Synthetic Aperture Radar

Jakob J. van Zyl
Yunjin Kim
Jet Propulsion Laboratory
California Institute of Technology

GLOSSARY

Antenna A device that is designed to radiate or receive electromagnetic waves.

Doppler frequency The change of frequency of a received signal due to the relative velocity of a transmitting antenna with respect to an illuminated object.

Polarization The property of a radiated electromagnetic wave describing the time-varying direction and amplitude of the electric field vector.

Radar An acronym for Radio Detection and Ranging. A radar system measures the distance to an object by transmitting an electromagnetic signal and receiving an echo reflected from the object.

Radar interferometer A radar system that determines the angle of arrival of a radar signal by phase comparison of the signals received at separate antennas.

Synthetic aperture radar A radar system that synthesizes the effect of a long antenna through the motion of a small antenna relative to the target.

THE TERM "RADAR" is an acronym for Radio Detection and Ranging. A radar measures the distance, or *range*, to an object by transmitting an electromagnetic signal 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 have measured the propagation time (t), we can easily calculate the range (ρ) as

$$\rho = \frac{1}{2}ct, \tag{1}$$

where c is the speed of light. The factor $1/2$ accounts for the fact that the radar signal actually traveled twice the distance; first from the radar to the object and then from the object to the radar.

Radars provide their own signals to detect the presence of objects. Therefore, radars are known as *active* remote sensing instruments. Because the radar provides its own signal, it can operate during day or night. In addition, radar signals 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 this reason, radars are often referred to as *all-weather* instruments. Imaging remote sensing radars such as synthetic aperture radars 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 post-processing techniques.

This article focuses on a specific class of implementation of radar known as *synthetic aperture radar*, or SAR, with particular emphasis on spaceborne SAR. As mentioned above, SAR is a way to achieve high-resolution *images* using radio waves. Here, we shall first describe the basics of radar imaging and follow that with a description of the synthetic aperture principle. We will then look at some advanced SAR implementations such as SAR polarimetry and SAR interferometry. We will also briefly discuss some examples of civilian spaceborne SAR missions.

I. 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, like that we are all familiar with when 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 Fig. 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 is, the higher the resolution and the smaller the details that can be recognized in the images. As the camera moves farther 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

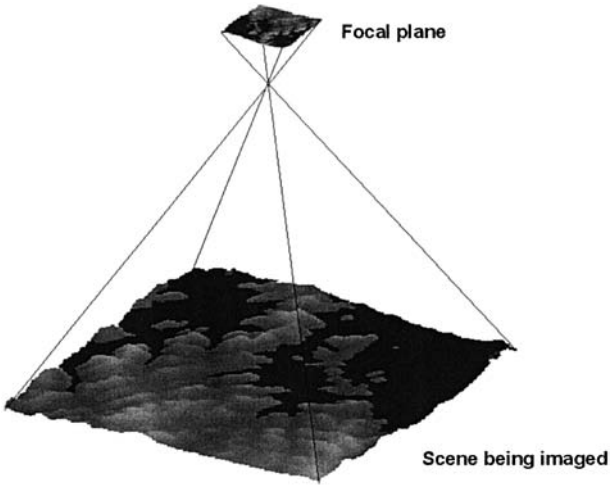

FIGURE 1 Passive imaging systems in the visible and infrared part of the electromagnetic spectrum conserve the angular relationships between objects in the scene and their images in the focal plane of the instrument.

are also quite different from that of visible and infrared images. There are two different ways radars can be used to produce images. These 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 article. In order to separate objects in radar images in the cross-track direction and the along-track direction, 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 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 SAR) 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

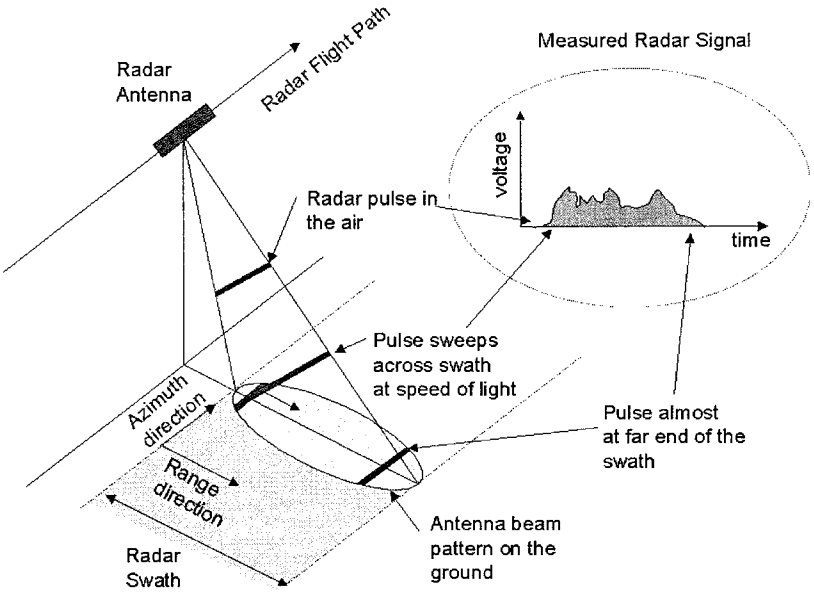

FIGURE 2 Imaging geometry for a side-looking radar system.

down or at least have no fundamental limitation that prevents them from taking pictures looking straight down from the spacecraft or aircraft.This is not so for imaging radars. To avoid so-called *ambiguities*, which we will discuss in more detail later, the imaging radar sensor has to use an antenna which illuminates the surface to one side of the flight track. Usually, the antenna has a fan beam which illuminates a highly elongated elliptical-shaped area on the surface, as shown in Fig. 2. The illuminated area across track 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 a proportionally later time (see Fig. 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, thus allowing strip mapping of the surface line by line. This is comparable to strip mapping with a so-called pushbroom 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 SARs 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 Fig. 2 in the way that pulses propagate. Slant ranges are measured directly from the radar. 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.

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, while the *incidence angle* is defined as the angle between the vertical direction and the radar wave propagation vector at the surface. 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, which 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 may vary in the radar image from pixel to pixel.

Consider the simple case of a single hill illuminated by a radar system as shown in Fig. 3. Also shown is the

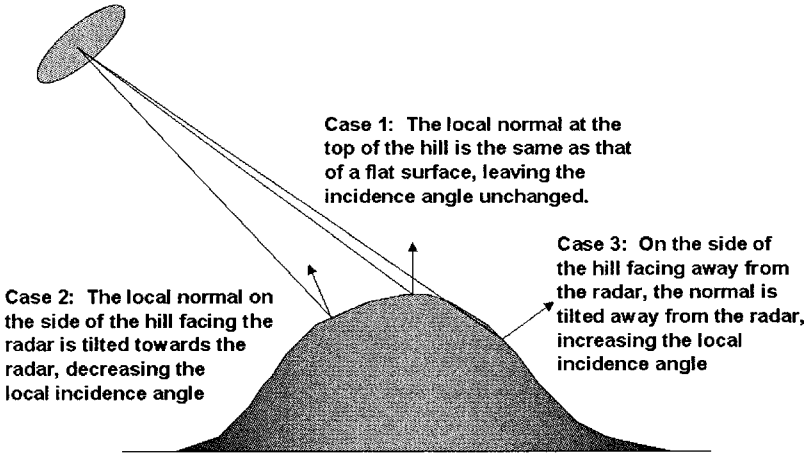

FIGURE 3 Topographic variations in the image will cause the local incidence angle to be different from that expected for a flat surface with no relief.

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 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 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 article, we shall use look angle and incidence angle to describe the imaging geometry.

A. Radar Resolution

The *resolution* of an image is defined as the 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

$$\Delta t = 2x_r/c, \tag{2}$$

where c is the speed of light and the factor 2 is included to account for the signal round trip propagation as described before. Radar waves are usually not transmitted continuously; instead, radar usually transmits short bursts of energy known as radar *pulses*. 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, as shown in Fig. 4. 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

$$2x_r/c = \tau \Rightarrow x_r = \frac{c\tau}{2}. \tag{3}$$

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

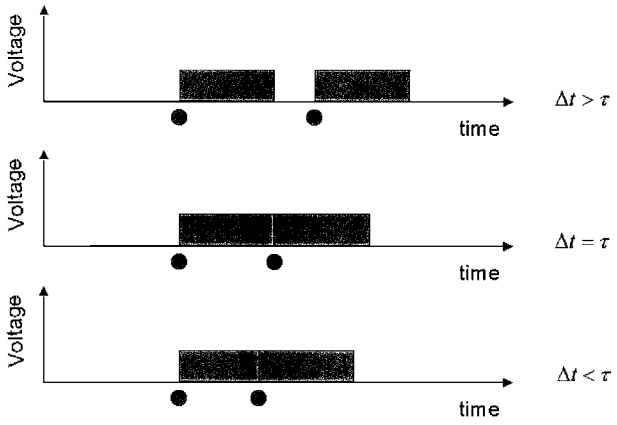

FIGURE 4 If the radar echoes from two point targets are separated in time by more than or equal to the length of the radar pulse, it is possible to recognize the echoes as those from two different scatterers, as shown in the top two panels. If the time difference between the echoes is less than the radar pulse length, it is not possible to recognize two distinct scatters, as in the case of the bottom panel.

$$\Delta t = 2x_r \sin\theta / c. \tag{4}$$

The angle θ in Eq. (4) 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

$$2x_r \sin\theta / c = \tau \Rightarrow x_r = \frac{c\tau}{2\sin\theta}. \tag{5}$$

In other words, the range resolution is equal to half the footprint of the radar pulse on the surface.

Sometimes 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,

$$\tau = 1/B. \tag{6}$$

The $\sin\theta$ term in the denominator of Eq. (5) means that the ground range resolution of an imaging radar will be a function of the incidence angle.

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, particularly in the case of spaceborne sensors, the peak power is usually strongly limited by the available power sources. On the other hand, an increased pulse length leads to a poorer range resolution [see Eq. (5)]. 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 such modulation scheme is the linear frequency modulation or *chirp*.

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

$$B = |(f_0 + \Delta f) - f_0| = |\Delta f|, \tag{7}$$

which is independent of the pulse length. 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

$$f(t) = f_0 + \frac{B}{\tau'}t \quad \text{for} \quad -\tau'/2 \le t \le \tau'/2, \tag{8}$$

and the corresponding signal amplitude is

$$A(t) \sim \Re\left\{\exp\left[-i2\pi \int f(t)\,dt\right]\right\} = \cos\left[2\pi\left(f_0 t + \frac{B}{2\tau'}t^2\right)\right], \tag{9}$$

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 that shown in Eq. (9) has a physical pulse length τ' and a bandwidth B. The product $\tau' B$ is known as the *time bandwidth product* of the radar system. In typical radar systems, time bandwidth products of several hundred are used.

At first glance it may seem that using a pulse of the form in Eq. (9) 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, which 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. (8), 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. Suppose we transmit a signal of the form described in Eq. (9). The signal received from a single point scatterer at a range ρ is a scaled replica of the transmitted signal delayed by a time $t = 2\rho/c$. The output of the matched filter for a 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,

$$V_0(t) = \tau' E_r \exp(iwt)\exp(-i4\pi\rho/\lambda) \times \frac{\sin(\pi B(t - 2\rho/c))}{\pi B(t - 2\rho/c)}. \tag{10}$$

This compressed pulse has a half power width of $1/B$, and its peak position occurs at time $2\rho/c$. Therefore, the achievable range resolution using a modulated pulse of the kind given by Eq. (9) 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 the visible and infrared

part of the spectrum. 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.

B. Radar Equation

One of the key factors that determine the quality of the radar imagery is the corresponding *signal-to-noise* ratio, 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 Fig. 5.

The total received power is equal to the power intercepted by the receiving antenna and is given by

$$P_r = \frac{P_t G_t}{4\pi\rho^2} s\sigma_0 \frac{\lambda^2 G_r}{(4\pi\rho)^2}, \tag{11}$$

where G_t is the gain of the transmitting antenna, G_r is the gain of the receiving antenna, and λ is the wavelength of the radar signal. σ_0 is the surface *backscattering cross section* which represents the efficiency of the surface in re-emitting back toward the sensor some of the energy incident on it. It is similar to the surface albedo at visible wavelengths, except that it is designed to scatter the energy over a sphere rather than a hemisphere as in the case of the albedo. Here, the area responsible for reflecting the scattered power is denoted by s. In Eq. (13) we explicitly show that the transmit and receive antennas may have different gains. This is important for the more advanced SAR techniques, like polarimetry, where antennas with different polarizations may be used during transmission and reception.

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

$$P_N = kTB, \tag{12}$$

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

$$\text{SNR} = P_r / P_N. \tag{13}$$

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.

Typical spaceborne SAR frequencies are shown in Table I. Notice that each frequency band has the frequency range allocated for remote sensing.

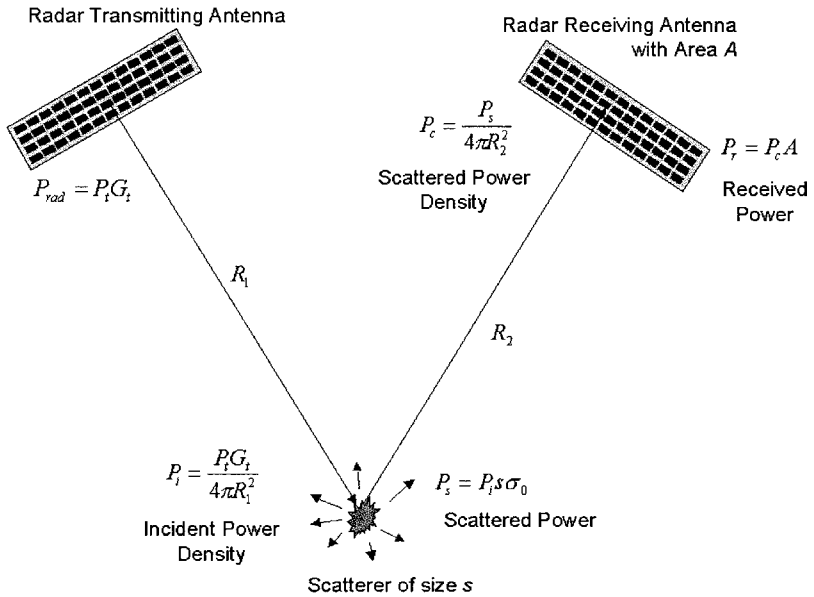

FIGURE 5 Schematic of the derivation of the radar equation.

TABLE I Frequency Allocation at Typical SAR Frequency Bands

Frequency band	Allocated frequency range for remote sensing (MHz)
L-band	1215–1300
S-band	3100–3300
C-band	5250–5460
X-band	8550–8650
	9500–9800

II. REAL AND SYNTHETIC APERTURE RADAR

A. Real Aperture Radar

The real aperture imaging radar sensor also uses an antenna which illuminates the surface to one side of the flight track. As mentioned before, the antenna usually has a fan beam which illuminates a highly elongated elliptical-shaped area on the surface, as shown in Fig. 2. As shown in Fig. 2, the illuminated area across the 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

$$\theta_r \approx \lambda / W, \tag{14}$$

and the resulting surface footprint or swath S is given by

$$S \approx \frac{h\theta_r}{\cos^2\theta} = \frac{\lambda h}{W\cos^2\theta}, \tag{15}$$

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. (15) ignores the curvature of the earth. For spaceborne radars, this effect should not be ignored, especially if the antenna beamwidth is large. In that case, 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 driven by the size of the antenna as well as the range to the scene. Assuming an antenna length of L, the antenna beamwidth in the along-track direction is

$$\theta_a \approx \frac{\lambda}{L}. \tag{16}$$

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

$$x_a \approx \rho\theta_a \approx \frac{\lambda\rho}{L} \approx \frac{\lambda h}{L\cos\theta}. \tag{17}$$

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 is 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 (about 1 μm), 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 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 by the projection of the pulse on the ground in the cross-track direction. This is shown in Fig. 2 for the pulses in the radar swath. The along-track dimension of the antenna pattern is given by Eq. (17). If the pulse has a length τ 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

$$l_g = \frac{c\tau}{2\sin\theta_i}. \tag{18}$$

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

$$P_r = \frac{P_t G_t G_r \lambda^2}{(4\pi)^3\rho^4}\frac{\lambda\rho}{L}\frac{c\tau}{2\sin\theta_i}\sigma_0. \tag{19}$$

This shows that for the real aperture radar, the received power decreases as the range to the third power. In terms of the physical antenna sizes, we can rewrite this expression as

$$P_r = \frac{P_t W^2 L c\tau\sigma_0}{8\pi\lambda\rho^3\sin\theta_i}. \tag{20}$$

This is the radar equation for a so-called distributed target for the real aperture radar case. From Eq. (20) 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 increases 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.

B. Synthetic Aperture Radar (SAR)

SAR 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 SAR, 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 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 techniques was originally known as Doppler beam sharpening, but later became known as 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 case.

As the radar moves along the flight path, it transmits pulses of energy and records the reflected signals, as shown in Fig. 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 Fig. 6. As the radar moves along the flight path, the distance between the radar 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 $-4\pi\rho/\lambda$. 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.

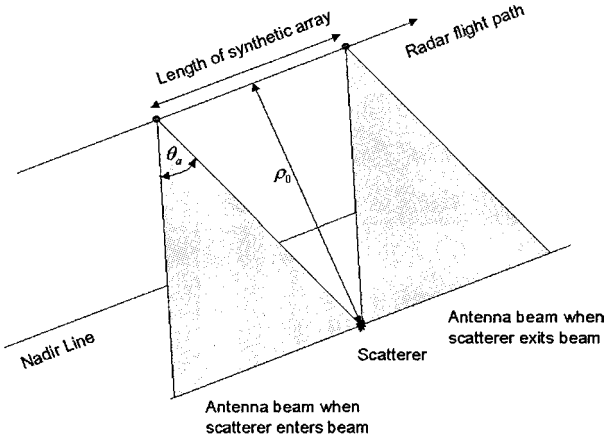

FIGURE 6 The SAR integrates the signal from the scatter for as long as the scatterer remains in the antenna beam.

The range between the radar and the scatterer as a function of position along the flight path is given by

$$\rho(s) = \sqrt{\rho_0^2 + v^2 s^2}, \tag{21}$$

where ρ_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 $vs \ll \rho_0$. (This may not be true for very high resolution radars, but the basic principle remains the same.) In this case, we can approximate the range as a function of slow time as

$$\rho(s) \approx \rho_0 + \frac{v^2}{2\rho_0} s^2. \tag{22}$$

The phase of the signal after range compression as shown in Eq. (10) then becomes

$$\phi(s) = -\frac{4\pi\rho(s)}{\lambda} \approx -\frac{4\pi\rho_0}{\lambda} - \frac{2\pi v^2}{\rho_0\lambda} s^2. \tag{23}$$

The instantaneous frequency of this signal is

$$f(s) = \frac{1}{2\pi}\frac{\partial\phi(s)}{\partial s} = -\frac{2v^2}{\rho_0\lambda} s. \tag{24}$$

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 horizontal beamwidth is $\theta_a = \lambda/L$, so that the scatterer at the range of closest approach ρ_0 is illuminated for a time

$$s_{tot} = \frac{\lambda\rho_0}{Lv}. \tag{25}$$

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. (24), which is the Doppler bandwidth of the SAR signal, is

$$B_D = \frac{2v}{L}. \tag{26}$$

If this signal is filtered using 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

$$x_a = \frac{v}{B_D} = \frac{L}{2}. \tag{27}$$

This result shows that the azimuth (or along-track) surface resolution for an SAR 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 is sampled on a repetitive basis. In order 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

$$\text{PRF} \geq 2B_D = \frac{2v}{L}. \tag{28}$$

Note that we used half the Doppler bandwidth as the highest frequency in the Doppler bandwidth in Eq. (28). 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 (28) 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/sec 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.

III. RADAR IMAGE ARTIFACTS AND NOISE

Radar images could contain a number of anomalies which 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 addresses the anomalies which are most commonly encountered in radar images.

A. Range and Azimuth Ambiguities

As mentioned earlier (see Fig. 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, and 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 covered) is determined by the antenna beamwidth 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 echo are 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

$$T_e \approx 2\frac{\rho}{c}\theta_r \tan\theta = 2\frac{h\lambda}{cW}\frac{\sin\theta}{\cos^2\theta}. \tag{29}$$

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

$$\text{PRF} < \frac{cW}{2h\lambda}\frac{\cos^2\theta}{\sin\theta}. \tag{30}$$

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 designer has 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 may 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 Fig. 7. To reduce the azimuth ambiguities, the PRF of an SAR has to exceed the lower limit given by Eq. (28).

In order to reduce both range and azimuth ambiguities, the PRF must therefore satisfy both the conditions expressed by Eqs. (28) and (30). Therefore, we must insist that

$$\frac{cW}{2h\lambda}\frac{\cos^2\theta}{\sin\theta} > \frac{2v}{L}, \tag{31}$$

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

$$LW > \frac{4vh\lambda}{c}\frac{\sin\theta}{\cos^2\theta}. \tag{32}$$

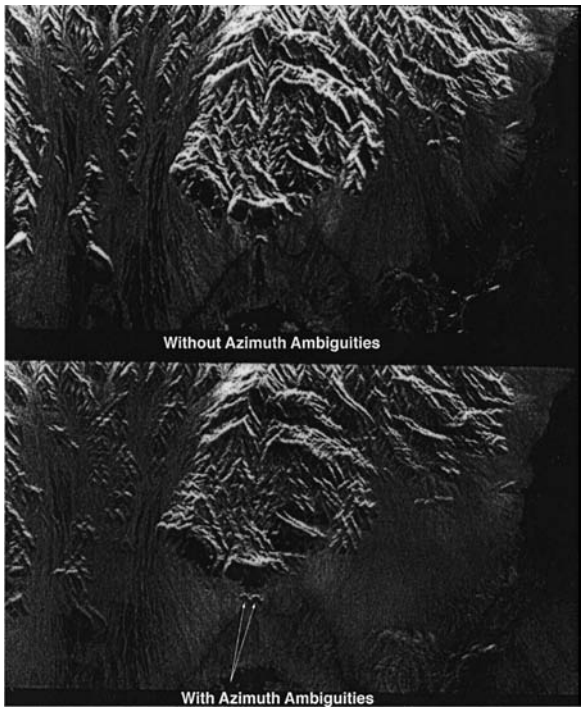

FIGURE 7 Azimuth ambiguities result when the radar pulse repetition frequency is too low to sample the azimuth spectrum of the data adequately. In this case, the edges of the azimuth spectrum fold over themselves, creating ghost images.

A word of caution about the use of Eq. (32). This expression is derived assuming the SAR processor uses the full Doppler bandwidth in the processing, and also a swath that covers the full antenna beam width is imaged. This may 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 may 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 are used, it is not appropriate to limit the antenna size as given by Eq. (32). In fact, in the case where both the swath is artificially limited and the resolution is increased, antennas significantly smaller that that given by Eq. (32) 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 the bright target, called side lobes, 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. They 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, whereas in the case of an SAR, the side lobe characteristics are determined mainly by the characteristics of the processing filters. In the radar case, the side lobes may be therefore reduced by suitable weighting of the signal spectra during matched filter compression. The equivalent procedure in optical systems is through 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.

B. Geometric Effects and Projections

The time delay/Doppler history basis of SAR image generation leads to an image projection different than 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 reprojected 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 which sometimes cannot be corrected.

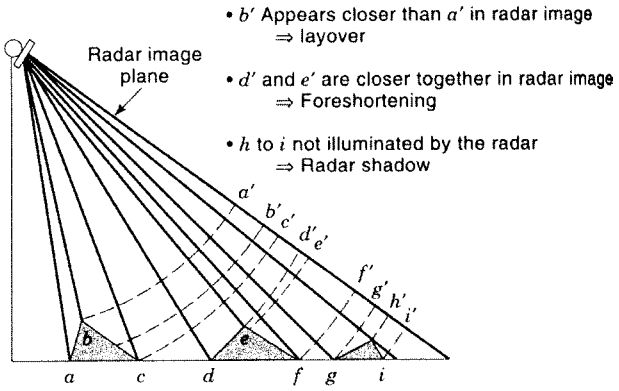

FIGURE 8 Radar images are cylindrical projections of the scene onto the image plane, leading to characteristic distortions.

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 Fig. 8. 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 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 morphologic patterns. Figure 9 contains some examples of foreshortening and shadowing.

C. Signal Fading and Speckle

A close examination of an SAR image shows that the brightness variation is not smooth, but has a granular texture which is called *speckle*. Even for an imaged scene which 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 noise-like radar signal has well-defined statistical properties.

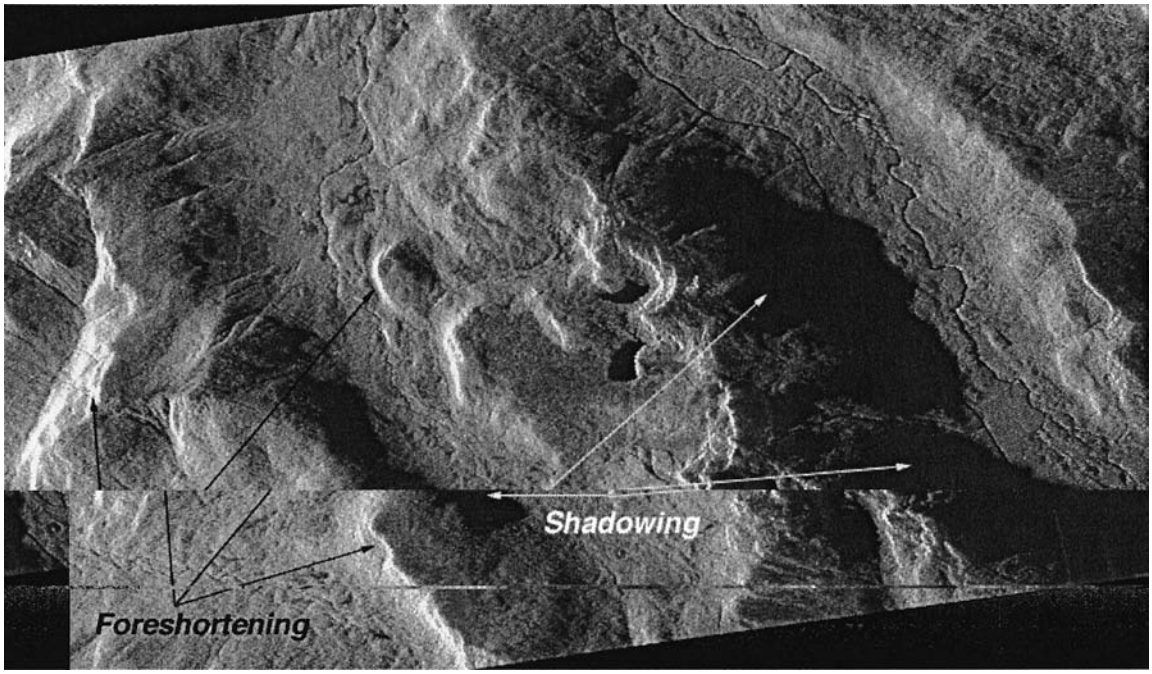

FIGURE 9 This NASA/JPL AIRSAR image shows examples of foreshortening and shadowing. Note that since the radar provides its own illumination, radar shadowing is a function of the radar look direction relative to the terrain and does not depend on the sun angle. This image was illuminated from the left.

The measured signal amplitude has a Rayleigh distribution, and the signal power has an exponential distribution. In order 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 reduce 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 the reason 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 pixels averaged is called the number of looks N. It can be shown that the signal standard deviation S_N is related to the mean signal power $\bar{P}$ by

$$S_N = \frac{1}{\sqrt{N}} \bar{P}. \tag{33}$$

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. It should be noted that for N larger than about 25, a 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 increase in the spatial resolution. For example, if one were to average 10 resolution cells in a 4-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 this loss in resolution is worth the reduction in speckle noise depends on both the aim of the investigation, and the kind of scene imaged.

Figure 10 shows the effect of multilook averaging. An image acquired by the NASA/JPL AIRSAR system is shown displayed at 1, 4, 16, and 32 looks, respectively. Figure 10 clearly illustrates the smoothing effect, as well as the decrease in resolution resulting from the multilook process. In one early survey of geologists, the results showed that even though the optimum number of looks depended on the scene type and resolution, the majority of the responses preferred 2-look images. However, this survey dealt with images that had rather poor resolution to begin with, and one may well find that with today's higher resolution systems, analysts may be asking for a larger number of looks.

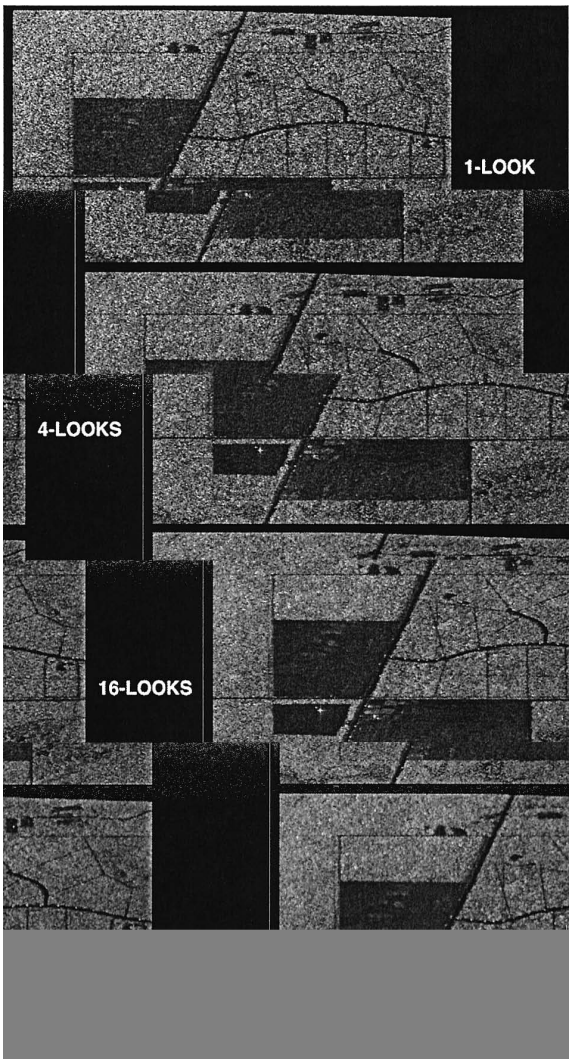

FIGURE 10 The effects of speckle can be reduced by incoherently averaging pixels in a radar image, a process known as multilooking. Shown in this image is the same image processed at 1 look, 4 looks, 16 looks, and 32 looks. Note the reduction in the granular texture as the number of looks increase, while at the same time the resolution of the image decreases. Some features, such as those in the largest dark patch, may be completely masked by the spackle noise.

IV. SAR POLARIMETRY AND INTERFEROMETRY

The field of SAR has advanced dramatically over the past two decades with the introduction of SAR polarimetry and SAR interferometry. Many airborne and spaceborne SAR instruments were developed to demonstrate various applications of both techniques. As an example, the NASA/JPL AIRSAR system has collected both polarimetric and interferometric SAR data as an airborne research tool from the 1980s. Two Shuttle Imaging Radar-C

flights in 1994 were the first spaceborne polarimetric SAR at L- and C-band. The Shuttle Radar Topographic Mission (SRTM) launched in 2000 was the first spaceborne implementation of a single pass SAR interferometer to produce a three-dimensional map of 80% of the earth's land surface. One of the most exciting applications of SAR interferometry is differential interferometry that measures minute surface changes (millimeter to centimeter scale) by subtracting two interferometric pairs separated in time.

A. SAR Polarimetry

Since an electromagnetic wave is polarized, electromagnetic wave propagation is a vector phenomenon. Both electric and magnetic fields associated with an electromagnetic wave are perpendicular to the propagation direction. Hence, we can decompose the wave polarization into a vector with two components perpendicular to each other. Mathematically, the electric field ($\vec{E}(t)$) of an electromagnetic wave can be written as

$$\vec{E}(t) = \hat{h}E_h(t) + \hat{v}E_v(t), \tag{34}$$

where $\hat{h}$ represents a unit vector in the horizontal direction that is parallel to the illuminated ground and $\hat{v}$ is a unit vector in the vertical direction that is perpendicular to both $\hat{h}$ and the propagation direction. If this electromagnetic wave is scattered by an object, the polarization of the scattered wave depends on the electrical and geometrical properties of the scattering object. One can consider the scatterer as a mathematical operator which takes one two-dimensional complex vector (incident wave) and changes it into another two-dimensional vector (scattered wave). Therefore, a scatterer can be characterized by a complex 2×2 scattering matrix.

The typical hardware implementation of a radar polarimeter involves transmitting a wave of one polarization and receiving echoes in both polarizations simultaneously. This is followed by transmitting a wave with a second polarization and again receiving echoes with both polarizations simultaneously. In this way, all four elements of the scattering matrix are measured. In order to use the polarimetric data derived from the scattering matrix for science applications, it is necessary to calibrate the data. Polarimetric calibration usually involves four steps: cross-talk removal, phase calibration, channel imbalance compensation, and absolute radiometric calibration.

The availability of calibrated polarimetric SAR data allowed researchers to perform quantitative analysis of the data. Significant progress has been made in the classification of different types of terrain using polarimetric SAR data. One of the most active areas of research in polarimetric SAR involves estimating geophysical parameters such as forest biomass, surface roughness, and soil moisture. Polarimetric SAR data may be optimal for flood monitoring and mapping, especially in the presence of vegetation canopies. Since the water surface is flat and a radar signal can penetrate into vegetation canopies, the inundated area can be identified using the copolarization ratio that depends on the strength of tree-water double bounce scattering.

B. SAR Interferometry

A conventional SAR measures the along-track and cross-track location of an object by projecting topographic relief information into a two-dimensional image plane. SAR interferometry is capable of measuring the third dimension to generate three-dimensional images. The basic principles of SAR interferometry can be explained using the geometry shown in Fig. 11. The slant range difference ($\delta\rho$) of two interferometric antennas can be calculated as

$$\delta\rho \approx -B_L \sin(\theta - \alpha), \tag{35}$$

where B_L is the interferometric baseline length and α is the baseline tilt angle shown in Fig. 11. An interferometric phase difference can be derived from $\delta\rho$ as

$$\delta\phi \approx \frac{a2\pi}{\lambda} B_L \sin(\theta - \alpha), \tag{36}$$

where $a = 1$ for the case where signals are transmitted by one antenna and received through both antennas simultaneously and $a = 2$ for the case where the signal is alternatively transmitted and received through one of the two antennas only. It can be noticed from Eq. (36) that all parameters are known except for θ. Therefore, the angle θ can be derived from the measured differential phase. Then, the elevation of the point being imaged is given by

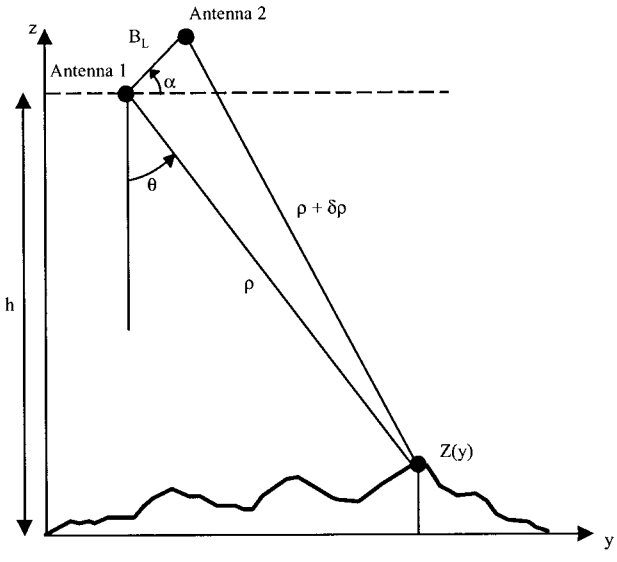

FIGURE 11 SAR interferometry geometry.

$$z(y) = h - \rho \cos\theta, \tag{37}$$

where h denotes the height of the imaging antenna. Since the measured phase varies between 0 and 2π, it must be unwrapped to retrieve the original phase by adding and subtracting multiples of 2π. Phase unwrapping algorithm development remains one of the most active areas of research in SAR interferometry, and many algorithms are under development.

Using differential SAR interferometry, minute surface deformations can be measured with unprecedented accuracy. This differential interferometry is implemented by subtracting two interferometric pairs separated in time. This technique was first demonstrated using SEASAT data to measure millimeter-scale ground motion in agricultural fields. Since then, this technique has been applied to measure co-seismic displacement and volcanic deflation. Differential SAR interferometry has been used for polar ice sheet research by providing information on ice deformation at great spatial details. SAR interferometry can also be used for mapping ocean surface movement by imaging the same area twice within ocean decorrelation time. In this case, the interferometer is implemented in such a way that one antenna images the scene a short time before the second antenna with the exact same viewing geometry. This technique is known as along-track SAR interferometry.

V. SPACEBORNE SAR EXAMPLES

The SEASAT SAR is the first civilian spaceborne SAR system designed to acquire high-resolution images of the earth surface. The SEASAT SAR was launched by NASA in 1978, and it operated successfully for 100 days and acquired approximately 50 h of radar data. The spare hardware of the SEASAT SAR system was modified and flown on the space shuttle in November 1981. This flight is known as the SIR-A (Shuttle Imaging Radar-A) experiment. After several modifications to the SIR-A system, the SIR-B system was flown in 1984, and it was capable of varying the look angle from 20° to 60°, while the SIR-A system operated at a fixed 50° look angle. The characteristics of SEASAT, SIR-A, and SIR-B SAR systems are summarized in Table II.

The SIR-C/X-SAR experiment is a joint U.S./German/Italian project. It is the first spaceborne multifrequency and multipolarization SAR system. The instrument, flown aboard the space shuttle Endeavour, collected valuable multiparameter SAR data in April and October 1994. Two flights provided the opportunity for assessment of change with time due to seasons and other factors. During the second flight, the orbit was trimmed for the last 3 days to demonstrate SAR interferometry to generate large-scale topographic maps. The SIR-C instrument used phased array antennas to provide the sensitivity to detect weak cross-polarization signals, while maintaining an acceptable transmit power level by using many transmit/receive modules close to the antenna radiating elements. SIR-C also demonstrated, for the first time, the concept of ScanSAR from space. The ScanSAR mode is implemented by moving the antenna beam sequentially in the across-track direction. A wider swath is accomplished at the expense of azimuth resolution. The SIR-C/X-SAR characteristics are shown in Table III.

The SRTM was launched in February 2000 to produce a topographic map of 80% of the earth's land surface in a single 10-day space shuttle flight. The mission was highly successful in collecting all the planned data. The addition of C-band and X-band receiver antennas, extended from the space shuttle bay on a 60-m mast and operating in concert with the existing SIR-C/X-SAR antennas, forms a single pass interferometric radar system. It is estimated that the height accuracy of the C-band SRTM data is

TABLE II SEASAT, SIR-A, and SIR-B Radar System Characteristics

SAR parameter	SEASAT	SIR-A	SIR-B
Frequency band	L-band	L-band	L-band
Bandwidth (MHz)	19	6	12
Altitude (km)	790	245	225
Polarization	Horizontal	Horizontal	Horizontal
Pulse length (μsec)	33.9	33.9	33.9
Peak power (W)	1000	1000	1000
Antenna width (m)	2.16	2.16	2.16
Antenna length (m)	10.7	9.36	10.7
Swath width (km)	100	38	30
Look angle (degrees)	20.5	50	20–60
Launch date	1978	1981	1984

TABLE III SIR-C/X-SAR Radar System Characteristics[a]

SAR parameter	SIR-C (L-band)	SIR-C (C-band)	X-SAR
Frequency band	L-band	C-band	X-band
Bandwidth (MHz)	10, 20, and 40	10, 20, and 40	10 and 20
Altitude (km)	225	225	225
Polarization	Fully polarimetric	Fully polarimetric	Vertical
Pulse length (μsec)	8.5, 17, and 34	8.5, 17, and 34	40
Peak power (W)	4000	1200	3300
Antenna width (m)	2.9	0.74	0.4
Antenna length (m)	12	12	12
Swath width (km)	15–80	15–80	15–80
Look angle (degrees)	20–60	20–60	20–60
Launch date	1994	1994	1994

[a] The fully polarimetric mode measures a complex 2×2 scattering matrix.

less than 16 m absolute and 11 m relative, where these values are the 90% linear error after processing both ascending and descending passes. The elevation posting of the SRTM data is 30 m.

The international use of spaceborne imaging SAR for long-term earth observation dramatically increased in the 1990s. While NASA missions were space shuttle based, four SAR imaging satellites were launched by Europe, Japan, and Canada. The European Space Agency (ESA) launched an SAR instrument aboard the European Remote-Sensing Satellite (ERS) in August 1991 and another in April 1995. Pairs of ERS images, including tandem ERS-1 & 2, have been used to generate interferometric products of surface topography and deformation. The Japanese National Space Development Agency (NASDA) launched an SAR instrument aboard the Japanese Earth Resources Satellite-1 (JERS-1) in February 1992. JERS SAR has provided a global land cover map, and its repeat pass interferometric data have been used to measure surface deformations. The Canadian Space Agency (CSA) launched an SAR instrument aboard RADARSAT in September 1995. The RADARSAT antenna is capable of electronic beam steering that enables very large area mapping using the ScanSAR mode. The characteristics of these four SAR systems are shown in Table IV.

Unlike earth-observing satellites, planetary radar operation is severely limited by available mass, power, and data downlink rate. The concept of an imaging radar mission to Venus was first considered in the late 1960s to penetrate the dense cloud cover of Venus. In 1989, Magellan was launched from the space shuttle Atlantis to image the Venus surface. The radar operates at S-band with low transmit peak power. The Magellan mission imaged 98% of the surface of Venus, far exceeding its requirement of 70% coverage. The characteristics of the magellan radar system are shown in Table V.

TABLE IV ERS-1, ERS-2, JERS-1, and RADARSAT Radar System Characteristics

SAR parameter	ERS-1	ERS-2	JERS-1	RADARSAT
Frequency band	C-band	C-band	L-band	C-band
Bandwidth (MHz)	15.5	15.5	15	11.6, 17.3, and 30
Altitude (km)	780	780	568	800
Polarization	Vertical	Vertical	Horizontal	Horizontal
Pulse length (μsec)	37	37	35	42
Peak power (W)	4800	4800	1500	5000
Antenna width (m)	1	1	2.2	1.5
Antenna length (m)	10	10	11.9	15
Swath width (km)	100	100	75	10–500
Look angle (degrees)	23	23	39	20–50
Launch date	1991	1995	1992	1995

TABLE V Magellan Radar System Characteristics

SAR parameter	Magellan
Frequency band	S-band
Bandwidth (MHz)	2.26
Altitude (km)	275–2100 (elliptical orbit)
Polarization	Horizontal
Pulse length (μsec)	26.5
Peak power (W)	350
Antenna diameter (m)	3.7
Swath width (km)	20–25
Look angle (degrees)	13–47
Launch date	1989

VI. CONCLUSIONS AND OUTLOOK

SAR is now a mature imaging technique that has found widespread application in studying the Earth and other planets. Several countries are planning to launch next-generation SAR satellites in the next decade or so. Among them are the ESA with the ASAR system on the Envisat satellite, the CSA with Radarsat-2, and the NASDA with the PALSAR on the ALOS spacecraft. In parallel, many scientists are starting to use SAR images more routinely in their analysis of the earth as an integrated system, particularly in the fields of plate tectonics and glaciology, where the use of differential interferometry has revolutionized our understanding of the processes involved. As science matures into a complete analysis utilizing data from complimentary sensors that cover a wide range of the electromagnetic spectrum, it is expected that radar will play an increasing role in remote sensing of the Earth and other planets.

SEE ALSO THE FOLLOWING ARTICLES

COMMUNICATION SATELLITE SYSTEMS • RADAR • REMOTE SENSING FROM SATELLITES • SATELLITE COMMUNICATIONS • SIGNAL PROCESSING, ANALOG

BIBLIOGRAPHY

Curlander, J. C., and McDonough, R. N. (1991). "Synthetic Aperture Radar System and Signal Processing," Wiley, New York.
Elachi, C. (1988). "Spaceborne Radar Remote Sensing: Applications and Techniques," IEEE Press, New York.
Henderson, F. M., and Lewis, A. J. (1998). "Principles & Applications of Imaging Radar," Wiley, New York.
Ulaby, F. T., and Elachi, C. (1990). "Radar Polarimetry for Geoscience Applications," Artech House, Norwood, MA.