CHAPTER 1

SYNTHETIC APERTURE RADAR (SAR) IMAGING BASICS

The word "radar" is an acronym for "radio detection and ranging." A radar measures the distance, or range, to an object by transmitting an electromagnetic signal to and receiving an echo reflected from the object. Since electromagnetic waves propagate at the speed of light, one only has to measure the time it takes the radar signal to propagate to the object and back to calculate the range to the object. The total distance traveled by the signal is twice the distance between the radar and the object, since the signal travels from the radar to the object and then back from the object to the radar after reflection. Therefore, once we measured the propagation time (t), we can easily calculate the range (R) as

$$R = \frac{1}{2}ct \tag{1-1}$$

where c is the speed of light in a vacuum. The factor $^1/_2$ accounts for the fact that the radar signal actually traveled twice the distance measured: first from the radar to the object and then from the object to the radar. If the electric property of the propagation medium is different from that of the vacuum, the actual propagation velocity has to be estimated for advanced radar techniques, such as synthetic aperture radar (SAR) interferometry.

Radars provide their own signals to detect the presence of objects. Therefore, radars are known as active, remote-sensing instruments. Because radars provide their own signal, they can operate during day or night. In addition, radar signals

Synthetic Aperture Radar Polarimetry, First Edition. Jakob Van Zyl and Yunjin Kim.

typically penetrate clouds and rain, which means that radar images can be acquired not only during day or night but also under (almost) all weather conditions. For these reasons, radars are often referred to as all-weather instruments. Imaging, remote-sensing radars, such as SAR, produce high-resolution (from submeter to a few tens of meters) images of surfaces. The geophysical information can be derived from these high-resolution images by using proper postprocessing techniques.

This book focuses on a specific class of implementation of synthetic aperture radar with particular emphasis on the use of polarization to infer the geophysical properties of the scene. As mentioned above, SAR is a way to achieve high-resolution images using radio waves. We shall first describe the basics of radar imaging. This shall be followed by a description of the synthetic aperture principle. Finally, we shall discuss some advanced SAR implementations, such as SAR polarimetry and polarimetric SAR interferometry.

1.1 BASIC PRINCIPLES OF RADAR IMAGING

Imaging radars generate surface images that are at first glance very similar to the more familiar images produced by instruments that operate in the visible or infrared parts of the electromagnetic spectrum. However, the principle behind the image generation is fundamentally different in the two cases. Visible and infrared sensors use a lens or mirror system to project the radiation from the scene on a "two-dimensional array of detectors," which could be an electronic array or, in earlier remote-sensing instruments, a film using chemical processes. The two-dimensionality can also be achieved by using scanning systems or by moving a single line array of detectors. This imaging approach—an approach with which we are all familiar from taking photographs with a camera—conserves the relative angular relationships between objects in the scene and their images in the focal plane, as shown in Figure 1-1. Because of this conservation of angular relationships, the resolution of the images depends on how far away the camera is from the scene it is imaging. The closer the camera, the higher the resolution and the smaller the details that can be recognized in the images. As the camera moves further away from the scene, the resolution degrades and only larger objects can be discerned in the image.

Imaging radars use a quite different mechanism to generate images, with the result that the image characteristics are also quite different from those of visible and infrared images. There are two different mechanisms by which radars can be used to produce images; the two types of radars are broadly classified as real aperture and synthetic aperture radars. We shall discuss the differences between these two types in more detail later in this chapter.

Radar images are typically acquired in strips as the satellite or aircraft carrying the radar system moves along its flight path. These strips are often referred to as swaths or tracks. To separate objects in the cross-track direction and the along-track direction within a radar image, two different methods must be implemented. The cross-track direction, also known as the range direction in radar imaging, is the direction perpendicular to the direction in which the imaging platform is moving. In

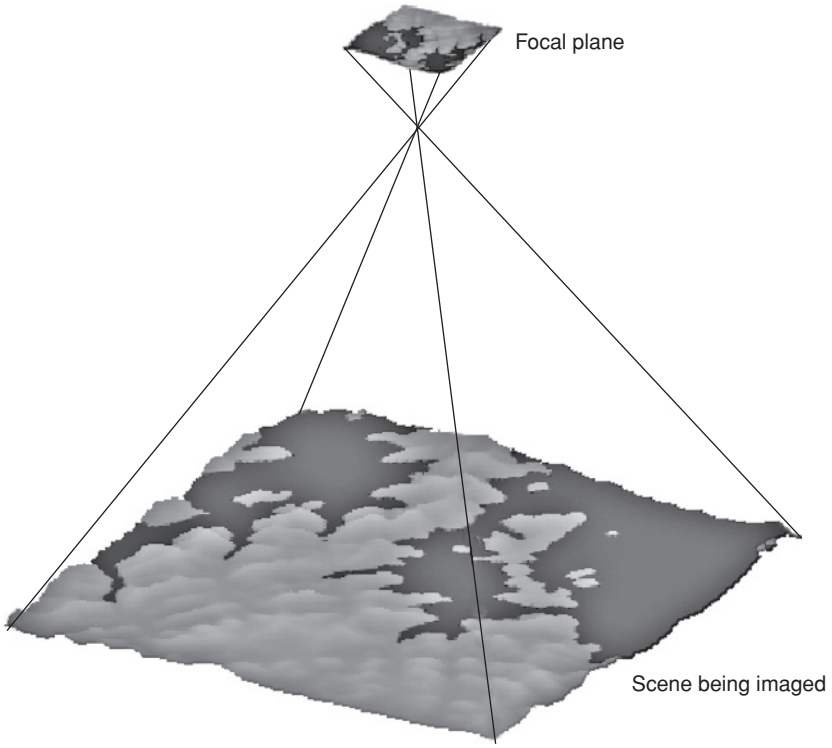

FIGURE 1-1 Passive imaging systems conserve the angular relationships between objects in the scene and their images in the focal plane of the instrument

this direction, radar echoes are separated using the time delay between the echoes that are backscattered from the different surface elements. This is true for both real aperture and synthetic aperture radar imagers. The along-track direction, also known as the azimuth direction, is the direction parallel to the movement of the imaging platform. The angular size (in the case of the real aperture radar) or the Doppler history (in the case of the synthetic aperture radar) is used to separate surface pixels in the along-track dimension in the radar images. As we will see later, only the azimuth imaging mechanism of real aperture radars is similar to that of regular cameras. Using the time delay and Doppler history results, SAR images have resolutions that are independent of how far away the radar is from the scene it is imaging. This fundamental advantage enables high-resolution, spaceborne SAR without requiring an extremely large antenna.

Another difference between images acquired by cameras operating in the visible and near-infrared part of the electromagnetic spectrum and radar images is the way in which they are acquired. Cameras typically look straight down, or at least have no fundamental limitation that prevents them from taking pictures looking straight down from the spacecraft or aircraft. Not so for imaging radars. To avoid so-called

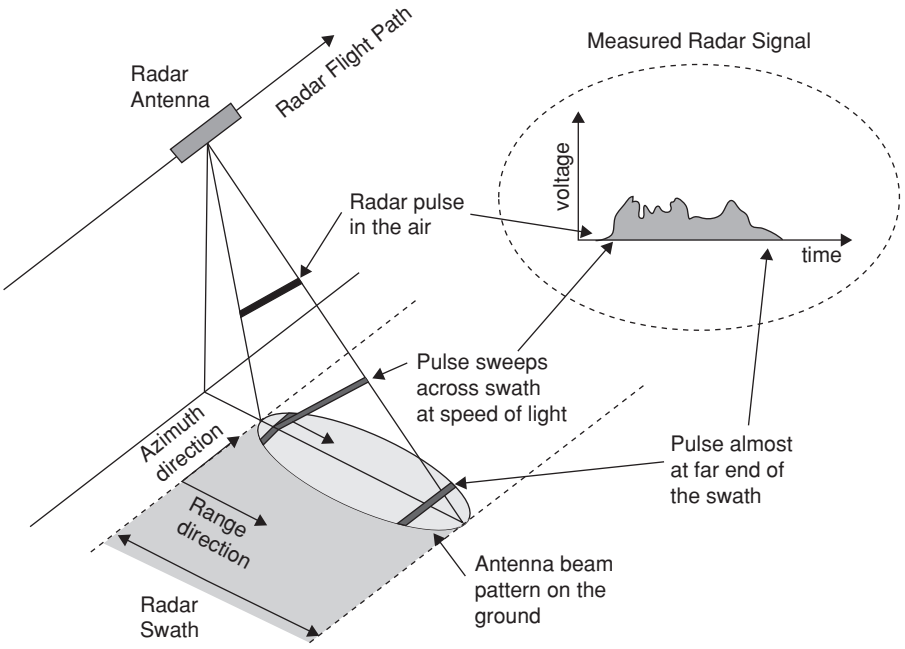

FIGURE 1-2 Imaging geometry for a side-looking radar system

ambiguities, which we will discuss in more detail later, the imaging radar sensor has to use an antenna that illuminates the surface to one side of the flight track. Usually, the antenna has a fan beam that illuminates a highly elongated, elliptically shaped area on the surface, as shown in Figure 1-2. The illuminated area across track generally defines the image swath.

Within the illumination beam, the radar sensor transmits a very short effective pulse of electromagnetic energy. Echoes from surface points farther away along the cross-track coordinate will be received at proportionally later times (see Figure 1-2). Thus, by dividing the receive time in increments of equal time bins, the surface can be subdivided into a series of range bins. The width in the along-track direction of each range bin is equal to the antenna footprint along the track x_a. As the platform moves, the sets of range bins are covered sequentially, thereby allowing strip mapping of the surface line by line. This is comparable to strip mapping with a so-called push broom imaging system using a line array in the visible and infrared part of the electromagnetic spectrum. The brightness associated with each image pixel in the radar image is proportional to the echo power contained within the corresponding time bin. As we will see later, the real difference between real aperture radars and synthetic aperture radars lies in the way in which the azimuth resolution is achieved.

This is also a good time to point out that there are two different meanings for the term "range" in radar imaging. The first is the so-called slant range and refers to the range along the radar line of sight, as shown in Figure 1-3. Slant ranges are measured

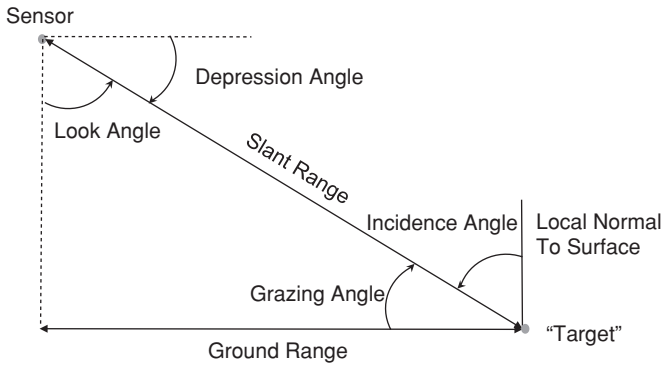

FIGURE 1-3 Definition of some common radar imaging angles

along the line connecting the radar and the object being imaged, often called the target or the scatterer. The second use of the term "range" is for the ground range, which refers to the range along a smooth surface (the ground) to the scatterer. The ground range is measured from the so-called nadir track, which represents the line described by the position directly underneath the radar imaging platform. One has to be careful to take topography into account when resampling radar images from slant range to ground range. This will be discussed in more detail in Section 1.6.2.

Before looking at radar resolutions, let us define a few more terms commonly encountered in radar imaging. The look angle is defined as the angle between the vertical direction and the radar beam at the radar platform. The incidence angle is defined as the angle between the vertical direction and the radar wave propagation vector at the surface (as shown in Figure 1-3). When surface curvature effects are neglected, the look angle is equal to the incidence angle at the surface when the surface is flat. In the case of spaceborne systems, surface curvature must be taken into account; this leads to an incidence angle that is always larger than the look angle for flat surfaces. It is quite common in the literature to find authors using the terms "look angle" and "incidence angle" interchangeably. This is only correct for low-flying aircraft, and only when there is no topography present in the scene. As we will see next, if topography is present (i.e., if the surface is not flat), the local incidence angle might vary in the radar image from pixel to pixel.

Consider the simple case of a single hill illuminated by a radar system, as is shown in Figure 1-4. Also shown in the figure are the local normal to the surface for several positions on the hill. Relative to a flat surface, it is clear that for points on the hill facing the radar, the local normal tilts more toward the radar; therefore, the local incidence angle will be smaller than for a point at the same ground range but on a flat surface.

A term commonly encountered in the military literature is "depression angle." This is the angle between the radar beam and the horizontal at the radar platform. The depression angle is, therefore, related to the look angle in that one is equal to 90° minus the other. A small look angle is equivalent to a large depression angle, and

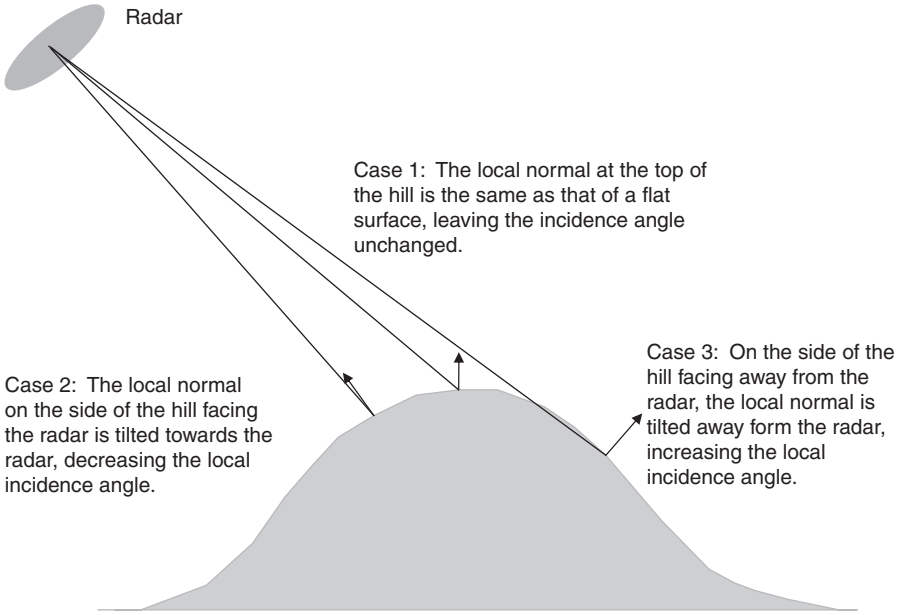

FIGURE 1-4 Topographic variations in the image will cause the local incidence angle to be different from that expected for a flat surface with no relief

vice versa. Similarly, one often finds the term "grazing angle" describing the angle between the horizontal at the surface and the incident wave in the military literature. The grazing angle is, therefore, related to the incidence angle in the same way that the depression angle is related to the look angle. In this text, we shall use "look angle" and "incidence angle" to describe the imaging geometry.

1.2 RADAR RESOLUTION

The resolution of an image is defined as that separation between the two closest features that can still be resolved in the final image. First, consider two point targets that are separated in the slant range direction by x_{r}. Because the radar waves propagate at the speed of light, the corresponding echoes will be separated by a time difference Δt equal to:

$$\Delta t = 2x_r/c \tag{1.2-1}$$

where c is the speed of light and the factor 2 is included to account for the signal round trip propagation, as was previously described. Radar waves are usually not transmitted continuously; instead, a radar usually transmits short bursts of energy known as radar pulses. The two features can be discriminated if the leading edge of

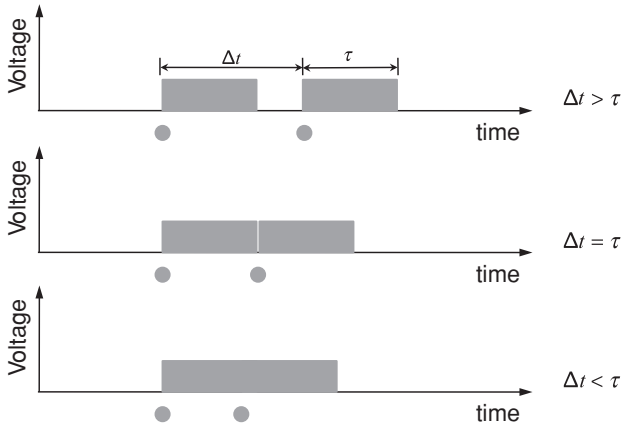

FIGURE 1-5 If the radar echoes from two point targets are separated in time by more than or equal to the length of the radar pulse, it is possible to recognize the echoes as those from two different scatterers, as shown in the top two panels. If the time difference between the echoes is less than the radar pulse length, it is not possible to recognize two distinct scatters, as in the case of the bottom panel

the pulse returned from the second object is received later than the trailing edge of the pulse received from the first feature, as shown in Figure 1-5.

Therefore, the smallest separable time difference in the radar receiver is equal to the effective time length τ of the pulse. Thus, the slant range resolution of a radar is:

$$2x_r/c = \tau \quad \Rightarrow \quad x_r = \frac{c\tau}{2} \tag{1.2-2}$$

Now let us consider the case of two objects separated by a distance x_g on the ground. The corresponding echoes will be separated by a time difference Δt equal to:

$$\Delta t = 2x_g \sin\theta/c \tag{1.2-3}$$

The angle θ in Eq. (1.2-3) is the local incidence angle. (This should actually be called the incident angle, or angle of incidence. Since "incidence angle" is used almost universally in the literature, we shall continue to use that term to avoid confusion.). As in the case of the slant range discussed above, the two features can be discriminated if the leading edge of the pulse returned from the second object is received later than the trailing edge of the pulse received from the first feature. Therefore, the ground range resolution of the radar is given by

$$2x_g \sin\theta/c = \tau \quad \Rightarrow \quad x_g = \frac{c\tau}{2\sin\theta} \tag{1.2-4}$$

In other words, the ground range resolution is equal to half the footprint of the radar pulse on the surface. Note that Eq. (1.2-4) implies that the ground range resolution is different for different incidence angles. This also means that the local slopes in the images will affect the ground range resolution. For slopes facing the radar, the ground range resolution will be poorer than that for slopes facing away from the radar.

Occasionally, the effective pulse length is described in terms of the system bandwidth B. As we will show in the next section, to a good approximation,

$$\tau = 1/B \tag{1.2-5}$$

A pulsed radar determines the range by measuring the round trip time by transmitting a pulse signal. In designing the signal pattern for a radar sensor, there is usually a strong requirement to have as much energy as possible in each pulse in order to enhance the signal-to-noise ratio. This can be done by increasing the transmitted peak power or by using a longer pulse. However, the peak power is usually strongly limited by the available power sources, particularly in the case of spaceborne sensors. On the other hand, an increased pulse length leads to a worse range resolution [see Eq. (1.2-4)]. This dilemma is usually resolved by using modulated pulses, which have the property of a wide bandwidth even when the pulse is very long. After so-called pulse-compression, a short effective pulse length is generated, increasing the resolution. One way to modulate the pulse is to vary the radar signal frequency linearly while the pulse is being transmitted. This waveform is known as chirp. Although other ways of modulating pulses are occasionally used, the chirp modulation is by far the most common; we shall therefore use this scheme to illustrate how signal modulation can be used to enhance radar range resolution.

In a chirp, the signal frequency within the pulse is linearly changed as a function of time. If the frequency is linearly changed from f_0 to $f_0 + \Delta f$, the effective bandwidth would be equal to:

$$B = |(f_0 + \Delta f) - f_0| = |\Delta f| \tag{1.2-6}$$

which is independent of the pulse length τ_p. Thus, a pulse with long duration (i.e., high energy) and wide bandwidth (i.e., high range resolution) can be constructed. The instantaneous frequency for such a signal is given by:

$$f(t) = f_0 + \frac{B}{\tau_p} t \quad \text{for} \quad -\tau_p/2 \le t \le \tau_p/2 \tag{1.2-7}$$

The corresponding signal amplitude is:

$$A(t) \sim \Re\left\{\exp\left[-i2\pi \int f(t)dt\right]\right\} = \cos\left[2\pi\left(f_0 t + \frac{B}{2\tau_p} t^2\right)\right] \tag{1.2-8}$$

where $\Re(x)$ means the real part of x. Note that the instantaneous frequency is the derivative of the instantaneous phase. A pulse signal such as shown in Eq. (1.2-8)

has a physical pulse length τ_p and a bandwidth B. The product $\tau_p B$ is known as the time bandwidth product of the radar system. In typical radar systems time bandwidth products of several hundreds are used.

At first glance, it might seem that using a pulse of the form shown in Eq. (1.2-8) cannot be used to separate targets that are closer than the projected physical length of the pulse (as shown in the previous section). It is indeed true that the echoes from two neighboring targets that are separated in the range direction by much less than the physical length of the signal pulse will overlap in time. If the modulated pulse and, therefore, the echoes have a constant frequency, it will not be possible to resolve the two targets. However, if the frequency is modulated as described in Eq. (1.2-7), the echoes from the two targets will have different frequencies at any instant of time and, therefore, can be separated by frequency filtering.

In actual radar systems, a matched filter is used to compress the returns from the different targets. Consider an example in which we transmit a signal of the form described in Eq. (1.2-8). The signal received from a single point scatterer at a range R is a scaled replica of the transmitted signal delayed by a time $t = 2R/c$. The output of the matched filter for such a point scatterer is mathematically described as the convolution of the returned signal with a replica of the transmitted signal. Being careful about the limits of the integration, one finds that for large time-bandwidth products,

$$V_o(t) = \tau_p E_r \ \exp(i\omega t) \ \exp(-i4\pi R/\lambda) \frac{\sin(\pi B(t - 2R/c))}{\pi B(t - 2R/c)} \tag{1.2-9}$$

where

$$\omega = 2\pi f_0$$

and

$$E_r = \text{electric field received by radar from a point source}$$

This compressed pulse has a half power width of approximately $1/B$, and its peak position occurs at time $2R/c$. Therefore, the achievable range resolution using a modulated pulse of the kind described in Eq. (1.2-8) is a function of the chirp bandwidth and not the physical pulse length. In typical spaceborne and airborne SAR systems, physical pulse lengths of several tens of microseconds are used, while bandwidths of several tens of megahertz are no longer uncommon for spaceborne systems, and several hundreds of megahertz are common in airborne systems.

So far, we have seen the first major difference between radar imaging and that used in passive imaging systems. The cross-track resolution in the radar case is independent of the distance between the scene and the radar instrument and is a function of the system bandwidth. Before looking at the imaging mechanisms in the along-track direction, we will examine the general expression for the amount of

reflected power that the radar receiver would measure. This is described through the so-called radar equation, which we will examine in the next section.

1.3 RADAR EQUATION

One of the key factors that determines the quality of the radar imagery is the corresponding signal-to-noise ratio (SNR), commonly called SNR. This is the equivalent of the brightness of a scene being photographed with a camera versus the sensitivity of the film or detector. Here, we consider the effect of thermal noise on the sensitivity of radar imaging systems. The derivation of the radar equation is graphically shown in Figure 1-6.

In addition to the target echo, the received signal also contains noise, which results from the fact that all objects at temperatures higher than absolute zero emit radiation across the whole electromagnetic spectrum. The noise component that is within the spectral bandwidth B of the sensor is passed through with the signal. The receiver electronics also generates noise that contaminates the signal. The thermal noise power is given by:

$$P_N = kTB \tag{1.3-1}$$

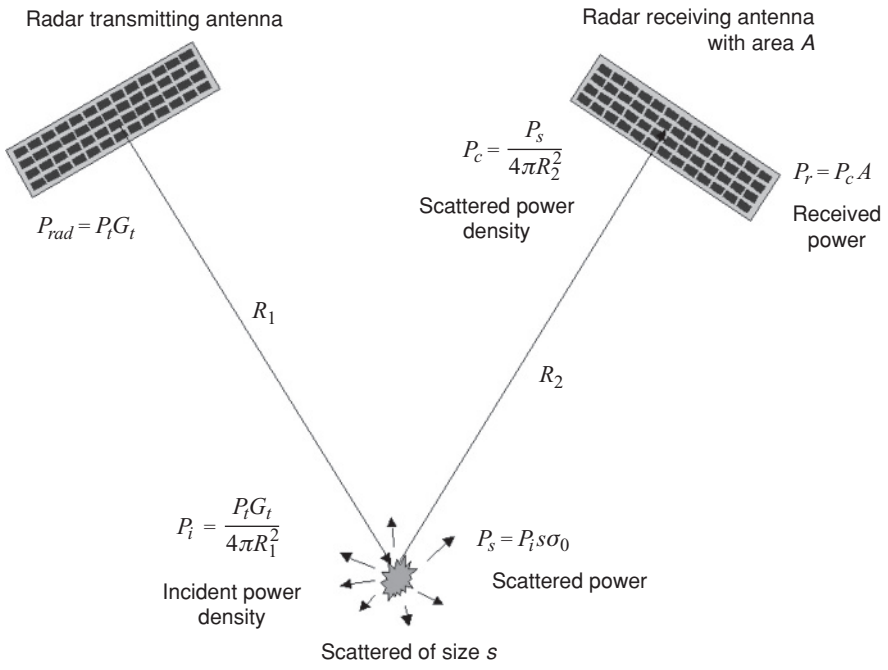

FIGURE 1-6 Schematic of the derivation of the radar equation

where k is Boltzmann's constant ($k = 1.6 \times 10^{-23}$ W/K/Hz) and T is the total equivalent noise temperature in kelvin. The resulting SNR is then:

$$SNR = P_r / P_N \tag{1.3-2}$$

Note that the noise bandwidth is usually larger than the transmit bandwidth, because of the hardware limitation.

One common way of characterizing an imaging radar sensor is to determine the surface backscatter cross section σ_N, which gives an SNR = 1. This is called the noise equivalent backscatter cross section. It defines the weakest surface return that can be detected and, therefore, identifies the range of surface units that can be imaged.

1.4 REAL APERTURE RADAR

The real aperture imaging radar sensor also uses an antenna that illuminates the surface to one side of the flight track. As mentioned before, the antenna usually has a fan beam that illuminates a highly elongated elliptical shaped area on the surface (see Figure 1-2). As shown in Figure 1-2, the illuminated area across track defines the image swath. For an antenna of width W operating at a wavelength λ, the beam angular width in the range plane is given by:

$$\theta_r \approx \lambda / W \tag{1.4-1}$$

and the resulting surface footprint or swath S is given by

$$S \approx \frac{h\theta_r}{\cos^2\theta} = \frac{\lambda h}{W \cos^2\theta} \tag{1.4-2}$$

where h is the sensor height above the surface, θ is the angle from the center of the illumination beam to the vertical (the look angle at the center of the swath), and θ_r is assumed to be very small. Note that Eq. (1.4-2) ignores the curvature of the Earth. For spaceborne radars, this effect should not be ignored. If the antenna beam width is large, one needs to use the law of cosines to solve for the swath width.

A real aperture radar relies on the resolution afforded by the antenna beam in the along-track direction for imaging. This means that the resolution of a real aperture radar in the along-track direction is determined by the size of the antenna as well as the range to the scene. Assuming an antenna length of L, the antenna beam width in the along-track direction is

$$\theta_a \approx \frac{\lambda}{L} \tag{1.4-3}$$

At a distance R from the antenna, this means that the antenna beam width illuminates an area with the along-track dimension equal to

$$x_a \approx R\theta_a \approx \frac{\lambda R}{L} \approx \frac{\lambda h}{L\cos\theta} \tag{1.4-4}$$

To illustrate, for $h = 800$ km, $\lambda = 23$ cm, $L = 12$ m, and $\theta = 20^\circ$, then $x_a = 16$ km. Even if λ is as short as 2 cm and h as low as 200 km, x_a will still be equal to about 360 m, which is considered to be a relatively poor resolution, even for remote sensing. This has led to very limited use of the real-aperture technique for surface imaging, especially from space. A real aperture radar uses the same imaging mechanism as a passive optical system for the along-track direction. However, because of the small value of λ ($\sim$1 μm) in the case of optical systems, resolutions of a few meters can be achieved from orbital altitudes with an aperture only a few tens of centimeters in size. From aircraft altitudes, however, reasonable azimuth resolutions can be achieved if higher radar frequencies (typically X-band or higher) are used. For this reason, real aperture radars are not commonly used in spaceborne remote sensing (except in the case of scatterometers and altimeters that do not need high-resolution data).

In terms of the radar equation, the area responsible for reflecting the power back to the radar is given by the physical size of the antenna illumination in the along-track direction and the projection of the pulse on the ground in the cross-track direction. This is shown in Figure 1-2 for the pulses in the radar swath. The along-track dimension of the antenna pattern is given by Eq. (1.4-4). If the pulse has a length τ_p in time, and the signal is incident on the ground at an angle θ_i, the projected length of the pulse on the ground is

$$l_g = \frac{c\tau_p}{2\sin\theta_i} \tag{1.4-5}$$

Therefore, the radar equation in the case of a real aperture radar becomes

$$P_r = \frac{P_t G_t G_r \lambda^2}{(4\pi)^3 R^4} \frac{\lambda R}{L} \frac{c\tau_p}{2\sin\theta_i} \sigma_0 \tag{1.4-6}$$

The normalized backscattering cross section (σ_0) is defined as

$$\sigma_0 = \lim_{R\to\infty} \frac{4\pi R^2}{A_0} \frac{|E_r|^2}{|E_i|^2} \tag{1.4-7}$$

where A_0 is the illuminated surface area, E_r is the reflected electric field, and E_i is the incident electric field.

This shows that when a real aperture radar images an extended area, the received power decreases as the range to the third power. In terms of the physical antenna

sizes, we can rewrite this expression as

$$P_r = \frac{P_t W^2 L c \tau_p \sigma_0}{8\pi\lambda R^3 \sin\theta_i} \tag{1.4-8}$$

This is the radar equation for a so-called distributed target for the real aperture radar case. From Eq. (1.4-8), it is clear that the received power increases as the square of the width of the antenna. However, increasing the antenna width also decreases the swath width. The received power only increases linearly with an increase in antenna length. Increasing the antenna length also improves the along-track resolution of the real aperture radar. For this reason, real aperture radars usually operate with antennas that are the longest that could be practically accommodated.

In summary, a real aperture radar uses the same imaging mechanism as passive imaging systems to achieve along-track resolution. The practically achievable resolutions are usually poorer than what is generally required for remote sensing applications. Real aperture radars are therefore not commonly used for remote sensing applications.

1.5 SYNTHETIC APERTURE RADAR

"Synthetic aperture radar" refers to a particular implementation of an imaging radar system that utilizes the movement of the radar platform and specialized signal processing to generate high-resolution images. Prior to the discovery of synthetic aperture radar, principle imaging radars operated using the real-aperture principle and were known as side-looking aperture radars (SLAR).

Carl Wiley of the Goodyear Aircraft Corporation is generally credited as being the first person to describe the use of Doppler frequency analysis of signals from a moving coherent radar to improve along-track resolution. He noted that two targets at different along-track positions will be at different angles relative to the aircraft velocity vector, resulting in different Doppler frequencies. (The Doppler effect is the well-known phenomenon that causes a change in the pitch of a car horn as it travels past a stationary observer.) Using this effect, targets can be separated in the along-track direction on the basis of their different Doppler frequencies. This technique was originally known as Doppler beam sharpening, but later became known as synthetic aperture radar (SAR).

The main difference between real and synthetic aperture radars is, therefore, in the way in which the azimuth resolution is achieved. The range resolution and radar equation derived previously for a real aperture radar are still valid here. The along-track imaging mechanism and the resulting along-track resolution are, however, quite different for the real and synthetic aperture radar cases.

As the radar moves along the flight path, it transmits pulses of energy and records the reflected signals, as shown in Figure 1-2. When the radar data are processed, the position of the radar platform is taken into account when adding the signals to integrate the energy for the along-track direction. Consider the geometry shown in Figure 1-7. As the radar moves along the flight path, the distance between the radar

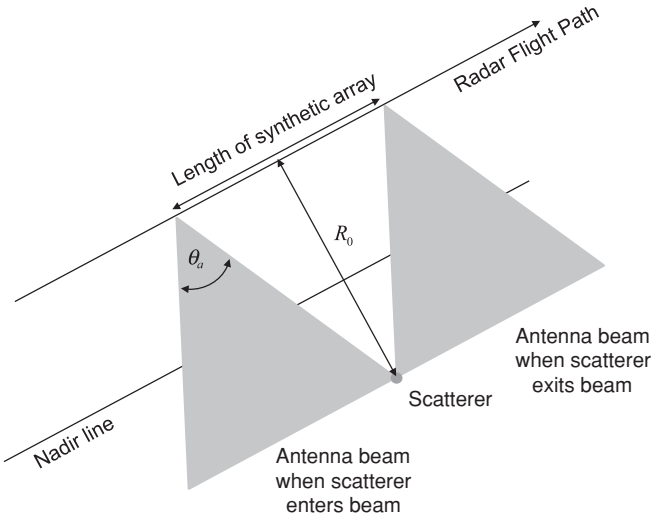

FIGURE 1-7 The synthetic aperture radar integrates the signal from the scatter for as long as the scatterer remains in the antenna beam

and the scatterer changes, with the minimum distance occurring when the scatterer is directly broadside of the radar platform. The phase of the radar signal is given by $-4\pi R(s)/\lambda$. The changing distance between the radar and the scatterer means that after range compression, the phase of the signal will be different for the different positions along the flight path.

This changing distance can be written as

$$R(s) = \sqrt{R_0^2 + v^2 s^2} \tag{1.5-1}$$

where R_0 is the range at closest approach to the scatterer, v is the velocity of the radar platform, and s is the time along the flight path (so-called slow time) with zero time at the time of closest approach. To a good approximation for remote-sensing radars, we can assume that $vs << R_0$. (This might not be true for very-high-resolution radars; however, the basic principle remains the same.) In this case, we can approximate the range as a function of slow time as

$$R(s) \approx R_0 + \frac{v^2}{2R_0} s^2 \tag{1.5-2}$$

The phase of the signal after range compression, as shown in the second exponential in Eq. (1.2-9), then becomes

$$\phi(s) = -\frac{4\pi R(s)}{\lambda} \approx -\frac{4\pi R_0}{\lambda} - \frac{2\pi v^2}{R_0 \lambda} s^2 \tag{1.5-3}$$

The instantaneous frequency of this signal is

$$f(s) = \frac{1}{2\pi}\frac{\partial\phi(s)}{\partial s} = -\frac{2v^2}{R_0\lambda}\, s \tag{1.5-4}$$

This is the expression of a linear frequency chirp. To find the bandwidth of this signal, we have to find the maximum time that we can use in the signal integration. This maximum "integration time" is given by the amount of time that the scatterer will be in the antenna beam. For an antenna with a physical length L, the half-power horizontal beam width is $\theta_a = \lambda/L$, so that the scatterer at the range of closest approach R_0 is illuminated for a time

$$s_{tot} = \frac{\lambda R_0}{Lv} \tag{1.5-5}$$

Half of this time occurs when the radar is approaching the range of closest approach and half of it is spent traveling away from the range of closest approach. Therefore, the bandwidth of the signal shown in Eq. (1.5-4), which is the Doppler bandwidth of the SAR signal, is

$$B_D = \frac{2v}{L} \tag{1.5-6}$$

If this signal is filtered with a matched filter, as described earlier under signal modulation, the resulting compressed signal will have a width in time of $1/B_D$. Since the radar platform moves at a speed of v, this leads to an along-track resolution of

$$\Delta_a = \frac{v}{B_D} = \frac{L}{2} \tag{1.5-7}$$

This result shows that the azimuth (or along track) surface resolution for a synthetic aperture radar is equal to half the size of the physical antenna and is independent of the distance between the sensor and the surface! At first glance, this result seems most unusual: It shows that a smaller antenna gives better resolution. This can be explained in the following way: The smaller the physical antenna is, the larger its footprint. This allows a longer observation time for each point on the surface (i.e., a longer array can be synthesized). This longer synthetic array allows a larger Doppler bandwidth and, hence, a finer surface resolution. Similarly, if the range between the sensor and surface increases, the physical footprint increases, leading to a longer observation time and larger Doppler bandwidth, which counterbalances the increase in the range.

As mentioned earlier, the imaging radar transmits a series of pulsed electromagnetic waves. Thus the Doppler history from a scatterer is not measured continuously but sampled on a repetitive basis. To get an accurate record of the Doppler history, the Nyquist sampling criterion requires that sampling occurs at least at twice the highest

frequency in the Doppler bandwidth. Thus the pulse repetition frequency, usually called PRF, must be larger than

$$PRF \geq 2\left(\frac{B_D}{2}\right) = \frac{2v}{L} \tag{1.5-8}$$

Note that we used half the Doppler bandwidth as the highest frequency in the Doppler signal in Eq. (1.5-8). The reason for this is that the Doppler frequency varies linearly from $-B_D/2$ to $+B_D/2$. Therefore, even though the total bandwidth of the signal is B_D, the highest frequency in the bandwidth is only $B_D/2$.

Equation (1.5-8) means that at least one sample (i.e., one pulse) should be taken every time the sensor moves by half an antenna length. As an example, for a spaceborne imaging system moving at a speed of 7 km/s and using an antenna 10 m in length, the corresponding minimum PRF is 1.4 kHz. As we will see in the next section, the requirement to cover a certain swath size provides an upper bound on the PRF. Interpreted in a different way, the requirement to adequately sample the signal bandwidth limits the size of the swath that could be imaged.

1.6 RADAR IMAGE ARTIFACTS AND NOISE

Radar images could contain a number of anomalies that result from the way imaging radars generate the image. Some of these are similar to what is encountered in optical systems, such as blurring due to defocusing or scene motion, and some (such as range and azimuth ambiguities) are unique to radar systems. This section summarizes the anomalies that are most commonly encountered in radar images.

1.6.1 Range and Azimuth Ambiguities

As mentioned earlier (see Figure 1-2), a radar images a surface by recording the echoes line by line with successive pulses. The leading edge of each echo corresponds to the near edge of the image scene; the tail end of the echo corresponds to the far edge of the scene. The length of the echo (i.e., swath width of the scene imaged) is determined by the antenna beam width or the size of the data window used in the recording of the signal. The exact timing of the echo reception depends on the range between the sensor and the surface being imaged. If the timing of the pulses or the extent of the echoes is such that the leading edge of one echo overlaps with the tail end of the previous one, then the far edge of the scene is folded over the near edge of the scene. This is called range ambiguity. The temporal extent of the echo is equal to:

$$T_e \approx 2\frac{R}{c}\theta_r \tan\theta = 2\frac{h\lambda}{cW}\frac{\sin\theta}{\cos^2\theta} \tag{1.6-1}$$

To avoid overlapping echoes, this time extent should be shorter than the time separating two pulses (i.e., 1/PRF). Thus, we must have

$$PRF < \frac{cW}{2h\lambda}\frac{\cos^2\theta}{\sin\theta} \tag{1.6-2}$$

In addition, the sensor parameters, specifically the PRF, should be selected such that the echo is completely within an interpulse period (i.e., no echoes should be received during the time that a pulse is being transmitted). The above equation gives an upper limit for the PRF as mentioned before. The SAR designers have to trade off system parameters to maximize the swath, while at the same time transmitting a high enough PRF to adequately sample the signal Doppler spectrum.

Another kind of ambiguity present in SAR imagery also results from the fact that the target's return in the azimuth direction is sampled at the PRF. This means that the azimuth spectrum of the target return repeats itself in the frequency domain at multiples of the PRF. In general, the azimuth spectrum is not a band-limited signal; instead, the spectrum is weighted by the antenna pattern in the azimuth direction. This means that parts of the azimuth spectrum might be aliased and high-frequency data will actually appear in the low-frequency part of the spectrum. In actual images, these azimuth ambiguities appear as ghost images of a target repeated at some distance in the azimuth direction, as shown in Figure 1-8. To reduce the azimuth ambiguities, the PRF of a SAR has to exceed the lower limit given by Eq. (1.5-8).

To reduce both range and azimuth ambiguities, the PRF must therefore satisfy both the conditions expressed by Eq. (1.5-8) and Eq. (1.6-2). Therefore, we must insist that

$$\frac{cW}{2h\lambda}\frac{\cos^2\theta}{\sin\theta} > \frac{2v}{L} \tag{1.6-3}$$

from which we derive a lower limit for the antenna size as

$$LW > \frac{4vh\lambda}{c}\frac{\sin\theta}{\cos^2\theta} \tag{1.6-4}$$

A word of caution about the use of Eq. (1.6-4): This expression is derived assuming that the SAR processor uses the full Doppler bandwidth in the processing, and a swath that covers the full antenna beam width is imaged. This might not always be the case. For many reasons, SAR images are sometimes processed to only a fraction of the achievable resolution, or the swath width might be limited artificially by recording only that section of the returned echo that falls in a so-called data window. When either of these conditions is used, it is not appropriate to limit the antenna size as given by Eq. (1.6-4). In fact, in the case where both the swath is artificially limited and the resolution is decreased, antennas significantly smaller that that given by Eq. (1.6-4) may be used with perfectly good results.

Another type of artifact in radar images results when a very bright surface target is surrounded by a dark area. As the image is being formed, some spillover from

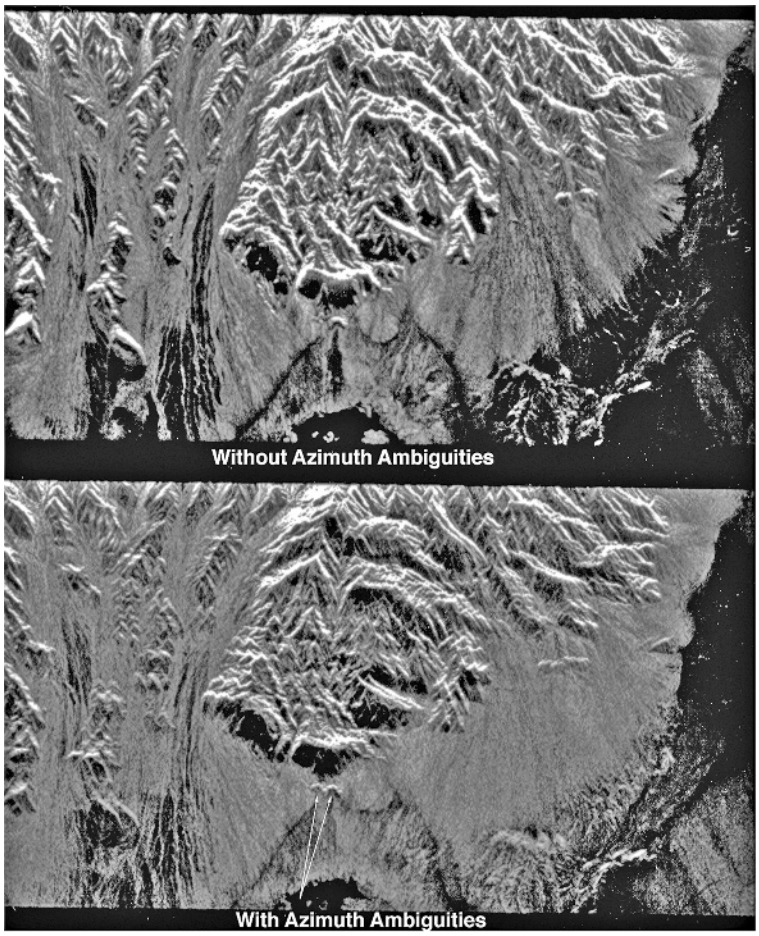

FIGURE 1-8 Azimuth ambiguities result when the radar pulse repetition frequency is too low to sample the azimuth spectrum of the data adequately. In this case, the edges of the azimuth spectrum fold over themselves, creating ghost images, as shown in this figure

the bright target, although weak, could exceed the background and become visible. It should be pointed out that this type of artifact is not unique to radar systems. These artifacts are common in optical systems, where they are known as the side lobes of the point spread function. The difference is that in optical systems the side lobe characteristics are determined by the characteristics of the imaging optics (i.e., the hardware). Conversely, in the case of a SAR, the side lobe characteristics are determined mainly by the characteristics of the processing filters. In the radar case, the side lobes might, therefore, be reduced by suitable weighting of the signal spectra during matched filter compression. The equivalent procedure in optical systems is called apodization of the telescope aperture.

The vast majority of these artifacts and ambiguities can be avoided with proper selection of the sensor and processor parameters. However, the interpreter should be aware of their occurrence, because in some situations they might be difficult, if not impossible, to suppress.

1.6.2 Geometric Effects and Projections

The time delay/Doppler history basis of SAR image generation leads to an image projection different from that in the case of optical sensors. Even though at first look radar images seem very similar to optical images, close examination quickly shows that geometric shapes and patterns are projected in a different fashion by the two sensors. This difference is particularly acute in rugged terrain. If the topography is known, a radar image can be re-projected into a format identical to an optical image, thus allowing image pixel registration. In extremely rugged terrain, however, the nature of the radar image projection leads to distortions that sometimes cannot be corrected without knowledge of the terrain elevations.

In the radar image, two neighboring pixels in the range dimension correspond to two areas in the scene with slightly different range to the sensor. This has the effect of projecting the scene in a cylindrical geometry on the image plane, which leads to distortions (as shown in Figure 1-9). Areas that slope toward the sensor look shorter in the image, while areas that slope away from the sensor look longer in the image than horizontal areas. This effect is called foreshortening. In the extreme case where the slope is larger than the incidence angle, layover occurs. In this case, a hill would look as if it is projected over the region in front of it. Layover cannot be corrected and can only be avoided by having an incidence angle at the surface larger than any expected surface slopes. When the slope facing away from the radar is steep enough such that the radar waves do not illuminate it, shadowing occurs and the area on that slope is not imaged. Note that in the radar images, shadowing is always away from

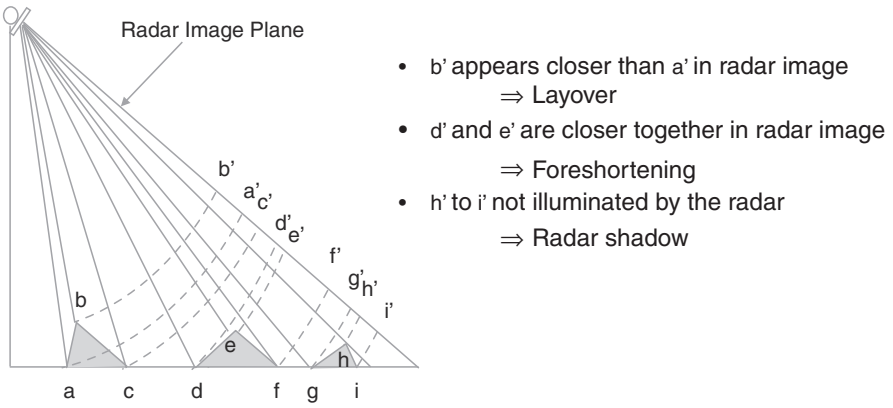

FIGURE 1-9 Radar images are cylindrical projections of the scene onto the image plane, leading to characteristic distortions

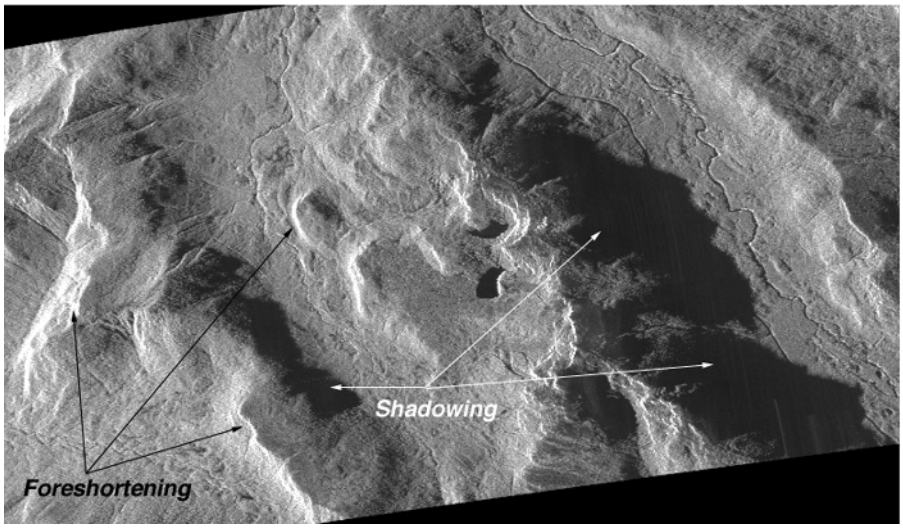

FIGURE 1-10 This NASA/JPL AIRSAR image shows examples of foreshortening and shadowing. Note that since the radar provides its own illumination, radar shadowing is a function of the radar look direction relative to the terrain and does not depend on the Sun angle. This image was illuminated from the left. (Image courtesy of Jet Propulsion Laboratory.)

the sensor flight line and is not dependent on the time of data acquisition or the Sun angle in the sky. As in the case of optical images, shadowing can be beneficial for highlighting surface morphological patterns. Figure 1-10 contains some examples of foreshortening and shadowing.

1.6.3 Signal Fading and Speckle

A close examination of a synthetic-aperture radar image shows that the brightness variation is not smooth but instead has a granular texture that is called speckle. Even for an imaged scene that has a constant backscatter property, the image will have statistical variations of the brightness on a pixel-by-pixel basis, but with a constant mean over many pixels. This effect is identical to when a scene is observed optically under laser illumination. It is a result of the coherent nature (or very narrow spectral width) of the illuminating signal.

Rigorous mathematical analysis shows that the noiselike radar signal has well-defined statistical properties. The measured signal amplitude has a Rayleigh distribution, and the signal power has an exponential distribution [1]. To narrow the width of these distributions (i.e., reduce the brightness fluctuations), successive signals or neighboring pixels can be averaged incoherently (i.e., their power values are added). This would lead to a more accurate radiometric measurement (and a more pleasing image) at the expense of degradation in the image resolution.

Another approach to reducing speckle is to combine images acquired at neighboring frequencies. In this case, the exact interference patterns lead to independent

signals but with the same statistical properties. Incoherent averaging would then result in a smoothing effect. In fact, this is why a scene illuminated with white light does not show speckled image behavior.

In most imaging SARs, the smoothing is done by averaging the brightness of neighboring pixels in azimuth or range, or both. The number of statistically independent pixels averaged is called the number of looks N. We can show that the signal standard deviation S_N is related to the mean signal power $\bar{P}$ by:

$$S_N = \frac{1}{\sqrt{N}}\bar{P} \tag{1.6-5}$$

The larger the number of looks N, the better the quality of the image from the radiometric point of view. However, this degrades the spatial resolution of the image. Note that for N larger than about 25, large increase in N leads to only a small decrease in the signal fluctuation. This small improvement in the radiometric resolution should be traded off against the large degradation in the spatial resolution. For example, if one were to average 10 resolution cells in a four-look image, the speckle noise will be reduced to about 0.5 dB. At the same time, however, the image resolution will be reduced by an order of magnitude. Whether the reduction in speckle noise is worth this loss in resolution depends on both the aim of the investigation and the kind of scene imaged.

Figure 1-11 shows the effect of multi-look averaging. An image acquired by the NASA/JPL AIRSAR system is shown displayed at one, four, 16, and 32 looks, respectively. This figure clearly illustrates the smoothing effect, as well as the decrease in resolution resulting from the multi-look process. In one early survey of geologists,

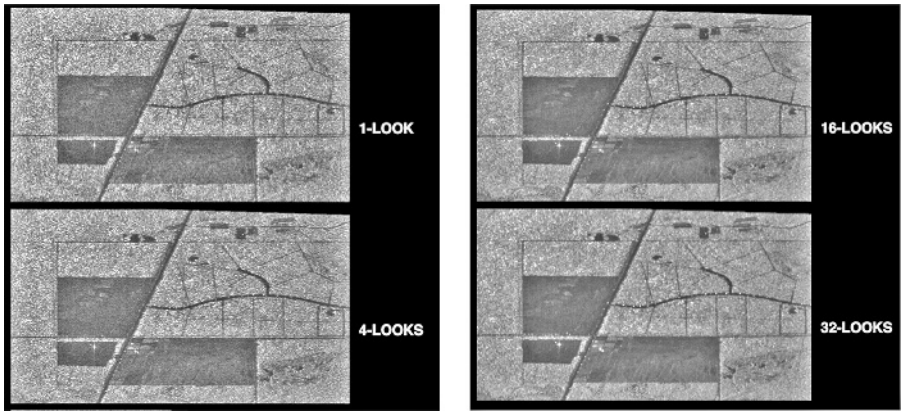

FIGURE 1-11 The effects of speckle can be reduced by incoherently averaging pixels in a radar image, a process known as multi-looking. Shown in this image is the same image, processed at 1 look, 4 looks, 16 looks, and 32 looks. Note the reduction in the granular texture as the number of looks increase, while at the same time the resolution of the image decreases. Some features, such as those in the largest dark patch, might be completely masked by the speckle noise. (Image courtesy of the Jet Propulsion Laboratory.)

the results showed that even though the optimum number of looks depended on the scene type and resolution, the majority of the responses preferred two-look images. However, this survey dealt with images that had rather poor resolution to begin with. One may well find that with today's higher-resolution systems, analysts might be asking for a larger number of looks.

1.7 SUMMARY

In this chapter we have introduced some of the terms required to understand SAR imaging as described in the rest of this text. For a more in-depth treatment of SAR imaging and system engineering, the reader is referred to the two texts in the bibliography [2,3]. In the rest of this text, we shall concentrate specifically on polarimetric SAR and its applications.

REFERENCES

1. F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing Active and Passive*, Vol. II, pp. 583–595, Artech House, Norwood, Massachusetts, 1986.
2. J. C. Curlander and R. N. McDonough, *Synthetic Aperture Radar System and Signal Processing*, John Wiley & Sons, New York, 1991.
3. C. Elachi, *Spaceborne Radar Remote Sensing: Applications and Techniques*, IEEE Press, New York, 1988.