced from the other by a certain distance or flown at a different time, and we refer to this geometry as an ***along-track interferometer***.

▶ As we see in later sections, the cross-track configuration provides topography (height) information, whereas the along-track configuration provides information about the movement of the surface between the observation times associated with the two images. ◀

The radar interferometric measurement technique relies on the use of phase information and it takes advantage of the ***parallel-ray approximation***, which requires ranges R_1 and R_2 in Fig. 15-4 to be different, but by only a very small amount. That is $|R_1 - R_2| \ll R_1$. The spacing between the two antennas—called the

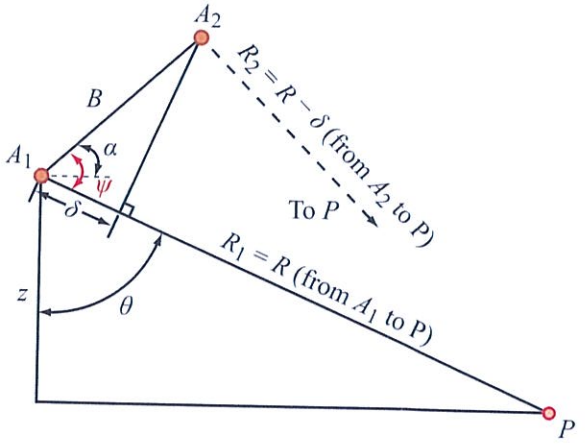


Figure 15-4: Parallel-ray approximation for InSAR geometry with $R_1 = R$ and $R_2 = R - \delta$. Note that R_1 is from A_1 to P.

baseline B—is greatly exaggerated in the figure relative to R_1 and R_2. For convenience, we label R_2 as $R-\delta$ and R_1 as R.

For the two-independent-radars configuration [Fig. 15-3(a)], the roundtrip path length between A_1 and P is $2R$ and for the second radar, it is $2R-2\delta$. The corresponding *interferometric phase difference* (of radar A_1 relative to radar A_2) is

$$\phi_{\text{int}} = \phi_2 - \phi_1 = -2k\delta = -\frac{4\pi\delta}{\lambda}. \qquad (15.2a)$$

(independent radars)

When the same transmit antenna is used, as in Fig. 15-3(b), the two-way path length for receive antenna A_2 is only $2R-\delta$, in which case the phase difference is

$$\phi_{\text{int}} = -k\delta = -\frac{2\pi\delta}{\lambda}. \qquad (15.2b)$$

(common-transmit antenna)

The two expressions for ϕ_{int} can be combined into

$$\phi_{\text{int}} = -\frac{2\pi n\delta}{\lambda}, \qquad (15.2c)$$

where

$$n = \begin{cases} 1 & \text{for common-transmit antenna,} \\ 2 & \text{for two independent radars.} \end{cases} \qquad (15.2d)$$

In Fig. 15-4, we denote the orientation of the baseline B with respect to the horizontal axis by the angle α. For triangle A_1A_2P, angle ψ is given by

$$\psi = \frac{\pi}{2} - \theta + \alpha, \qquad (15.3)$$

and from the law of cosines

$$(R-\delta)^2 = R^2 + B^2 - 2RB\cos\left(\frac{\pi}{2} - \theta + \alpha\right). \qquad (15.4)$$

However, since $\delta \ll R$, we can apply the parallel-ray approximation which leads to

$$\delta = B\sin(\theta - \alpha). \qquad (15.5)$$

For a point P at height z relative to the height of A_1,

$$z = R\cos\theta. \qquad (15.6)$$

Knowledge of the range R is available from the radar data. Through Eqs. (15.2a) and (15.2b), δ is determined from the phase difference measurement, which in turn leads to a determination of θ from Eq. (15.5), and finally to a determination of the topographic height z from Eq. (15.6).

15-2.2 Height Measurement Precision

A critical parameter for the design of a topographic mapping system is the *estimated height error* (precision) associated with the measurement of z, to which we assign the standard deviation symbol s_z. The precision s_z is related to the precision with which δ can be measured. Application of the chain rule

$$\frac{\partial z}{\partial \delta} = \frac{\partial z}{\partial \theta}\frac{\partial \theta}{\partial \delta}, \qquad (15.7a)$$

and using Eqs. (15.5) and (15.6) to compute the derivatives

$$\frac{\partial z}{\partial \theta} = -R\sin\theta, \qquad (15.7b)$$

$$\frac{\partial \delta}{\partial \theta} = -B\cos(\theta - \alpha), \qquad (15.7c)$$

leads to

$$\frac{\partial z}{\partial \delta} = \frac{-R\sin\theta}{B\cos(\theta - \alpha)}. \qquad (15.8)$$

Thus, a change in δ of s_δ leads (approximately) to a change in z of s_z given by

$$s_z = \frac{R\sin\theta}{B\cos(\theta - \alpha)}\, s_\delta. \qquad (15.9)$$

The quantities s_z and s_δ represent the standard deviations associated with the inferred values of z and δ, respectively. The minus sign in Eq. (15.8) was not carried over to Eq. (15.9) because standard deviations are positive quantities representing deviations relative to mean values.

For typical values of θ and α, the factor $\sin\theta/\cos(\theta-\alpha)$ seldom deviates from 1 by more than a factor of 2 (larger or smaller), so we may approximate Eq. (15.9) as

$$s_z \approx \frac{R}{B}\, s_\delta. \tag{15.10}$$

The precision with which δ can be measured depends on the precision with which the interferometric phase difference ϕ_{int} can be measured. Using Eq. (15.2c),

$$s_\delta = \frac{\lambda}{2\pi n}\, s_{\phi_{\text{int}}}, \tag{15.11a}$$

which leads to

$$s_z = \frac{R}{B}\,\frac{\lambda}{2\pi n}\, s_{\phi_{\text{int}}}. \tag{15.11b}$$

Here, $n = 1$ for the common-transmitter interferometer and $n = 2$ for the two-independent-radar interferometer. In practice, it is possible to measure ϕ_{int} with an accuracy of a few degrees. If, for demonstration purposes, we choose $n = 2$ and we set $s_{\phi_{\text{int}}} = 7.2^\circ = 0.1257$ rad, we obtain

$$s_\delta = \frac{\lambda}{100} \tag{15.12a}$$

and

$$s_z = \frac{R}{B}\cdot\frac{\lambda}{100}. \tag{15.12b}$$

By way of illustration, the relevant parameters for the ERS SAR are: $R = 800$ km, $B = 1$ km (for two orbital passes separated by 1 km), and $\lambda = 6$ cm. Using Eq. (15.12b), we obtain

$$s_z \approx \frac{R}{B}\,\frac{\lambda}{100} \approx \frac{1}{2}\ \text{m}.$$

In contrast, the measurement height precision when only R is measured and no phase information is used—as in the radar stereo case—is 7.2 km! While this is a rather ideal noise-free scenario for the interferometric radar, it does illustrate the power of the interferometric approach to topography estimation.

To illustrate with a real example, we show in Fig. 15-5 SAR images at various stages of interferometric processing, leading to a 3-D topography map. The images and associated processing steps are discussed in the following sections.

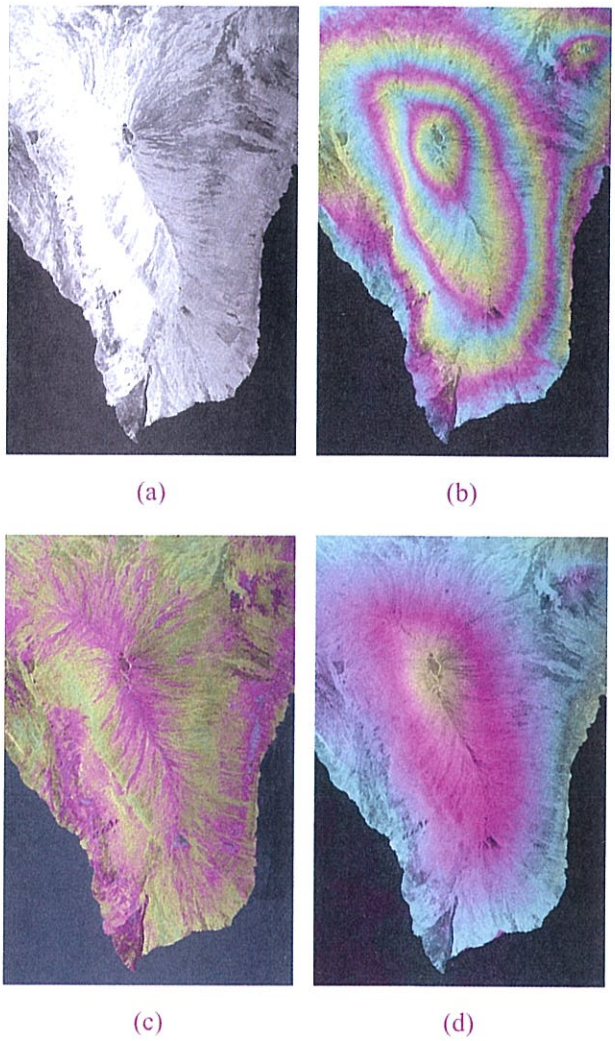


Figure 15-5: Example of interferometric SAR processing: (a) SAR magnitude image, (b) phase difference (interferogram) between two images with flat-Earth phase removed, (c) correlation map between interferogram component scenes, and (d) inferred surface topography.

15-2.3 The Role of SNR

Clearly, interferometric performance depends critically on how well the interferometric phase difference ϕ_{int} can be measured, which in turn depends on the signal-to-

noise ratio (SNR). We now consider how SNR impacts the phase-difference measurement uncertainty $s_{\phi_{\text{int}}}$.

The variance of the measured phase difference, $s^2_{\phi_{\text{int}}}$, is due to the random phase component introduced by the noise accompanying the signal, and it is approximately proportional to the inverse of the SNR,

$$s^2_{\phi_{\text{int}}} \approx \frac{1}{\text{SNR}} . \tag{15.13}$$

To secure the value $s_{\phi_{\text{int}}} = 0.1257$ rad that we used in the example leading to Eq. (15.12), it is necessary that SNR = 63.29, or, equivalently, 18 dB.

Collecting Eqs. (15.9), (15.11b), and (15.13) leads to

$$s_z = \frac{\lambda}{2n\pi} \cdot \frac{R\sin\theta}{B\cos(\theta-\alpha)} \cdot \frac{1}{\sqrt{\text{SNR}}} , \tag{15.14a}$$

$$n = \begin{cases} 1 & \text{for common transmitter,} \\ 2 & \text{for independent radar.} \end{cases} \tag{15.14b}$$

15-3 Cartographic Corrections

In order to further develop the interferometric technique for cartographic applications, we need to understand how topography affects the radar images. In the following, we briefly consider ***geocoding*** and ***georeferencing*** and their relationships to radar interferometry.

As described in Chapter 14, a typical ***slant range*** SAR image is presented in ***radar coordinates***, a coordinate system in which the range coordinate is the slant range and the azimuth coordinate corresponds to the along-track distance. In a ***ground range*** SAR format, the slant range is converted to ground range (Fig. 15-6). Surface topography often leads to foreshortening or even layover in the resulting image. This is caused by the nonlinear mapping of slant range into ground range, plus any effects from uncompensated topography.

Figure 15-7 illustrates various geometrical distortions that occur when the radar-coordinate data are mapped into a geographically corrected coordinate system. When the surface is flat, this transformation is straightforward, but it is more complicated when surface topography is present. As a consequence of the nonlinear mapping from slant to ground range, the shapes of objects are not preserved and ground range lines are not imaged as straight lines. This is illustrated in Fig. 15-8.

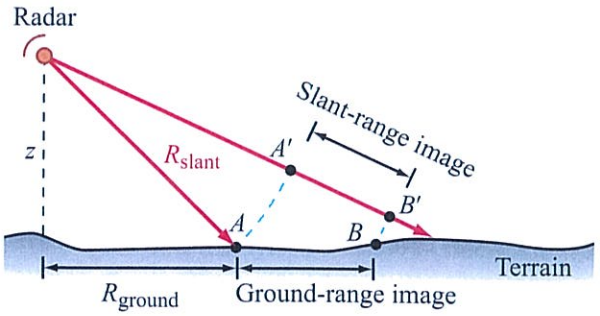


Figure 15-6: Relationship between slant-range and ground-range image presentations for a side-looking radar.

Multipath effects in radar reflection can lead to ghosting, artifacts, or replicated targets in the SAR image. Figure 15-9 illustrates how multipath can result in target reflections appearing at two different slant ranges. Consequently, the target appears at two different slant ranges in the image. Typically, the direct-path target image is well focused, but the longer multipath image is often less well focused due to the mismatch between the azimuth matched filter at the apparent range and the range migration of the target multipath.

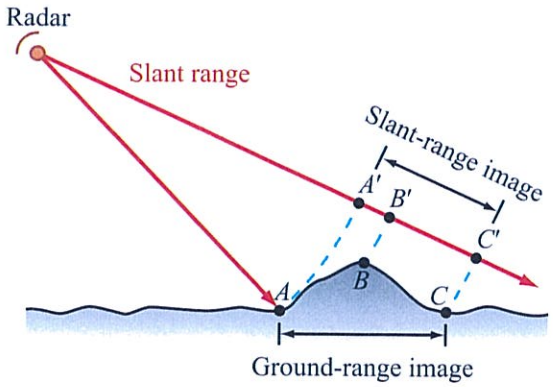
These distortions result from the two-dimensional imaging of conventional SAR. Fortunately, by using interferometric techniques, InSAR generates three-dimensional information that allows us to estimate the terrain height, which allows us to correct for these distortion effects. Note that slant range R_{slant} and ground range R_{ground} are related to height by (Fig. 15-6)

$$R^2_{\text{ground}} + z^2 = R^2_{\text{slant}} , \tag{15.15a}$$

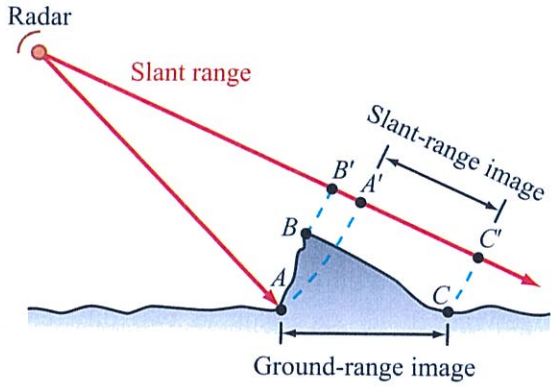
so if we resample a slant range image according to

$$R_{\text{ground}} = \sqrt{R^2_{\text{slant}} - z^2} , \tag{15.15b}$$

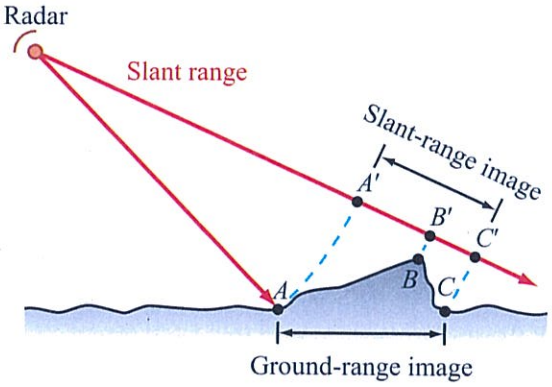
we can properly recover the undistorted ground-range image from the slant-range image. With the additional



(a) Foreshortening

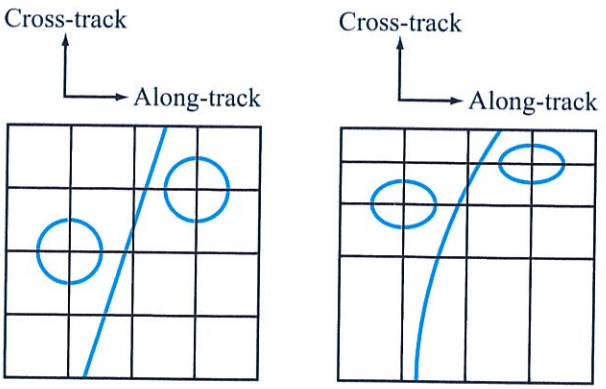


(b) Layover



(c) Shadow

Figure 15-7: Geometric distortions: (a) foreshortening, which occurs when the distance between A' and B' in the slant range is less than the ground-range distance between A and B, (b) layover, where B' is imaged in front of A', and (c) shadow, where the unilluminated mountain back yields a dark region between B' and C'.



(a) Slant-range format (b) Ground-range format

Figure 15-8: Distortions resulting from mapping (a) ground-range to (b) slant-range format. Note that slant-range images do not faithfully reproduce shapes and lines in ground-range images.

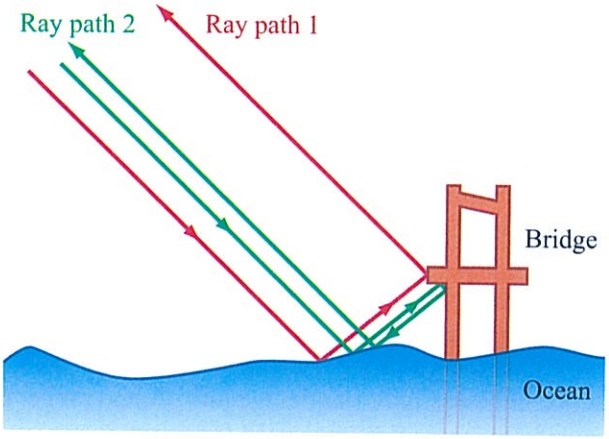


Figure 15-9: Multipath distortion. Two separate ray paths, with different total lengths, illuminate a bridge, causing it to appear twice in the radar image.

knowledge of height, a geometrically accurate ground range image can be formed from the slant range. This process is known as ***SAR georeferencing***, or SAR geocoding, although the latter is also concerned with properly locating each SAR image pixel in an absolute sense. Shadows and extreme foreshortening leading to layover still provide distortions in the georeferenced image. Backscatter variations due to topographic slope have to be corrected separately.

15-4 Forming the Radar Interferogram

A radar interferometer employs two antennas displaced from each other. In ***across-track interferometry***, the two antennas are displaced in the across-track direction, while in ***along-track interferometry*** they are displaced in the along-track direction. In ***single-pass interferometry***, the two antennas are mounted on a single platform and the data are collected simultaneously. In ***multiple-pass interferometry***, data collections are made at different times using either multiple platforms or multiple passes over the data collection area with the same platform.

In Section 15-2, we established that the phase difference between the signals received by the two antennas contains information about the surface topography. An image of the phase difference is known as an ***interferogram***. In the following, we examine in detail the formation of the interferogram.

15-4.1 Displacement versus Range

We start by restating Eqs. (15.5) and (15.6):

$$\delta = B\sin(\theta - \alpha), \tag{15.16}$$

and

$$z = R\cos\theta. \tag{15.17}$$

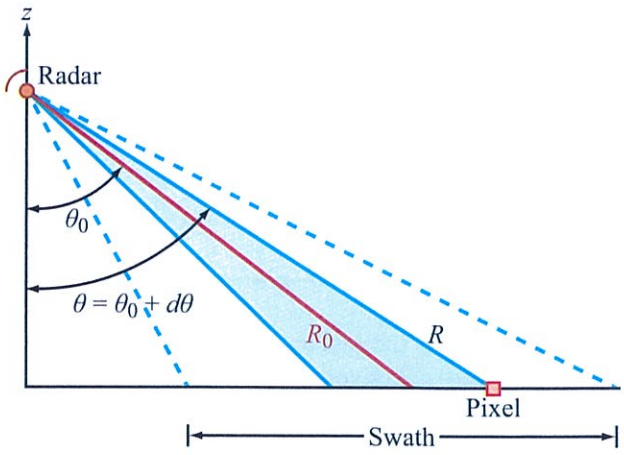


Figure 15-10: Illuminated swath is typically centered in angle about a mean look angle θ_0.

When creating and processing a radar interferogram, it is useful to reference the range location of a pixel to the center of the image swath. In Fig. 15-10, θ_0 represents the incidence angle along the center of the swath and the corresponding range is R_0. The elevation beamwidth of a satellite SAR is typically on the order of a few degrees (e.g., for ERS-1/2 SAR, θ extends from 20° to 26°, with θ_0 at 23°, and for JERS-1, θ extends from 32° to 38°, with θ_0 at 35°). This means that for the pixel P in Fig. 15-10, the angular deviation $d\theta$ from θ_0 is at most 3°. Replacing θ with $\theta_0 + d\theta$ in Eqs. (15.16) and (15.17) and using the approximations $\sin(d\theta) \approx d\theta$ and $\cos(d\theta) \approx 1$ gives

$$z = R\cos(\theta_0 + d\theta) \approx R[\cos\theta_0 - \sin\theta_0\, d\theta] \tag{15.18}$$

and

$$\begin{aligned}\delta &= B\sin(\theta_0 - \alpha + d\theta)\\ &= B[\sin(\theta_0 - \alpha) + \cos(\theta_0 - \alpha)\, d\theta].\end{aligned} \tag{15.19}$$

Solving Eq. (15.18) for $d\theta$ yields

$$d\theta = \frac{R\cos\theta_0 - z}{R\sin\theta_0}, \tag{15.20}$$

and when this is inserted in Eq. (15.19), the latter becomes

$$\delta = B\sin(\theta_0-\alpha)+B\cos(\theta_0-\alpha)\,\frac{\cos\theta_0 - z/R}{\sin\theta_0}$$
$$= B\sin(\theta_0-\alpha)+\frac{B\cos(\theta_0-\alpha)}{\tan\theta_0}$$
$$-\frac{B\cos(\theta_0-\alpha)}{\sin\theta_0}\,\frac{z}{R}\,. \qquad (15.21)$$

Expanding z/R of the last term in Eq. (15.21) about R_0, where $R = R_0 + dR$, gives

$$\frac{z}{R} = \frac{z}{R_0+dR} = \frac{z}{R_0}\left[\frac{1}{1+(dR/R_0)}\right]$$
$$\approx \frac{z}{R_0}\left(1-\frac{dR}{R_0}\right) = \frac{z}{R_0}-\frac{z\,dR}{R_0^2}\,. \qquad (15.22)$$

Incorporating Eq. (15.22) into Eq. (15.21) leads to

$$\delta = B\sin(\theta_0-\alpha)+\frac{B\cos(\theta_0-\alpha)}{\tan\theta_0}$$
$$-\frac{B\cos(\theta_0-\alpha)}{\sin\theta_0}\left(\frac{z}{R_0}-\frac{z}{R_0^2}\,dR\right). \qquad (15.23)$$

Upon using the approximation $z \approx R_0\cos\theta_0$, Eq. (15.23) simplifies to

$$\delta = B\sin(\theta_0-\alpha)+\frac{B\cos(\theta_0-\alpha)}{\tan\theta_0}\,\frac{dR}{R_0}\,. \qquad (15.24)$$

We can define the baseline at the nominal incidence angle θ_0 in Fig. 15-11 as having a ***parallel*** component $B_{\|}$ and a ***perpendicular*** component $B_{\perp}$ given by

$$B_{\|} = B\sin(\theta_0-\alpha),$$
$$B_{\perp} = B\cos(\theta_0-\alpha).$$

In other words, the total offset between the two images δ can be expressed as a function of the two ***InSAR baseline components*** $B_{\|}$ and $B_{\perp}$ as

$$\delta = B_{\|} + B_{\perp}\,\frac{dR}{R_0\tan\theta_0}\,. \qquad (15.25)$$

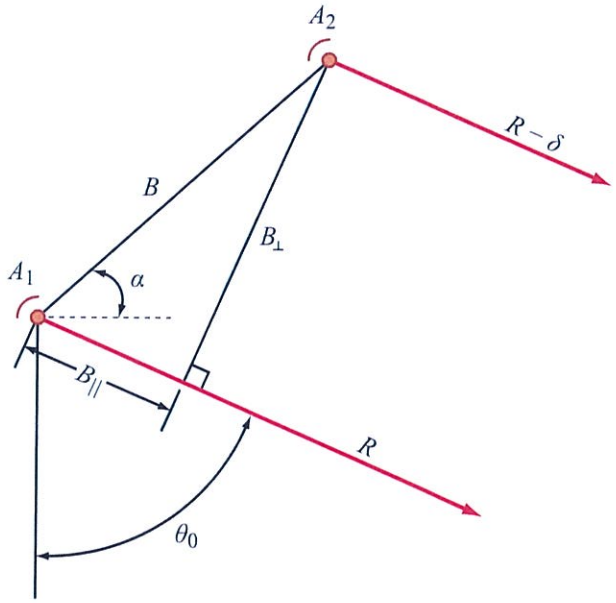
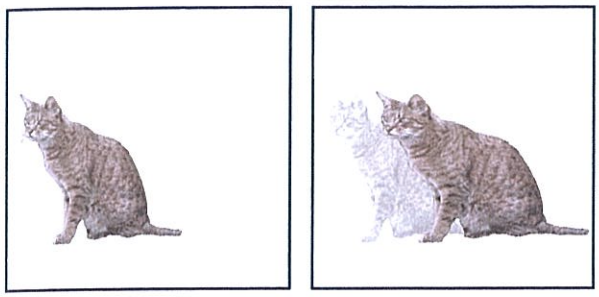


Figure 15-11: Parallel and perpendicular components of the interferometer baseline B as defined by incidence angle θ_0. Range R is from A_1 to target point.

This result suggests that δ has a constant shift equal to the parallel component of the baseline, and a linearly increasing shift (or ***stretch***) that varies inversely with the geometric factor $R_0\tan\theta_0$. Figure 15-12 schematically illustrates the spatial shift and scaling that occurs between the images created from the two antennas. Because of the scale and shift, a target observed at one pixel location in a particular image, is at a different pixel position in the second image

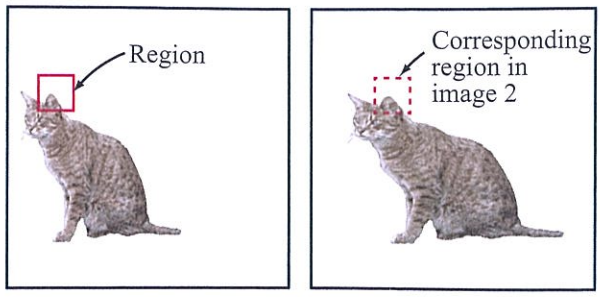
When forming the interferogram, the phase difference relates the phase in a pixel in image 1 to the corresponding phase of the same target in image 2, which may be at a different pixel location. Thus, it is necessary to establish a formatting relationship between each pixel in image 2 and its corresponding location in image 1.

In practice, this is accomplished by estimating the shift in image 2 relative to image 1 by identifying corresponding points in the two images and solving for



(a) Antenna 1 image (b) Antenna 2 image: shifted and scaled

Figure 15-12: (a) Image as seen by antenna 1 and (b) image as seen by antenna 2. The combination of angle-dependent mapping into slant range coordinates causes antenna 2 to see a shifted and scaled version of the image seen by antenna 1.



(a) Image 1 (b) Image 2

Figure 15-13: Schematic illustration of corresponding small regions in the two interferometer images: (a) image from antenna 1, with a particular small region identified, and (b) corresponding area in the image from antenna 2.

an equation for the offset. This is termed ***determining the interferometric offset field*** or ***image coregistration***.

15-4.2 Offset Determination

By comparing the positions of the same point scatterers in the two images, we can derive an empirical equivalent of the relation given Eq. (15.25). Because the radar data contain some noise, the offset determination process is imperfect, resulting in a somewhat noisy estimate. This is overcome by estimating the offsets for a large number of locations and solving for the parameters determining δ using a least-squares approach to find a well-fitting low-order approximate relation.

There are many ways to determine the offsets, but for the sake of brevity, only one recipe is considered here. Since in the general case there may be shifts both in range and azimuth, we adopt a 2-D approach to solving the shift estimation problem

► Offsets are found by choosing a small region in image 1, and cross-correlating it over a similar area in image 2 to find the best matching terrain. The position of the correlation peak then gives the offset at that location. ◄

The approach begins by identifying a small region within image 1, as shown in Fig. 15-13. From the imaging geometry, we know, approximately but not exactly, where the region should occur in image 2. To find the exact location, the cross-correlation of the region locations is computed. Let $\mathbf{C}_1(m,n)$ be the complex image in the selected region in image 1 at location (m,n), and let $\mathbf{C}_2(m_2,n_2)$ be the complex image of a region in image 2 centered at location (m_2,n_2). The ***cross-correlation function*** $\mathbf{K}_C(i,k)$ is computed using

$$\mathbf{K}_C(i,k) = \frac{\sum_i \sum_k \mathbf{C}_1(m,n)\, \mathbf{C}_2^*(m_2+i,\, n_2+k)}{\sqrt{\overline{\mathbf{C}_1}\;\overline{\mathbf{C}_2}}}, \tag{15.26}$$

with

$$\overline{\mathbf{C}_1} = \sum\sum \mathbf{C}_1(m,n)\, \mathbf{C}_1^*(m,n),$$

$$\overline{\mathbf{C}_2} = \sum\sum \mathbf{C}_2(m_2+i,\, n_2+k)\, \mathbf{C}_2^*(m_2+i,\, n_2+k),$$

and where the sums are computed over an area of the size of $\mathbf{C}_1$. In practice, a two-dimensional signal processing window is used in this computation. Because the initial offset is not known exactly, in practice $\mathbf{C}_2$ is typically selected to be larger than $\mathbf{C}_1$ as illustrated in Fig. 15-14.

The magnitude of the cross-correlation function $\mathbf{K}_C(i,k)$ describes how similar the two image regions are when region 2 is spatially shifted by (i,k). The

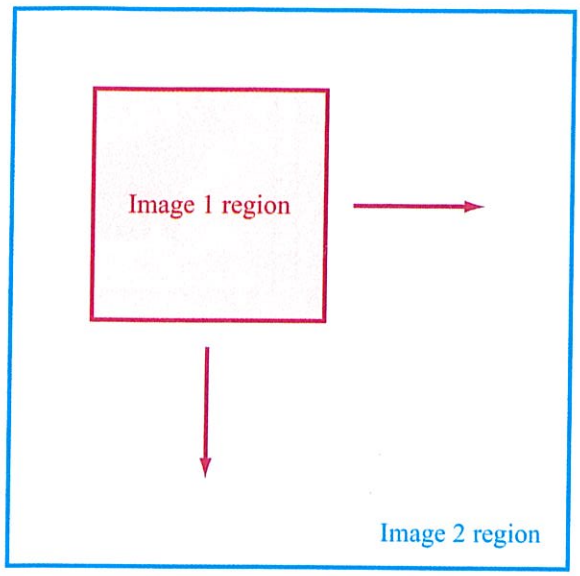


Figure 15-14: In computing the cross-correlation function, a small region from image 1 is cross-correlated against a larger piece of image 2 to enable a search range for the correlation peak.

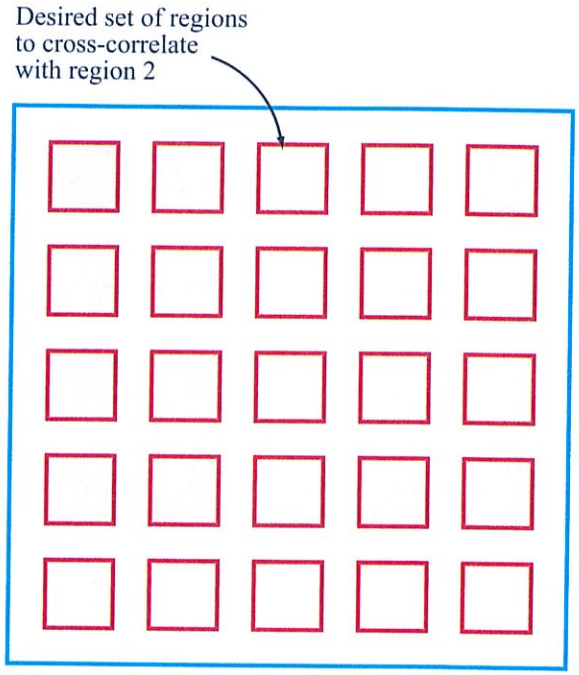


Figure 15-15: Grid of image 1 regions to cross-correlate with image 2.

cross-correlation is maximized at the shift that causes the complex regions to be the most similar. This peak, located at (i_0, k_0), is the image offset we seek.

Since the images are complex, the complex correlation described in Eq. (15.26) can be directly implemented to find the required shift. However, when the phase noise is large (due, for example, to low SNR in regions of the image with low σ^0), the complex correlation may lead to inaccurate estimates. It is useful to correlate the image powers (pixel magnitudes) rather than the complex pixel values. In this case the complex values in Eq. (15.26) are replaced with the pixel magnitude squared; i.e., by $|\mathbf{C}_1|^2$ and $|\mathbf{C}_2|^2$. This eliminates noise-related issues in regions with inadequate phase fidelity.

Note that because we are multiplying two detected images together when computing the cross-correlation, the bandwidth of the cross product (in the correlation process) effectively is twice the bandwidth of the original data. This can lead to aliasing if the original image data are not oversampled by at least a factor of 2. Hence, for best results we generally interpolate each region by at least a factor of 2 in each direction before performing the correlation.

The method described thus far is for a particular subregion of image 1. By analyzing multiple regions, the offset estimate can be improved. To obtain a good set of estimates for the next step, a good rule of thumb is to examine 10–20 locations in range and in azimuth as indicated in Fig. 15-15. For a given offset estimate, once the region correlations are computed, δ can be estimated empirically using a linear least-squares approach. Figure 15-16 is a typical example of a plot of the offset estimates for individual regions as a function of the range R. Also shown is the linear least-squares fit to the data. From the equation of the line, we can estimate both the parallel and perpendicular components of the baseline. The y intercept corresponds to $B_{\|}$ at the reference pixel defined by θ_0 and the slope of the line in Fig. 15-16 corresponds to $B_{\perp}$. The initial estimates of

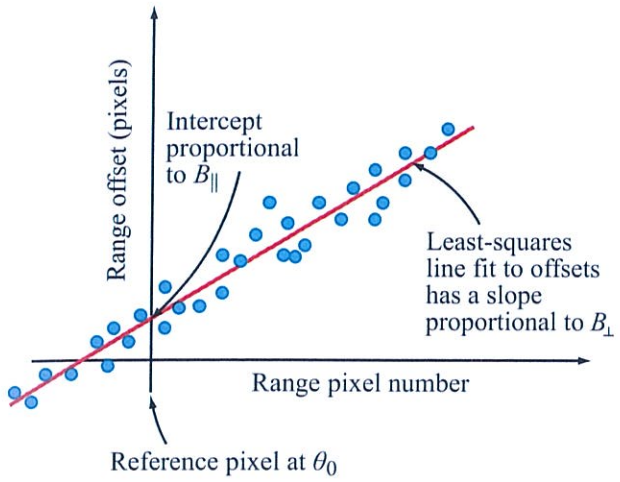


Figure 15-16: Plot of noisy measured offsets as a function of range location and the least-square fit of the data. The slope and intercept of the line yield the parallel and perpendicular baseline components.

$B_{\|}$ and $B_{\perp}$ can be refined later using known topographic corrections.

The next step is to resample image 2 to the coordinates of image 1 using the relation for δ given by Eq. (15.25) with the empirically determined values of $B_{\perp}$ and $B_{\|}$. In addition, if there is an azimuth shift due to timing considerations, this is removed at this time as well.

Finally, the interferogram $\mathbf{I}(x,y)$ is formed by computing

$$\mathbf{I}(x,y) = \text{image } 1(x,y) \\ (\text{image } 2(x+\delta,\ y+\alpha))^*, \qquad (15.27)$$

where the complex conjugate multiplication results in the desired phase-differencing operation. In this equation x corresponds to the pixel index in the range direction, y is the pixel index in the azimuth direction, and α is the azimuth offset.

In summary, the steps involved in formatting the interferogram include:

(1) Processing the two SAR images measured by the two receive antennas.

(2) Measuring the offsets by:

- (a) selecting a small region in image 1 and its approximate equivalent in image 2,
- (b) interpolating by 2 or more,
- (c) detecting complex values to obtain power images of several small regions,
- (d) cross-correlating and recording of offset peak locations,
- (e) repeating for many different locations,
- (f) obtaining a least-square fit for the offset values as a function of range; also determining azimuth offset, if any, and
- (g) solving for $B_{\perp}$ and $B_{\|}$ from the slope and intercept of the line.

(3) Resampling image 2 to image 1 coordinates.

(4) Cross-multiplying image 1 by the conjugate of image 2 to form the interferogram.

15-4.3 Multilooking

Having computed the interferogram, it is often desirable to ***multilook*** the data in order to improve the accuracy of the phase estimate, at the expense of spatial resolution. As described in Chapter 5, multilooking means that pixel intensities are averaged in one or both dimensions to reduce fluctuations. Multilooked data are much smoother than single-look imagery and can be downsampled without losing much information. This also reduces computation and the required data volume for storage.

For image data, recall from Chapter 5 that multilooking involves averaging power values rather than complex values. As a general rule, in most radar calculations, quantities representing second-order statistics are averaged in a multilook processor. For example, when considering the sum of two Gaussian variables, the variances, not the standard deviations, add. Thus when we average two pixels in a radar image, the mean power is preserved, not the mean standard deviation. It is this property that leads to the reduction of relative uncertainty for multilooked pixels. Because the

interferogram pixels are derived from the product of two field quantities, these are essentially powers, even though they have a distinct phase associated with each of them.

► Hence, averaging the interferogram pixels directly approaches the mean power while preserving the average phase value. ◄

To help understand the difference from another perspective, note that in the multilooking case (when a single image is involved), the pixel phases are uniformly distributed and are essentially independent from pixel to pixel, so when they are added together to reduce noise, the powers add. For the interferogram (which is the phase difference between two images), the interferometric phase changes slowly from pixel to pixel, so coherent integration to reduce noise is very helpful.

15-4.4 Correlation

An important measure of interferogram quality is the ***degree of correlation*** between the two images, or how closely the phase in one image tracks that of the other. This quantity varies between 0 for images that are completely independent of each other to 1 for images with identical phases. The image correlation can be computed when generating the multilooked image by applying the following definition:

$$\rho = \frac{\sum \text{image } 1_i \text{image } 2_i^*}{\sqrt{\sum \text{image } 1_i \text{image } 1_i^*}\sqrt{\sum \text{image } 2_i \text{image } 2_i^*}}, \tag{15.28}$$

where image 1_i and image 2_i are the complex values of a corresponding point in image 1 and image 2 after image 2 has been resampled according to the estimated shift. Note that the numerator is the interferogram while the denominator is the product of the image *amplitudes*, not powers.

In general, the degree of correlation is a function of SNR and the local target characteristics (to be described later).

15-5 Decorrelation

► ***Correlation*** is quite fundamental to understanding and interpreting interferometric measurements and is closely coupled with our ability to estimate interferometric phase accurately. ◄

In general, accurate phase measurements require high correlation. Unfortunately, there are a number of mechanisms that can cause signal decorrelation. These are considered in the following subsections.

15-5.1 Speckle

Radar echoes are considered to be "correlated" with each other if the measured phases and amplitudes coincide and hence represent the "same" interactions of the radar signal with the scattering terrain. In an imaging radar, this means that the observed scattering characteristics or "speckle" patterns are nearly the same. As introduced in Chapter 5, ***speckle*** is an inherent consequence of the coherent nature of radar. It is observable as the "graininess" evident in single-look radar images. It arises due to EM interactions with multiple scattering centers within a single resolution element (or ***cell***) of the radar rather than actual variations in σ^0 of the surface.

Consider the backscattered voltage V from a given resolution cell when viewed by two radars from slightly different orientations. Since V is the coherent sum of the signals backscattered from the individual scatterers contained in the resolution cell, a change in orientation causes changes in the ranges to the individual scatterers, and hence a change in the magnitude and phase of V. This leads to ***geometric decorrelation***, and the degree of decorrelation depends on the amount of angular offset between the two antennas relative to the target.

15-5.2 Decorrelation Model

In this section we quantify the effects of decorrelation in InSAR radar images. Consider two complex radar signals, $\mathbf{V}_1$ and $\mathbf{V}_2$, acquired by two interferometric antennas, located very near each other, and passing through separate receivers. The signals can be modeled

as consisting of a common signal component $\mathbf{C}$ and individual noise components $\mathbf{N}_1$ and $\mathbf{N}_2$:

$$\mathbf{V}_1 = \mathbf{C} + \mathbf{N}_1, \tag{15.29a}$$

$$\mathbf{V}_2 = \mathbf{C} + \mathbf{N}_2. \tag{15.29b}$$

Since the antennas are located very close to one another, the common signal component is essentially the same for both signals, but the receiver noise is different.

The received signal correlation is

$$\rho_{\text{thermal}} = \frac{\langle \mathbf{V}_1 \mathbf{V}_2^* \rangle}{\sqrt{\langle \mathbf{V}_1 \mathbf{V}_1^* \rangle \langle \mathbf{V}_2 \mathbf{V}_2^* \rangle}}, \tag{15.30}$$

where $\langle\ \rangle$ denotes ensemble averaging and the subscript "thermal" indicates that the noise is due to thermal properties.

Because $\mathbf{N}$ and $\mathbf{N}_2$ are uncorrelated random variables with zero means, it can be shown that

$$\rho_{\text{thermal}} = \frac{|\mathbf{C}|^2}{|\mathbf{C}|^2 + |\mathbf{N}|^2}. \tag{15.31}$$

Using $\text{SNR} = |\mathbf{C}|^2/|\mathbf{N}|^2$, the correlation is related to the measurement SNR by

$$\rho_{\text{thermal}} = \frac{1}{1 + \frac{1}{\text{SNR}}} = \frac{\text{SNR}}{1 + \text{SNR}}. \tag{15.32}$$

If $\text{SNR} \gg 1$, $\rho_{\text{thermal}} \approx 1$.

In InSAR, some geometric decorrelation occurs due to the difference in viewing angles of the two antennas. We assume for now that the viewing angle changes only slightly between the two images so that we can model the two signals as

$$\begin{aligned} \mathbf{V}_1 &= \mathbf{C} + \mathbf{D}_1 + \mathbf{N}_1, \\ \mathbf{V}_2 &= \mathbf{C} + \mathbf{D}_2 + \mathbf{N}_2, \end{aligned} \tag{15.33}$$

where $\mathbf{C}$ again is the correlated part of the signal, $\mathbf{N}_1$ and $\mathbf{N}_2$ are the thermal noise components, and $\mathbf{D}_1$ and $\mathbf{D}_2$ represent additive "noise" (which really is signal variability) due to the change in viewing direction. For reasons to be made apparent later, we call $\mathbf{D}_1$ and $\mathbf{D}_2$ the ***spatial decorrelation*** terms, since they involve spatial movement of the antennas.

Using the same arguments as previously, while ignoring thermal noise, leads to

$$\rho_{\text{spatial}} = \frac{|\mathbf{C}|^2}{|\mathbf{C}|^2 + |\mathbf{D}|^2}, \tag{15.34}$$

where the subscript "spatial" is used instead of "thermal."

If we include both the thermal and noise effects and treat them as uncorrelated random variables, the correlation is

$$\rho_{\text{spatial+thermal}} = \frac{|\mathbf{C}|^2}{|\mathbf{C}|^2 + |\mathbf{D}|^2 + |\mathbf{N}|^2}. \tag{15.35}$$

This result can be rewritten as

$$\rho_{\text{spatial+thermal}} = \frac{|\mathbf{C}|^2}{|\mathbf{C}|^2 + |\mathbf{D}|^2} \cdot \frac{|\mathbf{C}|^2 + |\mathbf{D}|^2}{|\mathbf{C}|^2 + |\mathbf{D}|^2 + |\mathbf{N}|^2}, \tag{15.36}$$

and, upon re-defining SNR as the ratio of all nonthermal to thermal powers, i.e.,

$$\text{SNR} = \frac{|\mathbf{C}|^2 + |\mathbf{D}|^2}{|\mathbf{N}|^2}, \tag{15.37}$$

we obtain the result

$$\begin{aligned} \rho_{\text{spatial+thermal}} &= \rho_{\text{spatial}} \cdot \rho_{\text{thermal}} \\ &= \frac{|\mathbf{C}|^2}{|\mathbf{C}|^2 + |\mathbf{D}|^2} \cdot \frac{1}{1 + \frac{1}{\text{SNR}}} \\ &= \frac{|\mathbf{C}|^2}{|\mathbf{C}|^2 + |\mathbf{D}|^2} \cdot \frac{\text{SNR}}{1 + \text{SNR}}. \end{aligned} \tag{15.38}$$

▶ Thus, the total correlation is the product of the two individual correlations, one due to signal correlation and the other due to thermal noise. ◀

Finally, we consider correlation when the two surfaces are imaged at different times. Even though the backscattered phase is a random quantity because it is the sum of echoes from n scatterers at distances R_1 to R_n, if the scene area is unchanged between observation

times and the radar is at the same location, then R_1 to R_n are unchanged and the total echoes are completely correlated. However, if some of the scatterers move in location, even slightly, the echoes become slightly different. This is known as ***temporal decorrelation***.

There are a number of ways in which a surface might change between observations. The surface can undergo erosion, or if it is vegetated, the plants may grow, or some natural hazard such as an earthquake can disturb the scatterers. On short time scales, wind can alter the positions of leaves and stems. While the causes for temporal decorrelation are many, any disturbance of the locations of the scattering centers within a resolution element can produce temporal decorrelation in the radar signal.

The temporal decorrelation from slightly changed surfaces can be modeled by introducing a term in the backscatter equation corresponding to surface change. A little thought and repetition of the previous arguments leads to an expression of the form

$$\rho_{\text{total}} = \rho_{\text{thermal}} \cdot \rho_{\text{spatial}} \cdot \rho_{\text{temporal}}, \tag{15.39}$$

where ρ_{temporal} represents decorrelation due to a change in the surface condition between observations, ρ_{spatial} refers to a change in sensor location, and ρ_{thermal} is due to thermal noise inside the radar system.

Equation (15.39) is an approximation that only holds exactly for fairly high correlations.

15-5.3 Calculation of Spatial Baseline Decorrelation

This section quantifies ρ_{spatial} and determines just how close the two interferometer antennas need to be located to achieve accurate phase estimation. We begin by deriving a Fourier transform relation between the radar impulse response and the correlation coefficient as a function of the difference in viewing angles using the imaging geometry shown in Fig. 15-17.

Consider the phase of the reflection from the scatterer at P compared with the phase of the reflection from the center of the resolution cell. Using the parallel-ray approximation, the phase from P to A_1 is advanced

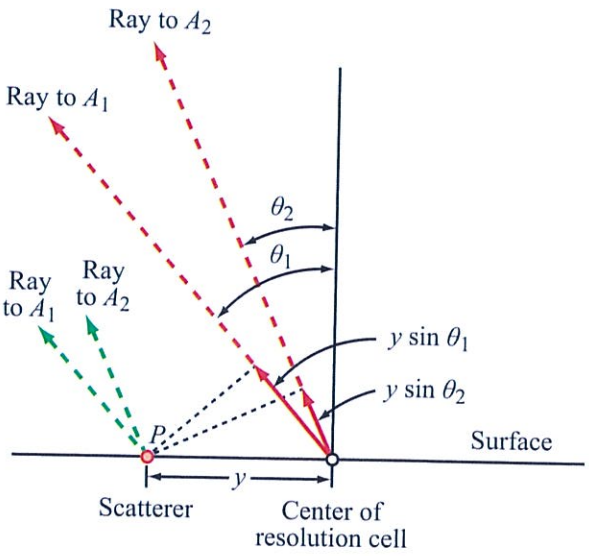


Figure 15-17: Imaging geometry for the decorrelation model. Scatterer P, at a distance y from the pixel center, is observed from two different incidence angles, so the phase of the wave reflected from P differs in each observation. The observed phase relative to the phase of the center of the pixel is that accumulated by a wave propagating a distance $y \sin \theta$ for each ray.

by $y \sin \theta_1$ relative to the ray from O to A_1, and a similar phase advance of $y \sin \theta_2$ applies to antenna A_2. Assuming there are many scattering elements in the resolution cell, the total signal $\mathbf{V}_1$ at antenna 1 can be modeled as the integral of the signals reflected from all of the scatterers in the resolution element. That is,

$$\mathbf{V}_1 = \iint \mathbf{S}(x - x_0,\, y - y_0) \exp\left[-j\frac{4\pi}{\lambda}(R + y \sin\theta_1)\right] \cdot \mathbf{p}(x,y)\, dx\, dy, \tag{15.40}$$

where $\mathbf{S}(x,y)$ represents the complex scattering amplitude from (x,y) on the surface and $\mathbf{p}(x,y)$ is the radar system impulse response function expressed in spatial coordinates. This is the sum of all of the contributions from the points in the resolution cell weighted by the impulse response.

Similarly, for antenna 2,

$$\mathbf{V}_2 = \iint \mathbf{S}(x-x_0,\ y-y_0) \exp\left[-j\frac{4\pi}{\lambda}\,(R+y\sin\theta_2)\right] \cdot \mathbf{p}(x,y)\,dx\,dy. \tag{15.41}$$

In both cases R is the range to the center of the resolution cell. In this analysis we assume R to be the same for both antennas. If this not the actual configuration, the mean phase of the return changes, but the conclusions derived below remain unchanged.

The interferogram is formed by computing $\mathbf{V}_1\mathbf{V}_2^*$:

$$\mathbf{V}_1\mathbf{V}_2^* = \iiiint \mathbf{S}(x-x_0,\ y-y_0)\,\mathbf{S}^*(x'-x_0,\ y'-y_0) \cdot \exp\left[-j\frac{4\pi}{\lambda}\,y(\sin\theta_1-\sin\theta_2)\right] \cdot \mathbf{p}(x,y)\,\mathbf{p}^*(x',y')\,dx\,dy\,dx'\,dy'. \tag{15.42}$$

Making the assumptions that the surface scatterers are arranged randomly, distributed uniformly, and uncorrelated, it follows that

$$\langle \mathbf{S}(x,y)\,\mathbf{S}^*(x',y')\rangle = \sigma^0\,\delta(x-x',\ y-y'), \tag{15.43}$$

which enables reduction of the four-fold integral in Eq. (15.42) to

$$\langle \mathbf{V}_1\mathbf{V}_2^*\rangle = \sigma^0 \iint e^{-j(4\pi/\lambda)y\cos\theta\ d\theta}|\mathbf{p}(x,y)|^2\,dx\,dy\,, \tag{15.44}$$

where θ is the average of θ_1 and θ_2 and $d\theta$ is their difference. Since the exponential kernel is linear in y, it can be viewed as a scaled Fourier transform relating the correlation function $\langle \mathbf{V}_1\mathbf{V}_2^*\rangle$ to the power impulse response $|\mathbf{p}(x,y)|^2$.

Our objective is to compute the spatial correlation

$$\rho_{\text{spatial}} = \frac{\langle \mathbf{V}_1\mathbf{V}_2^*\rangle}{\sqrt{\langle \mathbf{V}_1\mathbf{V}_1^*\rangle\langle \mathbf{V}_2\mathbf{V}_2^*\rangle}}\,. \tag{15.45}$$

To that end, we approximate the impulse response as a sinc function in both dimensions, which is a good representation for high time-bandwidth chirped systems. That is,

$$\mathbf{p}(x,y) = \frac{1}{\sqrt{r_x r_y}}\left[\frac{\sin(\pi x/r_x)}{(\pi x/r_x)}\cdot\frac{\sin(\pi y/r_y)}{(\pi y/r_y)}\right], \tag{15.46}$$

where r_x and r_y are the ground resolutions of the radar system along the azimuth and range directions, respectively. Use of Eq. (15.46) leads to

$$\langle \mathbf{V}_1\mathbf{V}_1^*\rangle = \langle \mathbf{V}_2\mathbf{V}_2^*\rangle = \sigma^0\,\delta(x-x',\ y-y'). \tag{15.47}$$

After evaluating the Fourier transform of Eq. (15.44) and using the results in Eq. (15.45), we obtain the expression

$$\rho_{\text{spatial}} = 1 - \frac{2r_y\cos\theta}{\lambda}\,d\theta. \tag{15.48}$$

The differential angle $d\theta = \theta_1 - \theta_2$ in Fig. 15-17 can be related to $B_\perp$ and R as

$$d\theta = \frac{B_\perp}{R}\,, \tag{15.49}$$

in which case Eq. (15.48) can be cast in the form

$$\rho_{\text{spatial}} = 1 - \frac{2B_\perp r_y\cos\theta}{\lambda R}\,. \tag{15.50}$$

Note that this is a linearly decreasing function of $B_\perp$; ρ_{spatial} extends from 1 at $B_\perp = 0$ (perfect correlation) to zero at a ***critical value of*** $B_\perp$ denoted B_c, with

$$B_c = \frac{\lambda R}{2r_y\cos\theta}\,. \tag{15.51}$$

If the perpendicular baseline $B_\perp$ approaches or exceeds B_c, no correlation is observed. This quantifies our earlier observation that the interferometer antennas must be located near each other in order for the two signals measured by them to be highly correlated.

Figure 15-18 shows the theoretical linear falloff to zero at the critical baseline. A sinc function spatial-response function is assumed. Measured data are often seen to fall off more quickly than predicted from the theoretical model. While the impulse response of the SEASAT radar used to collect the data of Fig. 15-18 was indeed a sinc function, its resolution may have been overestimated by the system model.

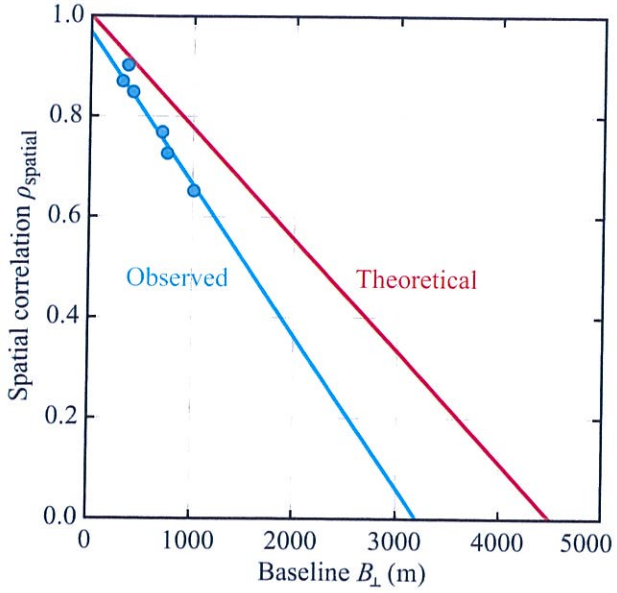


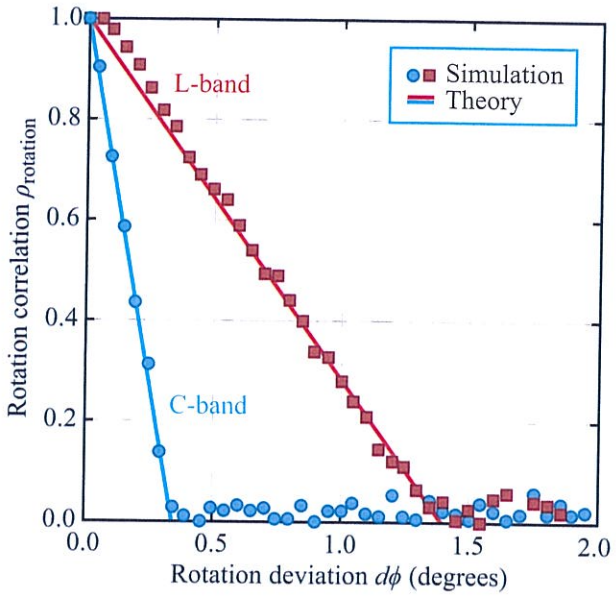
Figure 15-18: Baseline decorrelation for the sinc impulse function of the SEASAT SAR compared with actual data collected over Death Valley in 1978. The discrepancy in the rates of falloff between the curves is attributed to imprecise system models of the resolution. [Zebker and Villasenor, 1992.]

15-5.4 Rotational Decorrelation

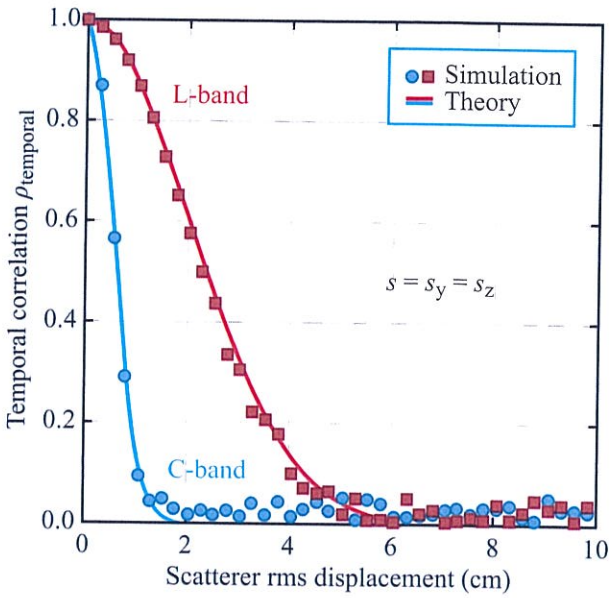
In multipass interferometry, if the radar flight track of one pass is rotated relative to the track of the second one (i.e., they are not parallel), a second source of decorrelation due to the imaging geometry is introduced. The derivation of this correlation term is not provided here, but can be found in Zebker and Villasenor (1992). The rotational correlation can be expressed as

$$\rho_{\text{rotation}} = 1 - \frac{2r_x\, d\phi \sin\theta}{\lambda}, \qquad (15.52)$$

where r_x is the azimuth resolution and $d\phi$ is the rotation angular deviation between the two tracks. When applicable, this term should be included as a multiplicative factor in the correlation product expression in Eq. (15.39). Figure 15-19(a) compares Monte Carlo simulations and the predicted rotational correlation versus the path rotation angle for radars



(a) Rotational decorrelation



(b) Motion decorrelation

Figure 15-19: Simulated and theoretical dependence of InSAR correlation on (a) track rotation as a function of rotation angle, and on (b) scatterer motion as a function of rms displacement of the scatterers. These curves are for a radar with 5 m resolution in both slant range and azimuth. [Zebker and Villasenor, 1992.]

operating at L- and C-band. Note that the correlation is high when the tracks are aligned (rotation angle zero) but falls off as the rotational angle grows. The fall-off with angle increases as the wavelength decreases.

Deviations from straight-line paths can also lead to decorrelation, but suitable models for evaluating the decorrelation are very specific to the details of the flight path difference, and must be computed in each instance following the methods described above.

15-5.5 Temporal Decorrelation

As previously noted, temporal decorrelation occurs if the scatterers move with respect to each other between observation times. The temporal decorrelation term can be modeled using 3-D motions of each of the individual scatterers, leading to the correlation product given by Zebker and Villasenor (1992) as:

$$\begin{aligned}\mathbf{V}_1\mathbf{V}_2^* = \iiiint \mathbf{S}(x-x_0,\, y-y_0,\, z-z_0) \\ \cdot \mathbf{S}^*(x'-x_0,\, y'-y_0,\, z'-z_0) \\ \cdot \exp\left[-j\frac{4\pi}{\lambda}\,(\eta_y \sin\theta + \eta_z \cos\theta)\right] \\ \cdot \mathbf{p}(x,y)\,\mathbf{p}^*(x',y')\,dx\,dy\,dz\,dx'\,dy'\,dz',\end{aligned} \tag{15.53}$$

where η_y and η_z are incremental changes in the y- and z-locations of a scatterer. The computed average of this integral can be expressed as

$$\begin{aligned}\langle \mathbf{V}_1\mathbf{V}_1^* \rangle = \sigma^0 \iint \exp\left[-j\frac{4\pi}{\lambda}\,(\eta_y \sin\theta + \eta_z \cos\theta)\right] \\ \cdot f_y(\eta_y)\, f_z(\eta_z)\, d\eta_y\, d\eta_z,\end{aligned} \tag{15.54}$$

where $f_y(\eta_y)$ and $f_z(\eta_z)$ are the probability distributions of motions in the y and z directions, respectively. By treating the pdfs as independent Gaussian processes, Zebker and Villasenor (1992) show that

$$\rho_{\text{temporal}} = \exp\left[-\frac{1}{2}\left(\frac{4\pi}{\lambda}\right)^2 (s_y^2 \sin^2\theta + s_z^2 \cos^2\theta)\right], \tag{15.55}$$

where s_y^2 and s_z^2 are variances of motions in y and z.

Figure 15-19(b) compares Monte-Carlo simulations with the predicted temporal correlation based on Eq. (15.55), with both plotted as a function of the rms displacement $s = s_y = s_y$ of the scatterers. The correlation function drops off at a much faster rate at the shorter wavelength, as expected.

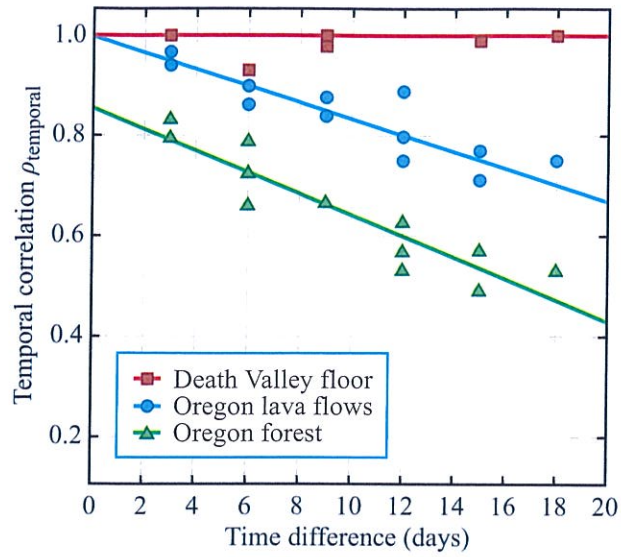
The observed change in correlation as a function of the time difference between repeat passes of the SEASAT SAR is displayed in Fig. 15-20(a) for three geographic locations. Because the vegetation-free floor of Death Valley changes very little over the 18-day observation period, the temporal correlation remained at approximately 1 throughout. For the lightly vegetated lava flows in Eastern Oregon, the correlation decreased approximately linearly with time. Finally, for the heavily forested area, the rate of decrease is only slightly faster, but significant decorrelation is observed after only 2 days.

In Fig. 15-20(b), reported by Rosen et al. (2000), we note that the correlation is the lowest for the water surfaces and highest for urban areas.

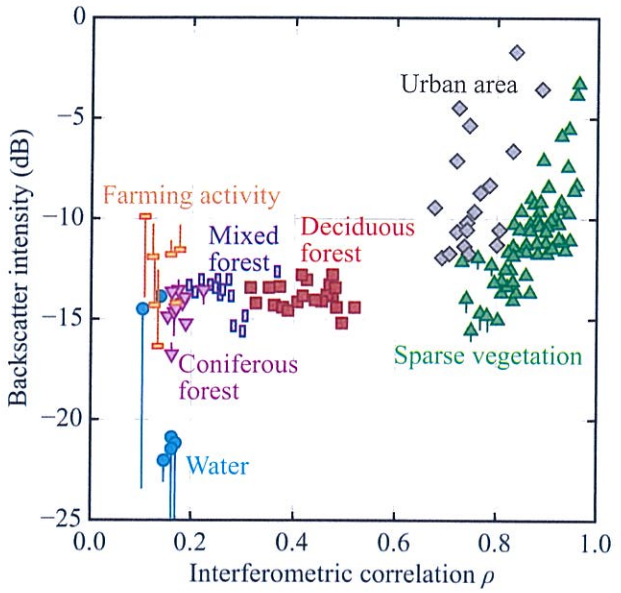
15-6 Measurement of Topography

Before we examine the InSAR topographic measurement process in detail, let us highlight the primary reasons for measuring topography:

(1) Most of the world's surface is not well mapped, especially in digital form and at high resolution. InSAR is particularly suited for addressing this deficiency. In fact, the ***Shuttle Radar Topography Mission*** (SRTM), which flew for only 11 days in 2000, managed to map the topography of the entire segment of the globe between 56° S and 60° N at a spatial resolution of 90 m and a vertical precision of 16 m. SRTM used a baseline of 60 m between its two antennas. In many areas around the world, the SRTM-generated topography data is the best available today. However, there is need for improved mapping capability, and for mapping areas that were not mapped by SRTM, because they have very high relief or very low backscattering coefficients (such as some desert areas).



(a)



(b)

Figure 15-20: (a) Observed decorrelation with time for three different geographic regions, as measured by the SEASAT SAR in 1978; (b) scatter plot of relative backscatter intensity versus interferometric correlation for a water surface and various types of terrain [Rosen et al., 2000].

(2) Historically, different parts of the world were mapped on different ***datums***, or coordinate systems, making it difficult to combine them for use in global studies such as global change. InSAR processing is location dependent and can provide ***digital elevation maps*** (DEMs) on a single grid. SRTM demonstrated this by generating global DEMs with a consistent projection.

(3) Earthquakes, landslides, floods, and other events modify the local topography. InSAR-derived DEMs can detect topographic changes and help us understand the effects of such events and how they modify the terrain.

15-6.1 Inferring Topography from Interferometric Phase

In Section 15-2, we used the parallel-ray approximation to demonstrate the relationship between the measured phase difference ϕ_{int} and the height z of an image pixel. Precise implementation of the topographic mapping process, however, calls for the development of exact solutions. To that end, we use the observation geometry shown in Fig. 15-21 and reintroduce Eq. (15.4):

$$\begin{aligned}(R-\delta)^2 &= R^2+B^2-2RB\cos\left(\frac{\pi}{2}-\theta+\alpha\right)\\ &= R^2+B^2-2RB\sin(\theta-\alpha),\end{aligned} \tag{15.56}$$

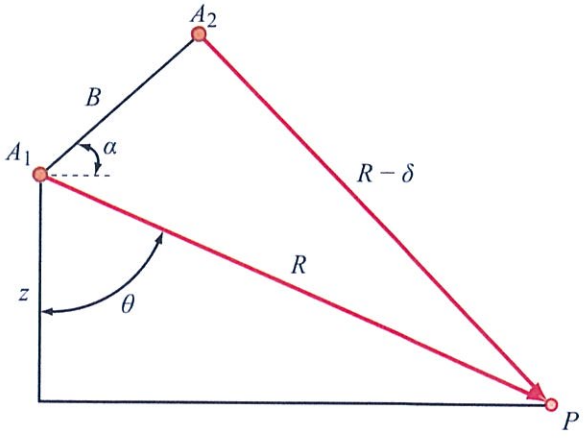


Figure 15-21: InSAR observation geometry.

from which it follows that

$$\sin(\theta-\alpha)=\frac{R^2+B^2-(R-\delta)^2}{2RB}\,. \tag{15.57}$$

Phase error

From Eqs. (15.2a) and (15.6),

$$\delta=-\frac{\lambda}{4\pi}\,\phi_{\text{int}} \tag{15.58}$$

and

$$z=R\cos\theta, \tag{15.59}$$

where ϕ_{int} pertains to the two-independent-radars configuration, with the understanding that if the interferometer uses a common transmitter instead, the factor of 4 should be replaced with a factor of 2.

Hence, given B, α, and R we can solve for θ and use it to estimate z. To estimate the associated error, we use derivatives. In this case, a slightly different chain rule is used:

$$\frac{\partial z}{\partial\phi_{\text{int}}}=\frac{\partial z}{\partial\theta}\frac{\partial\theta}{\partial\delta}\frac{\partial\delta}{\partial\phi_{\text{int}}}\,, \tag{15.60}$$

and we note that

$$\frac{\partial z}{\partial\theta}=\frac{\partial}{\partial\theta}\,(R\cos\theta)=-R\sin\theta\,. \tag{15.61}$$

To obtain an expression for $\partial\theta/\partial\delta$, we differentiate Eq. (15.57) while treating only θ and δ as variables:

$$\frac{\partial}{\partial\delta}[\sin(\theta-\alpha)]=\frac{\partial}{\partial\delta}\left[\frac{R^2+B^2-(R-\delta)^2}{2RB}\right],$$

which yields

$$\cos(\theta-\alpha)\,\frac{\partial\theta}{\partial\delta}=\frac{2(R-\delta)}{2RB}\approx\frac{1}{B}\,. \tag{15.62}$$

In the final step, we approximated $(R-\delta)/R$ as 1. Hence,

$$\frac{\partial\theta}{\partial\delta}=\frac{1}{B\cos(\theta-\alpha)}\,. \tag{15.63}$$

Finally, from Eq. (15.58),

$$\frac{\partial\delta}{\partial\phi_{\text{int}}}=\frac{-\lambda}{4\pi}\,,$$

so that we can write

$$\begin{aligned}\frac{\partial z}{\partial\phi_{\text{int}}}&=(-R\sin\theta)\left(\frac{1}{B\cos(\theta-\alpha)}\right)\left(\frac{-\lambda}{4\pi}\right)\\&=\frac{\lambda}{4\pi}\,\frac{R\sin\theta}{B\cos(\theta-\alpha)}\,,\end{aligned} \tag{15.64}$$

which is consistent with the result given by Eq. (15.14a) (but without the SNR part) on the basis of the parallel-ray approximation.

Baseline-length error

If the two receive antennas of the SAR interferometer are physically connected to one another, such as if they are mounted on the two wings of an aircraft or at the two ends of a long mast—as was the case for the SRTM instrument, then the baseline distance B is a constant quantity with no significant variability. For repeat-pass systems, on the other hand, the knowledge of the value of B may have an element of uncertainty, either in terms of its length or, more likely, in terms of the angle α. Errors in the values of the baseline length or orientation contribute to height errors.

The contribution of the baseline-length error to the height error can be obtained by computing the appropriate partial derivatives:

$$\frac{\partial z}{\partial B}=\frac{\partial z}{\partial\theta}\cdot\frac{\partial\theta}{\partial B}\,. \tag{15.65}$$

To derive an expression for $\partial\theta/\partial B$, we start with Eq. (15.57):

$$\frac{\partial}{\partial B}[\sin(\theta-\alpha)]=\frac{\partial}{\partial B}\left[\frac{R^2+B^2-(R-\delta)^2}{2RB}\right],$$

which leads to

$$\frac{\partial\theta}{\partial B}=\frac{B^2+\delta^2-2R\delta}{2RB^2\cos(\theta-\alpha)}\,. \tag{15.66}$$

Since the range $R\gg\delta$ and also $R\gg B$, Eq. (15.66) can be simplified to

$$\frac{\partial\theta}{\partial B}=\frac{-\delta}{B^2\cos(\theta-\alpha)}\,. \tag{15.67}$$

Combining this result with Eq. (15.61) leads to

$$\frac{\partial z}{\partial B} = \frac{\partial z}{\partial \theta} \cdot \frac{\partial \theta}{\partial B} = (-R\sin\theta)\left[\frac{-\delta}{B^2\cos(\theta-\alpha)}\right] = \frac{R\delta\sin\theta}{B^2\cos(\theta-\alpha)}. \quad (15.68)$$

Using the approximate relation $\delta = B\sin(\theta-\alpha)$, as given by Eq. (15.5), leads to

$$\frac{\partial z}{\partial B} \approx \frac{R}{B}\sin\theta\tan(\theta-\alpha). \quad (15.69)$$

Fortunately, the uncertainty in the length of B can be controlled to within a small value, so this error source usually is small for most practical systems.

Baseline-orientation error

We now consider the error due to the uncertainty associated with the baseline-orientation angle α, which depends on the knowledge uncertainty of the platform track or orbit and the platform altitude, plus any flexing of the platform itself. To compute $\partial z/\partial\alpha$, we use the chain rule

$$\frac{\delta z}{\partial \alpha} = \frac{\partial z}{\partial \theta} \cdot \frac{\partial \theta}{\partial \delta} \cdot \frac{\partial \delta}{\partial \alpha}. \quad (15.70)$$

The derivatives $\partial z/\partial\theta$ and $\partial\theta/\partial\delta$ are given by Eqs. (15.61) and (15.63), respectively. To compute $\partial\delta/\partial\alpha$, we differentiate Eq. (15.57) while treating only α and δ as variables:

$$\frac{\partial}{\partial\alpha}[\sin(\theta-\alpha)] = \frac{\partial}{\partial\alpha}\left[\frac{R^2+B^2-(R-\delta)^2}{2RB}\right],$$

which leads to

$$-\cos(\theta-\alpha) = \frac{2(R+\delta)}{2RB}\frac{\partial\delta}{\partial\alpha},$$

or

$$\frac{\partial\delta}{\partial\alpha} = \frac{-RB}{R+\delta}\cos(\theta-\alpha) \approx -B\cos(\theta-\alpha). \quad (15.71)$$

Using Eqs. (15.61), (15.63), and (15.71) in Eq. (15.70) gives

$$\frac{\partial z}{\partial\alpha} = (-R\sin\theta)\left[\frac{1}{B\cos(\theta-\alpha)}\right][-B\cos(\theta-\alpha)], \quad (15.72)$$

which simplifies to

$$\frac{\partial z}{\partial\alpha} = R\sin\theta. \quad (15.73)$$

▶ The fact that $|\partial z/\partial\alpha|$ and $|\partial z/\partial\theta|$ [as in Eq. (15.61)] are given by the same expression indicates that a tilt $d\alpha$ in the interferometer baseline is indistinguishable from a change $d\theta$ in the local slope of the surface. ◀

If we denote s_z, $s_{\phi_{\text{int}}}$, s_B, and s_α as the standard deviations associated with their respective parameters, we can summarize the phase, baseline-length, and baseline-angle errors as follows:

Phase error:

$$s_z = \frac{\lambda}{4\pi}\frac{R\sin\theta}{B\cos(\theta-\alpha)}s_{\phi_{\text{int}}}. \quad (15.74a)$$

Baseline-length error:

$$s_z = \frac{R\sin\theta\tan(\theta-\alpha)}{B}s_B. \quad (15.74b)$$

Baseline-angle error:

$$s_z = R\sin\theta\, s_\alpha. \quad (15.74c)$$

To gain insight into the magnitude of the error contributions, we consider the design parameters of the airborne NASA/JPL TOPSAR system:

$$R \approx 10\text{ km}, \quad \theta \approx 30°,$$
$$B = 1.5\text{ m}, \quad \alpha = 63°,$$
$$\lambda = 6\text{ cm}, \quad \text{SNR} \approx 100\ (20\text{ dB}).$$

From Eq. (15.13), $s_{\phi_{\text{int}}}$ for a 1-look image is

$$s_{\phi_1} = \frac{1}{\sqrt{\text{SNR}}} = 0.1\text{ rad}.$$

This particular radar calculates about 10 looks in the interferogram before estimating the height. For N looks (Rodriguez and Martin, 1992),

$$s_{\phi_N} \approx \frac{s_{\phi_1}}{\sqrt{2N}}. \tag{15.75}$$

Hence, using $N = 10$ for the TOPSAR,

$$s_{\phi_{10}} \approx 0.022 \text{ rad} \approx 1.25°.$$

Using this result in Eq. (15.74a), along with the system parameters, leads to an rms height error (due to phase error) of

$$s_z = \frac{\lambda}{4\pi} \frac{R \sin\theta}{B\cos(\theta - \alpha)} s_{\phi_{10}} = 0.42 \text{ m} \qquad \textbf{(phase error)}.$$

For the baseline-length error, which for this system is primarily due to ***aircraft deformation***, s_B is estimated to be about 10^{-4} m. Using this value in Eq. (15.74b) yields

$$s_z = \frac{R \sin\theta \tan(\theta - \alpha)}{B} s_B = 0.216 \text{ m},$$

(baseline-length error)

which is about one-half of the phase-associated error.

Finally, for this system the baseline-orientation error is due to uncertainty in the ***aircraft attitude***, for which $s_\alpha \approx 0.01°$. Using this value in Eq. (15.74c) gives

$$s_z = R \sin\theta \, s_\alpha = 0.88 \text{ m} \qquad \textbf{(baseline-angle error)},$$

which is about twice the phase-associated error. Thus we see that for this system the uncertainty in baseline orientation is the major source of error. This is true for most systems.

15-6.2 Phase Unwrapping

In the preceding sections, we developed a set of equations that relate the phase difference ϕ_{int} (measured by an interferometric SAR) for a particular pixel to the height z of that pixel. Hence, given an imaging geometry and absolute values of ϕ_{int} for each pixel, it should be possible to create a topographic map of the imaged scene. Phase, however, is only measured *modulo* 2π, which means that the phase must first be *unwrapped* before applying it to determine z. The measured phase modulo 2π is known as the ***wrapped phase***. The process of converting the wrapped phase into absolute phase is called ***phase unwrapping***. It entails determining how many multiples of 2π should be added to the wrapped phase to produce a continuous phase surface.

▶ Whereas the wrapped phase is limited to the range $[0, 2\pi]$, the unwrapped phase is not so limited. ◀

The problem of phase unwrapping for InSAR is fairly complicated, so in this chapter we detail only the separation of the total phase into the background, or flat-Earth, component, and the phase component due to topography or crustal deformation. The flat-Earth term is known from the imaging geometry, and its removal from the observed phase leaves only the more slowly varying signal of interest.

The phase unwrapping problem for InSAR was first addressed by Goldstein et al. (1988), who developed a theory for inconsistencies in the phase field, which they called residues. They showed that by properly joining residues with a series of curves that are not crossed during integration of measured pixel-to-pixel phase differences a consistent and physically plausible solution could be achieved. This was a conservative method that produced few incorrect unwrapped pixels, but often left large gaps in the resulting interferogram product. Ghiglia and Romero (1996) used variable-norm criteria to optimize the placement of the uncrossable boundaries and were able to retrieve more spatial coverage. Following this work, Costantini (1998) was able to show a duality between the placement of branch cut lines and the linear-programming methods used to optimize network accessibility problems, which yielded not only still greater coverage, but also incorporated network solver algorithms that permitted solution of complicated phase fields. The current state of the art in phase unwrapping follows Chen and Zebker (2001), who added weighting functions, based on statistics of radar signals from terrain slopes and deformations, to the network programming approach. Its combination of

accurately placed cut lines and robustness of solution leads to significant areal coverage and accuracy.

We start with Eq. (15.57), and then we simplify it as follows:

$$\begin{aligned}\sin(\theta-\alpha) &= \frac{R^2+B^2-(R-\delta)^2}{2RB} \\ &= \frac{R^2+B^2-R^2+2R\delta-\delta^2}{2RB} \\ &= \frac{B^2+2R\delta-\delta^2}{2RB} \\ &\approx \frac{B}{2R}+\frac{\delta}{B} \approx \frac{\delta}{B}, \end{aligned} \tag{15.76}$$

where we used the knowledge that $\delta \ll R$ and $B \ll R$. This is the same expression given by Eq. (15.5) obtained using the parallel-ray approximation.

Next, we expand Eq. (15.76):

$$\begin{aligned}\delta &= B\sin(\theta-\alpha) \\ &= B[\sin\theta\cos\alpha-\cos\theta\sin\alpha]. \end{aligned} \tag{15.77}$$

From Eq. (15.6), $\cos\theta = z/R$. It follows that

$$\sin\theta = \sqrt{1-\frac{z^2}{R^2}}$$

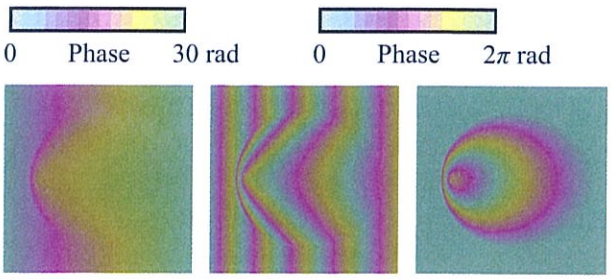
and

$$\delta = B\left[\sqrt{1-\frac{z^2}{R^2}}\cos\alpha-\frac{z}{R}\sin\alpha\right]. \tag{15.78}$$

For an InSAR with two independent radars, the (absolute) phase difference is

$$\begin{aligned}\phi_{\text{int}} &= -\frac{4\pi}{\lambda}\delta \\ &= -\frac{4\pi B}{\lambda R}\left[\sqrt{R^2-z^2}\cos\alpha-z\sin\alpha\right]. \end{aligned} \tag{15.79}$$

However, as we stated earlier, the *measured* phase is modulo 2π, whereas the true phase ϕ_{int} as given by Eq. (15.79) varies continuously with z/R. To illustrate, we simulate in Fig. 15-22(a) a continuous phase pattern that extends over a range of 30 radians. The measured phase difference, modulo 2π, is illustrated in part (b) of



(a) Unwrapped phase (b) Wrapped phase (c) Unwrapped topographic phase

Figure 15-22: Simulation of InSAR phase difference: (a) true phase ϕ, (b) measured phase modulo 2π, and (c) tropospheric component of ϕ after removal of flat-Earth component.

Fig. 15-22. The task of phase unwrapping is to convert the measured phase displayed in Fig. 15-22(b) into the true phase in part (a). An additional step of phase unwrapping involves segregating the two components of ϕ_{int}, namely the component ϕ_{topo} due to topographic (height) variations, from the component ϕ_{fe} due to the ***flat Earth*** reference. The latter is the phase pattern that the InSAR would measure if $z = z_0$ everywhere across the image, where z_0 is a constant. Figure 15-22(c) simulates the unwrapped phase after removal of the flat-Earth component. This is the true phase pattern ϕ_{topo} associated with the height variations that the interferometric process seeks to quantify.

To segregate the two phase components, we define z_0 as the mean height of the surface and dz as the deviation due to the topographic height relative to z_0. After replacing z with z_0+dz in Eq. (15.79), we have

$$\begin{aligned}\phi_{\text{int}} &= -\frac{4\pi B}{\lambda R}\left[\sqrt{R^2-(z_0+dz)^2}\,\cos\alpha-(z_0+dz)\sin\alpha\right] \\ &= -\frac{4\pi B}{\lambda R}\left[\sqrt{R^2-z_0^2-2z_0dz-(dz)^2}\cdot\cos\alpha\right. \\ &\qquad \left.-(z_0+dz)\sin\alpha\right]. \end{aligned} \tag{15.80}$$

Neglecting $(dz)^2$ relative to the other terms, and then factoring out $\sqrt{R^2 - z_0^2}$ leads to

$$\phi_{\text{int}} = -\frac{4\pi B}{\lambda R}\left[\sqrt{R^2 - z_0^2}\cdot\sqrt{1 - \frac{2z_0\,dz}{R_0^2 - z_0^2}}\cos\alpha - z_0\sin\alpha - dz\sin\alpha\right]. \tag{15.81}$$

The second square root in Eq. (15.81) is of the form $\sqrt{1-a}$ with $a \ll 1$. Using the Taylor series expansion $\sqrt{1-a} \approx 1 - a/2$ leads to

$$\phi_{\text{int}} = -\frac{4\pi B}{\lambda R}\left[\sqrt{R^2 - z_0^2}\left(1 - \frac{z_0\,dz}{R_0^2 - z_0^2}\right)\cos\alpha - z_0\sin\alpha - dz\sin\alpha\right]. \tag{15.82}$$

The ***flat-Earth component*** is obtained by setting $dz = 0$:

$$\phi_{\text{fe}} = -\frac{4\pi B}{\lambda R}\left[\sqrt{R^2 - z_0^2}\cos\alpha - z_0\sin\alpha\right], \tag{15.83}$$

and the ***topographic component*** is given by

$$\begin{aligned}\phi_{\text{topo}} &= \phi_{\text{int}} - \phi_{\text{fe}} \\ &= \frac{4\pi B}{\lambda R}\left[\frac{z_0\cos\alpha}{\sqrt{R^2 - z_0^2}} + \sin\alpha\right] dz \\ &= \frac{4\pi B}{\lambda R}\left[\frac{\cos\alpha}{\tan\theta_0} + \sin\alpha\right] dz,\end{aligned} \tag{15.84}$$

where θ_0 is the incidence angle relative to the mean surface, and

$$\cot\theta_0 = \frac{z_0}{\sqrt{R^2 - z_0^2}}.$$

► The linear relationship between ϕ_{topo} and dz indicates that constant-phase contours appear at constant heights on the imaged surface. ◄

The wrapped phase repeats every 2π radians, with each ***fringe*** (a 0 to 2π transition) corresponding to a height difference Δz known as the ***ambiguity height***. We obtain Δz by setting $\phi_{\text{topo}} = 2\pi$ in Eq. (15.84), which leads to

$$\Delta z = \frac{\lambda R}{2B\left(\dfrac{\cos\alpha}{\tan\theta_0} + \sin\alpha\right)}. \tag{15.85}$$

By way of an example, for the TOPSAR with the system parameters listed in the preceding subsection,

$$\Delta z = \frac{6\times 10^{-2}\times 10^4}{2\times 1.5\left(\dfrac{\cos 63^\circ}{\tan 30^\circ} + \sin 63^\circ\right)} = 119.24 \text{ m}.$$

This means that the phase contours repeat every approximately 120 m. We should note that this is based on single values for R and θ_0; in an exact topographic reduction process, the variations of R and θ_0 across the image should be properly incorporated. Examples of digital elevation model (DEM) maps acquired by airborne InSAR systems are shown in Fig. 15-23.

15-6.3 Curved-Earth Phase Pattern

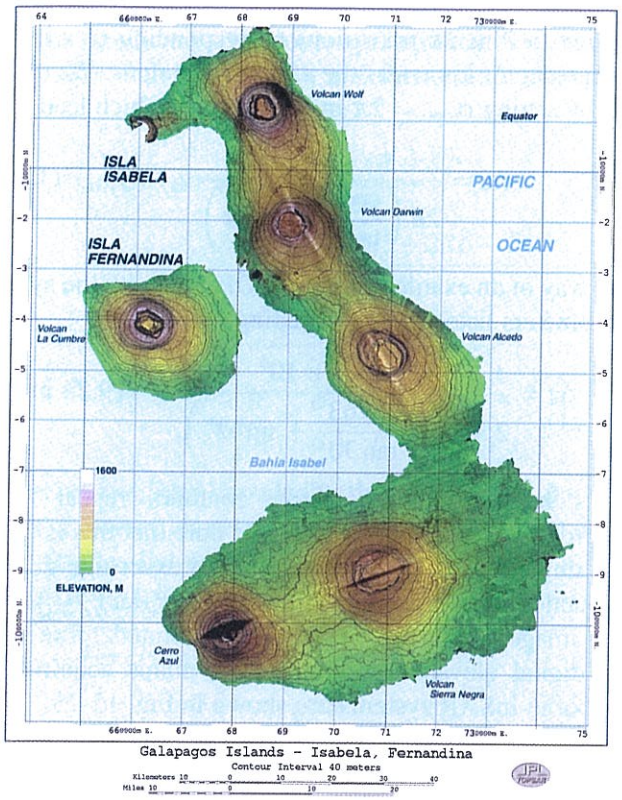
For airborne InSAR observations at moderate incidence angles, the height of the radar above the surface is such that it is reasonable to treat the Earth surface as flat. Such an assumption, however, is not applicable for spaceborne SAR applications, nor for airborne SARs with high incidence angles. To achieve useful height accuracies from the InSAR phase measurements, we should replace the reference flat-Earth model with a curved-Earth model. The relevant geometry is shown in Fig. 15-24. Now, the incidence angle at point P relative to the mean curved-Earth surface is θ', not θ. Angles θ and θ' can be computed by applying the law of cosines, which yields

$$\theta = \cos^{-1}\left[\frac{(h+R_0)^2 + R^2 - R_0^2}{2R(h+R_0)}\right] \tag{15.86a}$$

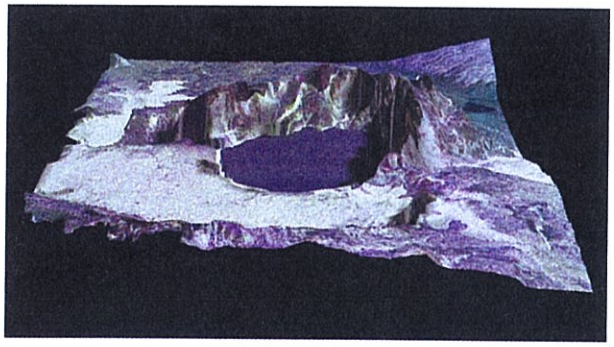
and

$$\theta' = \pi - \cos^{-1}\left[\frac{R^2 + R_0^2 - (h+R_0)^2}{2RR_0}\right], \tag{15.86b}$$

where R_0 is the Earth radius.



(a) Galápagos Islands: Fernandina and Isabela



(b) Askja, Iceland

Figure 15-23: (a) DEM of the Galápagos Islands Fernandina and Isabela, acquired by the NASA/JPL TOPSAR C-band interferometer. The swath width of the instrument is only 20 km, so many parallel flight lines were collected and combined. (b) DEM of Askja, Northern volcanic zone, Iceland, derived from the C-band EMISAR topographic mapping system of Denmark. The color variation in the image is derived from L-band EMISAR polarimetry [Christensen et al., 1998].

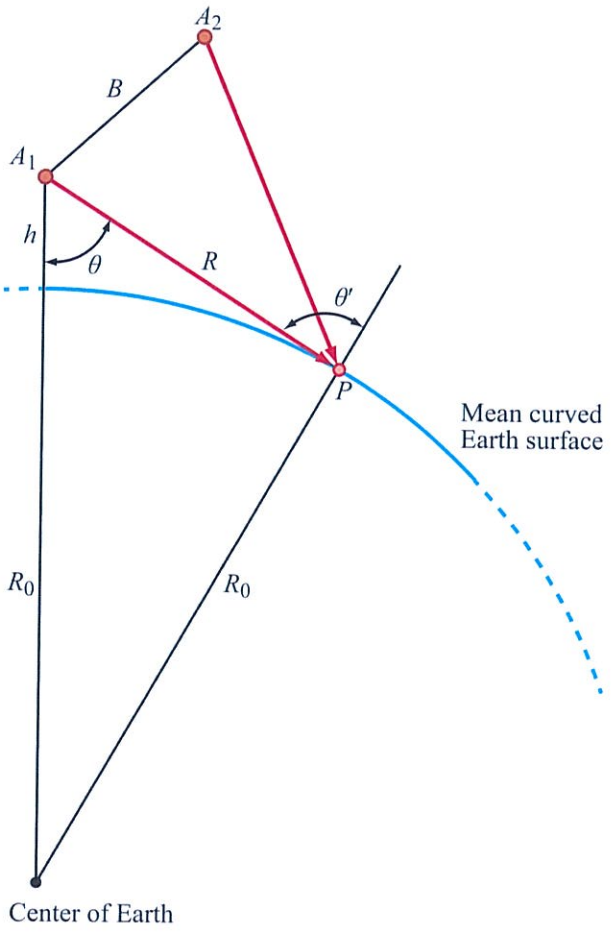


Figure 15-24: Geometry for computing the curved-Earth phase component.

15-7 Mapping Earth's Topography: The SRTM Mission

In February 2000, NASA, in partnership with the National Geospatial-Intelligence Agency and the German and Italian Space Agencies, launched and operated the Shuttle Radar Topography Mission (SRTM), producing the most complete, highest-resolution digital elevation model of the Earth to date. It used dual radar antennas attached to the space shuttle to acquire interferometric radar data, which were then processed to digital

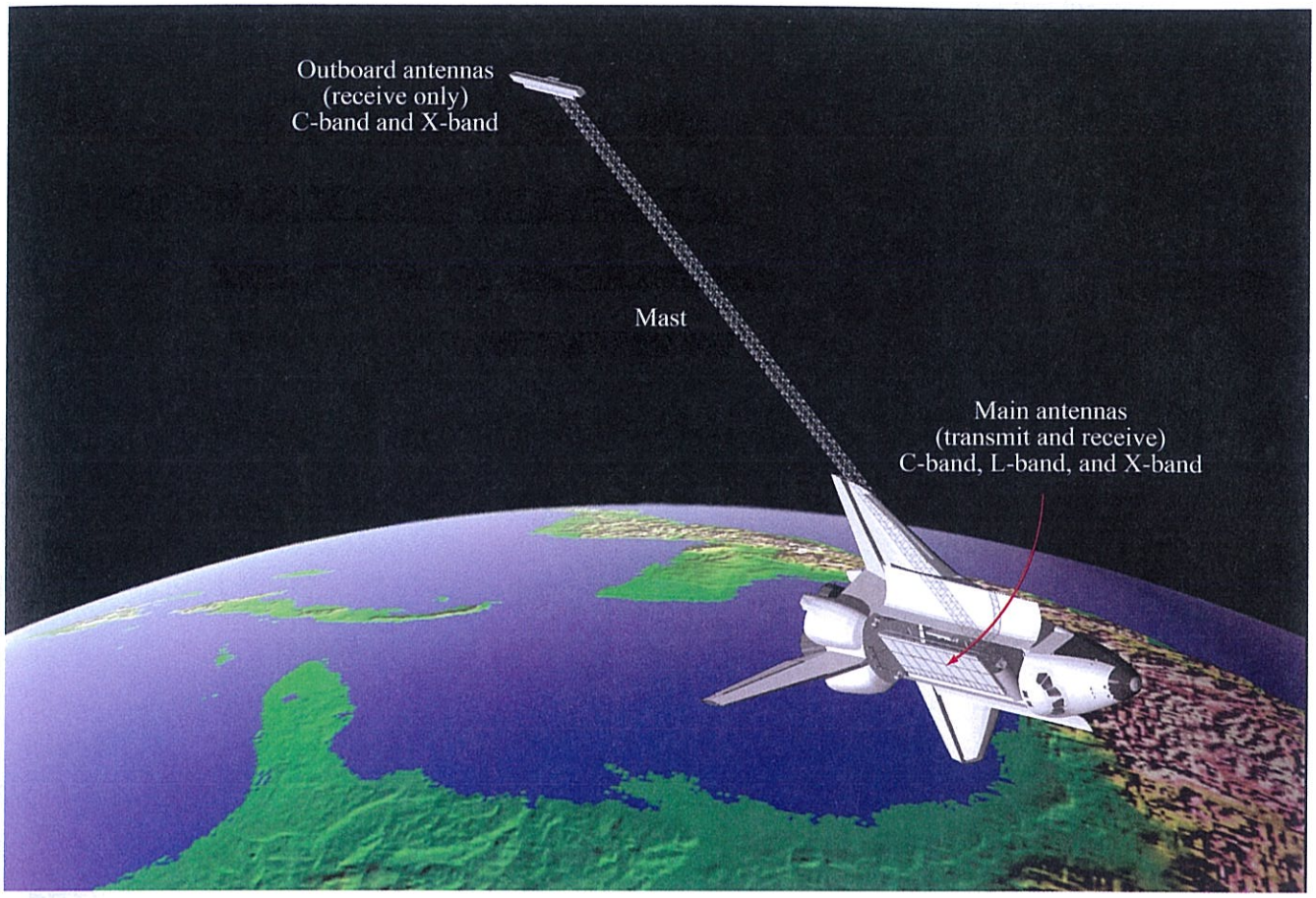


Figure 15-25: Cartoon illustrating the configuration of the SRTM instrument. Radar signals are transmitted from the main shuttle bay, and received simultaneously at that location and at an outboard antenna, forming an across-track baseline. [Courtesy NASA.] The mast is 60 m long.

topographic models at 1 arcsec resolution. SRTM was built on the hardware used by the previous NASA SIR-C mission, which included L-, C-, and X-band SAR systems. An illustration of the configuration of the instrument aboard the shuttle is shown in Fig. 15-25.

The mission objective was to acquire a digital elevation model of all land surfaces between about 60° north latitude and 56° south latitude, constituting about 80% of Earth's land surface. The data acquired by the system continue to be one of the most complete set of digital representations of the Earth available to date, and they are used by many different user communities.

The data-reduction process used a version of the steps outlined in the preceding sections. Because the system had a long boom holding the second set of antennas, and also because the shuttle itself exhibits significant attitude variations in flight, the main contributions to errors were from those components. Much of the processing was performed to address and compensate for attitude errors (Farr et al., 2007). Additional flight hardware systems were added to the shuttle avionics in order to better measure details of the craft's orbit and attitude.

The topographic products were derived from the C-band radar data, and were processed at JPL over a

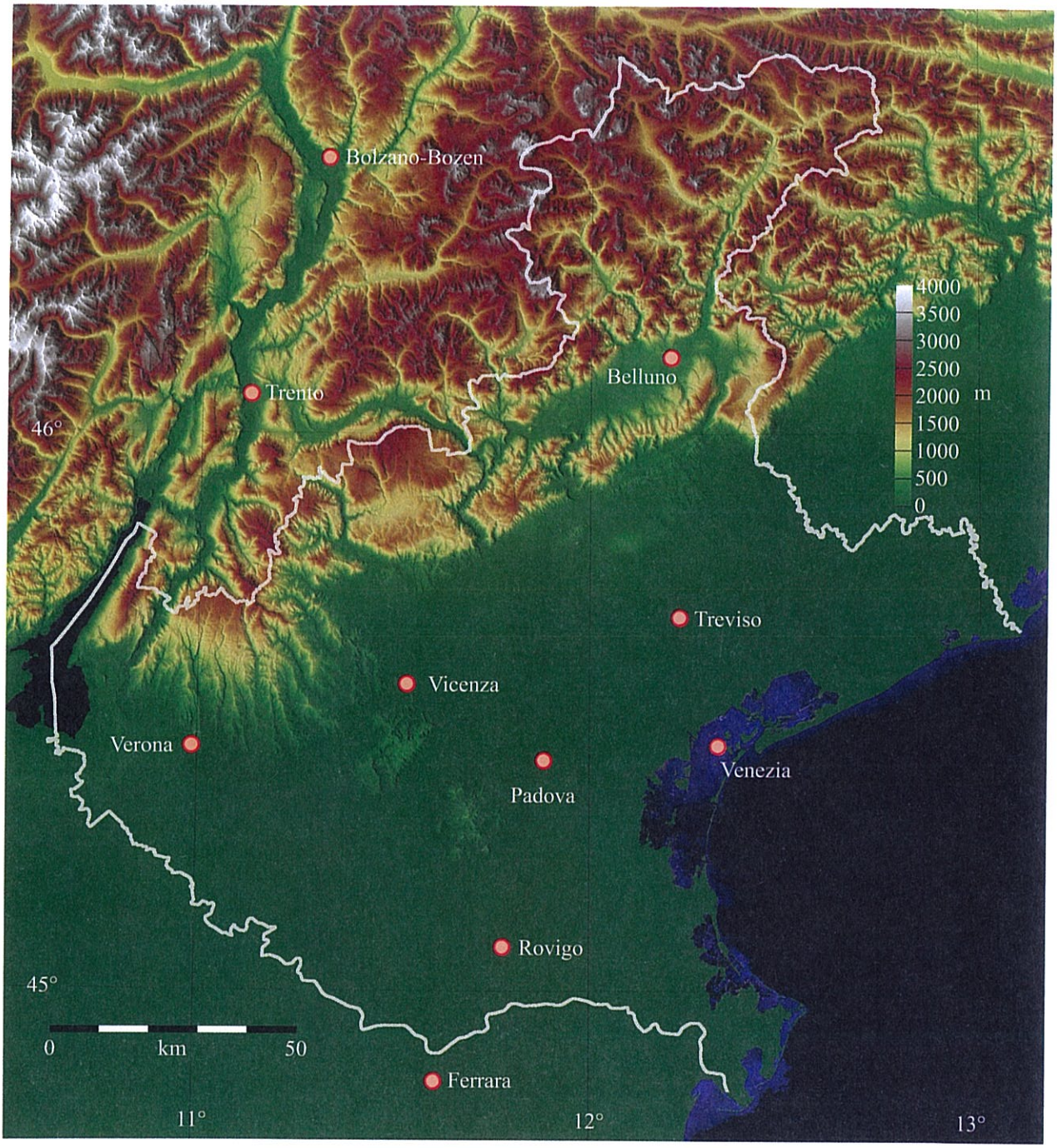


Figure 15-26: SRTM topographic map of part of Italy.

period of about 9 months. The terrain height data were produced in DTED-2 format at 1×1 arcsec spacing up to latitude 50° and 100 m (latitude) $\times$ 200 m (longitude) above 50°. All data are available free of charge from a server operated by the US Geological Survey. An example is shown in Fig. 15-26.

15-8 Along-Track Interferometry

The previous sections of this chapter were focused on the use of InSAR for measuring topography.

► By altering the imaging geometry to an ***along-track interferometer***, it is possible to use the InSAR to measure surface motion, rather than elevation. ◄

Many of the processing techniques described previously continue to be applicable, but in the new imaging geometry (Fig. 15-27) the baseline is described in terms of the temporal variable t, rather than in terms of spatial coordinates.

15-8.1 Temporal Baseline

The configuration depicted by Fig. 15-27 is an ***along-track radar interferometer*** in which antennas A_1 and A_2 are mounted on the front and back sections of the aircraft fuselage, rather than across-track on the two wings. The separation between the two antennas is the baseline B and the vehicle velocity is u. The two images generated by the two independent antennas are

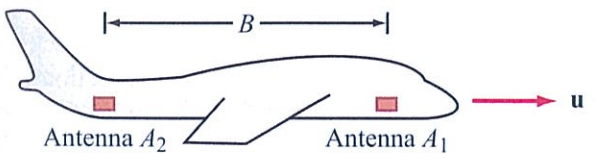


Figure 15-27: Interferometer geometry to measure along-track motions. Antennas mounted at the front and rear of an aircraft capture two radar images displaced in time by the interval it takes the plane to travel the baseline distance B.

essentially identical, except that the image generated by antenna A_2 is delayed in time by $\Delta t = B/u$, namely the amount of time it takes the aircraft to fly over the distance B. For a typical aircraft velocity, Δt is a fraction of a second. In Section 15-9 we consider the case where the two images are acquired by repeat pass satellite interferometry in which the same, almost identical, path is traversed by the same SAR at two times separated not by seconds, but by years! The complex voltage distributions $\mathbf{V}_1(x,y)$ and $\mathbf{V}_2(x,y)$ of the two images are related to each other by

$$\mathbf{V}_1(x,y) = \mathbf{V}_2(x+B,\, y), \tag{15.87}$$

where x and y are the along-track and cross-track coordinates, respectively. Alternatively, we can express the voltages in terms of (t,R), where t is time and R is the range to pixel (x,y). That is,

$$\mathbf{V}_1(t,R) = \mathbf{V}_2\left(t+\frac{B}{u},\, R\right). \tag{15.88}$$

The two images are identical, except for the shift in position in the along-track dimension. In Fig. 15-28, we show a top view of the imaging geometry, with particular interest in an object located at a distance y from the flight track. We first consider the interferogram for a stationary object. If image 1 is acquired at t_1, then when the second antenna acquires image 2 from the same position relative to the target, it acquires it at $t_2 = t_1 + B/u$. Since the ranges to the stationary object are identical, the interferometric phase difference is zero. But if the object is moving *away* from the flight path with a radial velocity component v_r, then

$$R_2 = R_1 + v_r \frac{B}{u} \sin\theta, \tag{15.89}$$

where θ is the incidence angle and $v_r \sin\theta$ is the component of v_r along the direction of the radar line of sight. The corresponding interferometric phase is

$$\begin{aligned}\phi_{\text{int}} = \phi_1 - \phi_2 &= -\frac{4\pi}{\lambda}(R_1 - R_2) \\ &= \frac{4\pi}{\lambda u} B v_r \sin\theta. \end{aligned} \tag{15.90}$$

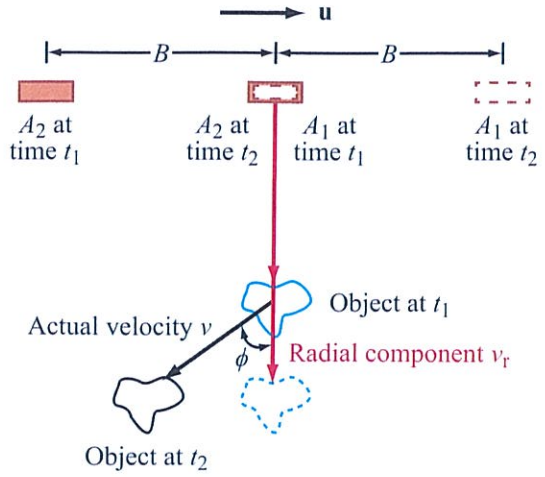


Figure 15-28: Observing a moving object receding from the flight path of the aircraft. In the time between observation, the object moves further away, and thus the phase of the radar signal increases.

The interferometric phase is directly proportional to the radial velocity of the object, v_r. Note that v_r is related to the true velocity v of the object by $v_r = v\cos\phi$ (Fig. 15-28).

15-8.2 Ocean Currents

In Section 15-5.5, we noted that in cross-track interferometry, the motion of scatterers in the resolution cell between the two observations leads to signal decorrelation, and therefore loss of interferometric phase information. In contrast, the along-track interferometry result encapsulated by Eq. (15.90) indicates that ϕ_{int} can be used to measure the radial velocity v_r, provided all scatterers in the resolution move together. This is the case for ocean currents (Fig. 15-29), which means that along-track interferometry can be used to measure v_r of the ocean surface.

To quantify how well the radial velocity v_r of the ocean surface can be measured, we replace ϕ_{int} in Eq. (15.90) with the phase measurement uncertainty $s_{\phi_{\text{int}}}$ and we also replace v_r with the radial velocity

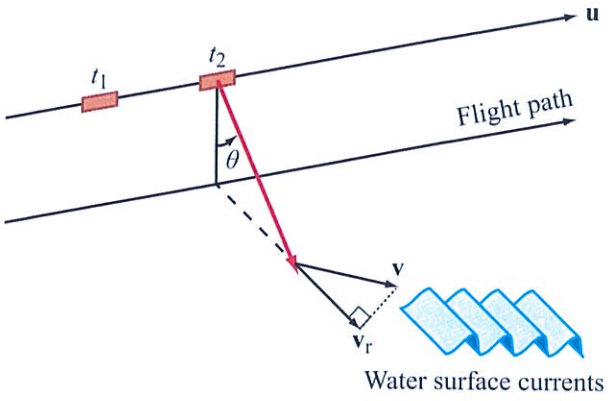


Figure 15-29: Along-track interferometer observing the velocity of the ocean surface.

measurement uncertainty s_{v_r}. The process gives

$$s_{v_r} = \frac{\lambda u}{4\pi B \sin\theta}\, s_{\phi_{\text{int}}}. \tag{15.91}$$

For an L-band SAR with $\lambda = 24$ cm, a baseline spacing $B = 20$ m along the belly of the aircraft, an aircraft velocity $u = 250$ m/s, and an incidence angle $\theta = 45°$,

$$s_{v_r} = \frac{0.24 \times 250}{4\pi \times 20 \times 0.707}\, s_{\phi_{\text{int}}} = 0.34 s_{\phi_{\text{int}}}.$$

If ϕ_{int} can be measured with a precision $s_{\phi_{\text{int}}} = 3° \approx 0.05$ rad, then $s_{v_r} \approx 1.7$ cm/s.

15-9 Measuring Surface Deformation

Satellites can be made to repeat their flight tracks with minimal cross-track deviation. The combination of two repeat-path SAR measurements constitutes an along-track interferometer. For a perfectly aligned flight track, such an interferometer loses all sensitivity to topography, but gains sensitivity to radial motions of the imaged surface.

Using a sequence of ERS-1 SAR images acquired with nearly identical paths over Antarctica, Goldstein et al. (1993) demonstrated that the flow velocity of the

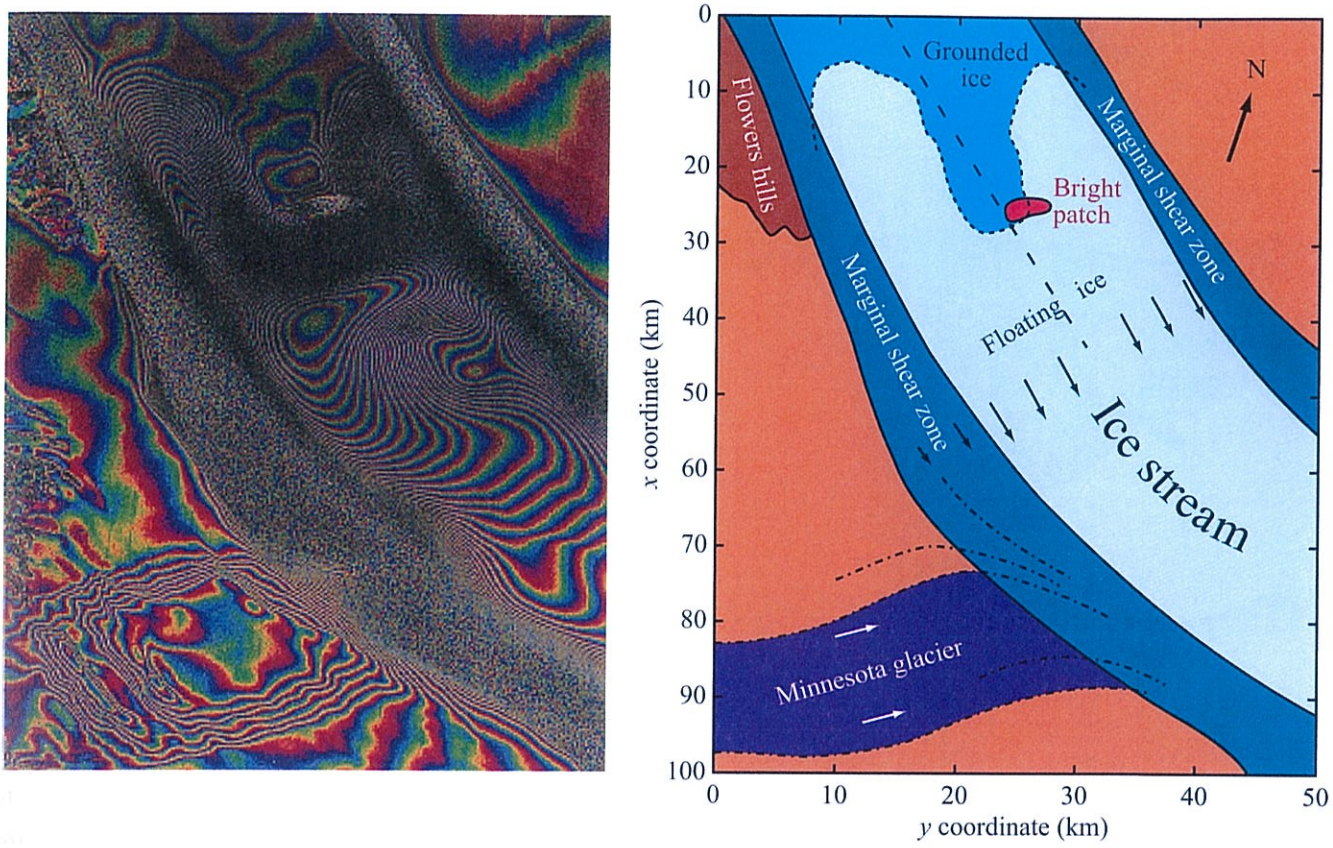


Figure 15-30: Radar interferogram of a portion of the Rutford ice stream in Antarctica, based on two ERS-1 images taken six days apart. The fringe pattern (color cycle) is essentially a map of ice-flow velocity, with one fringe representing 28 mm of range change along the radar line of site. [Image courtesy Jet Propulsion Laboratory, California Institute of Technology.]

ice stream can be measured with a precision of about 3×10^{-8} m/s. An interferometric image and associated ice-flow map are shown in Fig. 15-30.

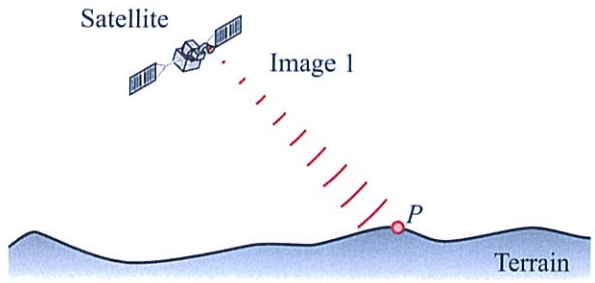
In Eq. (15.91), the ratio B/u represents the time delay Δt between the two SAR observations. Making such a substitution gives

$$s_{v_r} = \frac{\lambda}{4\pi \sin\theta\, \Delta t}\, s_{\phi_{\text{int}}}. \tag{15.92}$$

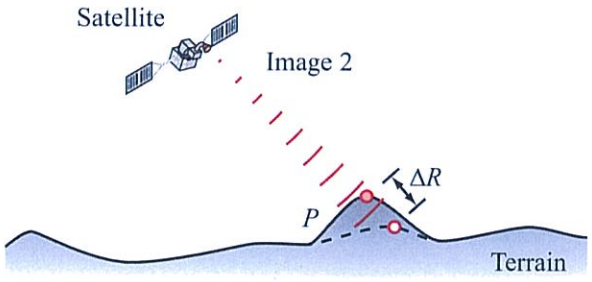
If the two observations are separated by $\Delta t = 30$ days, repetition of the L-band example that follows Eq. (15.92) leads to

$$\begin{aligned} s_{v_r} &= \frac{0.24}{4\pi \times 30 \text{ days}} \times 0.05 \\ &= \frac{0.24 \times 0.05}{4\pi \times 0.707 \times 30 \text{ days}} \\ &= 4.5 \times 10^{-5} \text{ m/day} = 1.6 \text{ cm/year}. \end{aligned}$$

This measurement precision is comparable to the speed of tectonic motions of the major plates on the surface of the Earth, which means that InSAR has the capability to map even such very slow-motion processes.



(a) Image 1 at time t_1



(b) Image 2 at time t_2

Figure 15-31: Surface deformation has displaced the location of point P by ΔR along the direction of the radar line of sight.

Actually, detecting displacements directly is a more practical approach than inferring velocities, in part because many tectonic motions are episodic and do not occur as smooth, continuous motions.

The scenario shown in Fig. 15-31 depicts a surface deformation that occurred between the times of acquisitions of two SAR images. The deformation causes the surface to change by ΔR along the direction of the radar line of sight. The phases of the pixel P in the two images are

$$\phi_1(t_1) = -\frac{4\pi}{\lambda} R_1 + \phi_{\text{scat}_1}, \tag{15.93a}$$

$$\phi_2(t_2) = -\frac{4\pi}{\lambda} (R_1 - \Delta R) + \phi_{\text{scat}_2}, \tag{15.93b}$$

where ϕ_{scat_1} is the phase of the scattering amplitude of the pixel at time t_1, and a similar definition applies to ϕ_{scat_2} at time t_2.

If the surface deformation moves the entire pixel, but does not change the relative locations of the individual scatterers contained in the pixel, then $\phi_{\text{scat}_1} = \phi_{\text{scat}_2}$. Consequently, the measured interferometric phase is

$$\phi_{\text{int}} = \phi_1 - \phi_2 = -\frac{4\pi}{\lambda} \Delta R.$$

Hence, in principle, by measuring the interferometric phase, we can quantify the line-of-sight deformation of the surface of every pixel in the image. This measurement capability, however, is compromised by two factors: (1) changes in the surface scatterers, causing $\phi_{\text{scat}_1} \neq \phi_{\text{scat}_2}$ and (2) deviation in the satellite orbit between the two observations. The first factor is related to the temporal decorrelation discussion of Section 15-5.5, and the second factor is associated with the role of topography. If the two tracks have a cross-track shift between them, ϕ_{int} becomes a sum of two components, one due to topography and another due to surface deformation. Through the application of more than two passes, together with the availability of precise track data, it is possible to separate the effects due to topography from those due to surface deformation.

The interferometric image in Fig. 15-32 shows surface displacements caused by the 6.3-magnitude L'Aquila earthquake in Italy, which occurred on April 6, 2009. The InSAR images were acquired by the ASAR system on ESA's ENVISAT. The indicated displacement values are the components of the true displacement along the direction of the radar line of sight.

15-10 Worldwide Dual Satellite InSAR Coverage: The TanDEM-X Mission

Most spaceborne InSAR experiments to date have been implemented using a single satellite in a repeat orbit to obtain the multiple images needed to form interferometric pairs. One major exception to this is SRTM, described in Section 15-7, in which two antennas on a single platform illuminate the ground simultaneously. Because the baseline length in this configuration is physically limited to structures that

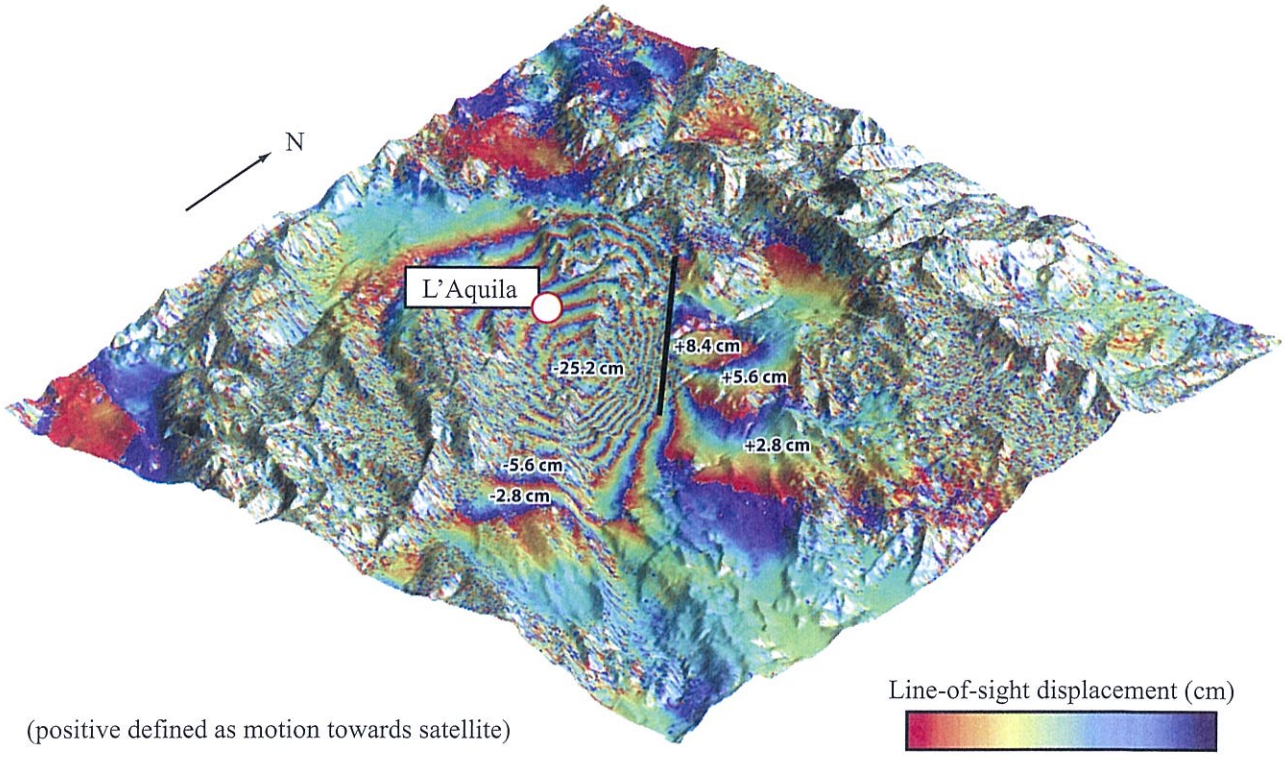


Figure 15-32: Interferogram due to surface deformation caused by the 6.3-magnitude L'Aquila earthquake in Italy [Walters et al., 2009].

can be reliably deployed in space, the performance of such a system is limited as described in Eq. (15.74). Mounting the two antennas on separate spacecraft with coordinated orbits so that simultaneity of observation is preserved, however, allows for a longer baseline separation in space (for topographic measurements) and in time (for velocity measurements, see Section 15-8). This approach was used by the German Aerospace Center, DLR, and Astrium consortium to build and fly the TanDEM-X (TerraSAR Digital Elevation Model - X) mission, which was launched in 2010. Over its expected three-year lifetime, the TanDEM-X mission will survey all 150 million square kilometers of Earth's land surface several times.

In this mission, each of two identical X-band radar satellites image the surface so that the same footprint is viewed by both satellites at the same time. Figure 15-33 is an artist's conception of the dual satellite configuration. This nearly eliminates the spatially and temporally varying atmospheric phase delay that is typically the largest statistical error source for InSAR. The ability to position the satellites independently allows for large spatial baselines, further reducing errors in the topographic product. The digital elevation models produced by the mission, with its 12 m spatial resolution and 2 m vertical accuracy, surpass even the best SRTM data sets acquired in 2000.

Furthermore, the two satellites may be positioned along track, rather than cross track, to form a temporal aperture and map surface velocities from space. This was first demonstrated using TanDEM-X data by Romeiser et al. (2010), who recorded images of currents

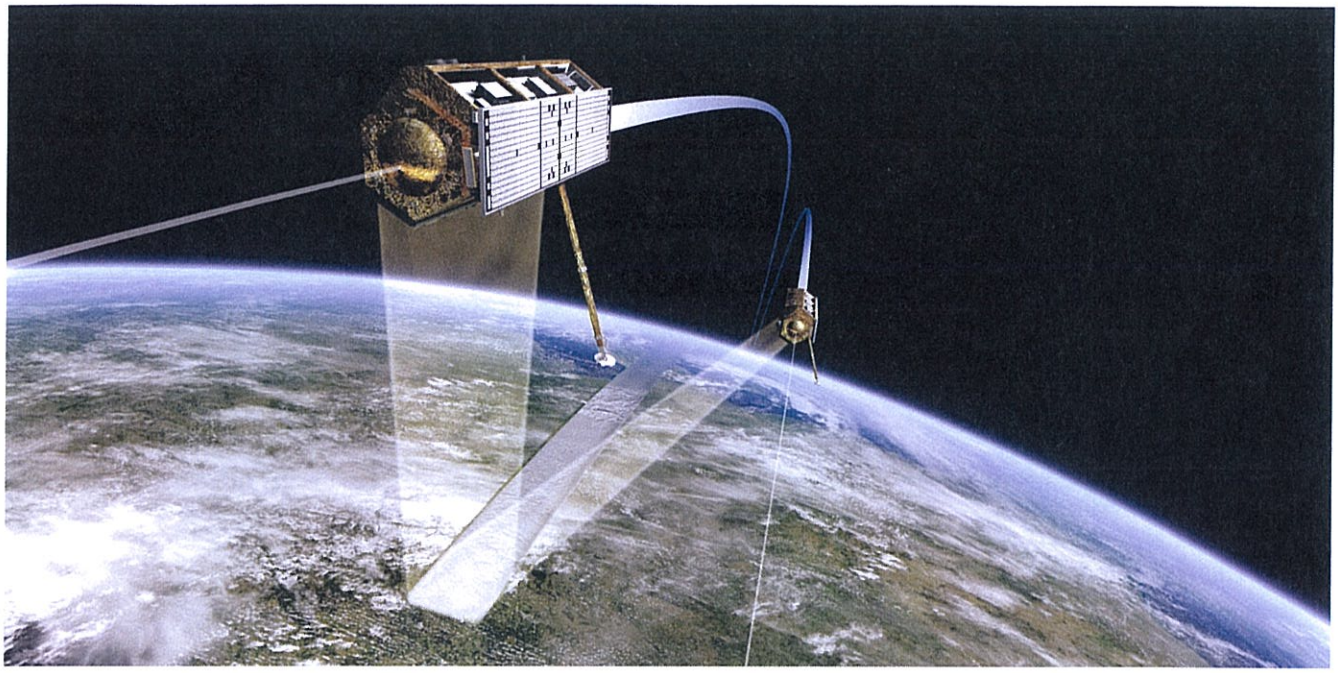


Figure 15-33: Artist's conception of twin TerraSAR satellites collecting simultaneous data to form the TanDEM-X system. The ability to achieve large cross-track baselines and greatly suppress atmospheric phase artifacts leads to digital elevation models with 2 m height accuracy at 12 m postings.

at the mouth of the Elbe river at varying tidal phases in order to visualize the flow and ebb of the water. They were able to achieve accuracies of several tenths of one m/s, as verified through ground measurements and predicted water velocities.

15-11 Time-Series InSAR Applications

Two-pass and three-pass interferometric methods for measuring surface deformation have evolved into ***multipass radar interferometry***. Additional applications are enabled if a long time series of observations from nearly repeating orbits are acquired. Multipass observations serve to improve the inferred deformation by averaging the predicted value, or to observe temporally changing deformation patterns, or both. Many existing SAR systems produce such data sets routinely. Through the use of time-series methods, such as ***stacking***, ***small baseline subset analysis*** (SBAS), and ***persistent scattering*** (PS), it is possible to improve the precision of the surface deformation estimate provided by a standard InSAR by an order of magnitude or more. Examples of InSAR analyses achieving millimeter-level precision with PS and SBAS techniques have been reported in numerous studies, including those by Ferretti et al. (2007), Johanson et al. (2009), Li et al. (2009), and Amelung et al. (2008).

15-11.1 Stacking

The stacking technique involves co-registering a ***stack of interferograms*** and adding them together (Lyons and Sandwell, 2003; Fialko, 2006). When implementing the stacking process, the interferograms may or may not be unwrapped in phase, and are often scaled by time separation in order to preserve deformation rates.

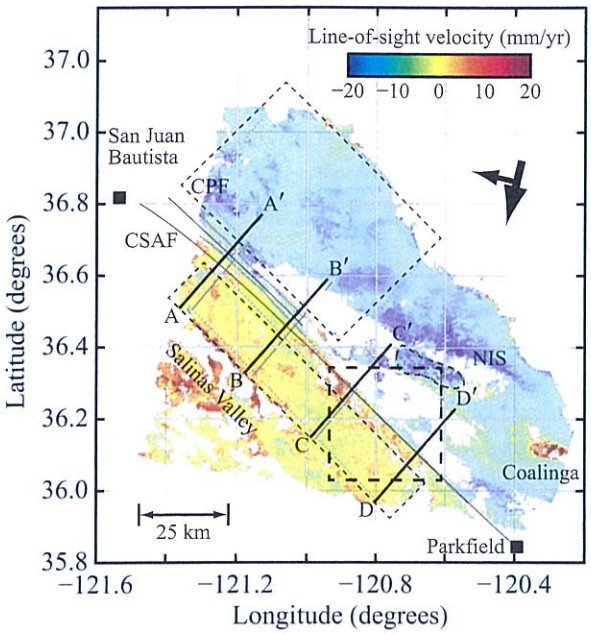


Figure 15-34: Stack of 12 descending ERS interferograms over the San Andreas Fault acquired between May 1992 and January 2001. The average creep rate in this section is 1.5–2.0 cm/yr, and it is easily seen in this image. Areas corrupted by spatial decorrelation, mainly because this is at a short, 6 cm wavelength, yield no useful displacement estimates and are white in this image. [Ryder and Bürgmann (2008).]

The main advantage of the stacking method is reduction of statistical errors, but it yields only a single average measurement and cannot easily retrieve time histories of deformation. An example of a stack is shown in Fig. 15-34, depicting average ***creep rate*** along the San Andreas Fault in California, from Parkfield northward to San Juan Batista. The fairly small displacement rate of about 1.5 cm/yr is easily discernible.

15-11.2 Small Baseline Subset Analysis (SBAS)

► The small baseline subset analysis (SBAS) technique can provide the *history* of the deformation of the surface crust, not just the mean deformation rate provided by the stacking technique. ◄

Successful implementation requires the selection of time-series interferograms with high correlation. The technique was first demonstrated by Berardino et al. (2002) who combined sets of high-correlation interferograms to obtain the time history for each correlated patch of ground over a site in southern Italy. The process, which includes spatial low-pass filtering and temporal high-pass filtering of the series of interferometric phase measurements for each imaged pixel, leads to the generation of an image displaying the measured rms deformation (Fig. 15-35) over the total observation time period (which in this case was close to 10 years), as well as a time series of the deformation at selected points in the image (Fig. 15-36). Other

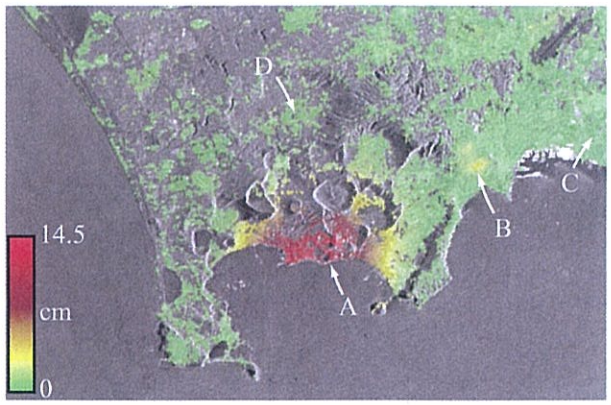


Figure 15-35: False-color map of the measured deformation rms superimposed on the SAR image amplitude of the investigated area. The temporal evolution of the deformations in the selected points identified by A, B, C are shown in Fig. 15-36(a)–(c), respectively [Berardino et al., 2002].

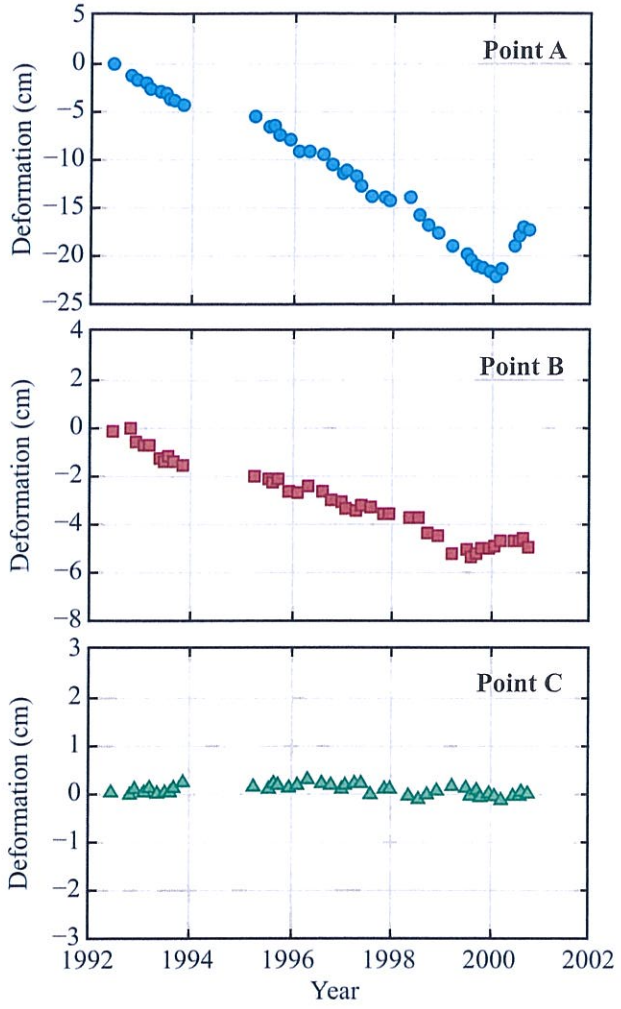


Figure 15-36: Time-series deformation measured at (a) point A, (b) point B, and (c) point C of Fig. 15-35 [Berardino et al., 2002].

examples of the SBAS technique are available in Pepe et al. (2005, 2011) and Reeves et al. (2011).

15-11.3 Persistent Scattering (PS)

The preceding techniques—stacking and SBAS—use a time series of conventional interferograms, and as such , they are subject to the decorrelation limitations of InSAR. Since decorrelation is dependent on the type of terrain and on changes that occur between one date and another, the SBAS technique can produce reliable deformation estimates over only those patches of the Earth's surface where the measured InSAR phase is reliable over some spatial extent. Consequently, decorrelation remains an obstacle to comprehensive coverage of surface deformation across the entire imaged scene.

One way to avoid the decorrelation problem is to identify individual pixels in the image that do not *scintillate*, thereby remaining highly correlated over the complete series of interferograms. Examples include buildings, bridges, and similar man-made structures. If a sufficient number of these ***persistent-scattering*** (PS) pixels are located in the imaged scene, a sufficiently dense geodetic network can be formed to allow deformation observations at small scales at fine (mm-level) precision. The PS technique was initially developed to track precise motions (subsidence) of man-made structures with mm-level resolution (Ferretti et al., 2000, 2001, 2004; Colesanti et al., 2003; Adam et al., 2003; Crosetto et al., 2003; Lyons and Sandwell, 2003; Werner et al., 2003).

The statistics of the field scattered from the many scatterers within a resolution cell are usually modeled by a Rayleigh fading model in which the power is exponentially distributed and the the phase is uniformly distributed (see Section 5-7). And whereas the Rayleigh fading model is appropriate for a cell containing a large number of randomly located scatterers, the model is not applicable when the total scattering is dominated by one or a few scatterers. Examples of dominant scatterers are corner reflectors formed by two walls or a wall and a ground surface, much like the corner reflectors used as radar calibration devices (see Section 13-11.4).

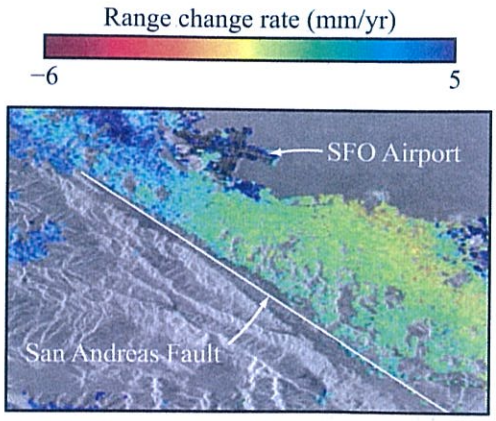
► Such persistent-scattering (PS) structures exhibit a *permanent* or nonscintillating behavior when viewed repeatedly by a SAR from the same perspective, which allows for very precise (mm-scale) measurements of their deformations. ◄

The PS technique was generalized and successfully applied to natural terrain to measure volcano deforma-

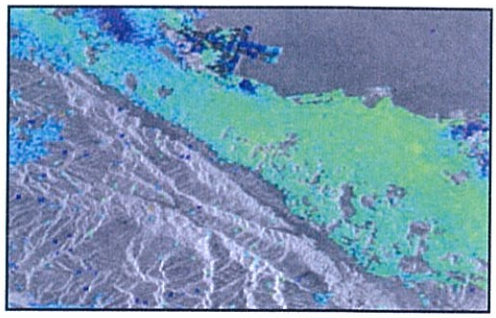
tion in vegetated areas (Hooper, 2006a; Hooper et al., 2007). This was accomplished by identifying persistent scatterers on the basis of their phase variations as a function of time, instead of their amplitude scintillation patterns.

The PS technique was further expanded in two major ways. First, by applying maximum-likelihood estimation methods, it is possible to improve the identification of PS points in the imaged scene, particularly in natural terrain (Shanker and Zebker, 2007). This leads to a denser geodetic network in space and time, which helps resolve more subtle and finer-scale deformations. The second expansion is related to the development of more accurate phase unwrapping methods (Shanker and Zebker, 2009, 2010). The improvement realized by these steps is illustrated in Fig. 15-37, which compares PS points identified using the original PS method (Ferretti et al., 2000), the more generalized approach introduced by Hooper (2006b), and the maximum-likelihood approach proposed by Shanker and Zebker (2007). Gray pixels in the images represent scintillating pixels and color pixels represent PS points. The maximum likelihood approach is able to identify not only all of the PS points identified by the other techniques, but also many others that are located in vegetated areas.

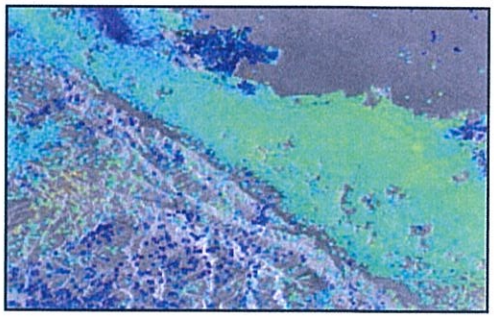
Application of the PS maximum-likelihood method to the San Andreas Fault area in California is demonstrated by the image in Fig. 15-38(a). The color of each identified PS point is associated with its displacement rate in mm/yr, measured along the line-of-sight of the radar. Hence, it is called the ***LOS displacement rate***. For the 20 km long segment noted in Fig. 15-38(a), the displacement rate as a function of distance from the fault is shown in Fig. 15-38(b). Pixels above the fault line exhibit an average displacement rate of about 6 mm/yr, compared with only about 1 mm/yr in the opposite direction for those pixels below the fault line.



(a) Permanent scatterers amplitude dispersion (Ferretti et al., 2000)

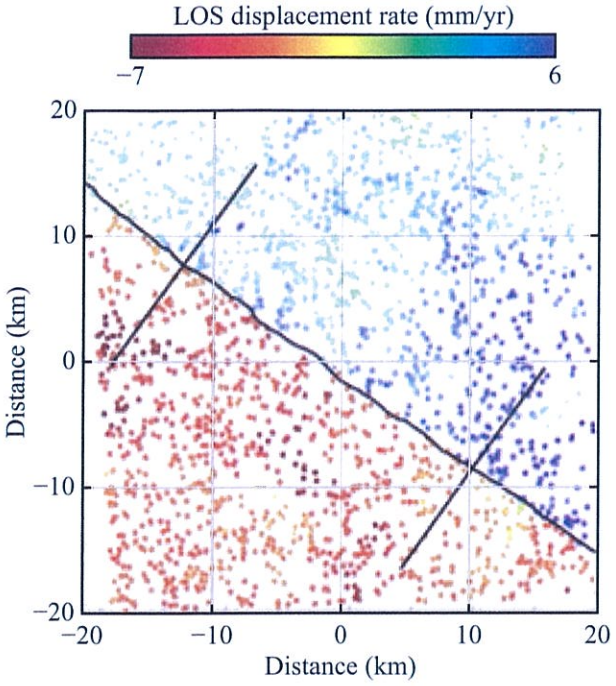


(b) Pixel phase and filtering (Hooper et al., 2004)

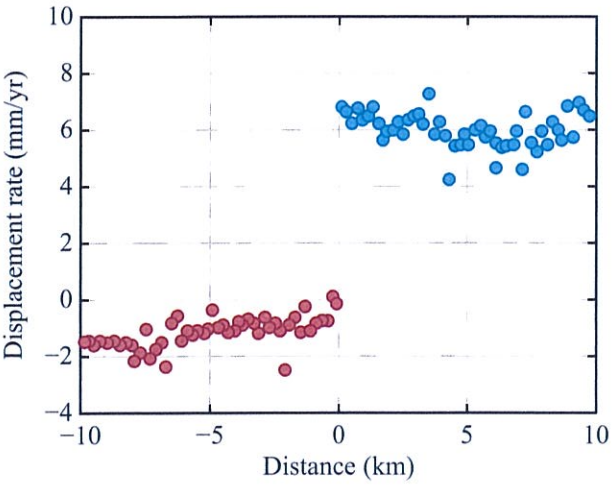


(c) Maximum likelihood (Shanker and Zebker, 2007)

Figure 15-37: Comparison of PS identification algorithms according to the methods by Ferretti et al. (2000), (b) Hooper et al. (2004), and (c) Shanker and Zebker (2007). Color pixels are persistent-scattering points.



(a) Average LOS displacement rate image over San Andreas Fault estimated by using the edgelist phase unwrapping algorithm



(b) Average LOS displacement rate as a function of distance from the fault

Figure 15-38: (a) Average LOS displacement rate image over the San Andreas Fault, (b) average LOS displacement rate as a function of distance from the fault. Assuming that all the displacement is purely due to ***strike slip motion*** across the fault, the estimated ***slip rate*** is 22 mm/yr. [Shanker, 2010.]

PROBLEMS

15.1 For an orbital topographic interferometer with an SNR of 15 dB, a perpendicular baseline length of 500 m, a range of 800 km, and operating at a wavelength of 6 cm, what is the error in the derived heights due to noise? The incidence angle is 23°.

15.2 Derive the slant to ground-range mapping function for a radar operating at a 35° incidence angle. Assume that the Earth is flat.

15.3 Repeat Problem 15.2 for a spherical Earth.

15.4 Given the following measurements of pixel offset in range versus pixel range deviation from the range to the pixel at the center of the beam of one of the images, what are the perpendicular and parallel baseline components for a radar with a range of 850 km, incidence angle of 25°, and slant-range pixel spacing of 7.5 m?

Range pixel deviation (m)	Measured offset (m)
100	−2.8
150	−1.8
200	−1.45
250	−1.1
300	−0.2
350	0.3
400	0.4
450	1.2
500	1.7
550	1.8
600	2.3
650	3.2
700	3.2
750	4.1
800	4.4

15.5 Calculate the critical baseline for a 12 cm wavelength radar with a sinc impulse response, a bandwidth of 80 MHz, a range of 700 km, and operating at 30° incidence angle?

15.6 A 1280 MHz radar operating from a range of 750 km has an SNR of 12 dB, an InSAR baseline of

2 km, and an incidence angle of 35°. Assuming that the baseline orientation angle is known to 0.01° and the baseline length to 2 m, what are the contributions to height error for topographic applications from phase, baseline length, and baseline orientation angle?

15.7 What is the uncertainty associated with the radial velocity of the surface when measured by a radar with 10 dB SNR, an along-track baseline of 10 m, and a platform velocity of 200 m/s, if the radar operates at 24 cm wavelength and an incidence angle of 45°?

For Problems 15.8–15.12, consider a radar system with the following description:

Orbit altitude:	696000 m
Wavelength:	0.236057 m
PRF:	2159.827 Hz
Sample rate f_s:	16.0 MHz
Chirp slope:	5.1851852×10^{11} Hz/s
Pulse length:	27.0 μs
Satellite velocity:	7179.4 m/s
R_0:	844768 m
f_D:	0 Hz

Copy two data files, image.1 and image.2, from *Images* in the book website mrs.eecs.umich.edu. Each image is a complex floating-point file consisting of 14336 lines of 3072 complex samples each. These files are used for the next five problems.

15.8 Form an interferogram by multiplying the two images together and display the phase, showing interferogram fringes. Calculate and submit a multilook version of the interferogram with 4 looks in range and 16 looks in azimuth.

15.9 Calculate the offsets over the original images, and plot the range offset versus range. Determine the average azimuth offset.

15.10 Calculate the baseline components $B_{\|}$ and $B_{\perp}$ from the offsets. Calculate the baseline length B and orientation angle α.

15.11 Resample image.2 to align with image.1 and form the new interferogram. If you need to resample in azimuth, be sure to account for the Doppler centroids of the images in some manner.

15.12 Calculate the correlation between the two images from step 1 and step 4, using 4×16 look averages for determining each correlation coefficient. Display the correlation images as byte files where zero correlation maps to 0, and unity correlation maps to 255.