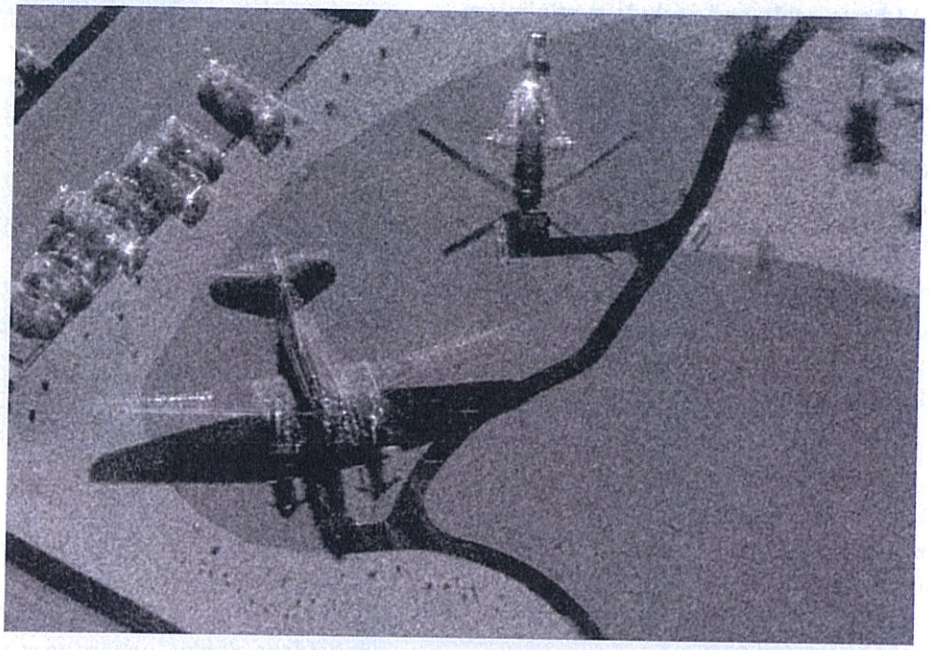
CHAPTER
14

Real- and Synthetic-Aperture Side-Looking Airborne Radar



High-resolution SAR image of an airport runway with a plane and helicopter (Courtesy Sandia National Laboratories)

CONTENTS

Overview

Remote sensing radars can be either imaging or nonimaging. Nonimaging radars were discussed in Chapter 13 and now we consider the family of imaging radars generically known as ***side-looking airborne radars*** (SLARs). Two different types of SLAR are considered: ***real-aperture SLAR*** and ***synthetic-aperture radar*** (SAR). These two imaging radars differ primarily in how the signal's Doppler spectrum is processed. While technically SAR is a type of SLAR, by convention the term SLAR is generally only used to refer to real-aperture radars. We will follow this convention in this book. The goal of both SLAR and SAR systems is to generate maps or images of the radar backscatter from the illuminated area.

The basic theories of SLAR and SAR are discussed in this chapter, along with considerations relating to system design and construction, but the details of transmitters, receivers, and other hardware components are not treated in this book. Processing algorithms for SAR are only briefly considered. More detailed treatment of SAR processing algorithms can be found in separate texts (e.g., Cumming and Wong, 2005; Curlander and McDonough, 1991).

14-1 Introduction

Radar systems are much more complex than radiometer systems due to the addition of a transmitter. However, this complexity provides the designer with additional flexibility via modulation of the transmitted signal and demodulation and possible Doppler processing of the received signal. Radar receivers are similar to the radiometer receivers discussed in Chapter 7 and many radiometer techniques can be used for very sensitive radars that operate with signal-to-noise ratios close to or less than unity. However, radar receivers are primarily designed for receiving the reflected signal and often do not have to be as sensitive as radiometer receivers. In many cases the transmit power level may be large enough that rather noisy receivers can be used without degradation in system performance.

In a remote sensing context, the term ***target*** describes the surface or volume that scatters the radar signal illuminating it back to the receiver. Originally in radar, the word *target* referred to a specific object such as an aircraft or a ship, and the target background was referred to as ***clutter***. However, in modern usage *target* can refer to anything that reflects the radar signal, including areas on the surface. In most remote-sensing radar imaging applications, the target is the ground. For ***radar imaging***, we want to resolve an extended area target into smaller areas, sometimes termed ***pixels*** or ***resolution elements***. For SLAR, this is done using a combination of the motion of the radar, a narrow antenna beamwidth, and the radar's range resolution. SAR achieves finer resolution in the along-track direction by also using Doppler processing.

The received radar signal amplitude is proportional to the target area's scattering coefficient, which depends on the incidence angle, the frequency of operation, and the polarization states of the transmit and receive antennas. We note that the variation of the scattering coefficient with angle of incidence is different for different classes of targets. This variation can be used in target analysis and identification. With all imaging radars, a single pass that produces a single image provides scattering measurements at a single angle of incidence at each location. Multiple passes along different flight lines or at different altitudes can give information about the variation in scattering as a function of angle of incidence. Some radars are specifically designed to gather this information by employing one or more beams pointed ahead of or behind the platform track so that the scattering coefficient for a particular target may be measured at a variety of angles as the radar advances past the target.

Some targets can be identified or mapped by their characteristic shape, texture, and/or context in an image. While variations due to amplitude differences among different target classes is important, the use of shape, texture, and context does not necessarily require accurate amplitude calibration. Thus, some applications of imaging radar do not require accurate calibration, which can simplify their design. Photo analysts, for example, are accustomed to working with

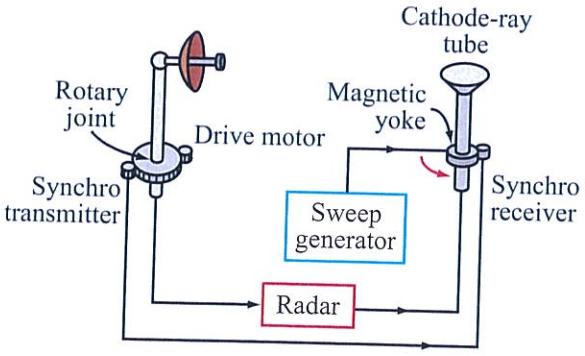


Figure 14-1: PPI radar system.

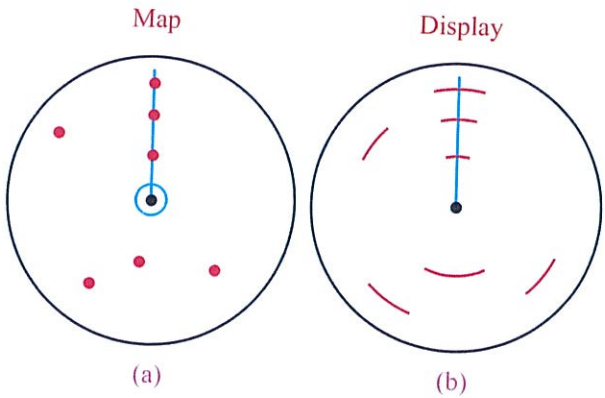


Figure 14-2: PPI display (360°) geometry: (a) surface view from above; the radar is at the center with bright targets indicated, and (b) corresponding PPI display.

aerial photographs in which calibration of amplitude is almost completely lacking. They are able to extract the information of interest from contrasting image brightness using shape, texture, and context, whether it is in photographs or radar images. In such cases the production of images with good geometric fidelity and good contrast may be more important than accurate calibration. On the other hand, radar imaging applications that relate the observed σ^0 to a geophysical property of interest may require accurate calibration on either an absolute or a relative basis.

The earliest imaging radar systems were based on real-aperture rotating antennas. The ***plan-position indicator*** (PPI) was invented early in World War II as a means of displaying rotating-antenna images. Figure 14-1 illustrates a PPI-type radar system. With PPI radar, the antenna is rotated through 360°, although in some systems the antenna rotates through a partial rotation sector. The PPI displays the radar echo power or voltage amplitude as a function of range and the angular position of the antenna. The rotation angle is sensed by a potentiometer, a two-phase "synchro" servo system, or other means. In an old cathode-ray tube system, the synchro receiver mechanically rotates the yoke used for magnetic deflection of the display tube in synchronism with the antenna. Thus, when the antenna is pointed straight ahead of the aircraft, the synchro positions the sweep so that the sweep line is vertical. When the antenna is pointed directly to the side, the synchro positions the sweep so that the sweep line is horizontal, and so on. By starting the sweep at the display center and finishing the sweep at the edge, the full 360° range may be displayed as indicated in Fig. 14-2. A simple target map is shown with a straight line directly ahead of the radar. With the PPI display, individual targets appear at approximately the correct geometric position, though they tend to appear on the display as segments of circles rather than as points. A PPI radar is capable of producing an image that represents reasonably faithfully the location of targets on the ground, and a PPI radar built with a large-enough antenna can achieve the same kind of results that can be achieved with a side-looking real-aperture radar having an antenna of the same length.

14-2 Real-Aperture SLAR

The general idea of side-looking airborne radar was introduced in Section 1-5.1. The aircraft (or spacecraft) carries a long, thin antenna that produces a fan beam pointed to the side of the flight track as illustrated in Fig. 14-3(a). In pre-1970 radar systems, the forward motion of the aircraft was synchronized with the motion of film in a recorder that passes the film in front of an intensity-modulated cathode-ray tube that produces a single line showing the radar echo amplitude versus

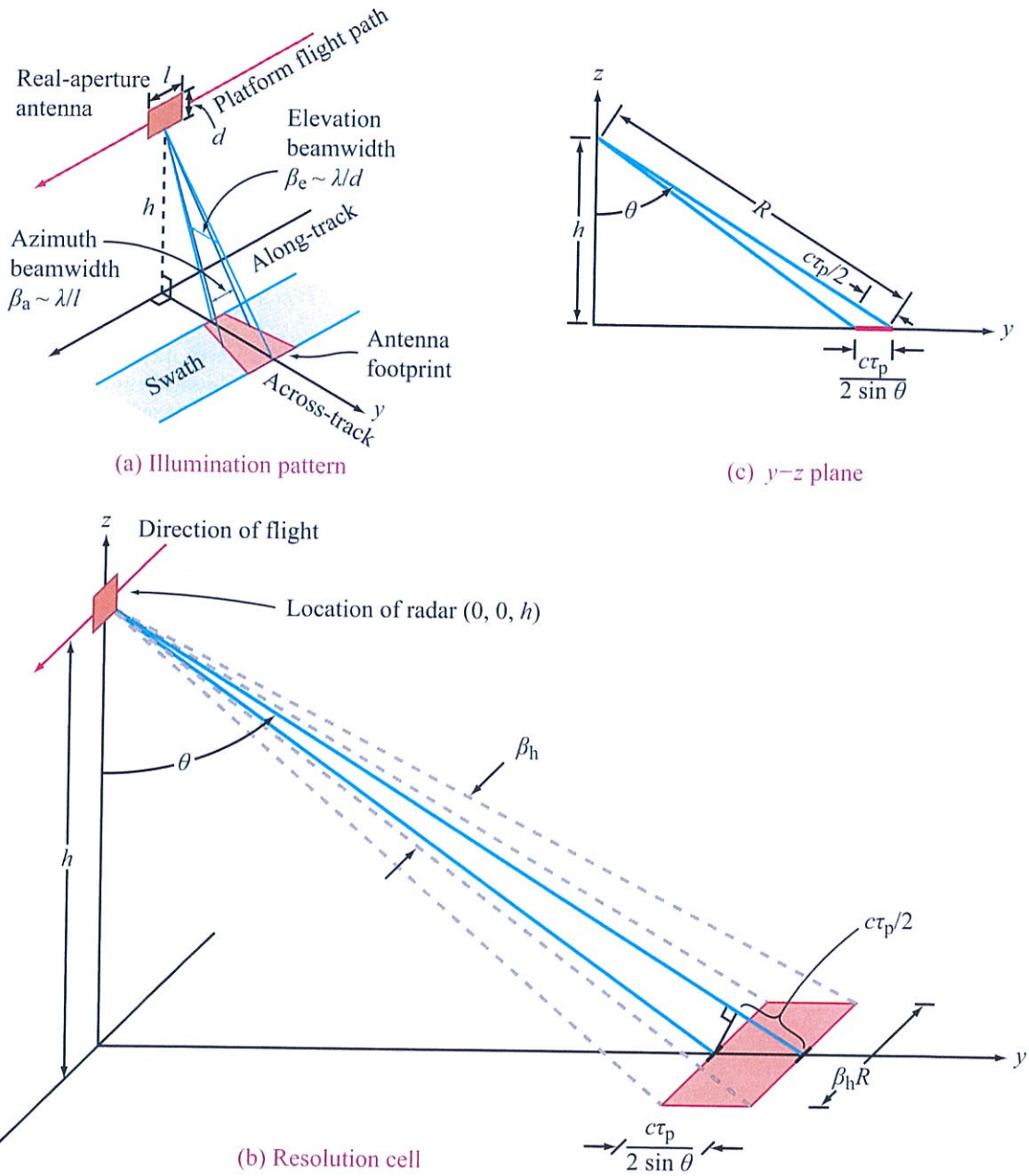


Figure 14-3: Geometry of real-aperture SLAR: (a) antenna illumination pattern on ground, (b) three-dimensional view of one resolution cell, and (c) view in the y–z (cross-track/vertical) plane.

time for each pulse transmitted. The motion of the film and aircraft produced an intensity image by this technique. In modern systems, cathode-ray tubes and film recordings are replaced by digital sampling and storage. This simple description applies to a real-aperture SLAR. For SAR, the recorded signal is further processed to obtain finer resolution in the along-track direction. In this section we focus on real-aperture SLAR; SAR systems are considered later.

14-2.1 SLAR Resolution

The spatial resolution of a real-aperture SLAR is dictated by the antenna beamwidth in the ***along-track*** or ***azimuth direction***, and by the effective pulse length (after range compression, if used—see Section 13-6) in the ***across-track direction***. Figure 14-3 gives the detailed geometry for resolution determination of the real-aperture SLAR.

The total round-trip time delay T experienced by the radar signal traveling from the radar to a point at ***slant range*** R is

$$T = \frac{2R}{c}. \tag{14.1}$$

For a pulse of length τ_p, the ***slant-range resolution*** r_r in the slant-range direction is

$$r_r = \frac{c\tau_p}{2}. \tag{14.2}$$

If range compression is used, τ_p is the effective pulse length after range compression. The ratio of the slant-range resolution to the total slant range is the same as the ratio of the pulse duration τ_p to the total time delay T. We are more interested in the resolution distance on the ground (i.e., the ground resolution) rather than in the slant-range resolution. The relation between the slant-range and ground-range resolutions can be determined with the aid of Fig. 14-3. The ***ground-range resolution*** r_y is given by

$$r_y = \frac{c\tau_p}{2\sin\theta}, \tag{14.3}$$

where θ is the angle of incidence.

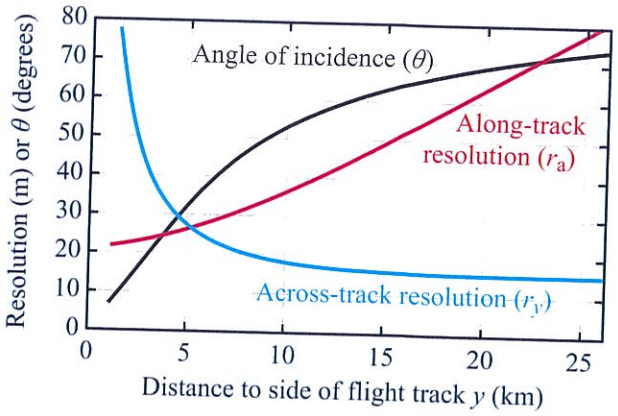


Figure 14-4: Example of SLAR resolution versus incidence angle for altitude $h = 7.5$ km, pulse length $\tau_p = 100$ ns, and along-track horizontal beamwidth $\beta_h = 3$ mrad (0.17°).

The ***along-track resolution*** r_a of a ***real-aperture*** SLAR is dictated by the antenna beamwidth at the ground; i.e., by the arc length corresponding to the ***horizontal beamwidth*** β_h:

$$r_a = \beta_h R = \frac{\beta_h h}{\cos\theta}, \tag{14.4}$$

where h is the height of the radar above the ground.

Because of the factor $\sin\theta$ in the denominator of the expression given by Eq. (14.3) for the ground-range resolution and the factor $\cos\theta$ in the denominator of the along-track resolution expression of Eq. (14.4), the along-track and ground-range resolutions are functions of the angle of incidence and, consequently, of the distance to the side of the flight track, as shown in Fig. 14-4.

▶ The across-track resolution r_y is degraded at short distances (corresponding to low incidence angles), and the along-track resolution r_a is degraded at long distances. ◀

Even with very long antennas, the along-track resolution degrades very quickly with incidence angle. For this

reason, real-aperture SLARs are seldom used to produce wide-swath images and are rarely used from space. As an illustration, a real-aperture X-band SLAR used on a spacecraft at a 600 km altitude with a 10 m long antenna (beamwidth ~3 mrad or 0.17°) achieves a best-case along-track resolution of only 1.8 km.

Note that in general the SLAR image resolution cells (pixels) are not square. Figure 14-4 reveals that, for a given real-aperture SLAR, a pixel is square at only one range (in this example, at 5 km). At other ranges a pixel is rectangular, being elongated in the across-track direction at short ranges and in the along-track direction at long ranges. Thus, in specifying the resolution of a real-aperture SLAR, one should use both the beamwidth and slant-range resolution to describe the performance of the system, rather than a single value. From Eqs. (14.3) and (14.4), the pixel area $r_a r_y$ is a minimum at an incidence angle of 45°.

14-2.2 The SLAR Radar Equation

For a SLAR with the geometry of Fig. 14-3, the radar equation given by Eq. (5.40) becomes

$$P_r = \frac{P_t G^2 \lambda^2 \sigma^0}{(4\pi)^3 R^4} (\beta_h R) \left(\frac{c\tau_p}{2 \sin\theta} \right) \tag{14.5}$$

or

$$P_r = \frac{P_t G^2 \lambda^2 \sigma^0 \beta_h c \tau_p}{2(4\pi)^3 R^3 \sin\theta}, \tag{14.6}$$

where P_r and P_t are the received and peak transmitted powers, G is the antenna gain, σ^0 is the scattering coefficient, and the illuminated area is written in terms of the product $r_a r_y$, using Eqs. (14.3) and (14.4). Proper correspondence between P_r and σ^0 requires P_r to be the ensemble average over many samples of the distributed target.

Often, the ***average transmitted power*** $P_{t_{av}}$, rather than the ***peak power*** P_t, is used. For a ***pulse-repetition rate*** f_p (in pulses/s), the average power $P_{t_{av}}$ is

$$P_{t_{av}} = P_t \tau_p f_p, \tag{14.7}$$

where τ_p is the pulse length. Accordingly, Eq. (14.6) may be written as

$$P_r = \frac{P_{t_{av}} G^2 \lambda^2 \sigma^0 \beta_h c}{2(4\pi)^3 R^3 f_p \sin\theta}. \tag{14.8}$$

From Chapter 7, the receiver noise power is

$$P_n = kT_0 BF, \tag{14.9}$$

where k is Boltzmann's constant, T_0 is the reference temperature (290 K), B is the bandwidth, and F is the receiver noise figure. Consequently, the ***signal-to-noise ratio*** S_n is

$$S_n = \frac{P_r}{P_n} = \frac{P_t G^2 \lambda^2 \sigma^0 \beta_h c \tau_p}{2(4\pi)^3 k T_0 B F R^3 \sin\theta}. \tag{14.10}$$

(SLAR SNR equation)

Note that, as discussed in Chapter 13, for a pulse with a realistic shape, one must use the actual form of the pulse $P_t(t)$ and the actual shape of the antenna pattern, so that Eq. (14.6) should be written as

$$P_r(t) = \iint_{\text{total area illuminated}} \frac{P_t(t-T)\, G^2(x,y) \lambda^2 \sigma^0(x,y)}{(4\pi)^3 R^4} dx\, dy. \tag{14.11}$$

For a fan-beam antenna, the gain pattern of the antenna can be separated into independent components in the θ (cross-track) and ϕ (along-track) directions, we can write

$$G(\theta,\phi) = G_\theta(\theta)\, G_\phi(\phi) = G_0\, g_\theta(\theta)\, g_\phi(\phi), \tag{14.12}$$

where $g_\theta(\theta)$ and $g_\phi(\phi)$ are pattern factors with maximum value of unity and G_0 is the peak gain. Note that for the narrow-beam case, the along-track differential distance may be expressed as

$$dx = d(R\phi) = R\, d\phi, \tag{14.13}$$

and $\sigma^0(\theta,\phi) \approx \sigma^0(\theta)$. Substituting Eqs. (14.12) and (14.13) in (14.11) leads to

$$P_r(t) = \int \frac{P_t(t-T)\, G_0\, g_\theta^2(\theta)\, \lambda^2\, \sigma^0(\theta)}{(4\pi)^3 R^4} dy \int g_\phi^2(\phi)\, d\phi. \tag{14.14}$$

An alternative form of this equation can be derived by noting that

$$R = cT/2 \tag{14.15}$$

and

$$P_t(t) = P_{t_{max}}\, p(t), \tag{14.16}$$

where $P_{t_{max}}$ is the maximum power value of the pulse and the pulse envelope $p(t)$ has a maximum value of one. Also note from the geometry of Fig. 14-3 that

$$\theta = \cos^{-1}\left(\frac{h}{R}\right) = \cos^{-1}\left(\frac{2h}{cT}\right). \tag{14.17}$$

Since

$$y = \sqrt{R^2 - h^2},$$

dy in Eq. (14.11) can be replaced with

$$dy = dR\,\frac{R}{\sqrt{R^2 - h^2}}. \tag{14.18}$$

Making all of these substitutions in Eq. (14.14), we obtain

$$P_r(t) = \frac{P_{t_{max}} G_0^2 \lambda^2}{(4\pi)^3 (c/2)} \cdot \int \frac{p(t-T)\, g_\theta^2\left(\cos^{-1}\left(\frac{2h}{cT}\right)\right) \sigma^0\left(\cos^{-1}\left(\frac{2h}{cT}\right)\right)\, dT}{T^2\sqrt{\left(\frac{cT}{2}\right)^2 - h^2}} \cdot \int g_\phi^2(\phi)\, d\phi, \tag{14.19}$$

which may be expressed in the form

$$P_r(t) = \frac{2P_{t_{max}} G_0^2 \lambda^2}{(4\pi)^3 c} \int p(t-T)\, b^2(T)\, dT, \tag{14.20}$$

where

$$b^2(T) = \frac{g_\theta^2\left(\cos^{-1}(2h)\, cT\right) \sigma^0\left(\cos^{-1}\left(\frac{2h}{cT}\right)\right)\, dT}{T^2\sqrt{\left(\frac{cT}{2}\right)^2 - h^2}} \cdot \int g_\phi^2(\phi)\, d\phi. \tag{14.21}$$

In Eqs. (14.14) and (14.19) the along-track variation of the antenna beam (i.e., in its narrow direction) is in a separate integral, which is a great convenience because it means that the effect of the shape of the antenna pattern in the along-track direction is similar for all ranges. This simplifies processing and image pixel calibration.

The expressions given by Eqs. (14.19) and (14.20) describe a convolution of the pulse shape, the shape of the antenna pattern, the shape of the scattering-coefficient variation with angle, and the time delay or distance. For many applications, one can assume that the antenna pattern can be written in the form of Eq. (14.6) with half-power beamwidth β_h in the azimuth direction and equivalent half-power pulse duration τ_p in the range direction; however, for precise measurements, and in most cases near vertical, Eqs. (14.19) and (14.20) should be used instead.

14-2.3 SLAR Systems

► Unlike SAR systems, SLAR systems do not necessarily require pulse-to-pulse coherence. ◄

In this regard they are similar to pulse radars used for aircraft tracking. The primary differences between SLAR and aircraft tracking radars are associated with the method of scanning the antenna (by moving the aircraft rather than merely rotating the antenna) and the methods used to record and present the information. In the past, many SLAR systems were built using standard pulse radar receiver-transmitter units developed for other purposes such as marine navigation.

Figure 14-5 shows a conceptual block diagram of a 1960s SLAR that employs a film recorder for data storage. The entire system is controlled by a timing synchronizer that develops pulses used in controlling the transmitter pulse time, the gain control of the receiver, and the deflection circuits for the display. The synchronizer triggers a modulator that produces a pulse that modulates the power transmitter. In the past, magnetron oscillators were widely used for SLAR transmitters due to their ready availability, compact size, and reliability. While magnetrons produce short pulses with high peak power, their pulses do not always start with the same phase. Consequently, magnetrons are not suitable sources for fully coherent

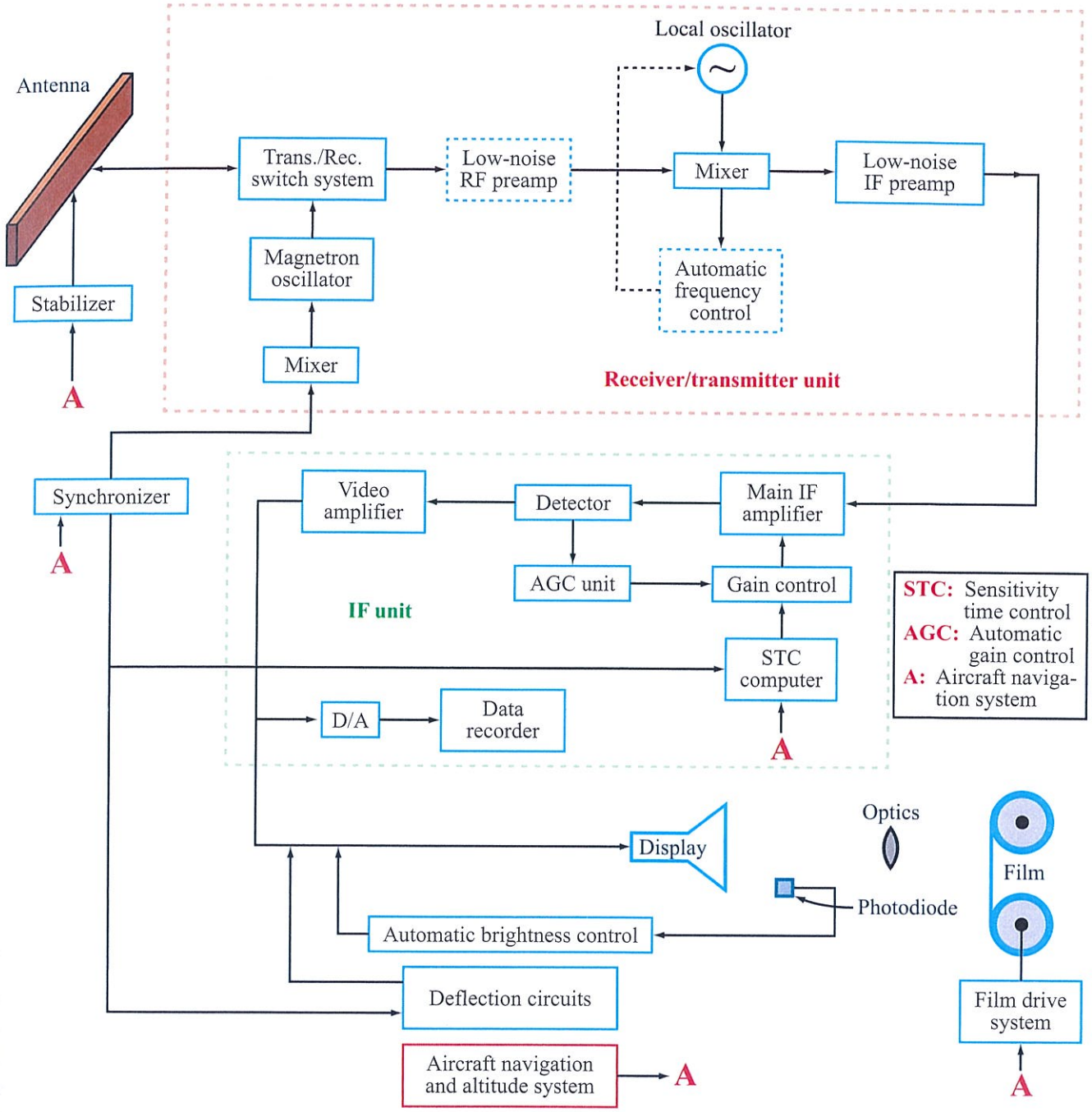


Figure 14-5: Example block diagram of a SLAR system.

radar systems. Modern systems often use ***traveling wave tube amplifiers*** (TWTAs) or ***solid-state transmitters***.

The output of the transmitter goes to a ***transmit-receive switching system***. This system involves ***circulators***, nonlinear receiver-protection devices, and so on, but its primary purpose is to provide a low-loss path between the transmitter and antenna during transmission while still protecting the receiver from the high peak power of the transmitter. During reception it must provide a low-loss path from the antenna to the receiver and reduce the amount of the quiescent noise output of the oscillator that reaches the receiver to prevent deterioration of the signal-to-noise ratio.

► A SLAR antenna is typically long and narrow so that it can achieve a narrow beam in the along-track direction and a wide beam in the across-track (or elevation) direction. ◄

The antenna may be rigidly mounted to the side of an aircraft when the application can permit avoidance of turbulent atmospheric conditions, but to compensate for attitude changes (of the aircraft) induced by ***atmospheric turbulence***, the antenna may be stabilized mechanically or electronically. When the antenna must be stabilized, the usual procedure is to allow for small motions about both the vertical axis and the horizontal axis. The small motions about the vertical axis compensate for minor variations in the roll of the aircraft, and those about the horizontal axis compensate for minor variations in yaw. Compensation for pitch variations is not usually attempted. The effect of yaw motions is to move the beam about on the ground and consequently cause the points imaged at a particular instant to be different from those assumed if a fixed antenna position relative to the flight line is assumed. Geometric variations in the elevation direction caused by roll of the aircraft are less significant, but aircraft roll affects the gain and image brightness across the swath. Since the antenna is typically long, mechanical compensation for yaw is more difficult than compensation for roll. A moving antenna must be enclosed in an aerodynamic ***radome*** and space must be left for the antenna's movement.

The transmitted signal travels to the target and returns to the antenna, and from there to the transmit-receive switching system, and on to the receiver. We note that radar design involves a trade-off between performance, on the one hand, and cost and complexity on the other. The receiver may or may not contain a low-noise RF preamplifier, depending upon the requirements for signal-to-noise ratio at the output, the expected value of the scattering coefficient, the gain of the antenna, and the power of the transmitter. Typically mixer-IF preamplifier systems may achieve noise figures of 8 to 10 dB, and sometimes better. Since the improvement associated with using a low-noise preamplifier in this situation is not likely to be more than about 6 dB, a preamplifier may not be required. The superheterodyne receiver is similar to the radiometer receivers discussed in Chapter 7. The primary differences between the radar receiver and the radiometer receiver relate to the greater need for a low noise-figure and greater gain-stability in the radiometer receiver, and the need for automatic frequency control and time-variable gain in the radar receiver.

When Doppler processing is not used, the transmitter need not be phase-coherent with the receiver. However, if transmit/receive IF frequency drift can occur, either ***excess receiver bandwidth*** or ***automatic frequency control*** (AFC) may be necessary to ensure capturing the full bandwidth of the received signals. To minimize RF losses, RF components are mounted in a receiver-transmitter unit close to the antenna. Other components can be located further away since losses in the transmission line connecting the low-noise IF preamplifier with the main IF amplifier have little impact on system performance. The IF unit contains the main IF amplifier (which usually has an adjustable gain in at least one stage), the detector, and the data recording system. Originally, data recording was done on film using cathode ray tubes but it is now done digitally by digitizing the signal versus time using an ***analog-to-digital*** (A/D) converter and storing the digital data.

Historically, to produce good real-time images with a SLAR, great care was taken to match the transfer characteristics of the different components of the system. SLARs have traditionally incorporated

automatic-gain-control (***AGC***) and/or sensitivity-time-control (***STC***) circuitry to minimize dynamic range requirements in signal processing and recording. There is a trade-off in system dynamic range and complexity: dynamic range can be reduced by using antenna pattern and/or receiver gain correction at the expense of complexity.

14-3 Synthetic-Aperture Radar (SAR)

► More commonly referred to merely as a synthetic-aperture radar (SAR), a synthetic-aperture side-looking airborne radar system combines radar and signal processing to form high-resolution backscatter images. ◄

Multichannel SARs that make simultaneous (or nearly simultaneous) measurements at different polarizations are known as ***quad-pol*** or ***polarimetric*** SARs. SARs that simultaneously employ multiple antennas with specific geometries are termed ***interferometric*** SARs, or InSARs for short. InSARS are covered in the next chapter. In this section we consider the fundamental theory of SAR operation and associated trade-offs. We refer to other books for more details on SAR image formation processing (e.g., Cumming and Wong, 2005; Curlander and McDonough, 1991). We discuss the ambiguity problem, which is of extreme importance for spacecraft radars, along with motion compensation and image distortion. Signal-to-noise considerations for SARs are similar to those for real-aperture radars, but are not identical, and so the differences are highlighted.

14-3.1 Ways to Consider SAR

Fundamentally, SAR is based on movement of the radar antenna over a stationary target area. ***Inverse SAR*** fixes the radar in place and moves the target. An example is a target on a rotating platform. In the operation of a SAR, pulses are issued and received as the radar passes over the target area. The recorded received signals for each pulse are combined to generate a high-resolution radar image.

► The key to SAR is coherent recording and processing of the received signal over multiple pulses. ◄

SAR can be examined from several different points of view that can lead to different insights into its operation and implementation. Careful analysis of the mathematical descriptions of these views reveals that they are mathematically equivalent, as they must be. However, the fact that multiple methods can be used to describe the same process is an indication of its complexity.

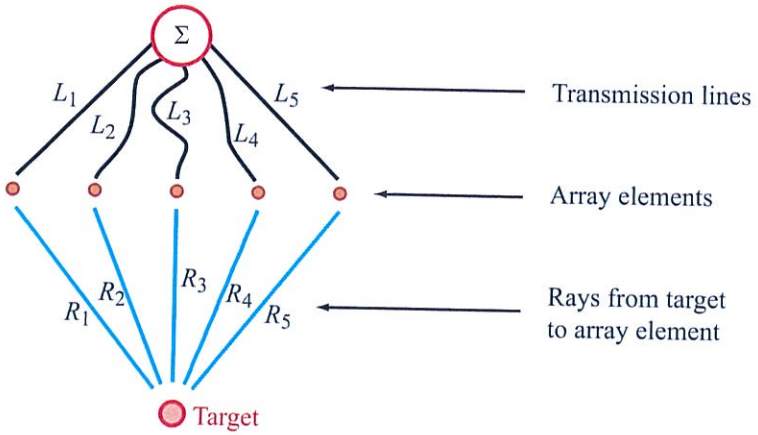
The points of view listed below are explored in more detail in the subsections that follow:

1. Synthesized antenna aperture.
2. Doppler beam sharpening.
3. Correlation or matched filtering with a reference point-target response.
4. Dechirping of Doppler frequency shift.
5. Optical-focusing equivalent.

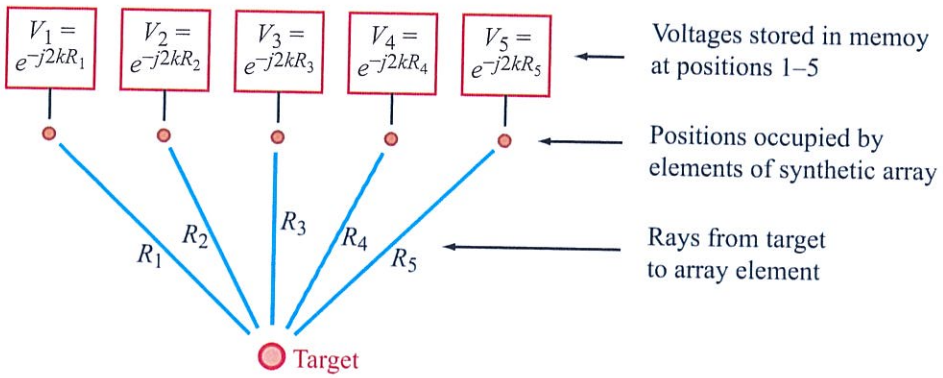
14-3.2 Synthesized Aperture

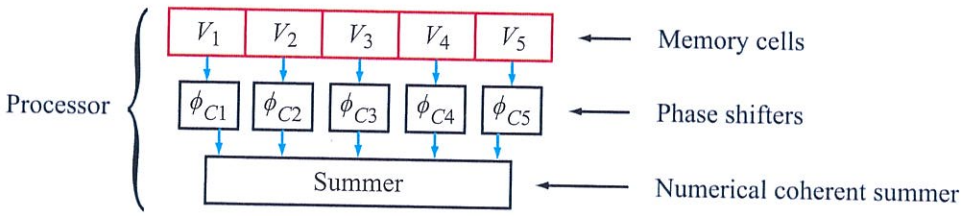
As suggested by its name, synthetic-aperture radar operates by "synthesizing" a long antenna aperture to produce finer resolution in the along-track direction. The synthesis occurs by recording the signal for each position of the antenna along a linear track, then combing the recorded signals as if they were simultaneously collected by an array antenna; in effect, SAR can be thought of as using time-storage techniques and the platform motion to create a long array antenna. Array-antenna analysis techniques similar to those developed in Chapter 3 can be applied to analyze the resolution of the SAR system.

Figure 14-6 illustrates the concept of a 1960s ***synthesized antenna aperture*** focused on a particular target. Note the difference between a focused antenna and the usual unfocused antenna. In unfocused antennas, the transmitted beam is confined as nearly as possible



(a) Real-aperture equivalent array focused on target
$L_1 + R_1 = L_2 + R_2 = L_3 + R_3 = L_4 + R_4 = L_5 + R_5$





(b) Synthetic-aperture equivalent of (a)

Figure 14-6: Concept of synthetic aperture.

to parallel rays. For reception we regard the source of the signal to be so far away that the rays are essentially parallel. Antennas for communication, for real-aperture radar (RAR) systems, and for most other radars are designed with this in mind. However, a synthetic-aperture radar usually has an aperture that is so long that the beam must be focused at a point where the target is located, much as the lens in a camera is focused for nearby objects. In an analogy with the camera, we may say that in real-aperture systems, the "lens" (antenna) is focused at infinity. For a synthetic aperture the lens, or the "array" of antenna positions, is focused on a nearby object.

For clarity and intuition, Fig. 14-6(a) illustrates how one might build a real-aperture array of five antenna elements focused on the target T. Each element in the array is connected to a summing point, and the voltages from all the different elements are added at that point. The lengths of the transmission lines connecting each of the elements to the summing point are based on the distance to the target and the position of the particular element in the array.

▶ The total phase delay for a signal arriving at the summing point from the target must be the same for each element if the contributions of the different elements are to add in phase; when they do add in phase, the array is said to be focused at the target point. ◀

This means that the equivalent total distance from target to summing point must be the same for every element, so that

$$L_1 + R_1 = L_2 + R_2 = L_3 + R_3 = L_4 + R_4 = L_5 + R_5.$$

The lengths L_i are the "equivalent free-space lengths" of the transmission lines; that is, we choose the line length so that the time delay of the signal in the transmission lines corresponds to the travel time of the wave in space. On transmission, the signals all originate at the location that is the summing point for reception, so that for transmission the phase shift along each path is also the same as that along every other path. Consequently, for the location of target T, the total round-trip phase shift is the same for each element, and the electric fields from the different elements all add in phase. For other locations, the distances between target and summing point differ so that the contributions from the different array elements differ in phase. The summed output is smaller than it is for a target located at the focal point T. Note that the actual distances can be adjusted so long as the phase relationship is maintained.

Figure 14-6(b) shows the synthetic-aperture equivalent of the real-aperture focused array of Fig. 14-6(a). In SAR, each of the "elements" in the synthetic array is occupied at a different time as the SAR moves. When the element is occupied, both amplitude and phase of the received signal are recorded. When data from all element positions have been collected, the recorded signals can be summed to synthesize the aperture output.

For simplicity in our intuitive explanation here, we assume that the amplitudes are the same for each position and consider only the phase. The phases are as indicated in the boxes at the top of Fig. 14-6(b), with the round-trip phase shift for wave travel for element i being $2kR_i$, where $k = 2\pi/\lambda$. After all the target phases have been recorded—that is, after each position in the synthetic array has been occupied by the radar—the contents of memory cells 1 through 5 are phase-shifted by the amounts needed to compensate for the differing distances R_1 to R_5. In essence, V_1 is multiplied by e^{+j2kR_1}, and so on for the other elements. This phase correction $+2kR_1$ is indicated in the figure as ϕ_{C1}. Note that the phasor sum of all the corrected signals is 5 times the value for one signal.

It is significant to recognize that in this view of SAR, there is no requirement for continuous motion from array elements 1 to 2 to 3 to 4 to 5, that the elements be uniformly spaced, or that element positions be occupied in any particular order. The only requirement is that the position of each element relative to the focal point T be known and the appropriate phase correction applied in the processor (the summer in this case). It should be noted that this phase-correction scheme is applicable only when the system bandwidth is a small fraction of the operating frequency.

Mathematically we can designate the total phase shift

on path 1 of Figure 14-6(a) as ϕ_{T1}, the phase shift in space on path 1 as ϕ_{S1}, and the phase shift on the transmission line as ϕ_{L1}. Thus we may write

$$\phi_{T1} = \phi_{L1} + \phi_{S1} = \phi_{Ti} = \phi_{Li} + \phi_{Si} = \phi_T \qquad (14.22)$$

for $i = 1$ to 5, and the phase shift in space is

$$\phi_{Si} = 2kR_i, \qquad (14.23)$$

and the phase shift on the transmission line is

$$\phi_{Li} = 2k_L R_i, \qquad (14.24)$$

where $k_L = 2\pi/\lambda_L$ is the transmit signal wave number for the transmission line. [Note: The wavelength of the signal in free space (λ) may be different from the wavelength (λ_L) of the same signal in the transmission line.]

The required phase shift in the transmission line is therefore

$$\phi_{Li} = \phi_T - \phi_{Si}. \qquad (14.25)$$

From this and Eq. (14.24), the length of the transmission line can be determined as

$$L_i = \frac{\phi_T - 2kR_i}{2k_L}. \qquad (14.26)$$

The voltage received at element i is

$$V_i = e^{-j2kR_i} = e^{-j\phi_{Si}}. \qquad (14.27)$$

Therefore, we need a phase-shift correction

$$\phi_{Ci} = \phi_{Li} = \phi_T - \phi_{Si} \qquad (14.28)$$

to apply to the signal, as noted in Fig. 14-6(b). If we apply this correction, the resulting voltage at the summing point is

$$V = \sum_{i=1}^{5} e^{-j\phi_T} = 5e^{-j\phi_T} \qquad (14.29)$$

for a target at T. Since the absolute phase is essentially arbitrary, we set $\phi_T = 0$ by choosing $\phi_{Ci} = -\phi_{Si}$. More generally, to focus an N-element synthetic aperture on a target at position T, the expression

$$V = \sum_{i=1}^{N} a_i e^{j2kR_i} V_i \qquad (14.30)$$

is used, where V_i is the received signal for the ith aperture position, R_i is the distance between the ith aperture position and the target, and a_i is an optional weighting function that can be applied to control azimuth sidelobes.

14-3.3 Doppler Beam-Sharpening Approach

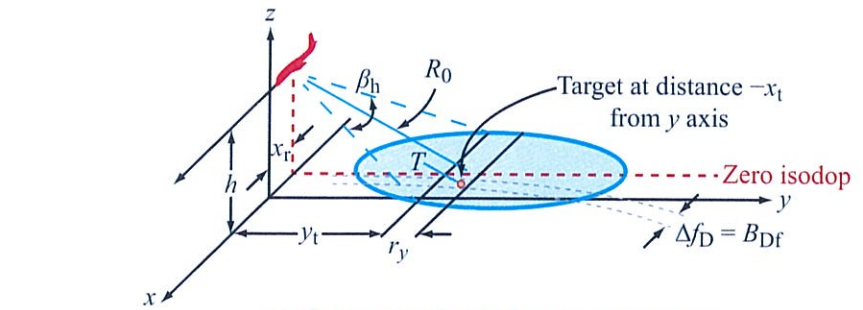
The original synthetic-aperture radar developed by Wiley in the early 1950s was called a "Doppler beam-sharpener." The use of the term "synthetic aperture" was adopted in the late 1950s. Figure 14-7 illustrates the view of a SAR as a focused Doppler beam-sharpening system, with Fig. 14-7(a) describing the geometry.

In Fig. 14-7(a) the radar is shown on an aircraft at a distance x_r back from the origin of coordinates at a height h; i.e, the aircraft coordinates are $(-x_r, 0, h)$ and the coordinates of the target T are $(-x_t, y_t, 0)$. The ellipse shown in the figure represents the half-power contour of the antenna illumination on the ground of the real antenna carried on the aircraft. The width of the ellipse is exaggerated for clarity; typically, it is a narrow fan beam. The beamwidth in the along-track direction is β_h. The effective ground-range resolution r_y is indicated. The shaded area surrounding the target, and bounded in the x direction by the edges of the beam, shows the area contributing to the return at a particular instant of time. The dashed lines are constant Doppler frequency values (isodops) surrounding the target. Their spacing Δf_D is the bandwidth B_{Df} of the Doppler filter used for beam sharpening.

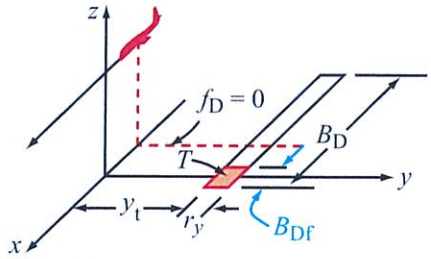
The Doppler frequency for the target in this case is given by

$$f_{dT} = \frac{2u(x_r - x_t)}{\lambda R}. \qquad (14.31)$$

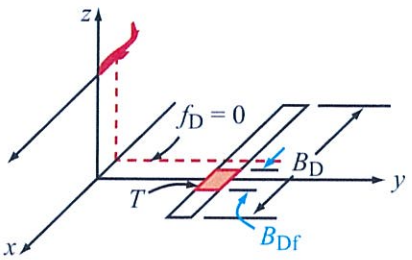
Note that the target location is fixed, whereas the radar location x_r is varying. For the illustration in Fig. 14-7(a), the Doppler frequency is positive because both values



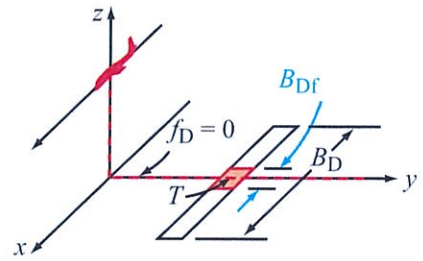
(a) Geometry of SAR Doppler calculations



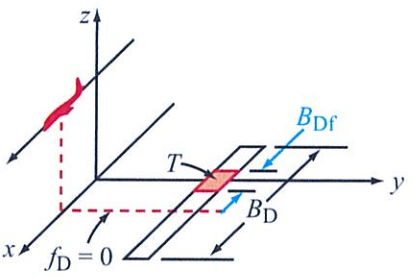
(b) High positive Doppler frequency



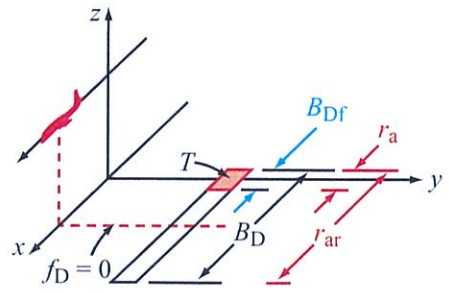
(c) Medium positive Doppler frequency



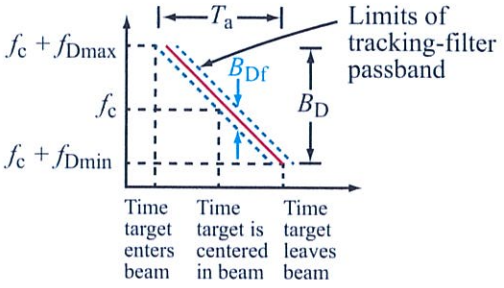
(d) Zero Doppler frequency



(e) Medium negative Doppler frequency



(f) High negative Doppler frequency



(g) Frequencies involved

Figure 14-7: Doppler beam-sharpening SAR concept for a narrow-band Doppler filter.

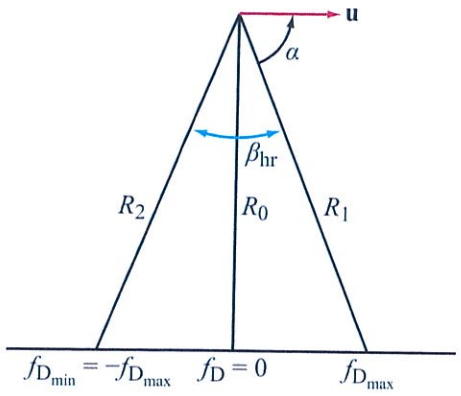


Figure 14-8: Real antenna with azimuth beamwidth β_{hr} traveling at speed **u** generates maximum and minimum Doppler shifts along R_1 and R_2.

of x are negative, with x_r having a larger magnitude. As x_r increases (it initially becomes less negative), the Doppler frequency decreases. If we neglect the variation in R (the slant distance from the radar to the target) during the passage of the beam over the target, the frequency decrease is linear as indicated in Fig. 14-7(g).

Figures 14-7(b)–(f) illustrate the effect of the "focused" ***Doppler beam-sharpening*** system as the target is successively illuminated first by the forward edge of the beam and later by other portions of the beam, until finally the beam passes it by. For illustration, the geometry of Fig. 14-7(a) is simplified by moving the coordinate system so that the target is located on the y axis at $(0, y_t, 0)$. The along-track length of the illuminated area on the ground is the resolution r_{ar} of a real-aperture radar of the same beamwidth; i.e.,

$$r_{ar} = \beta_{hr} R_0. \tag{14.32}$$

For a real-antenna aperture with azimuth beamwidth β_{hr} and velocity **u** (Fig. 14-8), the Doppler frequency shift is maximum for a target along the direction R_1 corresponding to the front edge of the beam,

$$f_{D_{max}} = \frac{2u}{\lambda} \cos \alpha, \tag{14.33}$$

where α is the angle between the velocity direction and the unit vector between the radar antenna and the ground. Upon replacing α with $(\pi/2 - \beta_{hr}/2)$, we have

$$f_{D_{max}} = \frac{2u}{\lambda} \sin\left(\frac{\beta_{hr}}{2}\right) \approx \frac{2u}{\lambda}\left(\frac{\beta_{hr}}{2}\right) = \frac{u\beta_{hr}}{\lambda}, \tag{14.34}$$

where we used the small angle approximation $\sin x \approx x$. The Doppler shift along R_2 is

$$f_{D_{min}} = -f_{D_{max}} = -\frac{u\beta_{hr}}{\lambda}. \tag{14.35}$$

Hence, the total Doppler bandwidth is

$$B_D = f_{D_{max}} - f_{D_{min}} = \frac{2u}{\lambda}\beta_{hr} = \frac{2ur_{ar}}{\lambda R_0}, \tag{14.36}$$

where r_{ar} is the azimuth resolution of the *real* antenna beam.

Doppler beam-sharpening is accomplished by using a narrow-bandwidth filter with a time-varying center frequency that "tracks" the Doppler-shifted signal frequency from the target. The limits on the ground defined by this filter are shown by the cross-hatched areas in Fig. 14-7(b)–(f) and have width r_a. The filter bandwidth is B_{Df}.

In Fig. 14-7(b), the beam begins to illuminate the target. The Doppler frequency is at its positive maximum, and the filter picks out a small section at the forward edge of the illuminated area. In Fig. 14-7(c), the aircraft has moved ahead about a quarter of the distance on the ground covered by the beam, and the Doppler filter has tracked the target signal down to a medium positive Doppler frequency. In Fig. 14-7(d), the aircraft has moved abreast of the target, whose Doppler frequency is now zero, so the filter is also centered on zero. In Fig. 14-7(e), the aircraft has moved forward to a point about another quarter beamwidth along the flight line, and the filter has tracked down to a medium negative Doppler frequency. Finally, in Fig. 14-7(f), the aircraft has moved forward to a point where the beam is about to leave the target. The filter is now centered on its maximum negative excursion.

Figure 14-7(g) shows this Doppler frequency shift as a function of time as the radar flies past the target. In

this figure the dotted line shows the limits of the tracking filter passband as the tracking filter follows the Doppler shift associated with the point target. Radar returns from other parts of the beam are filtered out by the tracking filter. By analogy with Eq. (14.36), the bandwidth of the tracking filter is related to the along-track resolution r_a [Fig. 14-7(f)] by

$$B_{Df} = \frac{2ur_a}{\lambda R_0}. \tag{14.37}$$

This expression can be inverted to obtain the SAR azimuth resolution in terms of the Doppler filter bandwidth:

$$r_a = \frac{\lambda R_0}{2u} B_{Df}. \tag{14.38}$$

Note that a different center frequency-versus-time tracking filter is required for each cross-range image pixel. Although the use of a tracking filter, or a bank of them, may be a conceptually simple way to implement Doppler beam-sharpening, practical considerations often lead to alternative implementations. In addition, the width of the Doppler filter may be reduced, which impacts r_a, as discussed later.

14-3.4 Correlation or Matched Filtering with Reference Point-Target Response

SAR can be viewed as a two-dimensional correlation or matched-filtering operation. One dimension is range or cross-track. ***Range compression*** is done in ***fast-time***, the time scale of an individual pulse. The other dimension is cross-range or along-track, also known as the "azimuth dimension." By analogy with range compression, azimuth processing is called ***azimuth compression***. Azimuth compression uses multiple pulses. This time scale is termed ***slow-time***. Range compression was considered in Section 13-6. Here we consider azimuth compression by correlation or match filtering.

We note that the received signal for each pulse is a continuous-time signal. As a result, range compression is a continuous-time problem, even if the signal is digitally sampled and processed. On the other hand, the azimuth signal is a fundamentally discrete-time problem due to the fact that the radar uses discrete pulses as it travels in azimuth. This requires that the sample rate due to the PRF satisfies the Nyquist sampling criterion. This is addressed in greater detail later. While the azimuth signal is actually a discrete-time signal, for the purposes of comparison with range compression, we treat it as a continuous-time signal in this section.

The basic scheme for a continuous-time correlation device is indicated in Fig. 14-9(a). The phasor output of the receiver $V_s(t)$ is fed to a correlator along with a phasor reference function $V_r(t)$. When the two signals are coincident in time and have the same waveform, the correlation is maximum. Hence the correlator output is maximized for a target located at a point for which the reference function is designed, and targets located elsewhere are discriminated against by the filter.

To illustrate how this operates, suppose that the phasor received signal is

$$V_s(t) = e^{j\phi(t)}.$$

The phase of the received signal, which is the time integral of the angular frequencies associated with the carrier and the Doppler shift, is given by

$$\phi(t) = \int (\omega_c + \omega_D)\, dt = \omega_c t + \int \frac{4\pi ux}{\lambda R_0}\, dt, \tag{14.39}$$

where x is the along-track displacement between the radar and the target. Since this displacement varies linearly with time, that is,

$$x = ut,$$

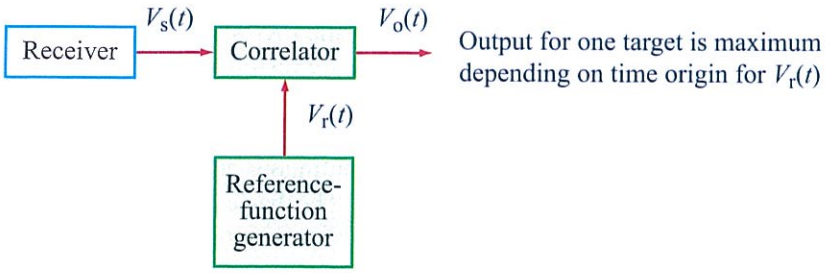
where u is the aircraft velocity, the resulting expression for the phase is

$$\phi(t) = \omega_c t + \frac{2\pi u^2 t^2}{\lambda R_0}. \tag{14.40}$$

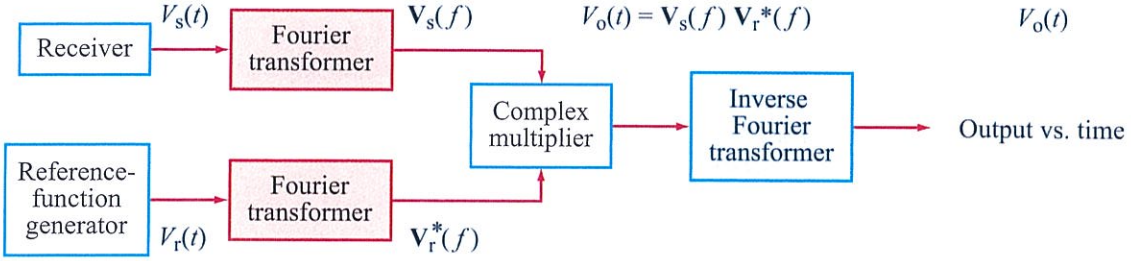
Ignoring the carrier-frequency term, the signal is

$$V_s(t) = e^{j2\pi u^2 t^2/\lambda R_0}. \tag{14.41}$$

This is the signal for a target whose x coordinate is zero; that is, the time origin is chosen as the instant that the radar comes abreast of the target. This radar position is known as the ***point of closest approach*** (PCA). The



(a) Basic correlator



(b) Frequency-domain correlator

Figure 14-9: Correlation processor concept for SAR.

appropriate reference function for a point target at this location is given by

$$V_r(t) = e^{j2\pi u^2 t^2/\lambda R_0}, \tag{14.42}$$

which, of course, is the same as $V_s(t)$. The output of the correlator is

$$V_o(t_r) = \int_{-T_a/2}^{T_a/2} V_s(t)\, V_r^*(t+t_r)\, dt, \tag{14.43}$$

where t_r is the offset between the reference used and the reference for the target at $x = 0$. Substituting the values from Eqs. (14.41) and (14.42) into Eq. (14.43) yields

$$V_o(t_r) = \int_{-T_a/2}^{T_a/2} e^{-j(2\pi u^2/\lambda R_0)(2t_r t + t_r^2)}\, dt,$$

which may be simplified to

$$V_o(t_r) = e^{-j(2\pi u^2 t_r^2/\lambda R_0)} \int_{-T_a/2}^{T_a/2} e^{-j(4\pi u^2 t_r/\lambda R_0)t}\, dt. \tag{14.44}$$

Note that the limits that have been placed on the integral correspond to the ***time of illumination*** of the target by the antenna beam, with the total duration of this illumination being T_a. The duration of the reference function is assumed to be long enough that it does not affect the limits of the integral. When this integral is evaluated, the result is

$$V_o(t_r) = e^{-j(2\pi u^2 t_r^2/\lambda R_0)} \left[\frac{\sin\left(2\pi u^2 T_a t_r/\lambda R_0\right)}{2\pi u^2 T_a t_r/\lambda R_0}\right] T_a. \tag{14.45}$$

Thus the output of the correlator is a maximum when $t_r = 0$, which occurs when the reference function and the target are aligned with each other. Note that the

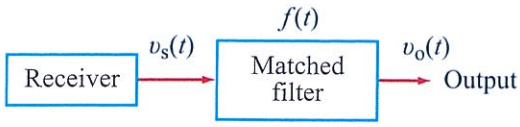


Figure 14-10: Matched-filter processor concept for SAR.

magnitude of the correlator output varies as $(\sin x)/x$. Hence. the output for a target located well away from the origin is small.

Clearly, if the SAR is to track multiple targets, many reference functions must be provided. A way to avoid this is to perform the equivalent of the correlation in the frequency domain, as indicated in Fig. 14-9(b). Here the reference function and receiver output are both Fourier-transformed, the Fourier transforms are multiplied, and the resulting frequency-domain function is converted by an inverse Fourier transform into the output waveform. For a single target, the time at which the output waveform has its maximum depends upon its location; for multiple targets, the output of the inverse transformer is a sequence, with the time for each target corresponding to its location. This frequency-domain approach is the basis of the ***range-Doppler algorithm*** (RDA) for processing SAR data, which we discuss in more detail in Section 14-9.4.

A similar but alternative view to correlation is ***matched filtering***. Figure 14-10 shows the elementary block diagram of the matched-filter processor. Here a filter, whose frequency response is the complex conjugate of the signal Doppler spectrum for a point target, is inserted in series with the input signal to obtain the output. The output signal is then the convolution of the input signal and the filter impulse response $f(t)$:

$$v_o(t) = v_s(t) * f(t), \tag{14.46}$$

where the filter frequency response is designed such that

$$\mathbf{F}(\omega) = \mathbf{V}_s^*(\omega). \tag{14.47}$$

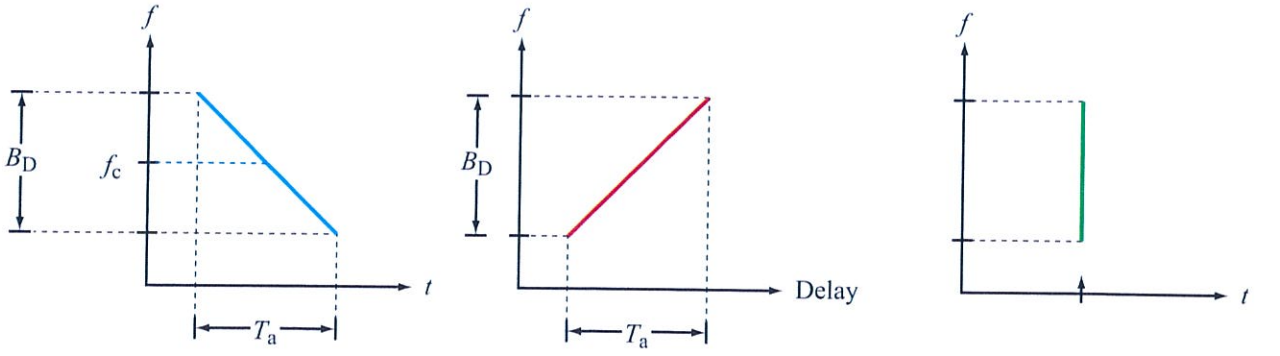
Recall that the matched filtering and correlation approaches are identical for the target location when the matched-filter and correlation reference function are properly defined.

14-3.5 Dechirping Comparable with Range Pulse-Compression Dechirping

As discussed in Chapter 13, range compression (matched filtering) of a linear frequency-modulated transmit signal can be accomplished using a dechirping operation. The same concept can be applied to the Doppler signal. Recall the linear change in Doppler with time shown in Fig. 14-7. The dechirping view of SAR is illustrated with the aid of Fig. 14-11. Figure 14-11(a) shows the decreasing Doppler frequency from the target, which lasts for a time T_a during which the target is within the area illuminated by the antenna. Since the highest Doppler frequency is received first and the lowest Doppler frequency is last, the idea of the frequency selective delay filter is to delay the higher-frequency signal by an amount corresponding to the time required for the actual Doppler frequency to change to its lowest value. Thus, the desired frequency-dependent time delay characteristic of the filter is as shown in Fig. 14-11(b). The resulting idealized output superimposes the outputs corresponding to all the Doppler frequencies to a single frequency output at a single time as indicated in Fig. 14-11(c). In reality, the output has a sinc function-like envelope because the dechirped signal has a finite duration and bandwidth, similar to the output for the range pulse-compression signal. Note that for range dechirping, the signal typically has a bandwidth of many MHz and a pulse duration of microseconds; for the Doppler frequency the bandwidth is in kHz or hundreds of Hz, and the time delay is measured in tenths of seconds or seconds.

14-3.6 Optical-Focusing Equivalent of SAR

While SAR processing is now performed digitally, the original SAR processors used optics for processing the radar signal. An exception was Wiley's first Doppler sharpening system. ***Optical processing*** begins with recording the SAR data on photographic film, as shown in Fig. 1-16. To record the signal phase (i.e., to record both positive and negative signal amplitudes), the signal zero level is set to be midway in the film's grayscale characteristic, with positive signals being recorded towards saturation on the film and negative signals being



(a) Doppler frequency variation (b) Delay characteristic of filter (c) Output for infinite bandwidth

Figure 14-11: Delay-filter azimuth dechirp concept for SAR.

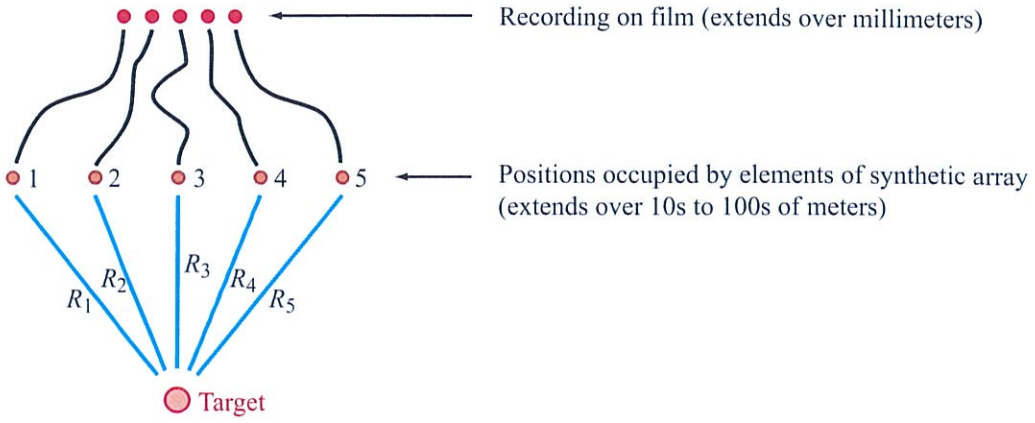
recorded toward the toe of the film response curve. The resulting recording on film is thus a grayscale rendition of the phase history of the received signal. Typically, a single pulse was recorded as a line on the film, as in the SLAR diagram illustrated in Fig. 1-16. Pulses were recorded sequentially as adjacent lines.

Figure 14-12 illustrates the principle of optical SAR processing. Recall the SAR concept shown in Fig. 14-6 and consider Fig. 14-6(a) for the real-antenna equivalent of the synthetic antenna as a model also for Fig. 14-12. In Fig. 14-12 the storage of the phase-coherent information is on film, rather than in a computer memory. The positions occupied by the elements of the actual antenna synthetic array extend over tens to hundreds of meters but are recorded over millimeters on the film and so the optics are closely spaced.

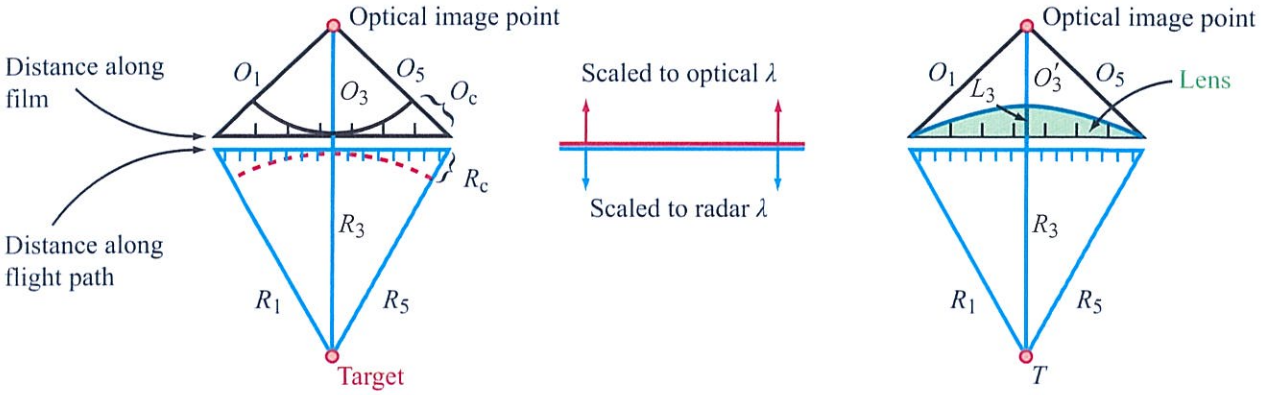
The optical processor is similar to building a real aperture as indicated in Fig. 14-6(a), except that the transmission lines are replaced by optical paths. This is indicated in Fig. 14-12(b) and (c). The bottom half of Fig. 14-12(b) is scaled with distances appropriate to the radar wavelength and radar geometry. The top half is scaled with distances appropriate to the optical wavelength and the geometry of the optical processor. Since the wavelength ratio is on the order of 10,000 or more in a typical case, the geometry is also scaled accordingly, although the actual optical wavelength is less important than the distances involved in recording on film. The radar signal goes from a point on the array out to the target and back, and then is recorded on film. The film is illuminated with collimated laser light, and the light passing through the film reproduces the strong and weak points of the equivalent phase front for the radar.

The signal is then focused at the optical image point with a lens as indicated in Fig. 14-12(c). Figure 14-12(b) illustrates the requirements for the lens geometry. The total phase shift between target and optical image point is the phase shift along path R_1 at the radar frequency and along path O_1 optically, and the same applies to all other path pairs R_i and O_i. The amount of excess radar distance at the ends of the aperture that must be corrected in the focusing is R_c. The corresponding amount of optical distance that must be corrected in the focusing is O_c. This focusing correction is achieved by the use of a convex lens, as indicated in Fig. 14-12(c). The shorter wavelength, or lower phase velocity, of the light in the glass of the lens causes additional phase shift where the lens is thick, relative to that where the lens is thin. The shape of the lens is thus selected to provide the desired correction,

The physical layout of the optical processor in its simplest form is shown in Fig. 14-12(d), where the

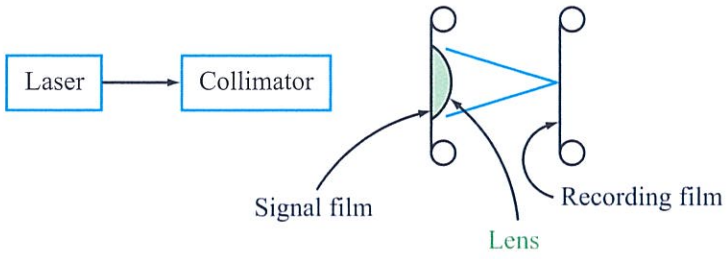


(a) Basic recording geometry



(b) Principle of required optical phase corrections

(c) Optical phase correction by lens



(d) General setup of processor

Figure 14-12: Optical processing for SAR.

laser light is collimated, fed through the film to the lens, and then focused on the film. Unfortunately, life is not this simple with optical processing, for the very process of focusing this signal on the film has the effect of defocusing in the range direction, so an additional lens must be added to compensate for that. Moreover, the correction computed for the center of the lens applies only for the particular radar range. Since the radar images a wide swath of the ground at varying slant ranges, the processor must accommodate these varying slant ranges by providing different corrections for different ranges for which the signal is recorded on film. As a result, optical processors were extremely complex to design and operate in practice. Optical processing was further complicated by the time delay required for processing the film, and the difficulties in accurately controlling the film processing.

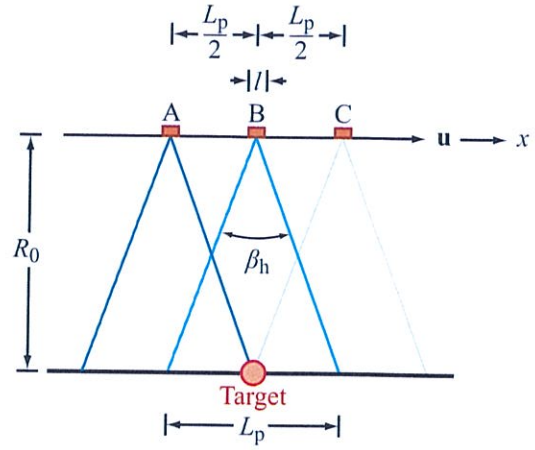
14-4 SAR Resolution

The previous section introduced various points of view of synthetic-aperture radar. Here, we derive the SAR resolution equation by considering two of these points of view: namely, the synthesized aperture and the Doppler beam sharpener points of view. These were selected because of their importance in the understanding of the SAR image formation process; however, the equations for resolution are the same no matter what point of view is used.

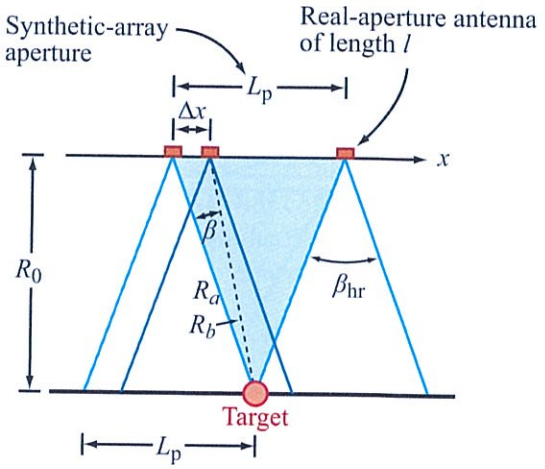
14-4.1 Synthesized-Aperture Point of View

Figure 14-13(a) illustrates the along-track geometry of the SAR measurement of a target T. The radar antenna on the airborne vehicle is shown at three different locations where echo data is collected, and they are labeled A, B, and C. At location A, the forward edge of the beam is first to intercept the target; at location B, the antenna is abreast of the target; and at location C, the aft edge of the beam is just leaving the target. The ***total possible length of the synthetic aperture*** L_p defines the finest resolution that can be obtained with the real antenna in a real-aperture SLAR:

$$L_p = \beta_{hr} R_0, \tag{14.48}$$



(a) Antenna beams illuminating target



(b) Two adjacent elements in synthetic array

Figure 14-13: Synthesized-array concept for SAR.

where β_{hr} is the horizontal or along-track beamwidth of the real antenna and the subscript r is used to distinguish the real-aperture antenna beamwidth from the beamwidth of the synthetic array antenna.

In considering the beamwidth of a synthetic antenna, one must take into account the fact that the phase shift associated with a given target is due to two-way, rather than one-way, travel on the path between the antenna and the target. Figure 14-13(b) illustrates the ranges for

two adjacent elements in an array. The beamwidth is determined by the relative phase shift along the paths between the target and the two elements. Two adjacent elements in the synthetic array, separated by a (small) distance Δx, are shown, with portions of the rays R_a and R_b going from these elements to the target. Since the two elements are close together, the rays may be considered essentially parallel. Hence, the difference in distance along the two rays is

$$R_a - R_b = \Delta x \sin\beta. \tag{14.49}$$

For the real-aperture antenna the phase shift associated with this distance is

$$\Delta\phi_r = k(R_a - R_b) = k\,\Delta x \sin\beta. \tag{14.50a}$$

(real aperture)

For the synthetic-aperture array, the phase shift enters twice: once from antenna to target and again as it returns from target to antenna. Thus, in this case

$$\Delta\phi_s = 2k(R_a - R_b) = 2k\,\Delta x \sin\beta. \tag{14.50b}$$

(synthetic aperture)

For a real (or synthetic) antenna of length $\Delta x = L$, where L is long enough so that the beamwidth is small, the first ***null*** in the azimuth pattern occurs at $\pm\beta_{nr}$ (or β_{ns}) such that

$$kL\sin\beta_{nr} \approx kL\beta_{nr} = \pi \qquad \text{(real aperture)}, \tag{14.51a}$$

$$2kL\sin\beta_{ns} \approx 2kL\beta_{ns} = \pi \qquad \text{(synthetic aperture)}. \tag{14.51b}$$

The corresponding null-to-null beamwidths are

$$\beta_{null_r} = 2\beta_{nr} = \frac{2\pi}{kL} \qquad \text{(real aperture)}, \tag{14.52a}$$

$$\beta_{null_s} = 2\beta_{ns} = \frac{\pi}{kL} \qquad \text{(synthetic aperture)}. \tag{14.52b}$$

Hence, the effective length of the synthetic aperture is $2L$, not just L. Accordingly, *the beamwidth for the synthetic aperture is half that for a real aperture having the same length.*

The along-track beamwidth β_{hs} for the synthetic aperture is therefore

$$\beta_{hs} = \frac{\lambda a_{hs}}{2L}, \tag{14.53}$$

where a_{hs} is the aperture-illumination taper factor associated with the synthetic aperture. If the synthetic aperture has uniform weighting all along its length, then $a_{hs} = 0.88$, as discussed in Chapter 3 for real antennas. However. most antennas, both real and synthetic, have nonuniform aperture illumination; so the value for a_{hs} is usually greater than 0.88. While the simplifying assumption that a_{hs} is 1 is common, we should keep in mind that actual antennas, both real and synthetic, have a factor closer to 1.5.

For the synthetic aperture, the along-track resolution is

$$r_a = \beta_{hs} R_0 = \frac{\lambda R_0}{2L} a_{hs}. \tag{14.54}$$

This expression applies regardless of the actual length of the synthetic aperture used. If the entire ***potential length*** L_p is used, the ***finest possible resolution*** r_{ap} for the SAR is given by

$$r_{ap} = \frac{\lambda R_0}{2L_p} a_{hs} = \frac{\lambda a_{hs}}{2\beta_{hr}}, \tag{14.55}$$

where we used the relationship $L_p = \beta_{hr} R_0$ relating the beamwidth of the real aperture β_{hr} to the along-track width L_p at the range R_0. Thus, *the resolution attainable is inversely proportional to the beamwidth of the real antenna used in the SAR.* This beamwidth of the real antenna is

$$\beta_{hr} = a_{hr}\frac{\lambda}{l}, \tag{14.56}$$

where a_{hr} is the aperture-illumination taper factor for the real antenna and l is its length. This value of β_{hr} may be substituted in Eq. (14.55) to obtain the potential synthetic-aperture resolution in terms of the length of the real antenna and the aperture taper factors as

$$r_{ap} = \frac{\lambda a_{hs} l}{2a_{hr}\lambda} = \left(\frac{a_{hs}}{a_{hr}}\right)\frac{l}{2}. \tag{14.57}$$

Since the aperture-illumination taper factors are often about the same for the both real and the synthetic aperture, so that $a_{hs}/a_{hr} \approx 1$, the usual result is

$$r_{ap} \approx \frac{l}{2}. \tag{14.58}$$

(optimal focused resolution)

▶ Thus, the finest potential synthetic resolution is one half the length of the real aperture. Moreover, the SAR azimuth resolution is independent of range and also independent of wavelength. ◀

The range independence is a consequence of the fact that the length of the synthetic aperture L_p is directly proportional to the range, so a larger range results in a larger aperture, which results in a finer beamwidth for the synthetic aperture and maintains the effective resolution. The reason the resolution is independent of the wavelength is similar: if the wavelength is longer, the synthetic-aperture length in wavelengths is also longer, so that the net result is that the resolution remains the same because the equivalent synthetic beamwidth remains the same. The expression for the resolution of the synthetic aperture in Eq. (14.58) should be contrasted with the real-aperture along-track (azimuth) resolution

$$r_{ar} = \beta_{hr} R_0 = a_{hr} \frac{\lambda R_0}{l}, \tag{14.59}$$

where the resolution increases if either λ or R_0 increases.

The SAR azimuth resolution r_{ap} given by Eq. (14.57) is the best resolution attainable with a synthetic-aperture radar having an antenna fixed to an aircraft or spacecraft flying in a straight line. As described later, finer resolution can be obtained if a scanning antenna is used so that it may look ahead at the target, keeping the target ***spotlighted*** (i.e., the antenna is pointed at the target) as the radar flies by, and continues to look back at the target for some time. This lengthens the synthetic aperture. Of course, an antenna scanned in this fashion achieves finer resolution for a particular target at the expense of neglecting other targets. The scanning motion results in gaps in coverage along the flight path. As described later, flying in a circular path concentric with the target can also achieve fine resolution, though such a path eliminates the possibility of wide-area coverage.

Although the finest possible resolution is achieved by using the entire available synthetic aperture L_p, the designer may choose to reduce the effective resolution to reduce speckle (discussed in Section 14-8) or system complexity. If a shorter aperture is used, the resolution is that given by Eq. (14.54), rather than that given by Eq. (14.57).

14-4.2 Unfocused SAR

One can shorten the synthetic aperture to the point where the phase correction discussed in Section 14-3.2 can be neglected. When this is done, the radar is said to be ***unfocused***. For this situation, all the signals received at the different points in the synthetic aperture can be added in phase without correction. Figure 14-6 illustrates the basic idea for synthetic-aperture corrections. The corrections are needed to reach a focus at the target point T because, as illustrated in that figure, the different path lengths between the radar and the target result in in-phase shifts along the paths that are significantly different. The amount of correction needed at a particular point is

$$\phi_{Ci} = \phi_T - \phi_{Si}. \tag{14.60}$$

For unfocused SAR, we neglect corrections less than $\pi/4$ radians. That is, if

$$\phi_{Ci} \leq \pi/4,$$

ϕ_{Ci} is ignored and all contributions are added in phase. This simplifies processing at the expense of precise focus and resolution.

To derive the conditions under which the unfocused SAR operates, consider Fig. 14-14, where R_{mp} is the ***maximum possible slant range*** to the target, R_m is the ***maximum slant range to the target in the unfocused SAR***, and R_0 is the ***minimum range to the target***. The phase shift associated with the maximum possible range is

$$\phi_{Cmp} = 2k(R_{mp} - R_0). \tag{14.61}$$

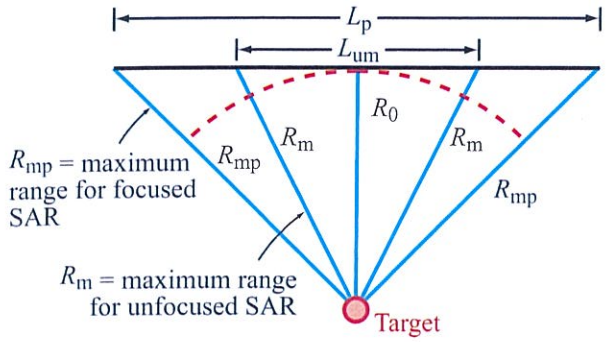


Figure 14-14: Illustration of needed phase compensation.

For the unfocused case,

$$\phi_{Cm} = 2k(R_m - R_0). \tag{14.62}$$

Setting $\phi_{Cm} = \pi/4$ and substituting $2\pi/\lambda$ for k yields

$$R_m - R_0 \le \frac{\pi/4}{2(2\pi/\lambda)} = \frac{\lambda}{16}. \tag{14.63}$$

From Fig. 14-14, it can be seen that

$$R_m^2 = R_0^2 + (L_{um}/2)^2.$$

However, because R_m and R_0 are comparable in length, we can apply the Taylor series

$$R_m = R_0\sqrt{1+\left(\frac{L_{um}}{2R_0}\right)^2} \approx R_0\left[1+\frac{1}{8}\left(\frac{L_{um}}{R_0}\right)^2\right],$$

which leads to

$$R_m = R_0 \approx \frac{L_{um}^2}{8R_0}. \tag{14.64}$$

Using the limiting case of Eq. (14.63) in Eq. (14.64), we find that the ***maximum length of the unfocused synthetic aperture***, L_{um}, is given by

$$L_{um}^2 = \frac{8R_0\lambda}{16} = \frac{R_0\lambda}{2}.$$

Hence

$$L_{um} = \sqrt{R_0\lambda/2}. \tag{14.65}$$

(maximum unfocused length)

If we neglect the effect of taper in the unfocused synthetic aperture, the ***unfocused best possible resolution*** r_{apu} is given by

$$r_{apu} = \frac{\lambda R_0}{2L_{um}} = \sqrt{R_0\lambda/2} = L_{um}. \tag{14.66}$$

(optimal unfocused resolution)

Thus, the azimuth resolution for the unfocused SAR is equal to the length of the synthetic aperture involved. This can result in some simplification of processing for an unfocused system. On the other hand, the resolution for the unfocused case is dependent on wavelength and distance, while for the focused resolution case the resolution is independent of range and wavelength.

14-4.3 Doppler Point of View

From the Doppler beam-sharpening point of view, the along-track resolution of the SAR is given by Eq. (14.38), which can be written as

$$r_a = \frac{\lambda R_0}{2u} B_{Df}, \tag{14.67}$$

where B_{Df} is the Doppler sharpening filter bandwidth. To determine the limitations on the resolution, one must consider how narrow the filter bandwidth B_{Df} can be without loss of information.

The target is visible during the time T_a required for the antenna to travel the distance L_p. This is the time over which the filter operates. As an approximate rule of thumb, the filter bandwidth is greater than the inverse of its integration length, which suggests that

$$B_{Df} \ge 1/T_a. \tag{14.68}$$

The ratio between the aperture length L_p and the time T_a required to travel that length is related to the speed u of the platform:

$$L_p = uT_a. \tag{14.69}$$

Combining these results, the ***finest obtainable resolution*** is

$$r_{ap} = \left(\frac{\lambda R_0}{2u}\right)\left(\frac{u}{L_p}\right) = \frac{\lambda R_0}{2L_p}. \tag{14.70}$$

Using Eq. (14.48) for L_p and Eq. (14.56) for β_{hr} we can write

$$r_{ap} = \frac{l}{2a_{hr}}. \tag{14.71}$$

If the taper of the aperture distribution for the synthetic array is also considered, this becomes

$$r_{ap} = \left(\frac{a_{hs}}{a_{hr}}\right)\frac{l}{2}, \tag{14.72}$$

which, as expected, is the same as Eq. (14.57) derived for a synthesized aperture.

14-4.4 Comparison of Real-Aperture and Synthetic-Aperture Resolution

The along-track resolutions attained when the aperture taper factors are assumed to be 1 are given by Eq. (14.59) for SLAR, by Eq. (14.66) for unfocused SAR, and by Eq. (14.72) for focused SAR. That is,

$$r_{ar} = \frac{\lambda R_0}{l} \quad \textbf{(real-aperture radar)}, \tag{14.73a}$$

$$r_{apu} = \sqrt{\frac{R_0\lambda}{2}} \quad \textbf{(optimal unfocused SAR)}, \tag{14.73b}$$

$$r_{ap} = \frac{l}{2} \quad \textbf{(optimal focused SAR)}. \tag{14.73c}$$

Note the differences in the variation of resolution with range, wavelength, and antenna length for each type of radar. For the real-aperture radar the finest resolution is attained with a small λ, a large l, and the smallest distance to the target R_0. For unfocused SAR, the antenna length l does not matter but the azimuth resolution variation with distance and wavelength is like that of real-aperture SLAR. In both cases, finer resolution comes with a longer antenna. For focused SAR, neither distance nor wavelength matters and the antenna should be made short to obtain the finest resolution. To illustrate, we provide a numerical comparison. Consider an X-band radar with $\lambda = 0.03$ m at two different antenna lengths, $l = 2$ m and $l = 10$ m. Three different ranges are considered: 5 km (a short-range airborne radar), 50 km (a long-range airborne radar), and 500 km (a spacecraft radar). The resulting values for the resolutions are given in Table 14-1.

Table 14-1: Comparison of azimuth resolutions for SLAR, fully focused SAR, and unfocused SAR.

	$l = 2$ m			$l = 10$ m		
Range (km)	5	50	500	5	50	500
Fully focused SAR	1	1	1	5	5	5
Unfocused SAR	8.7	27.4	86.8	8.7	27.4	86.6
Real-aperture radar	75	750	7500	15	150	1500

14-5 Ambiguity Considerations in SAR

► For aircraft SAR, timing ambiguity is seldom a limitation, but for spacecraft SAR systems it is a serious constraint. ◄

The problem is that the pulses must be far apart in time, and hence in along-track distance, to prevent range ambiguities; however, the large Doppler frequency spread for a high-speed spacecraft requires a high pulse repetition rate to be able to sample the Doppler bandwidth without aliasing. These conflicting requirements make ambiguity a problem.

Recall that the ***interpulse period*** T_p is defined as the reciprocal of the ***pulse repetition frequency*** (PRF) f_p:

$$T_p = 1/f_p. \tag{14.74}$$

Figure 14-15 illustrates the situation for a side-looking radar over a plane Earth. The maximum and minimum

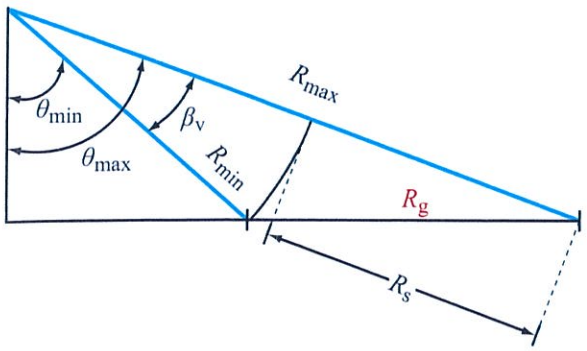


Figure 14-15: Range ambiguity diagram, side view. Radar is moving into or out of the page. Slant-range swath is R_s and ground-range swath is R_g.

ranges within the vertical beam (R_{max} and R_{min}) are shown, along with the vertical beamwidth β_v and the angles of incidence associated with the outer and inner edges of the beam (θ_{max} and θ_{min}). The slant swath width is the difference between maximum and minimum range and is given by R_s. The repetition period is determined by the condition that one cannot have two pulses contributing simultaneously to the received signal. The signals received from the ground from the two pulses would be from different ranges and consequently from different points on the ground. Mathematically this requirement is given by

$$cT_p/2 \geq R_s, \tag{14.75}$$

or in terms of the pulse repetition frequency,

$$f_p \leq c/2R_s. \tag{14.76}$$

Thus, the swath width places a constraint on the maximum possible pulse repetition frequency f_p.

A minimum constraint is placed on f_p by the requirement to sample the Doppler frequency properly. The magnitude of this constraint depends upon the specifics of processing used: ***zero-offset processing*** or ***azimuth-offset processing*** (Fig. 14-16). In zero-offset processing the carrier frequency is down-converted to zero frequency, so that the maximum video frequency is half the total bandwidth of the Doppler spectrum. For this type of processing, the constraint set by the Nyquist sampling theorem is that f_p must be twice this maximum bandwidth, or

$$f_{p0} \geq B_D. \tag{14.77}$$

The bandpass sampling theorem imposes this constraint if the sampling is actually done at a higher intermediate frequency.

Another type of processing, known as ***azimuth-offset processing***, down-converts the carrier frequency to approximately $B_D/2$, so that the maximum negative excursion of the Doppler frequency at baseband occurs at the zero-frequency (Fig. 14-16). For this situation the maximum video frequency is B_D, so the requirement on the pulse repetition frequency is

$$f_{pa} \geq 2B_D. \tag{14.78}$$

That is, the minimum sampling rate is $2B_D$.

We can combine Eqs. (14.77) and (14.78) by defining a ***Doppler processing constant*** a_D as 1 for zero-offset and 2 for azimuth-offset processing:

$$f_p \geq a_D B_D. \tag{14.79}$$

When zero-offset processing is used, both in-phase and quadrature channels are needed so that positive and negative Doppler frequencies may be distinguished from one another. Since two separate channels are used, one must collect two values [one ***in-phase*** (I) and one ***quadrature-phase*** (Q)] per sample at a sampling rate of B_D complex samples per second. Hence, the constraint on the PRF is different from that for the azimuth-offset processing, but in both cases $2B_D$ sample values per second are required.

Substituting Eq. (14.36) into Eq. (14.79) gives

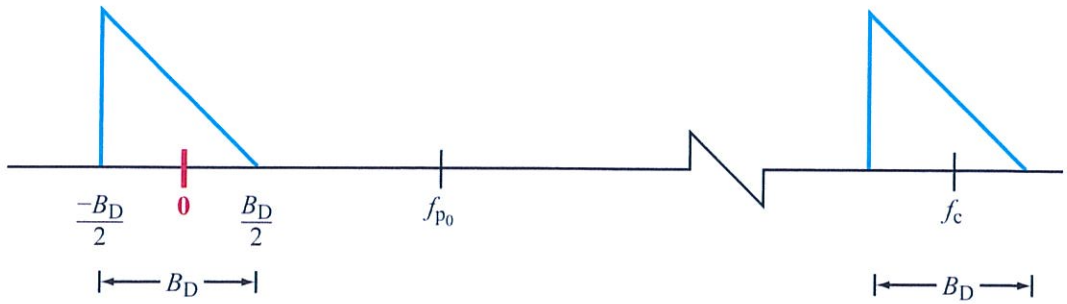
$$f_p \geq 2a_D \frac{u}{\lambda} \beta_{hr}. \tag{14.80}$$

The expressions given by Eqs. (14.76) and (14.80) can be combined into

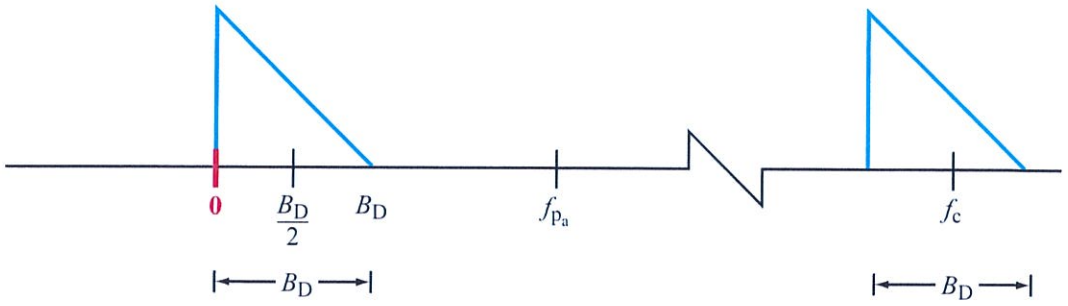
$$\frac{2a_D u}{\lambda}\beta_{hr} \leq f_p \leq \frac{c}{2R_s}. \tag{14.81}$$

Since the real-aperture azimuth beamwidth is

$$\beta_{hr} \approx (\lambda/l)a_{hr},$$



(a) Zero offset



(b) Azimuth offset

Figure 14-16: Doppler bandwidths for (a) zero-offset processing and (b) azimuth-offset processing.

Eq. (14.81) can be written as

$$a_{hr}\left(\frac{2a_D u}{l}\right) \le f_p \le \frac{c}{2R_s} . \qquad (14.82)$$

► The ratio l/u is the amount of time it takes the radar to fly one antenna length. Hence, for $a_{hr} = 1$, the sampling constraint in Eq. (14.82) states that at least one sample (i.e., one pulse) must be taken for each half antenna length of travel for zero-offset processing ($a_D = 1$), and at least one sample for each quarter antenna length of travel for azimuth-offset processing ($a_D = 2$). ◄

Equation (14.82) may be considered a limitation on the slant swath width. That is,

$$R_s \le \frac{cl}{4a_D u a_{hr}} . \qquad (14.83)$$

As a numerical example, consider an aircraft radar with $l = 2$ m and $u = 100$ m/s (about Mach 1). The slant swath is limited to

$$R_s \le 500 \text{ km } (a_D = 1) \quad \text{or} \quad 250 \text{ km } (a_D = 2)$$

for $a_{hr} = 1$. Since the aircraft cannot fly high enough to create a swath this wide, ambiguity is not a concern. For a spacecraft radar, however, the situation is quite different. A typical velocity for a low Earth-orbiting spacecraft is $u = 7.5$ km/s. For the antenna of length $l = 2$ m, we find a swath-width limitation of

$$R_s \le 20 \text{ km } (a_D = 1) \quad \text{or} \quad 10 \text{ km } (a_D = 2),$$

while for the antenna of length 10 m we find

$$R_s \leq 100 \text{ km } (a_D = 1) \qquad \text{or} \qquad 50 \text{ km } (a_D = 2),$$

which represent significant limitations on the swath width that must be considered in the design. (Typically a safety factor is added.) Even without safety factors, the 2 m antenna gives a swath width that is too narrow to be usable from spacecraft altitudes, and the 10 m antenna gives one that is barely usable. Thus, the ambiguity constraint on spacecraft radars is severe.

Safety factors are necessary in the ambiguity constraints because antenna beams do not have perfect edges. In the range direction illustrated in Fig. 14-15, the slant swath shown is limited by the edges of the beam. If one uses the half-power points to represent the edges, this means that significant signals come back to the radar from pulses separated by a time somewhat greater than T_p calculated from Eq. (14.75). If one pulse strikes the ground near the outer edge of the illustrated swath, and another one strikes the ground at a value of θ slightly smaller than θ_{min}, the return from the ground associated with this second pulse may be quite strong because it is only a few dB down in the antenna pattern. Thus one must apply a safety factor so that the antenna pattern can be reduced far enough before another pulse strikes the ground. We call this ***slant-range safety factor*** k_r, and it modifies Eq. (14.76) to

$$f_p \leq c/(2k_r R_s), \qquad k_r \geq 1. \tag{14.84}$$

Similarly, one may have problems in the along-track direction. In Fig. 14-16 we assume that the horizontal antenna pattern of the real aperture drops instantly to zero at the edge of the beam, so that the Doppler spectrum is sharply limited. In fact, the antenna pattern does not drop off so sharply, and consequently the Doppler spectrum is not so sharply limited. Therefore, a safety factor is also needed to assure adequate sampling of the Doppler spectrum. We call this ***along-track safety factor*** k_a. With this safety factor, Eq. (14.79) becomes

$$f_p \geq a_D k_a B_D, \qquad k_a \geq 1. \tag{14.85}$$

Equation (14.82) therefore becomes

$$\frac{2a_D u a_{hr} k_a}{l} \leq f_p \leq \frac{c}{2k_r R_s}. \tag{14.86}$$

Using this result, the constraint on the slant swath width becomes

$$R_s \leq \frac{cl}{4a_D u a_{hr} k_r k_a}. \tag{14.87}$$

Since the factors k_r and k_a are always greater than 1 (typical values range from 1.3 to 1.6), the slant swath width is even less than the values given for the example above. The swath-width limitation problem may be alleviated by using a longer antenna, but a longer antenna is harder to use in space.

To see how the slant swath width relates to the ground swath width for the plane-Earth case (a spherical Earth should be used with spacecraft, but for simplicity here, we use a flat Earth), refer to Fig. 14-15 and note that

$$R_s = h(\sec\theta_{max} - \sec\theta_{min}), \tag{14.88}$$

which may be written as

$$R_s = h[\sec(\theta_{min} + \beta_v) - \sec\theta_{min}]. \tag{14.89}$$

For a small vertical beamwidth (required by the narrow swath in the spacecraft case), this result can be approximated by

$$R_s = \beta_v R_{av} \tan\theta_{av}, \tag{14.90}$$

where θ_{av} is the average value of θ_{max} and θ_{min} and R_{av} is the slant range at θ_{av}. Because of the flat-Earth approximation, the expression given by Eq. (14.90) should be regarded only as a first-order approximation for spacecraft radars.

We may use this expression along with Eq. (14.87) to find a ***minimum acceptable area*** for an antenna, from an ambiguity point of view. Substituting Eq. (14.90) in Eq. (14.87) gives

$$\beta_v R_{av} \tan\theta_{av} \leq \frac{cl}{4a_D u a_{hr} k_r k_a}, \tag{14.91}$$

but the vertical beamwidth may be expressed as

$$\beta_v = \frac{\lambda}{H} a_{vr}, \tag{14.92}$$

where H is the height of the antenna aperture and a_{vr} is the ***real-aperture vertical taper factor***. Substituting

Eq. (14.92) into Eq. (14.91), using $f = c/\lambda$, and rearranging, the antenna area may be expressed by

$$\text{area} = lH \geq \frac{4a_D uR_{av} \tan\theta_{av}}{f} (a_{vr} a_{hr} k_r k_a). \quad (14.93)$$

► Thus, there is a minimum area for the radar antenna. Any smaller antenna will not satisfy the ambiguity constraints. ◄

Note that the ambiguity constraints can be satisfied by antennas with differing ratios of H to l. That is, the area of the antenna is constrained, but not its shape, and consequently not the resolution and swath width—the resolution is specified by l and the swath width by H.

The minimum area constraint rarely impacts aircraft radar, but is important for spacecraft radar. As a numerical example, consider the following parameters:

$$u = 7500 \text{ m/s}, \qquad R_{av} = 1000 \text{ km},$$
$$\theta_{av} = 45°, \qquad f = 10 \text{ GHz},$$
$$a_{vr} = a_{hr} = 1, \qquad k_r = k_a = 1.4.$$

Using Eq. (14.93), the minimum area of the antenna is

$$lH \geq 5.88 \text{ m}^2 \qquad (a_D = 1).$$

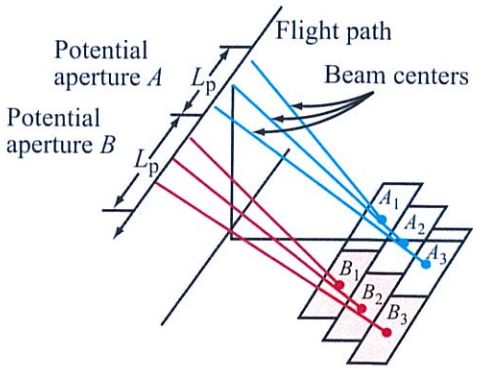
If the antenna length is 10 m, we find that the maximum slant swath width is $R_s \leq 51$ km. The ground swath width is then given by $R_g \approx 51/\sin 45° = 72$ km. This is not a very wide swath; for a smaller angle of incidence, however, the swath is wider. For $\theta_{av} = 20°$ with $R_{av} = 1000$ km, a smaller antenna can be used; i.e.,

$$lH \geq 2.14 \text{m}^2 \qquad (a_D = 1).$$

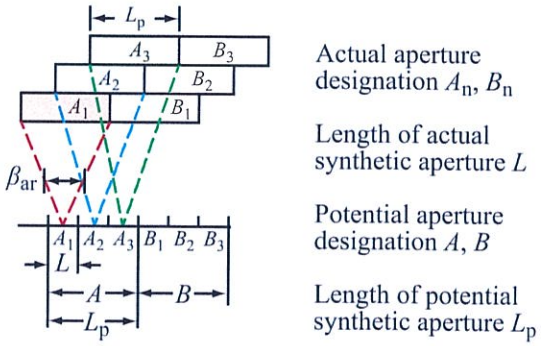
At the smaller angle of incidence, R_s is still 51 km, but R_g is now $51/\sin 20° = 149$ km, which is a more useful swath width.

14-5.1 Scanning Synthetic-Aperture Radar

Because of the ambiguity problem described in the previous section, the swath width for a spacecraft SAR is often less than desirable for some applications. One solution is to use a ***scanning synthetic-aperture radar*** (scanSAR), as illustrated in Fig. 14-17. Figure 14-17(a) illustrates a three-dimensional view of a scanSAR system in which the beam is scanned in elevation to three different positions. During the passage by the potential synthetic-aperture length L_p, one-third of this length is used for each of the three scan positions. In each scan position (subswath), Eq. (14.87) determines the swath width.



(a) Three-dimensional view



(b) Sizes of apertures for scanning SAR

Figure 14-17: Scanning synthetic-aperture radar geometry: (a) three-dimensional view, (b) sizes of aperture for scanning SAR.

► For a given real-aperture antenna length, the swath width obtained by scanning in this example is three times as great as that without scanning. The penalty paid for this is that only one-third of the potential synthetic aperture may be used to obtain along-track resolution so the possible resolution r_a for this three-step scanning SAR is three times as large as that for a single-swath nonscanning SAR. That is, $r_a = 3r_{ap}$. ◄

Figure 14-17(b) illustrates the situation in a two-dimensional plot that allows better presentation of the dimensions. Two potential synthetic apertures A and B are shown, each one subdivided into three segments. The limits of the real-aperture beam are also shown. During the time for passing aperture A_1, the swath A_1 is imaged. At the end of this time the beam scans outward and the radar starts to image swath A_2, which it does during the aperture interval indicated as A_2. At the end of that time the beam scans out farther and swath A_3 is imaged. By the time the beam has completed imaging swath A_3, the radar is at the end of the first potential synthetic aperture A. This means that the beam must be swung back to the original position immediately so that the imaging of swath B_1 may start; otherwise a gap is created between swath A_1 and swath B_1. The procedure continues as the radar goes past the potential aperture B, and so on, so that a complete image is built up with three times the swath width that would be achieved without scanning.

One can see by analogy with the three-step example of Fig. 14-17 that the total distance available for the length of a synthetic aperture for a scanSAR with N_{sc} scan positions is

$$L \leq L_p/N_{sc}. \tag{14.94}$$

Consequently, the along-track resolution obtained with the scanSAR is

$$r_a \geq N_{sc} r_{ap}. \tag{14.95}$$

If one has a fixed antenna length and wishes to image a very wide swath, N_{sc} must be large and a significant sacrifice in resolution must be made. For some applications, more-frequently repeated coverage is more important than resolution, but the trade-off between swath width and resolution should be kept in mind when utilizing scanSAR.

14-5.2 Other SAR Observation Geometries

So far we have only considered SARs that look directly to the side of the spacecraft or the aircraft flight track, which is nominally straight as shown in Fig. 14-18(a). Such systems can create a continuous image strip and are known as ***stripmap-mode*** SAR systems. ScanSAR is a type of stripmap SAR that can be used for multiple swaths parallel to the flight path or diagonal paths as illustrated in Figs. 14-18(b) and (c). As suggested by the discussion in Section 14-4.1, other SAR observation geometries can be used. We consider a few of the most important of these in the following.

Squint-mode SAR

Many SAR systems, particularly aircraft radars for military purposes, look at some other azimuth angle, usually ahead of the aircraft as shown in Fig. 14-19. This is known as ***squint-mode SAR***. SARs operated in squint mode can be stripmap or scanSAR systems. In the example in Fig. 14-19, the antenna beam has been tilted ahead of the side-looking direction, or "squinted," so that its pattern on the ground is an ellipse with its major axis at about 45° to either of the coordinate axes. For comparison, a side-looking antenna is also shown. In both cases the iso-range contours associated with the range resolution are concentric circles that are locally orthogonal to the major axis of the ellipse. Note that in the squinted case the isodops and the iso-range lines are not perpendicular to each other. As a result the shape of the resolution cell is no longer approximately rectangular as it is for the side-looking case. Hence, the range and azimuth resolutions become coupled to one another in the squint-mode case. The ground-range resolution is denoted by the ***pulse range resolution*** r_p rather than r_y, since the ground range is no longer in the y direction. Similarly, the ***Doppler resolution*** r_d, which is in the direction perpendicular to the isodops and associated with the Doppler filter bandwidth B_{Df}, is now used rather than r_a.

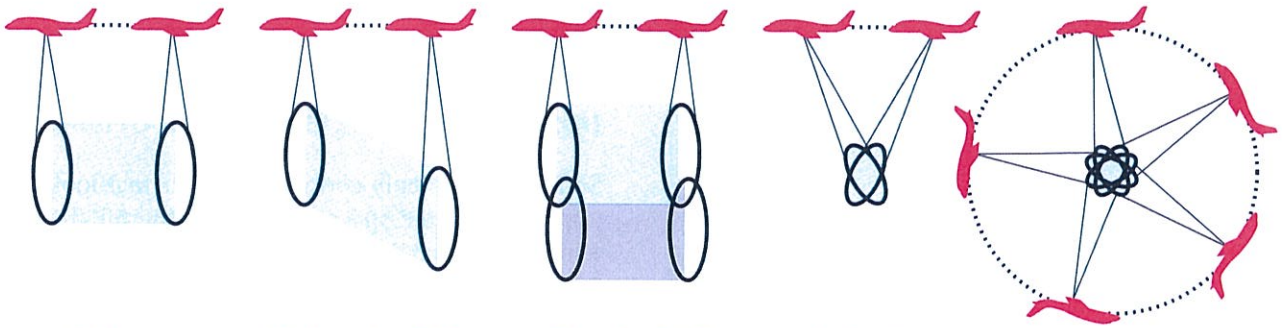


Figure 14-18: Comparison of SAR geometry modes: (a) stripmap, (b, c) scanning SAR, (d) spotlight, and (e) circular.

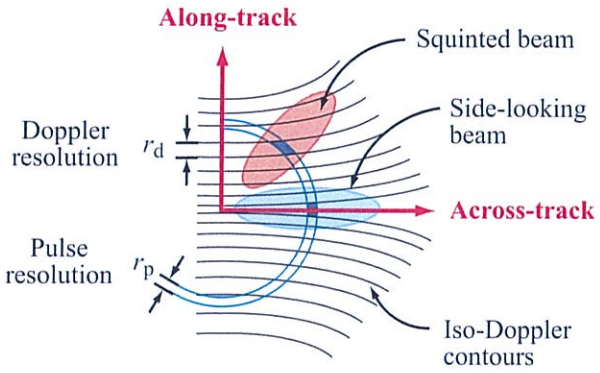


Figure 14-19: Squint-mode geometry.

Because the iso-Doppler and iso-range lines are not orthogonal, the commonly used range-Doppler processing algorithm cannot be used, and more sophisticated algorithms that explicitly account for the range and Doppler coupling must be used instead. We note that "***range walk***" is particularly important in squint-mode SARs. As discussed in Section 14-9.2, range walk occurs when the distance to the target migrates between range resolution bins or cells.

Spotlight-mode SAR

In ***spotlight-mode SAR***, a narrow-beam antenna remains steered toward the target as the aircraft passes by the target so that the target area is illuminated for an extended period, as shown in Fig. 14-18(d). This greatly lengthens the effective aperture, and thus improves the possible resolution at the expense of coverage. The narrow antenna beam has a higher gain than a fixed antenna that would produce the same synthetic aperture and thus provides higher signal-to-noise ratio. Otherwise, the analysis of spotlight SAR is similar to stripmap SAR. In both cases, the length of the synthetic array is limited by coherency over the azimuth angles the target is observed.

Circular SAR

The final SAR imaging geometry we consider is the ***circular SAR*** depicted in Fig. 14-18(e). In circular SAR, the aircraft flies a circular path about the target, which provides continuous observation of the target, which is a form of persistent surveillance. Figure 14-20 shows the geometry of circular SAR in more detail. For an aircraft moving in a circle about the origin $(0,0,0)$ at a speed u, radius r_0, and altitude h, the radar's angular position as a function of time is

$$\theta(t) = ut/2\pi r_0.$$

The time-varying x and y components of the distance from the target point (x_0, y_0) are

$$x(t) = r_0 \cos\theta(t),$$

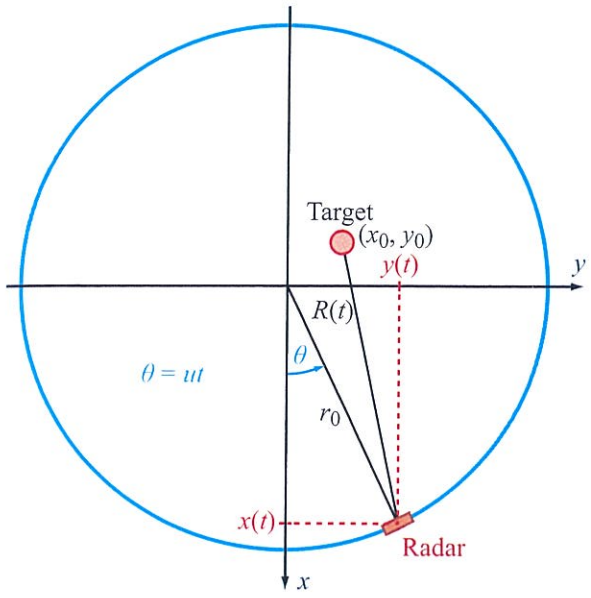


Figure 14-20: Spotlight-mode geometry.

and

$$y(t) = r_0 \sin\theta(t).$$

The distance R between the target and radar is

$$R(t) = \sqrt{(x(t)-x_0)^2 + (y(t)-y_0)^2 + h^2}\,. \qquad (14.96)$$

For simplicity, we consider only a small arc of the circle, which, without loss of generality, is centered at $\theta = 0$. For a small arc, $x(t) \approx r_0$ and $y(t) \approx r_0\,\theta(t) = ut/2\pi$ so that Eq. (14.96) becomes

$$R(t) = \sqrt{(r_0 - x_0)^2 + \left(\frac{ut}{2\pi} - y_0\right)^2 + h^2}. \qquad (14.97)$$

When the circle radius r_0 is significantly larger than the height h, we can use the Taylor's approximation $\sqrt{a^2+b} \approx a + b/2a$ valid for $a^2 \gg b$, which leads to

$$R(t) \approx (r_0 - x_0) + \frac{1}{2}\left(\frac{\left(\frac{ut}{2\pi} - y_0\right)^2 + h^2}{r_0 - x_0}\right)$$

$$\approx r_0 - x_0 + \frac{1}{2}\left(\frac{\left(\frac{u}{2\pi}\right)^2 t^2 - \left(\frac{uy_0}{\pi}\right) t + y_0^2 + h^2}{r_0}\right), \qquad (14.98)$$

where in the last step we used in the denominator the approximation $r_0 - x_0 \approx r_0$ for small angle θ. Grouping constants, this expression has the form

$$R(t) = h_0' - \frac{uy_0}{2\pi r_0}\, t + \frac{u^2}{8\pi^2 r_0}\, t^2, \qquad (14.99)$$

where $h' = r_0 - x_0 + (y_0^2 + h^2)/2r_0$. The expression given by Eq. (14.99) is a quadratic function of time. The phase of the received signal is $4\pi R(t)/\lambda$, so the received signal will have a quadratic phase chirp whose DC value (h') is related to the x position while the linear phase term (frequency) is a function of the y position. This is similar to the azimuth chirp generated by stripmap-mode operation, and matched filtering can be employed to create an image of the targets within the illuminated area. Note that for a target at the center of the circle, the range is constant and so the matched filter has a constant impulse response. Specialized SAR processing algorithms that exploit the geometry of spotlight-mode and circular SAR have been developed (e.g., Jakowatz et al., 1996; Carrara et al., 1995), but they are not considered in this book.

14-6 SAR Power Considerations

Power considerations for synthetic-aperture radar are similar to those for real-aperture radar, although one can express them in many other ways. In terms of the along-track and across-track resolutions r_a and r_y, the power received by a real-aperture radar for a single pulse can be expressed as

$$P_r = \frac{P_t G^2 \lambda^2 \sigma^0 r_a r_y}{(4\pi)^3 R^4}\,. \qquad (14.100)$$

Since the SNR S_n is often of more interest than the received power *per se*, we equate P_r to the noise in the receiver as was done with SLAR in Eq. (14.10),

obtaining

$$S_1 = \frac{P_t G^2 \lambda^2 \sigma^0 r_a r_y}{(4\pi)^3 R^4 k T_0 B F} . \tag{14.101}$$

Note the additional subscript here on S_1, which is used to refer to the SNR for a single pulse. In synthetic-aperture radar, coherent integration takes place over multiple pulses so one must also consider the signal-to-noise ratio for the summation of all of the pulses used in building the synthetic aperture.

The number of pulses in the synthetic aperture is N_s. Since the signals add coherently (and if we temporarily ignore the antenna gain variation with azimuth angle), the signal voltage after summing N_s pulses is related to the voltage received for one pulse V_1^s by

$$V_{N_s}^s = N_s V_1^s . \tag{14.102}$$

Because the noise voltages are random, they do not add coherently and we must use the ensemble average of the square of the noise voltage. Assuming the noise samples are independent, after adding N_s random squared noise voltages, the ensemble noise average is given by

$$\langle (V_{N_s}^n)^2 \rangle = N_s \langle (V_1^n)^2 \rangle . \tag{14.103}$$

▶ Hence, the signal-to-noise ratio after coherently adding N_s pulses is improved over the signal-to-noise ratio for one pulse by a factor N_s:

$$S_{N_s} = \frac{P^s}{P^n} = \frac{(V_{N_s}^s)^2}{\langle (V_{N_s}^n)^2 \rangle} = \frac{N_s^2 (V_1^s)^2}{N_s \langle (V_1^n)^2 \rangle} = N_s S_1 . \tag{14.104}$$

◀

It is useful to express N_s in terms of radar-system parameters. First note that N_s is the ratio of the length of the synthetic aperture to the distance traveled between two pulses. That is,

$$N_s = \frac{L}{u T_p} + 1 \approx \frac{L}{u T_p} , \tag{14.105}$$

where L is the length of the aperture, u is the speed, and T_p is the pulse repetition period. Since L is not a fundamental parameter of the system, we use Eq. (14.54) to relate it to r_a to obtain

$$N_s = \frac{\lambda R a_{hs}}{2 r_a u T_p} . \tag{14.106}$$

Note that this expression includes the weighting term a_{hs}, which accounts for the antenna gain pattern.

Combining Eq. (14.106) with Eqs. (14.104) and (14.101) gives the signal-to-noise ratio after N_s pulses:

$$S_{N_s} = \frac{P_t G^2 \lambda^3 a_{hs} \sigma^0 r_y}{2(4\pi)^3 R^3 k T_0 B F u T_p} . \tag{14.107}$$

This equation suggests that the signal-to-noise ratio for a SAR pixel is independent of the along-track (azimuth) resolution r_a. We may process for the finest possible resolution r_{ap} or for some coarser resolution without changing S_{N_s}. This means that the power requirement for the synthetic-aperture radar is independent of r_a, and is determined only by the across-track (range) resolution r_y. On the other hand, the processor complexity is determined by the fineness of r_a. Thus, if one wishes to obtain a constant-area pixel as indicated in the discussion on speckle in Section 14-8, one can make a trade-off between r_y and r_a so long as their product remains the same.

Pulse compression is always used in SAR, so the average power may be more important than the peak power. For a rectangular pulse, the effective pulse duration is τ_p. For another pulse shape or type, we find an equivalent τ_p by noting that the energy in the pulse for a given shape of pulse may be equated to the energy in an equivalent rectangular pulse by

$$P_{t_{max}} \tau_p = \int P_t(t)\, dt . \tag{14.108a}$$

Using this definition for τ_p, one may replace the peak power P_t in Eqs. (14.100), (14.101), and, more particularly, (14.107) by

$$P_t = P_{av} / \tau_p f_p = \frac{T_p P_{av}}{\tau_p} , \tag{14.108b}$$

where f_p is the pulse repetition frequency, P_{av} is the average power, and $f_p T_p \equiv 1$. We may relate the

bandwidth in Eq. (14.107) to τ_p through a pulse-taper factor a_B that relates B and T_p:

$$B = \frac{a_B}{\tau_p}. \tag{14.109}$$

Substituting these relations into Eq. (14.107), the SNR becomes

$$S_{N_s} = \frac{P_{av} G^2 \lambda^3 a_{hs} \sigma^0 r_y}{2(4\pi)^3 R^3 k T_0 F u a_B}. \tag{14.110}$$

14-6.1 SAR SNR Equation

According to Eq. (14.110), the SNR is strongly affected by the antenna gain, the wavelength, and the range to the target R.

Sometimes it is important to find the SNR for a system with a given set of specifications, in which case Eq. (14.110) is the most convenient form of the radar equation. However, in many cases it may be necessary to find the required average power for specified values of the SNR and other system parameters. For a synthetic-aperture radar, it is quite often more convenient to use the area of the antenna rather than the gain, because the individual components of the area relate to the resolution and the swath width. The effective area A_e for an antenna with gain G is

$$A_e = G\frac{\lambda^2}{4\pi}.$$

Using this expression in Eq. (14.110) and then solving for the average power P_{av} leads to

$$P_{av} = \frac{S_{N_s} 8\pi R^3 k T_0 F u \lambda a_B L_F}{A_e^2 a_{hs} \sigma^0 r_y}. \tag{14.111}$$

One additional important factor is included in Eq. (14.111), namely L_F, the ***loss factor*** that accounts for RF losses and the fact that the gain is not the same everywhere in the beam as it is at the beam center.

As a numerical example consider two radars, one on an aircraft and the other on a spacecraft, both with the same losses, noise figure, signal-to-noise ratio, and wavelength:

$$S_{N_s} = 4 \text{ (6 dB)}, \qquad F = 4 \text{ (6 dB)},$$
$$L_F = 4 \text{ (6 dB)}, \qquad \lambda = 0.03 \text{ m},$$
$$a_{hs} = a_B = 1, \qquad \sigma^0 = 10^{-2}.$$

The different speeds and altitudes of the aircraft and spacecraft radars lead to significantly different minimum area and power requirements:

	Aircraft	Spacecraft
R	20 km	1000 km
u	300 m /s	7.5 km /s
A_e	0.1 m^2	6 m^2
r_y	3 m	30 m
P_{av}	1.5 W	134 W

Note that the area of the spacecraft antenna, 6 m^2, is that required by ambiguity considerations in a middle angular range. Thus, the spacecraft's antenna cannot be as small as that of the aircraft radar. On the other hand, the aircraft radar can use an antenna as large as the spacecraft radar if, for some reason, one wanted to do so. If the aircraft antenna were to have an area of 6 m^2 and resolution r_y of 30 m, the resulting power requirement would be only 43 μW.

Aircraft radars, which operate at shorter ranges and lower speeds, usually use relatively small antennas to obtain fine resolutions, whereas spacecraft radars, operating at longer ranges and higher speed, are required to use larger antennas by both power and ambiguity considerations. Because of the high speed of the spacecraft and the consequent coverage rate for the radar mapping, a coarser resolution is also normally used.

14-6.2 Radiometric Resolution

The radiometric resolution of a SAR is often specified in terms of the ***noise equivalent*** σ^0 (NES). The NES is defined as the σ^0 value that yields an image pixel SNR of 1 (0 dB). Setting $S_{N_s} = 1$ in Eq. (14.110) and solving for σ^0, the NES is given by

$$\text{NES} = \frac{2(4\pi)^3 R^3 k T_0 F u a_B}{P_{av} G^2 \lambda^3 a_{hs} r_y}. \tag{14.112}$$

Note that the NES may vary in range since the SNR is range dependent.

While the requirements for the NES are application dependent, typical design values of the NES are −25 dB to −30 dB. Targets with σ^0 less than the NES cannot be reliably imaged since the pixel intensity value, which ideally is proportional to the σ^0 of the corresponding ground cell, is instead dominated by the system noise floor.

Two other parameters of interest in SAR system design are the spatial ***impulse response function***, sometimes abbreviated as IPR, and the ***integrated sidelobe ratio*** (ISLR). The IPR is the normalized system measurement response function for σ^0, and is a function of the SAR processing applied to the data, including any azimuth and range windows employed. The effective resolution of the SAR image is defined as the 3 dB dimensions of the IPR for one pixel. Note that signal processing windowing used in the SAR processing can be used to trade-off IPR sidelobes and the width of the IPR mainlobe: as sidelobes are reduced, the width of the mainlobe grows.

Typically, the IPR is computed in a small area centered over a particular pixel. It is typically normalized to a maximum value of one. The SAR processor processes and accumulates the pulses that illuminate the area. The pixel value is thus a weighted sum of the radar returns. The IPR is the result of the effective weighting in the SAR processor, the properties of the transmitted signal, and the measurement geometry.

► As the SAR processor impulse response, the IPR can be visualized as a two-dimensional plot of the output of the SAR processor for the pixel when the surface σ^0 is a delta function at the pixel center, as illustrated in Fig. 14-21. ◄

The grid shown in the figure represents the SAR image pixel grid, with the red dot in the center being the location of a corner reflector. Below it is an example IPR function for this area.

In practice, the IPR is often measured by placing a corner reflector or an active radar calibrator in a low-backscatter region of the image and computing a SAR image centered about the reflector. A sufficiently large

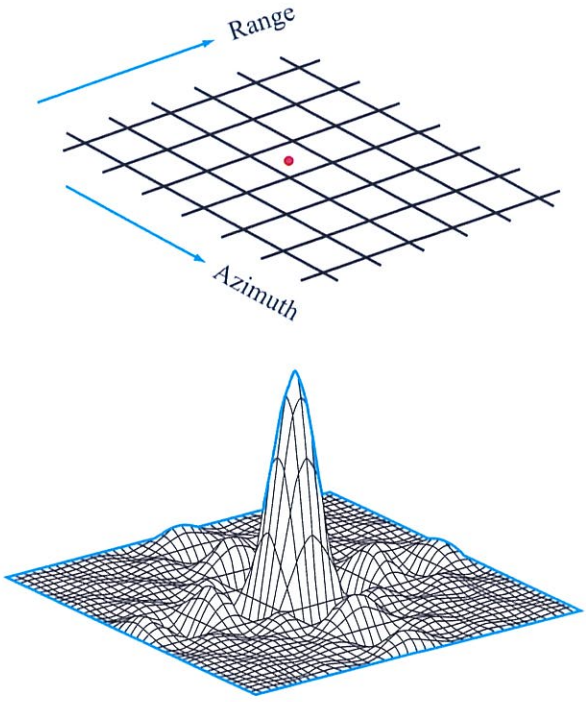


Figure 14-21: Illustration of an impulse response function for a SAR image.

reflector provides a very bright point-like radar target. The image response is the IPR, which often is computed at a higher than normal resolution or interpolated to reveal details as shown in Fig. 14-22.

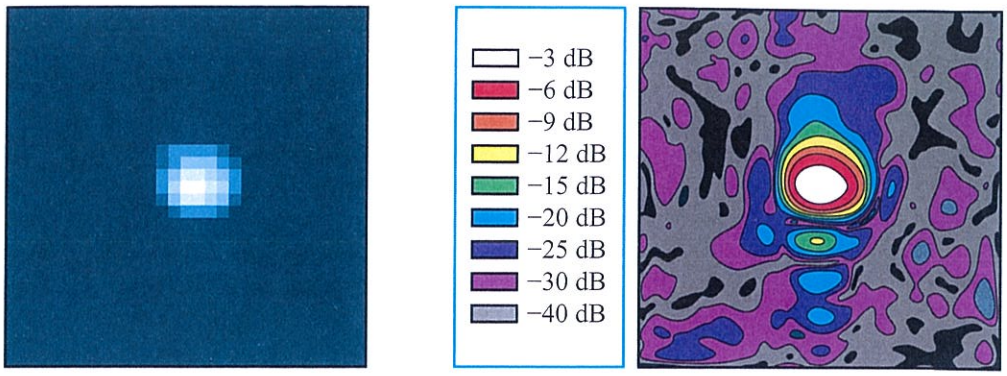
The ***integrated sidelobe ratio*** is defined as the ratio of the system gain over the sidelobes of the pixel response to the total integral of system response:

$$\text{ISLR} = \frac{\int_{\text{sidelobes}} \text{IPR}(x,y)\,dx\,dy}{\int_{\text{everywhere}} \text{IPR}(x,y)\,dx\,dy}. \qquad (14.113)$$

In a well-designed SAR, the ISLR is −13 dB or lower.

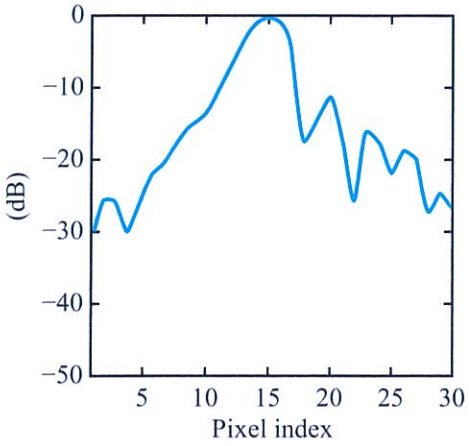
14-7 SAR System Configurations

SAR systems require coherent operation over the length of the synthetic aperture. The block diagram of a SAR

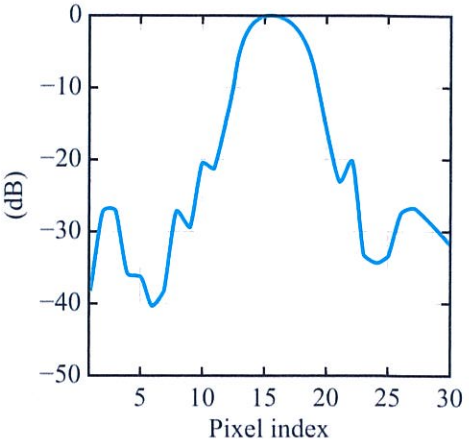


(a) Magnitude SAR image

(b) Contoured and interpolated magnitude image



(c) Vertical slice through peak response



(d) Horizontal slice through peak response

Figure 14-22: Example of a nonideal SAR IPR computed from a corner reflector in a dark scene: (a) magnitude image of area around reflector in a linear scale with the peak normalized to one, (b) contour plot of IPR, (c) plot of estimated IPR through a vertical (range) cut through the peak, which is the effective range compression response, (d) plot of estimated IPR through a horizontal cut through the peak, which is the effective azimuth response.

is similar to other pulse-Doppler radars such as ***moving target indicator*** (MTI) systems. Pulse compression is used for most applications of SAR so that the peak power may be kept down while obtaining a very fine range resolution. The key difference between SARs and other coherent radars lies in the processor rather than in the radar hardware.

Figure 14-23 shows a conceptual block diagram of a modern SAR. A key requirement of the hardware is maintaining phase and frequency coherence. All frequencies in the system are locked to a very stable oscillator at the intermediate frequency. For transmitting, a chirp generator produces a complex (I/Q) pulse with the desired timing and amplitude envelope. The pulse amplitude envelope may be shaped (windowed) or it may be a rectangular pulse. The chirp is fed to a ***single-sideband*** (SSB) mixer, where it is mixed with the local-oscillator signal to produce a pulse

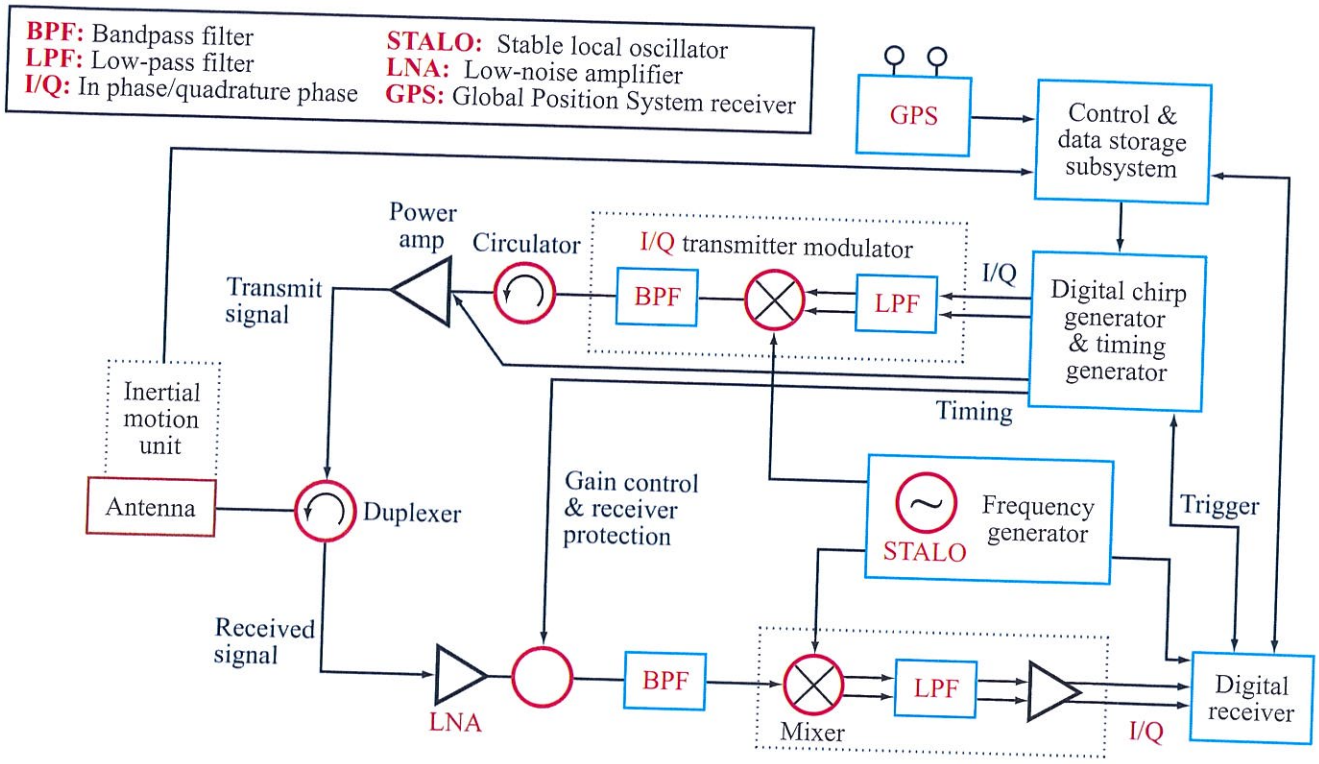


Figure 14-23: Basic block diagram of a SAR.

at the carrier frequency. The carrier-frequency output of the SSB mixer is then amplified to the required level in the transmitting RF power amplifier and sent to the duplexer. It passes through the duplexer to the antenna, and the signal is transmitted.

The received signal picked up by the antenna passes through the duplexer to the low-noise amplifier at the carrier frequency, and is mixed down to the intermediate frequency in the first mixer. The IF signal is amplified and digitally sampled. Information on antenna attitude and position is also collected.

The system shown is a zero-IF system, so it has SSB mixing with two mixers—an in-phase or *I*-mixer, and a quadrature or *Q*-mixer. These mixers combine the IF output of the receiver with an intermediate-frequency LO signal to beat the carrier signal down to zero frequency and the Doppler shifts down to frequencies near zero. In this regard the SAR system is just like any other coherent pulse radar. However, some SARs include circuitry for pre-correcting the IF due to motion errors, while others perform the correction as part of the processing. A more detailed SAR block diagram that includes such an error-correction system is shown in Fig. 14-24. The error-correction system operates on the second local oscillator (at IF) unit of the receiver. Such a correction system can also operate on the first receiver LO, and frequently the correction is accomplished in the processor.

For the system shown in Fig. 14-24, an error-frequency oscillator (typically at a very low frequency) has its output mixed with the IF signal from the stable IF oscillator. The result is an IF signal that has been shifted so that errors in velocity and attitude are converted to zero. The error-offset oscillator receives signals associated with both attitude and velocity errors. The velocity errors are obtained from integrating

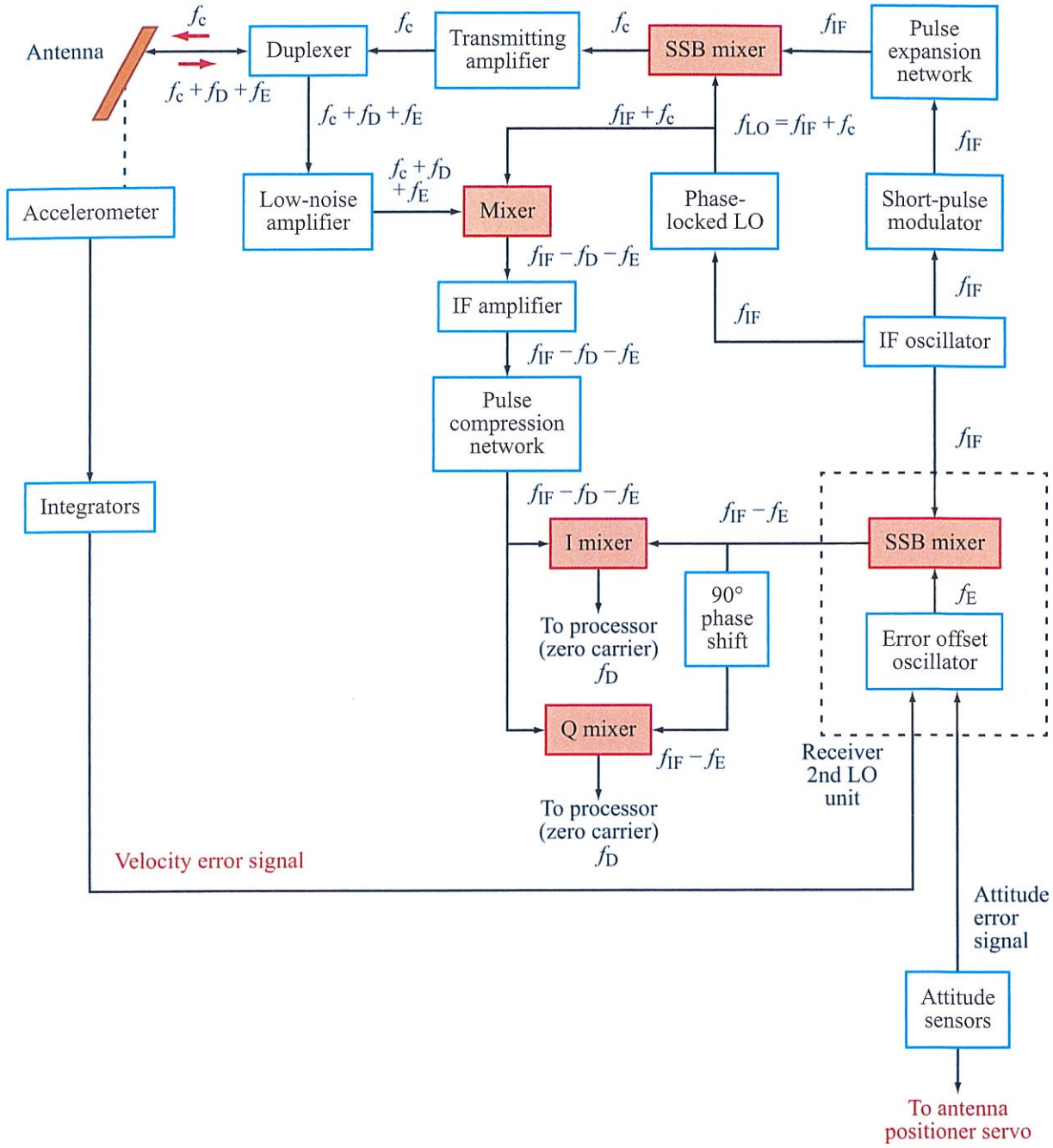


Figure 14-24: Block diagram of a SAR including IF error offset correction.

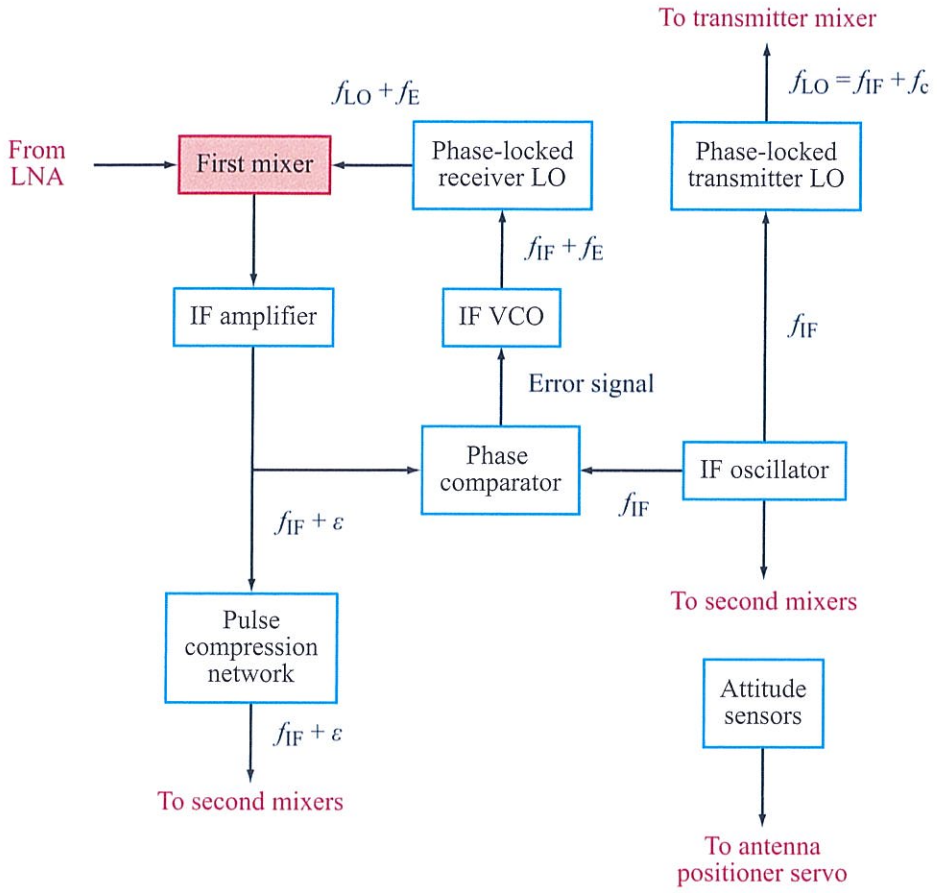


Figure 14-25: Clutter-lock system—IF implementation. f_E is the error offset frequency due to velocity and attitude errors, and ε is the frequency error in the phase-locked loop. For high loop gain, ε should be very small.

accelerometers mounted on the antenna, and the attitude errors are obtained from the aircraft navigation system. The attitude-error sensors produce signals that can make rapid corrections through the error-offset oscillator and slower corrections by repositioning the antenna using the antenna-positioner servo. Thus, the antenna-positioner servo serves the purpose of keeping the antenna attitude errors small, whereas the error-offset oscillator completes the correction. Note that the SSB mixer and error-offset oscillator are shown as the receiver second LO unit. When a clutter-lock system as described below is used, this second LO unit is not required.

The ***clutter-lock*** system uses the centroid of the received Doppler spectrum to adjust the LO frequency so that the IF carrier is located at the center of the IF Doppler spectrum rather than offset from it. Figure 14-25 illustrates such a system. The difference from the radar shown in Fig. 14-23 is in the form of the first local oscillator of the receiver. In Fig. 14-23 the receiver first LO is the same as that providing the transmitter signal. In Figure 14-25 a different receiver LO is used whose frequency is adjusted by a phase-locked loop so that the center of the IF Doppler spectrum is at the frequency of the IF oscillator, f_{IF}. This is accomplished by feeding the output of the IF amplifier to

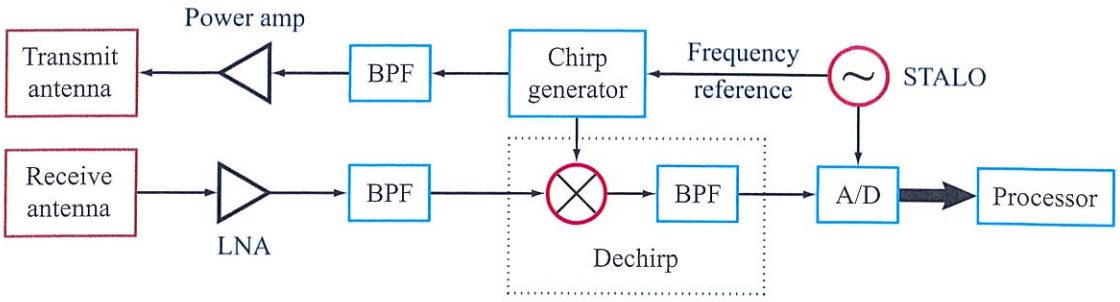


Figure 14-26: Functional block diagram of an LFMCW SAR.

a phase comparator along with the signal from the stable IF oscillator. An error signal is produced that adjusts the frequency of the IF ***voltage-controlled oscillator*** (VCO), which, in turn, adjusts the higher frequency of the phase-locked receiver LO. Adjusting this frequency in the proper manner reduces the difference between the centroid of the IF Doppler spectrum and the frequency f_{IF} to a minimum. Note that the clutter-lock phase comparator occurs before the pulse compression network, thereby allowing the relatively long pulse that contains information from many different ranges, to be used for adjusting the local oscillator. If the output of the pulse-compression network were used, the result would have to be integrated over many range gates to achieve the same stability and independence of individual-range target returns. If the pulse compression ratio is not large enough, such integration may be necessary even with the system shown in Fig 14-25. Of course, the time constant for operation of the phase-locked loop has to be adjusted to a reasonable level so it does not jitter with small variations in target-signal characteristics. A clutter-lock system can also be implemented completely in the processor. Details of the subsystems contained in the boxes of Figs. 14-23 through 14-26 are beyond the scope of this book.

For a low-cost, modest-performance SAR, a linear frequency-modulated continuous wave (LFMCW) system offers a simple design. A block diagram of an LFMCW SAR is shown in Fig. 14-26. Unlike other SAR designs, an LFMCW SAR continuously transmits and receives signals, though this complicates the receiver design due to leakage of the transmit signal into the receiver. Typically, separate transmit and receive antennas are used to maximize the transmit/receive isolation. By employing an LFMCW signal, the SNR is maximized, which can permit the use of a low-power transmitter. In Fig. 14-26 the transmitted LFM signal is mixed with the received signal to dechirp it. While intermediate stages can provide better performance, this simple homodyne design simplifies the hardware by lowering the baseband sampling frequency.

We note that most SAR systems operate in a pulsed mode in which the SAR transmits a short pulse, then waits for the echo. For an aircraft SAR, typically only a single pulse is in flight at a time. A spacecraft SAR, however, may have multiple pulses in flight. In both of these cases, due to the short pulse length, the movement of the SAR during signal transmission or during echo reception from the target is so small that it can be neglected. Thus, the system can be treated as stationary during the pulse (except for computation of the Doppler shift). This is known as the ***stop-and-go*** or ***stop-and-hop*** approximation. The stop-and-go approximation does not apply to LFMCW systems where transmission and reception are continuous and a more detailed analysis is required.

In a chirped LFMCW transmit signal, the frequency of the signal increases from a starting frequency f_0 and spans the bandwidth B at the ***chirp rate*** $\alpha = 2B/T_R$, where T_R is the chirp repetition period. The frequency then ramps back down, as shown in Fig. 14-27. This up-down cycle is repeated at the pulse repetition interval

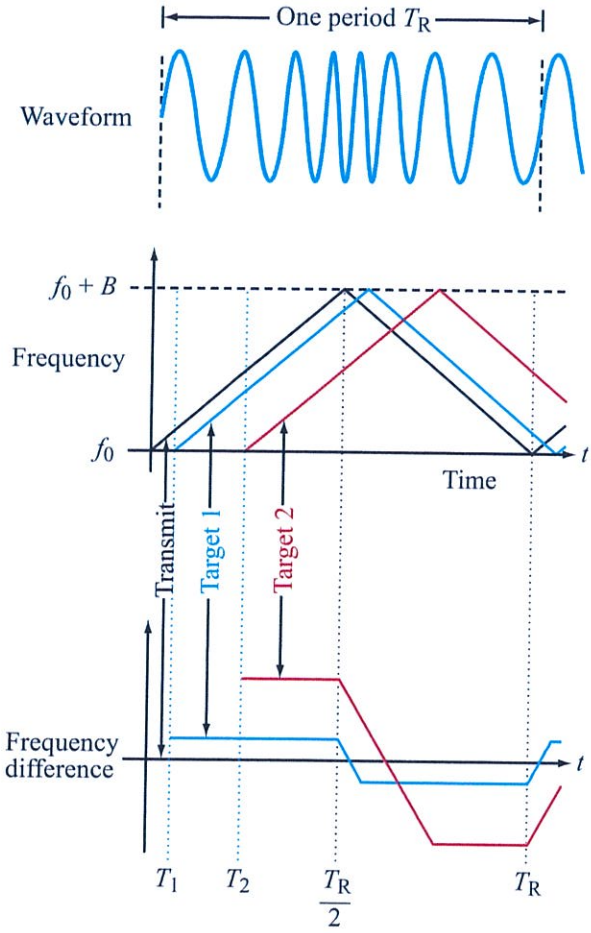


Figure 14-27: Frequency versus time for an LFMCW SAR.

of T_R. Using complex notation, the transmitted up-chirp signal can be expressed as

$$V_t(t) = e^{j(\phi + 2\pi f_0 t + \pi\alpha t^2)}, \tag{14.114}$$

where t is fast time and ϕ is the initial phase. This same signal is repeated for each up-chirp. The down-chirp signal can be expressed similarly to the up-chirp signal, but with $f_0 + B$ as the starting frequency and $-\alpha$ as the chirp rate.

The range R to the target is

$$R(t,\eta) = \sqrt{R_0^2 + u^2(t+\eta)^2}, \tag{14.115}$$

where R_0 is the range of the closest approach to the target, $\eta = NT_R$ with N being the slow-time chirp number, T_R the chirp interval, and u the along-track velocity. The received signal is delayed by the time-of-flight delay T:

$$T = \frac{2R(t,\eta)}{c} \tag{14.116}$$

so that the received signal for a single chirp can be written as

$$V_r(t,\eta) = e^{j(\phi + 2\pi f_0(t-T) + \pi\alpha(t-T)^2)}, \tag{14.117}$$

while remembering that T is a function of time.

The transmit signal is mixed with the received signal and low-pass filtered. This is equivalent to multiplying Eq. (14.117) by the complex conjugate of Eq. (14.114). This dechirping operation removes the carrier and results in the ***dechirped signal***

$$\begin{aligned} V_{dc}(t,\eta) &= e^{-j(2\pi f_0 T - \pi\alpha T^2)} \cdot e^{-j2\pi\alpha T t} \\ &= e^{-j\phi'} \cdot e^{-j2\pi f_{dc} t}, \end{aligned} \tag{14.118}$$

which is a sinusoid whose frequency $f_{dc} = \alpha T$ is proportional to range. Range compression is completed by taking the Fourier transform of the dechirped signal. This is typically done with a Fast Fourier Transform (FFT) after applying a signal processing window to suppress range sidelobes. Each output bin of the FFT corresponds to one range resolution. The range resolution is proportional to the inverse of the chirp bandwidth $1/\alpha T_p$. Motion of the platform during the time of flight delay has the effect of shifting the target in range, which can be corrected by applying a slow-time frequency chirp (de Wit et al., 2006).

14-8 Speckle in Radar Images

As discussed in Section 5-6 and illustrated in Section 5-8, fading and speckle are fundamental issues in radar. It is particularly important in radar imaging. The effect of the random fluctuations in the return signal observed from a distributed target is to produce noise-like speckle in the image. As the radar moves past a target pixel, it may only obtain a single

independent "look" at the pixel, or multiple looks spaced far enough apart so that they represent independent samples. If only a single look is obtained at each pixel, the pixel brightness has a Rayleigh or exponential distribution, depending upon whether amplitude or square-law equivalent detection is used. Further, if the receiver is nonlinear in some way other than square-law, the distribution may take some other form. If only one independent sample is observed, the speckle makes it difficult to determine with any confidence the value of the scattering coefficient for that particular pixel. As demonstrated in Section 5-8, the adverse effects of speckle can be ameliorated to some extent by averaging nearby pixels, though this has the side effect of reducing the image resolution. If a target area is covered by multiple pixels, the pixel average has less variance, which results in a better scattering coefficient estimate. Pixel averaging or low-pass filtering can improve the image interpretability. In the following, we separately consider speckle in SLAR and SAR images.

14-8.1 Speckle in SLAR Images

Since SLAR images tend to have low resolution, in most cases an individual pixel in a SLAR image represents an average of several independent samples. As more independent samples are averaged for each pixel, the degree of speckle decreases. If enough independent samples are averaged into the reported pixel value, an individual pixel yields a good estimate of the scattering coefficient for the underlying terrain. Note that averaging employed to reduce the effects of speckle also averages the noise, which improves the SNR. Hence, precise estimates of the scattering coefficient can be made in the presence of noise if enough independent samples are averaged together, even if the noise level is high.

In this section, we treat two problems in SLAR: calculation of the number of independent samples per pixel and methods for increasing the number of independent samples averaged. To show the number of independent samples per pixel for a SLAR image, we note that the Doppler frequency shift for the backscatter from a target strip at a location x from the side-looking position (Fig. 14-28) is

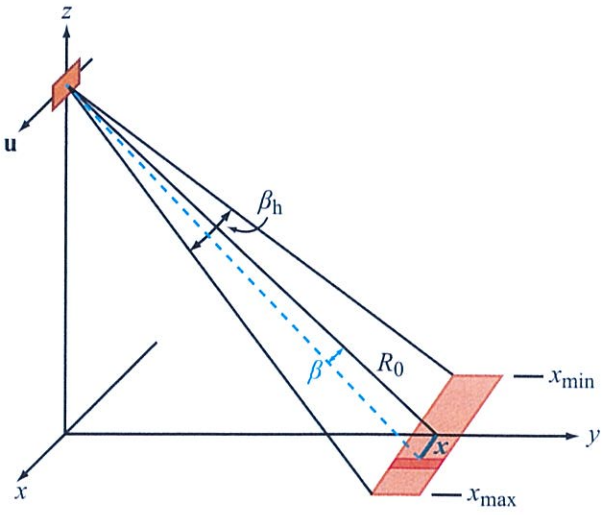


Figure 14-28: Geometry of SLAR fading.

$$f_D = \frac{2u}{\lambda}\sin\beta = \frac{2ux}{\lambda R_0}, \tag{14.119}$$

where u is the velocity of the aircraft, λ is the wavelength, and R_0 is the slant range. The maximum value taken on by x within the beam is

$$x_{max} = \beta_h R_0/2,$$

where β_h is the beamwidth of the antenna in the horizontal direction. Substituting this value into Eq. (14.119), the maximum Doppler frequency is

$$f_{Dmax} = \frac{u\beta_h}{\lambda} = -f_{Dmin}. \tag{14.120}$$

Note that this is also the magnitude of the minimum Doppler frequency f_{Dmin}, since the minimum value is equal in magnitude but opposite in sign to the maximum value. Hence, the ***Doppler bandwidth*** associated with the SLAR is

$$\Delta f_D = 2f_{Dmax} = \frac{2u\beta_h}{\lambda}. \tag{14.121}$$

The time for observing this fading signal, T, is the time that the target remains within the beam for a

given point target, which is related to the along-track resolution r_a,

$$T = \frac{r_a}{u} = \frac{\beta_h R_0}{u} . \tag{14.122}$$

Since, as shown in Chapter 13, the number of independent samples is approximately equal to the product of the Doppler bandwidth and the time of observation, the number N_a of independent along-track samples is

$$N_a = T\,\Delta f_D = \frac{2\beta_h^2 R_0}{\lambda} . \tag{14.123}$$

This expression can be written as

$$N_a = \left(\frac{2\beta_h}{\lambda}\right)(\beta_h R_0) \approx \left(\frac{2\lambda}{\lambda l}\right)(r_a) = \frac{2r_a}{l} = \frac{r_a}{l/2} , \tag{14.124}$$

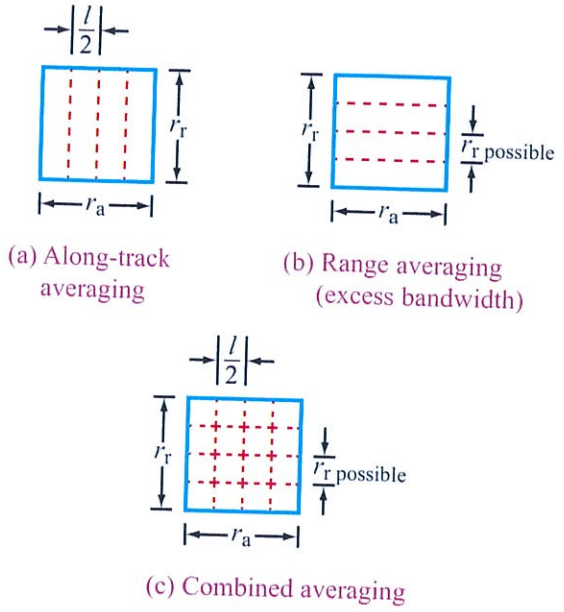
where for simplicity we have used $\beta_h = \lambda/l$, though we note that the effective beamwidth is a function of the antenna taper.

Equation (14.124) is very fundamental and applies to synthetic-aperture radar as well.

► This result states that the number of independent samples available to be averaged in the along-track direction is essentially equal to the ratio of the along-track resolution r_a to half the antenna length $(l/2)$ in the along-track dimension. ◄

Thus, if one can achieve a 30 m resolution with a 3 m antenna, the (approximate) number of independent samples N_a is 30/1.5, or 20. Figure 14-29(a) illustrates this averaging. For a pixel whose dimensions are $r_a \times r_r$, one can think of the along-track resolution r_a as being achieved by superimposing single independent-sample subpixels of width $l/2$, indicated by the vertical dashed lines in the figure.

As an example of the number of independent samples obtainable and the resulting relative standard deviation for the measurement for each pixel, consider the



(a) Along-track averaging

(b) Range averaging (excess bandwidth)

(c) Combined averaging

Figure 14-29: Ways to obtain averaging.

following SLAR numerical design example that operates at two frequencies with the following characteristics:

$$l = 5 \text{ m}$$

$$\lambda_1 = 1 \text{ cm}, \qquad \lambda_2 = 3 \text{ cm},$$
$$\beta_{h1} = 0.002 \text{ rad}, \qquad \beta_{h2} = 0.006 \text{ rad},$$
$$r_{a1} = 2R_{km} \text{ m}, \qquad r_{a2} = 6R_{km} \text{ m},$$

where R_{km} is the slant range measured in kilometers. Using these values in Eq. (14.124), we find that

$$N_{a1} = \frac{2R_{km}}{2.5} = 0.8R_{km}$$

and

$$N_{a2} = \frac{6R_{km}}{2.5} = 2.4R_{km}.$$

From Eq. (5.85), the standard deviation-to-mean ratio of an N_a-look intensity image is given by

$$\frac{s}{\mu} = \frac{1}{\sqrt{N_a}} .$$

We may use these equations to calculate the resolution, the number of independent samples, and the relative standard deviation for the two radar wavelengths operating at two distances, 5 and 20 km:

	$\lambda_1 = 1$ cm		$\lambda_2 = 3$ cm	
R	5 km	20 km	5 km	20 km
r_a	10	40	30	120
N_a	4	16	12	48
s/μ	0.5	0.25	0.29	0.14

These values assume that all possible samples have been averaged. Note that the longer-wavelength system has a poorer along-track resolution than the shorter-wavelength system, but that the normalized standard deviation, and consequently the precision of measurement, is significantly better for the longer-wavelength system because of the larger number of independent samples.

Speckle can be further reduced by averaging independent samples in the range direction.

► ***Range averaging*** can be done after image formation, or the number of independent samples may be increased by transmitting more bandwidth than required for the final range resolution. This is known as ***excess range bandwidth***. ◄

Suppose a pulse of length $\tau_r = 1/B_r$ is transmitted. The possible slant-range resolution is then

$$r_{r\ \text{possible}} = \frac{c\tau_r}{2} = \frac{c}{2B_r}.$$

Now consider integrating N_r returns with this resolution, thereby obtaining a poorer resolution in range but averaging N_r independent samples. The equivalent pulse length associated with the integrated result is

$$\tau = N_r \tau_r.$$

Since the actual range resolution is $c\tau/2$, this resolution can be written as

$$r_r = \frac{cN_r}{2B_r}.$$

This situation is illustrated in Fig. 14-29(b), where the four possible range resolution cells are integrated to form one larger pixel with resolution r_r. It follows that one can express the number of independent samples in the range direction in terms of either the effective pulse duration or the bandwidth:

$$N_r = \frac{r_r}{r_{R\ \text{possible}}} = \frac{\tau}{\tau_r} = \frac{B_r}{B}, \tag{14.125}$$

where B is the reciprocal of τ, the pulse length after integration. The bandwidth needed to achieve the resolution r_r is B, and the actual bandwidth used is B_r. Thus, N_r can be described as equal to the ***excess-bandwidth ratio***. Transmitting the excess bandwidth allows increasing the number of independent samples averaged.

The excess bandwidth ratio may be achieved by one of three methods:

1. Transmitting a short pulse of duration τ_r and then averaging multiple pixels in the range direction.
2. Transmitting a longer pulse of duration τ but modulating the pulse and detecting without correlation or dechirping.
3. Transmitting a long pulse with frequency modulation and dechirping only partially by using a filter whose delay-versus-frequency characteristic does not completely compensate for the chirp. This method is known as ***stretch processing*** (Skolnik, 1980).

In each of these methods, the averaging to achieve better standard deviation, and hence reduced speckle, is achieved by averaging in range (or, equivalently, range frequency) rather than in the along-track direction. If the averaging in frequency is performed continuously, as with methods 2 and 3 above, the resulting standard deviation depends upon the correlation on the frequency axis of the fading spectrum and upon the actual spectrum of the pulse.

Averaging in range and in azimuth may be combined to give a total number of independent samples

$$N = N_r N_a. \tag{14.126}$$

This situation is illustrated in Fig. 14-29(c), where a total of 16 samples is obtained, 4 by averaging along-track and 4 by averaging in range. In reality, one would not need to explicitly compute the pixels of the size of the subcells indicated in Fig. 14-29(c). The azimuth averaging is achieved naturally because of the wide beamwidth of the SLAR, and the range averaging may be attained by transmitting a wide bandwidth on a relatively long pulse and detecting it without dechirping. It should be noted that the transmission of excess bandwidth requires that the receiver have excess bandwidth; and if the receiver has excess bandwidth, the noise level is increased in proportion to the excess-bandwidth ratio. Hence, use of range averaging requires more power in the transmitter to overcome this noise, whereas use of along-track averaging requires no excess power.

14-8.2 Speckle in SAR Images

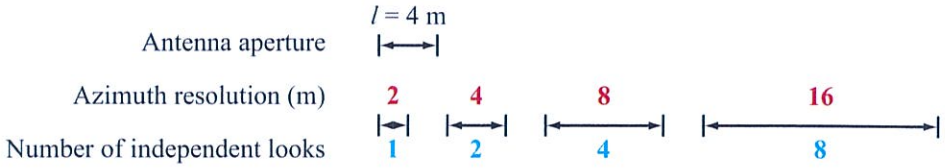
The problem of speckle is usually more severe with synthetic-aperture radars because SAR inherently produces pixels with one independent sample per pixel, which results in the worst-case speckle. To reduce speckle, the designer of the SAR processing system must specifically take into account the need for averaging multiple independent samples.

The speckle reduction schemes discussed previously for SLAR can be used for SAR. However, SAR processing provides additional options for speckle reduction, including high-resolution pixel averaging and ***subaperture processing***. Figure 14-30 illustrates some examples of the two approaches. In Fig. 14-30(a), a scale drawing compares the length of the effective antenna aperture and the along-track resolutions, ranging from 2 m for one independent look to 16 m for eight independent looks. In Fig. 14-30(b) the method of processing for the finest resolution r_{ap} and then averaging is illustrated. To do this, multiple along-track apertures must be processed simultaneously, each one resulting in a 2 m pixel for one of the four pixels shown. In Fig. 14-30(c) the synthetic aperture L_p is divided into four pieces or subapertures. Because the subapertures are shorter, the resolution of their pixels are coarser. Each of the subapertures is processed separately to obtain a pixel with an 8 m along-track dimension. The intensities (magnitude power) for the four pixels are then averaged. The processing involved in the method of (b) uses $4N_p$ pulses, whereas each subaperture look in Fig. 14-30(c) uses $N_p/4$ pulses for a total of N_p pulses. Thus, less processing is required for method of (c) than that of (b). Furthermore, the subaperture processing in (b) might be complicated by dealing with range-curvature and it is more susceptible to range-walk problems than (c). Finally, there is a difference in the ***depth of focus*** for the two cases (discussed in Section 14-9.3). Thus, one normally prefers to use shorter synthetic apertures and add the independent samples, rather than using long apertures and adding the smaller-size pixels. Some users, however, prefer to obtain the smaller pixels on the assumption that for certain applications the highest-resolution result may be superior. It allows for end users to make their own trade-off between resolution and speckle without returning to the raw SAR data.

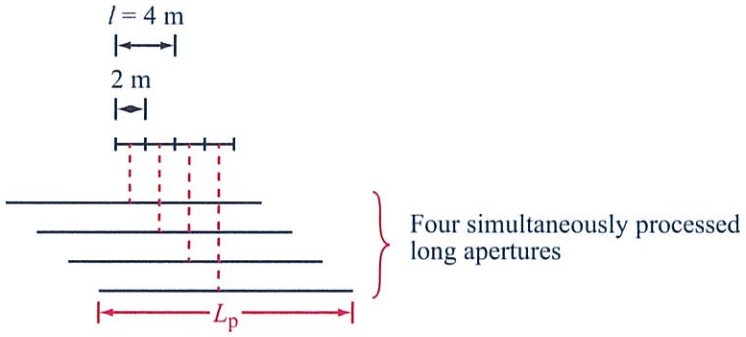
Figure 14-31 illustrates subaperture processing from the Doppler point of view. A full resolution requires the full bandwidth B_d. In subaperture images, separate images are created at lower resolution using bandpass filters that span the bandwidth of each subaperture. In this case, four nonoverlapping bands are used to cover the Doppler bandwidth. Incoherently summing them reduces speckle with $N_a = 4$.

14-9 Introduction to SAR Processing

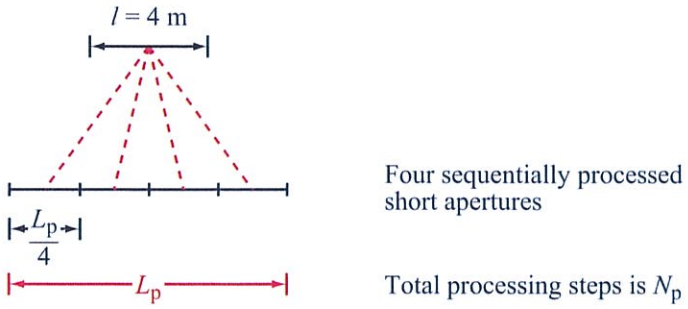
SAR processing is very complicated, partly because of the variety of options available and partly because of the complexity of the required calculations. In this section we describe a few of the processing algorithms, along with background information helpful in understanding SAR processing, notably the spectra that arise in the SAR system and the concept of depth of focus. As most SAR processing is now done digitally in computers, we consider only digital SAR processing algorithms. For the purpose of illustration, we present simplified examples of two SAR processing algorithms: the frequency-domain ***range-Doppler algorithm*** (RDA) and the ***time-domain backprojection*** (TDBP) algorithm.



(a) Example of 4 m antenna showing potential number of independent looks



(b) Obtaining 4 independent samples by processing to finest possible resolution and then averaging. Number of pulses used per aperture is N_p. Total processing steps is proportional to $4N_p$.



(c) Obtaining 4 independent samples by processing directly to final resolution. Number of pulses used per aperture is $N_p/4$ and total number of processing steps is proportional to N_p.

Figure 14-30: Obtaining independent looks with SAR.

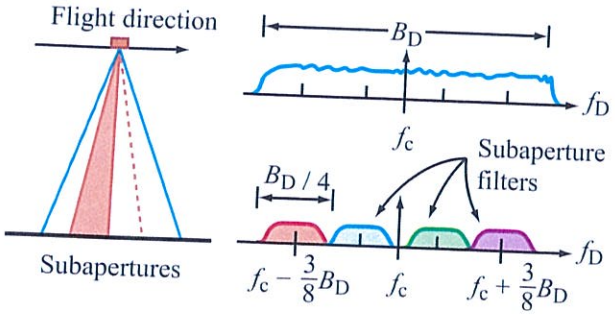


Figure 14-31: Subaperture processing from Doppler point of view.

Additional details on SAR processing are available in books dedicated to SAR processing algorithms (e.g., Cumming and Wong, 2005; Curlander and McDonough, 1991).

In line with common usage, we use the term azimuth instead of "along-track" to refer to the along-track direction and the processing associated with the formation of the synthetic beam. The term ***range processing*** refers to cross-track or perpendicular to along-track and ***azimuth processing*** refers to along-track or azimuth.

Figure 14-32 provides a block-diagram of the steps used in SAR processing. Here part (a) shows a general block diagram for the type of processing in which range compression occurs before azimuth compression, which is most common, but the alternative block diagram of (b) is also used with certain forms of processing. In (a), pulse compression is done first. Following range compression, and possibly using the Doppler-centroid information, the signal undergoes azimuth compression in which some method is used to process the phase-coherent signals to achieve fine azimuth resolution. After azimuth compression, the resulting complex images may be detected by computing the magnitude-square or amplitude of each pixel and then multilook averaging. In the alternative block diagram of Fig. 14-32(b), the only difference is that azimuth compression takes place before range compression.

14-9.1 SAR Signal Spectra

SAR has been described as the "marriage of radar and signal processing." Hence, it is instructive to understand the signal spectra in the synthetic-aperture system.

Different processors operate with different offsets of the carrier frequency from zero. Some processors beat the carrier frequency down to zero, and these must have means for distinguishing positive and negative frequencies. Others use a carrier frequency converted to near zero (***azimuth offset***), and still others use a carrier frequency somewhat higher (***range offset***).

The transmitted radar signal voltage is a sequence of repeating pulses that can be written in the form

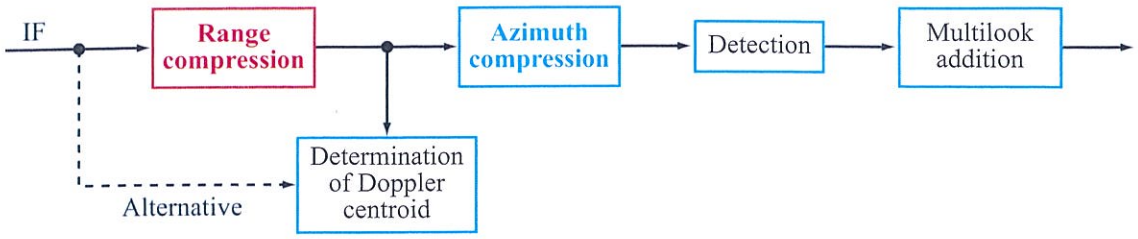
$$v_t(t) = \Re e\left\{\sum_{i=-\infty}^{\infty} V_0\, p(t - iT_p)\, e^{j\omega_c t}\right\}, \qquad (14.127)$$

where V_0 is the peak magnitude and $p(t)$ is the modulation or shape factor (envelope) for the transmitted pulse. The carrier frequency f_c corresponds to an angular carrier frequency $\omega_c = 2\pi f_c$. In the frequency domain this time series may be represented by

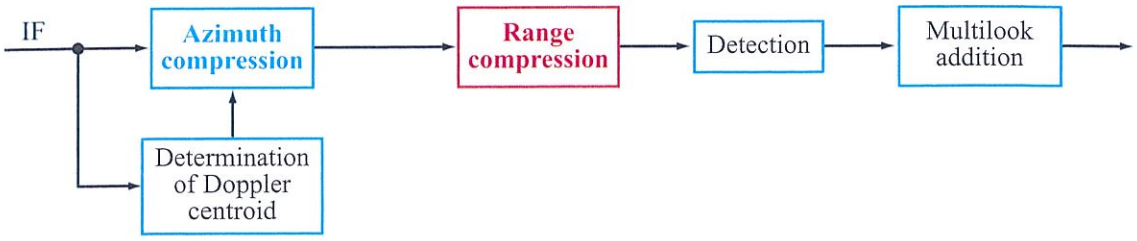
$$\mathbf{V}_t(f) = V_0 \sum_{k=-\infty}^{\infty} \mathbf{P}(kf_p - f_c)\, e^{-j2\pi k/T_p}, \qquad (14.128)$$

where f_p is the repetition frequency $(1/T_p)$ and $\mathbf{V}_t(f)$ and $\mathbf{P}(k)$ are the Fourier transforms of $v_t(t)$ and $p(t)$, respectively. The line spectrum is illustrated in Figure 14-33(a), where both the baseband and RF spectra are shown. The amplitude of the individual lines is governed by the spectrum of the envelope of the transmitted signal $\mathbf{P}(f)$. The baseband signal is shown on the left (centered at zero), and the RF signal is shown on the right as a translation of $\mathbf{P}(f)$ from center frequency zero up to center frequency f_c. For an infinite number of pulses, each line in this spectrum is a continuous sine wave at a particular frequency. If a finite number of pulses are transmitted, each line is then representable by a sinc function $[(\sin x)/x]$ whose width is the reciprocal of the total duration of the pulse train.

Since the transmit signal has a relatively narrow bandwidth compared with the carrier frequency, the Doppler shift of the received signal is essentially the



(a) General block diagram



(b) Alternative general block diagram

Figure 14-32: Steps in SAR processing.

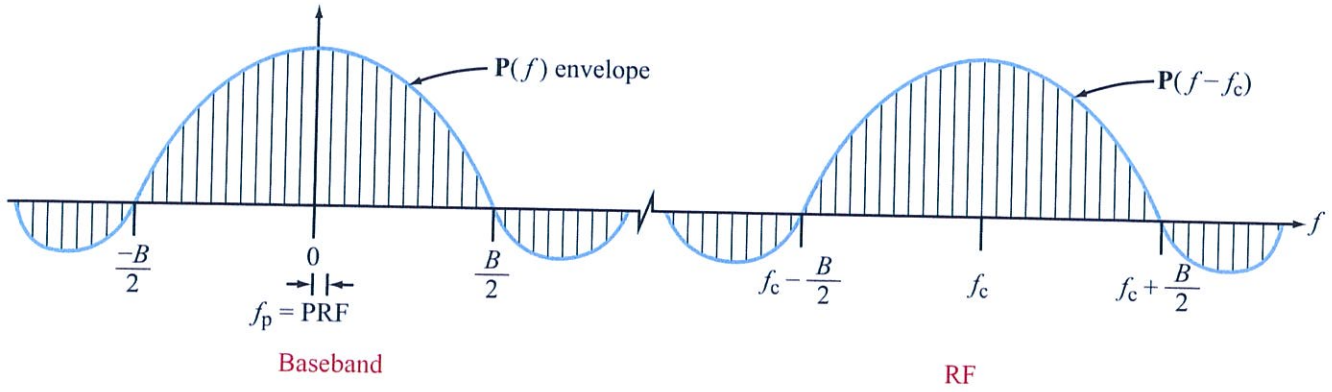
same for each of the frequencies of interest. (In reality, the Doppler spectra for the different lines differ slightly, but this effect is generally negligible.) Figure 14-33(b) presents an expanded view of the received signals in the neighborhood of the carrier frequency. An asymmetric Doppler spectrum is indicated symbolically by triangles, since one normally expects that there will be different Doppler amplitudes ahead of the sideways direction (positive Doppler shift) and behind the sideways direction (negative Doppler shift).

The SAR data may be collected with different frequency offsets in hardware, but is frequently processed with either one of two different frequency offsets: zero-offset and azimuth-offset. Range-offset is yet another type of offset. We describe each briefly.

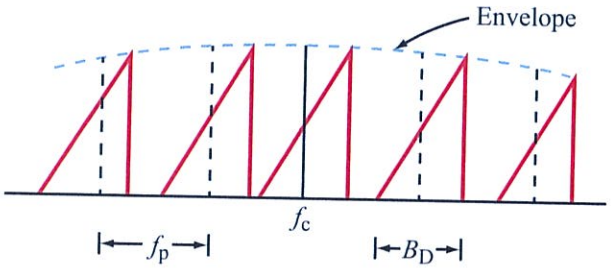
Zero-offset processing

Zero-offset processing is commonly used in digital SAR processors and requires complex (I and Q channel) signals and computation. For this case the signal is mixed with a local oscillator at the carrier frequency and the signal centered at zero frequency is retained as indicated in Figure 14-34(a). Zero-offset processing has the advantage of minimizing the sampling rate required for digitizing the data. In Figure 14-34(a) the Doppler-frequency spectra are now centered at zero compared with those in Figure 14-33(b). Each of these asymmetric spectra consists of summations of all of the different target Doppler chirps that are present at any particular instant. In zero-offset processing, normally only the spectral components centered about zero frequency are used, and low-pass filters are used to eliminate all of the components centered about harmonics of the PRF.

The PRF must be greater than or equal to the Doppler bandwidth, as indicated in the discussion on ambiguities in Section 13-9.2. This ensures that the spectral regions associated with the different harmonics of the PRF do not overlap and the Nyquist criterion is satisfied. Unfortunately, Doppler frequencies associated with sidelobes of the horizontal antenna pattern may extend to such high frequencies that they overlap the



(a) General view of baseband and RF signals



(b) Expanded view of RF received signals

Figure 14-33: Spectra of transmitted and received signals.

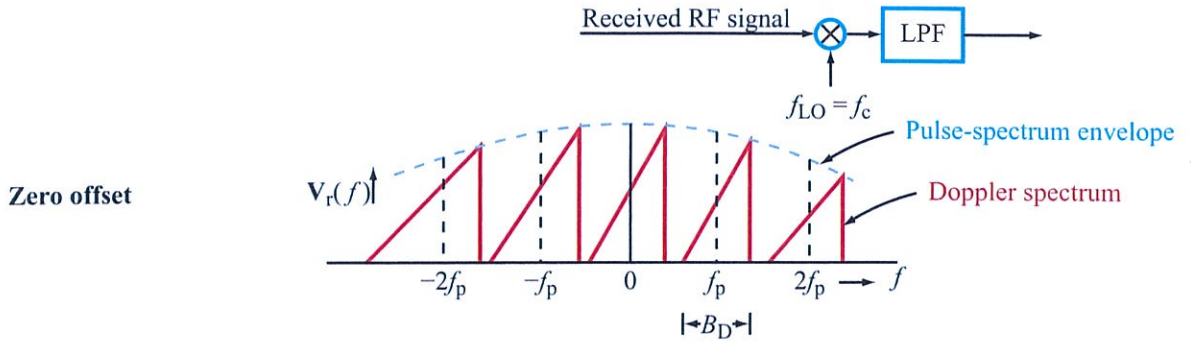
frequency range of the mainlobe below and above multiples of f_p. This results in signal "aliasing"; aliased signals produce artifacts in product images.

Azimuth-offset processing

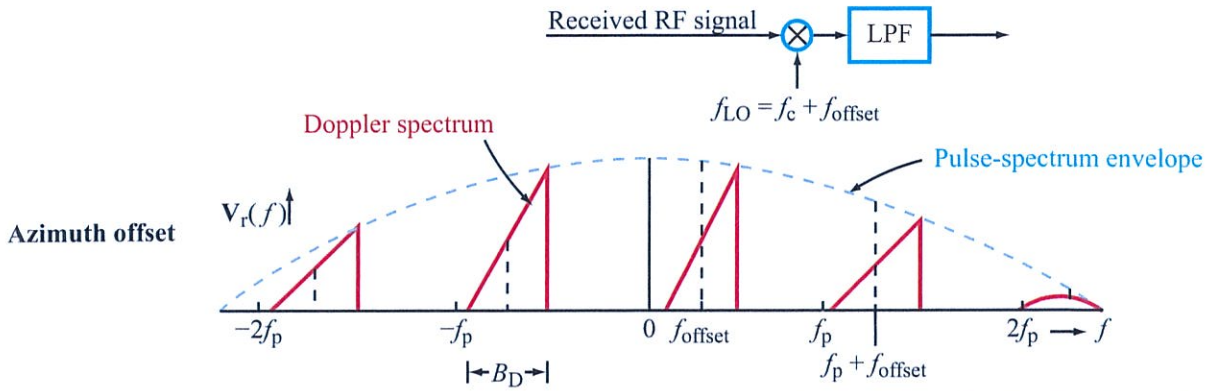
Figure 14-34(b) illustrates the case of azimuth-offset processing. In this situation the offset frequency is just high enough so that the entire Doppler spectrum associated with the carrier is above zero frequency. That is, the offset frequency is selected so that the maximum negative excursion of the Doppler spectrum just touches the zero-frequency point. Hence the offset frequency in this case is of the order of half the Doppler bandwidth. Since higher frequencies are present if the central Doppler spectral region is used, the sampling rate must be twice as high as with zero offset. That is, the Nyquist criterion requires a sampling rate of f_p to accommodate the maximum Doppler frequency, as indicated in Fig. 14-34(b). This type of processing has often been used for optical processing.

Range-offset processing

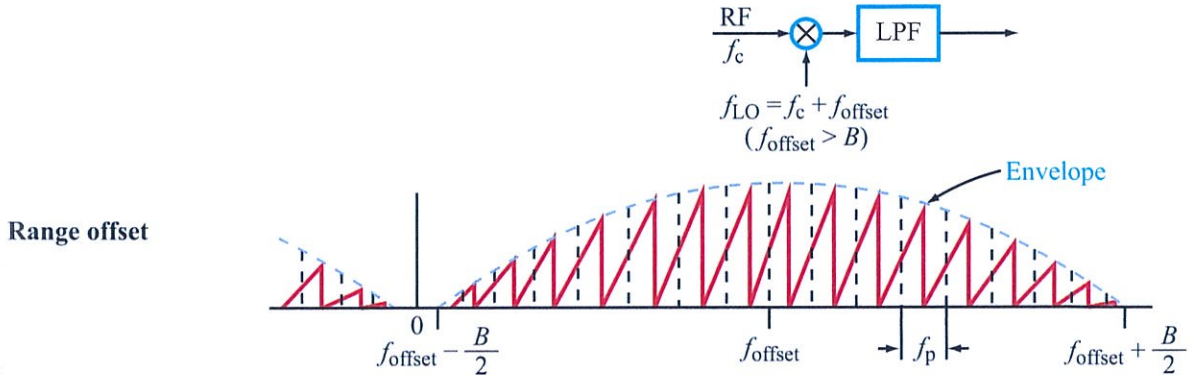
Figure 14-34(c) shows the third type of offset: range-offset processing. Here the offset carrier frequency is selected at approximately half the bandwidth of the transmitted pulse, so that the entire envelope of the transmitted pulse signal is above zero frequency. Range-offset processing is used when simultaneous processing of range and azimuth chirps is to be accomplished, as is often done in optical processing, and also when the azimuth processing is to be done in a range-sequential manner. The bandpass sampling theorem states that the PRF required to adequately sample this situation need only equal the Doppler bandwidth B_D, provided proper



(a) Zero-offset spectrum after mixing. *I*- and *Q*-channels are needed to distinguish positive and negative Doppler frequencies. [$f_p \geq B_D$ to satisfy Nyquist criterion.]



(b) Azimuth offset spectrum after mixing. Only one channel is needed. [$f_p \geq 2B_D$ to satisfy Nyquist criterion.]



(c) Range-offset spectrum after mixing. The need for *I*- and *Q*-channels depends on processing method. [$f_p \geq B_D$ to satisfy Nyquist criterion.]

Figure 14-34: Spectra associated with frequency offsets.

relations are maintained between the sampling rate and the carrier and pulse-repetition frequencies. Thus the PRF for range offset may be the same as that for zero offset, whereas that for azimuth-offset processing must be twice as great.

14-9.2 Range Migration

Range migration is an effect important to both SLAR and SAR. In stripmap SAR, as the platform flies by the target, the slant range to the target varies with time. Referring to the geometry in Fig. 14-35(a), the range to the target as a function of time is

$$R(t) = \sqrt{y_t^2 + h^2 + x_t^2(t)} \,. \tag{14.129}$$

Thus, $R(t)$ is a hyperbolic function of time; i.e., $R(t)$ versus time traces one side of a hyperbola as illustrated in Fig. 14-35(b). This curve is sometimes called the "***range smile***." The amount of change in range to the target over the period of time of the synthetic aperture is known as the ***range curvature***.

In processing SAR data, we are primarily interested in the portion of time over which the mainlobe of the antenna observes the target. Two cases are shown in Fig. 14-35: side-looking and squint mode. The range varies from R_1, which is the range at the edge of the beam, to R_0, which is the point of closest approach, to R_2, which is the range at the other beam edge. The change in range is R_C for the side-looking mode and R_w for the squint-looking mode. These are the ***range curvatures*** for these cases. For narrow beamwidths or coarse range resolution, R_C may be less than r_y, which means that the target remains in the same range resolution bin. When R_C or R_w are greater than the effective range resolution r_y, the target migrates through multiple range resolution elements or bins. This is termed ***range migration*** or ***range walk***. While range migration affects all SLAR and SAR systems, it is most significant in SAR when the synthetic aperture is long compared with the distance to the target, and also in squint-looking antenna systems.

Some SAR processing algorithms compensate for range migration, while others do not. If the SAR processor assumes that the target remains in a single range bin but there is significant range migration, the image quality is degraded since the target becomes spread over multiple pixels in range.

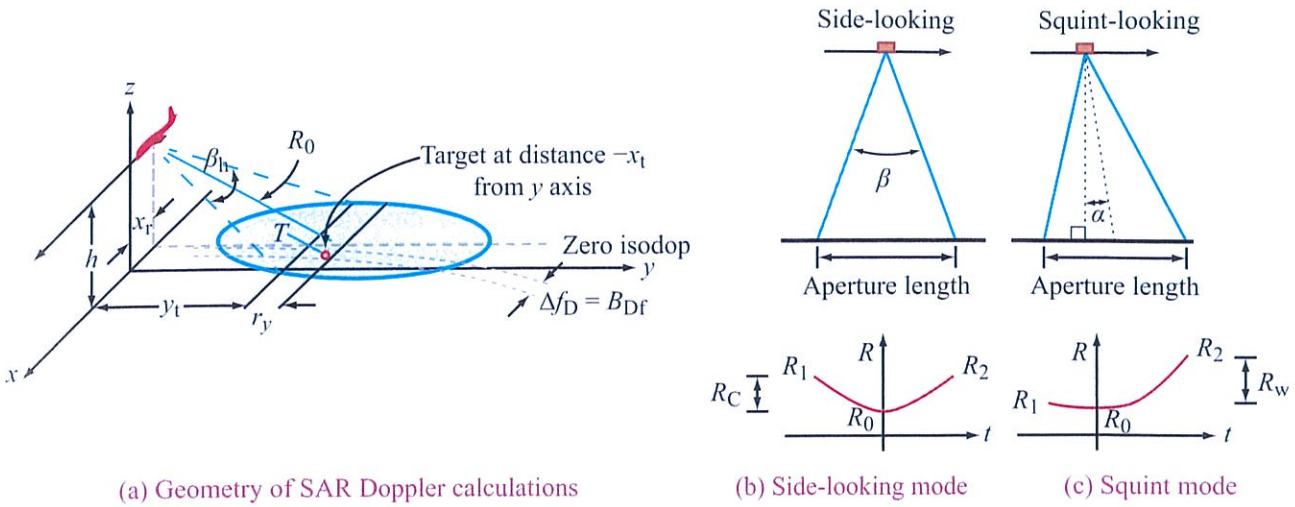
14-9.3 Depth of Focus

The ***depth of focus*** describes the distance over which an image remains in adequate focus. In synthetic-aperture processing, it describes mismatch tolerance between the assumed target range and its actual range. Within the depth of focus, a fixed reference function corresponding to a point target at a particular range can be used in the processor.

> ▶ The minimum number of distinct reference functions required to focus the SAR image is given by the ratio of the swath width to the depth of focus. ◀

If the depth of focus is great enough, only one reference may be required for all the ranges used, which could greatly simplify the processor: if only a single reference function is required, the phase shifters used in Fig. 14-6(b) may be the same for all swath locations or only a single Doppler sharpening filter or correlation reference function is required. However, if the depth of focus is small, many different reference functions must be used in the processor to ensure the image is properly focused over the full swath.

The depth of focus is also useful for evaluating the sensitivity of the image to errors or uncertainty in the motion of the platform. If the change in range due to nonideal platform motion is less than the depth of focus, the image remains in focus. The idea behind the depth of focus is to define a level of error tolerance in the processing. The error tolerance is often defined in terms of the range curvature R_C, the change ΔR in range between the radar and target over the synthetic aperture. The depth of focus is a span of target ranges ΔR_D such that the difference between the actual target range curvature and the range curvature of the reference target is less than a phase tolerance threshold. The most common criterion is that the phase error over the synthetic aperture be less than $\lambda/4$ or $\lambda/8$.



(a) Geometry of SAR Doppler calculations (b) Side-looking mode (c) Squint mode

Figure 14-35: Range curvature with side-looking and squint modes.

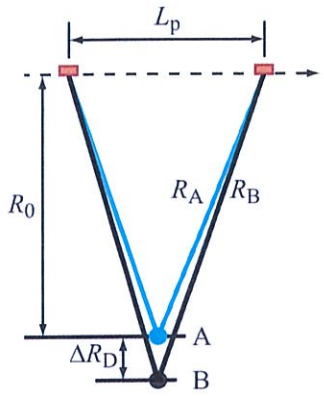


Figure 14-36: Depth of focus geometry. The actual target is at point A. The reference function is computed for the target at point B.

Figure 14-36 illustrates the depth of focus geometry, where one target (A) represents the actual target and the other (B) represents the reference target position. The worst-case range error occurs at the edges of the synthetic aperture. The range curvature for target A is R_{CA}:

$$R_{CA} = R_A - R_0, \tag{14.130}$$

where R_0 is the range of closest approach. In Fig. 14-36, L_p is the maximum possible length of the synthetic antenna. For a side-looking SAR, the point of closest approach occurs halfway through the synthetic-aperture length, so that

$$\begin{aligned} R_{CA} &= \left[\sqrt{\left(\frac{L_p}{2}\right)^2 + R_0^2}\right] - R_0 \\ &= \sqrt{R_0^2\left[\left(\frac{L_p}{2R_0}\right)^2 + 1\right]} - R_0 \\ &= R_0\left[\left(\sqrt{\left(\frac{L_p}{2R_0}\right)^2 + 1}\right) - 1\right]. \end{aligned} \tag{14.131}$$

Using a first-order Taylor series approximation for the square root,* it follows that

$$R_{CA} \approx \frac{L_p^2}{8R_0}. \tag{14.132}$$

*For $c \ll 1$, the first-order Taylor series expansion of $\sqrt{1+c}$ is $\approx 1 + (c/2)$.

The range to the point of closest approach for target B is $R_0 + \Delta R_D$, so R_{CB}, the range curvature for target B, is

$$R_{CB} = R_B - (R_0 + \Delta R_D). \tag{14.133}$$

Then, by the same approximation that led to Eq. (14.132),

$$R_{CB} \approx \frac{L_p^2}{8(R_0 + \Delta R_D)}. \tag{14.134}$$

To write this in a more convenient form, we first write

$$R_{CB} = \frac{L_p^2}{8R_0} \frac{1}{(1 + \Delta R_D/R_0)} \tag{14.135}$$

and use the first-order binomial-series approximation† for the second term, which then leads to

$$R_{CB} \approx \frac{L_p^2}{8R_0} - \frac{L_p^2 \Delta R_D}{8R_0^2}. \tag{14.136}$$

The difference in range curvatures between the actual and assumed target ranges is then

$$R_{CA} - R_{CB} = \frac{L_p^2 \Delta R_D}{8R_0^2}. \tag{14.137}$$

If the phase error threshold is set at $\lambda/4$, then $R_{CA} - R_{CB}$ should be $\leq \lambda/4$, or

$$\frac{L_p^2 \Delta R_D}{8R_0^2} \leq \frac{\lambda}{4}. \tag{14.138}$$

The depth of focus is the maximum value of ΔR_D satisfying Eq. (14.138):

$$\Delta R_D = \frac{2\lambda R_0^2}{L_p^2}. \tag{14.139}$$

Using Eqs. (14.48), (14.53), and $r_a \approx l/2$, where l is the length of the real antenna, this can be written as

$$\Delta R_D = \frac{2\lambda}{\beta_{hr}^2} = \frac{2l^2}{\lambda a_{hs}^2} = \frac{8r_a^2}{\lambda a_{hs}^2}. \tag{14.140}$$

(depth of focus)

Thus the depth of focus is greater for longer real antenna lengths and shorter wavelengths. Degrading the azimuth resolution, which is equivalent to lengthening the antenna, increases the depth of focus.

As a numerical example, suppose that $r_a = 2$ m or 20 m and that $\lambda = 10$ cm, with $a_{hs} = 1$. The depth of focus is 320 m at 2 m resolution and 32 km at 20 m resolution.

14-9.4 SAR Image Processing: The Range-Doppler Algorithm

The range-Doppler algorithm (RDA) is perhaps the simplest and most commonly used algorithm for stripmap SAR processing. It is remarkably robust to motion and velocity errors, though these degrade and distort the image. Figure 14-37 shows a diagram of the signal processing for the RDA. The algorithm separately performs matched filtering in range and azimuth. By implementing the matched filtering using frequency-domain multiplication, the algorithm is computationally

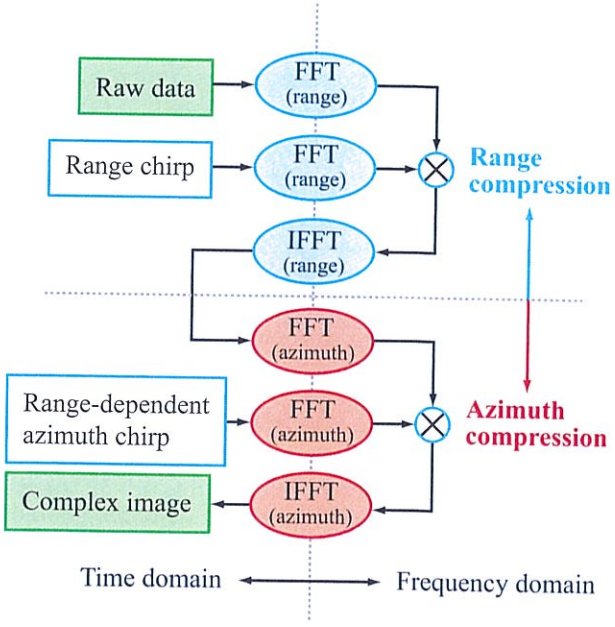


Figure 14-37: Signal-processing diagram for the range-Doppler algorithm.

†The first-order binomial expansion of $(1+x)^{-1}$ (with $x^2 < 1$) is $\approx 1 - x$.

more efficient than time-domain methods. RDA is an example of the generalized frequency domain processing algorithm (Zaugg and Long, 2009).

As shown in Fig. 14-37, RDA consists of two major steps: range compression and azimuth compression. Range compression is computed in fast time (i.e., within a single pulse) while azimuth compression is computed in slow time (i.e., across pulses). RDA is typically implemented on a block or sequence of pulses. Each pulse is first range-compressed by matched-filtering the digitized receive signal. This is done in the frequency domain by computing the Fast Fourier Transform (FFT) of the received signal and the range reference pulse (signal processing windows may be applied), which are then multiplied in the frequency domain and the inverse FFT (IFFT) is computed. This is equivalent to convolving the received signal and range reference pulse. The resulting signal contains the range-compressed signal magnitude and phase at each range bin or cell.

To illustrate RDA processing, Fig. 14-38(a) contains simulated raw SAR data collected over a scene containing two point targets at different cross-track ranges but the same azimuth position. For display purposes, the signal has been grossly over-sampled. Figure 14-38(b) illustrates the result of range compression. Note that in the range-compressed image, the signal is concentrated along two lines in azimuth at different ranges. Along these "smiles" the azimuth chirps are visible. The zero Doppler signal occurs about halfway along the azimuth axis. The range curvature or range migration is visible as the variation in the energy versus range along the azimuth axis.

RDA can be adapted for LFMCW SAR processing by modifying the range compression step. For dechirped LFMCW, range compression is accomplished by merely computing the FFT of the data.

For the azimuth compression step, RDA implicitly assumes that the pulses are evenly spaced in time and distance so that an FFT can be used. Figure 14-39 diagrams the isodop lines for a side-looking antenna at a given instant. These lines move with the radar as the platform flies in the along-track direction. The horizontal line reveals the track of a fixed target through

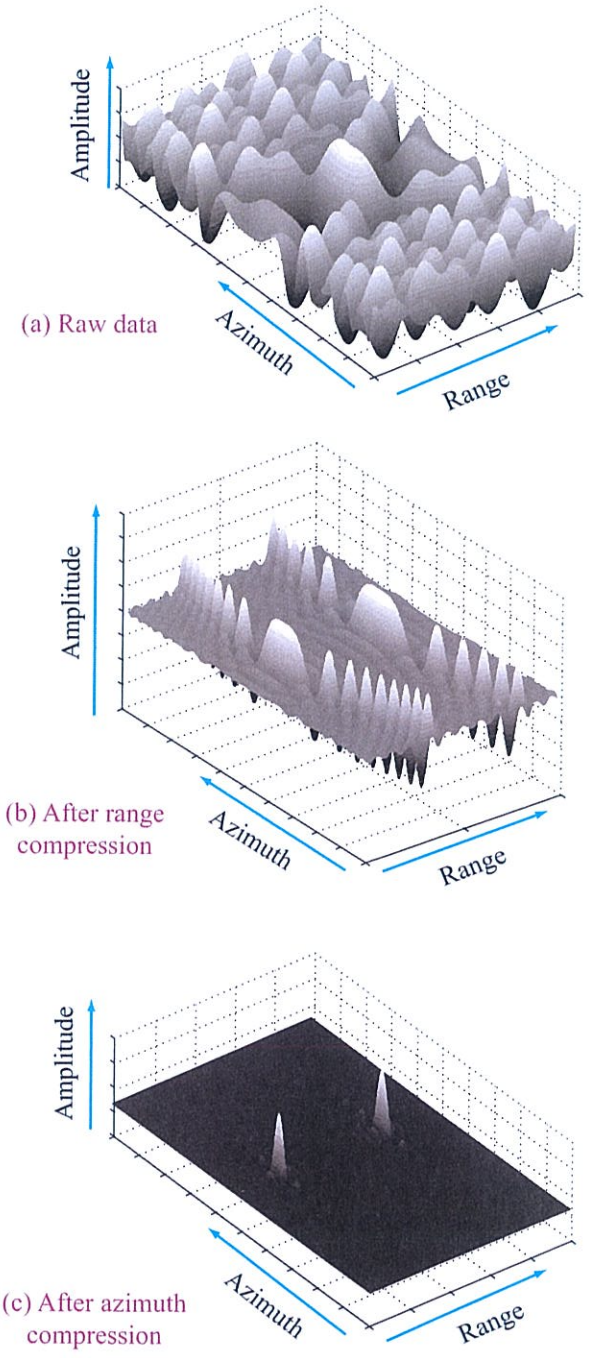


Figure 14-38: Illustration of range and azimuth compression showing the real part of the SAR signal from two point targets in (a) the raw data, (b) after range compression, and (c) after azimuth compression.

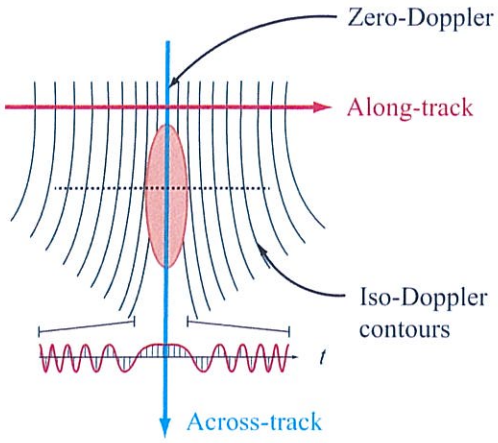


Figure 14-39: Azimuth sampling of the along-track Doppler chirp. The vertical lines in the bottom signal occur at the pulse sample times.

the isodop lines as the platform flies past. The diagram at the bottom shows a time-expanded plot of the real part of the Doppler-shifted signal as a function of time with the carrier removed. If there were only a single target and the signal were CW, this would be the received Doppler signal. Since the signal is pulsed, however, each pulse only provides a sample of this signal. The sample locations are illustrated with the vertical lines in the lower waveform, which shows the real part of the Doppler signal. The signal sample is the magnitude and phase (or I and Q values) of the range-compressed data corresponding to the range between the target and radar at each sample instant. Thus, the range-compressed data provides discrete-time samples of the Doppler signal. Assuming these samples are evenly spaced in time, frequency-domain techniques can be used to efficiently compute the azimuth-matched filter. This is the basis for RDA azimuth compression.

In RDA azimuth-compression, the reference azimuth chirp is computed for each range bin. The reference chirp is the expected Doppler signal for a point target in that range bin. (The number of reference chirps can be reduced by computing only one reference for each depth of focus.) The FFT of the reference chirp is then computed. For each range bin, the range-compressed data in the azimuth direction at the same bin is extracted and its FFT computed. The two frequency-domain signals are multiplied and the IFFT computed. This yields the azimuth-compressed SAR image. The distance between azimuth pixel spacing corresponds to the distance traveled between pulses, while the distance between range pixels is in the slant range and corresponds to the range bin spacing. Conversion to a ground-range image is considered in Section 14-10.2.

Note that the simple implementation of the RDA azimuth compression described above does not account for the range curvature, which limits the image quality when range walk is significant. Methods for accounting for the range curvature, known as ***range migration corrections***, are described in texts dedicated to SAR processing (e.g., Cumming and Wong, 2005).

Figure 14-38(c) shows the result of azimuth compression for the simulated signal example. Note that the energy spread in azimuth in Fig. 14-38(b) is now localized to the location of the point targets. We point out that the signals in the various panels have been normalized for display purposes. Due to the compression steps, the actual signal amplitude in Fig. 14-38(c) is much larger than (b), which is larger than (a).

The FFTs of reference range and azimuth chirps need to be computed only once, and can often be computed beforehand. The computational load for the remaining steps can be estimated from the number of range samples and azimuth pulses in a block. If there are N range samples and M pulses, the range-compression computation is proportional to $MN\log N$, while computation of azimuth compression is proportional to $NM\log M$ for a total of $MN(\log M + \log N)$. The log terms result from the computational efficiency of the FFT, which is $N\log N$ for a length N signal. Time domain methods for convolution would require computation proportional to $MN(M+N)$. The difference is significant for large M and N.

14-9.5 SAR Image Processing: the Backprojection Algorithm

The computational power of modern computer hardware makes time-domain SAR processing algorithms, such

as ***backprojection***, practical. Though computationally more intense than RDA, the backprojection algorithm implements a two-dimensional (range and azimuth) matched filter and can produce higher-quality images.

The discrete backprojection algorithm is a deceptively simple algorithm for generating images from synthetic-aperture radar data. It can be applied to stripmap, circular, or spotlight-mode SAR data as well as 3-D SAR imaging. It requires accurate knowledge of the relative positions of the SAR antenna and the pixel locations on the ground for each pulse, and it intrinsically compensates for nonideal platform motion during the computational process. Though computationally taxing and requiring high-precision antenna position information, it can generate high-quality SAR images even under nonideal motion conditions so long as the antenna positions are known.

The algorithm computes the radar cross section over a grid of pixels on the surface. Each pixel is separately computed. The idea is to compute time-Doppler history that the radar should observe for a point target at that pixel. This is the reference function for the matched filter used to evaluate the matched-filter output for the particular pixel from the SAR data. Accurately computing the reference function requires knowing the distance between the antenna and the pixel over the length of the synthetic aperture. When the "stop-and-go" approximation applies, the antenna location at the start of the transmit pulse can be used.

Consider a point target centered in the pixel. After range-compression, the received signal at time $t_i = 2R_i/c$, corresponding to range R_i between the radar and the point target, is given by

$$V_i(t_i) = a_i e^{-j2kR_i}, \tag{14.141}$$

where i is the pulse index and a_i is the signal amplitude from the radar equation. While precision processing accounts for variations in a_i, here we ignore the amplitude for the reference function and write the ***phase-only reference function*** as

$$V_{\mathrm{r}}\left(\frac{2R_i}{c}\right) = V_i^*\left(\frac{2R_i}{c}\right) = e^{j2kR_i}. \tag{14.142}$$

The matched-filter output is the azimuth convolution of the this reference function with the observed range-compressed data. This is completed by correlating the range-compressed data evaluated at the pixel range with the phase reference function. The pixel value A is

$$A = \sum_i V_i\left(\frac{2R_i}{c}\right) V_{\mathrm{r}}\left(\frac{2R_i}{c}\right) = \sum_i a_i, \tag{14.143}$$

where the sum is computed over the pixels that define the synthetic aperture, i.e., those in which the pixel is in the mainlobe of the antenna beam. More generally, a pulse number-dependent signal processing window may be included in the summation to reduce sidelobes. The computation in Eq. (14.143) is computed for each pixel location in the SAR image. Note that this formulation of backprojection forms the image directly in the ground plane. Topographic variations can be explicitly accounted for by including the local pixel height when computing its range to the radar.

While the backprojection pixels may be arbitrarily spaced, a single-look image is formed with the pixels spaced $l/2$ in azimuth. In range, the pixels are spaced by the slant-range resolution of the SAR. A uniform ground-range grid has variable spacing in slant range.

The backprojection algorithm can also be used with LFMCW SARs. As noted previously, such systems differ from conventional pulse-mode SARs in that the pulse length is very long, so transmission and reception occur simultaneously. For an LFMCW SAR the stop-and-go approximation is not generally valid. The motion during the chirp can be compensated for by including an additional range-dependent chirp during range compression (Zaugg and Long, 2012).

The backprojection algorithm explicitly deals with range migration and can handle nonideal motion so long as it is accurately measured and reported. However, experience suggests it is much less robust than RDA when there are unknown position or motion errors. Because each pixel is computed independently, pixel computation can be done in parallel to speed computation. The required computation can be estimated from the number of range samples and azimuth pulses. For the simple version of the backprojection algorithm outlined above, range compression can be performed the same way as RDA. If there are N range samples and M pulses, computation

in the range compression is then proportional to $MN\log N$. For an $M \times N$ image, the backprojection computation requires evaluating MN^2 terms that include computation of a square root (unless the R_i terms can be precomputed), a complex exponential, and a complex multiply and sum. There is also the expense of interpolating the range compressed data, which must be considered as well. Hence, the computational load for backprojection is substantially higher than it is for RDA.

14-10 Geometric Distortion in Radar Images

Some distortions occur naturally in SLAR and SAR images because of the radar imaging geometry, and others occur either due to nonideal motions of the vehicle or due to imperfect attempts to correct for geometric distortions or vehicle motion. We briefly consider some of these in the following subsections.

14-10.1 Elevation Distortion

Because a radar measures distance to the side of the aircraft by measuring slant range rather than ground range, a distortion occurs when the point being imaged is at a height other than the mean ground level. Consider the example shown in Fig. 14-40, where a tall flagpole target is used as an illustration. The radar is at height h above mean ground level, and consequently above the bottom of the flagpole, but the height of the radar from the top of the flagpole is, of course, different because of the flagpole height H. As seen in the figure, the slant range to the bottom of the flagpole is R_B, which is larger than the range to the top of the flagpole, R_T. Thus the top of the flagpole is closer to the radar than the bottom of the flagpole. As a result, in the radar image the top of the flagpole appears closer than the bottom. The angle of incidence associated with the top of the flagpole, θ_T, is larger than the angle associated with the bottom, θ_B. The top of the flagpole appears at a position to the side of the flight line by an amount y_T whereas its actual location to the side of the flight line is y.

Note that a ***radar shadow*** exists for the flagpole, or any other elevated object. This is the region behind the

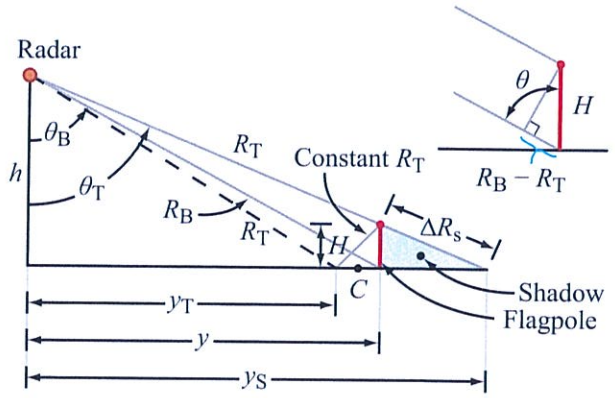


Figure 14-40: Radar imaging height distortion and shadow geometry.

object that is not illuminated by the radar (Fig. 14-40). The shadow extends from the bottom of the flagpole, y, to the end of the shadow at y_S. Thus the effect of the target height is both to distort the location of the top (and other elevated parts of the target) and to create a shadow region that cannot be imaged by the radar.

Optical imaging sensors determine the distance by the angle θ; that is, in an optical image, the top of the flagpole appears to be at a point y_s, whereas in a radar image it is much closer, at y_T. Thus, the image distortion due to height for radar is the opposite of that associated with an optical sensor. In both cases, the target shift is known as ***parallax*** displacement.

The magnitude of the displacement of the top of the flagpole is given by

$$R_B - R_T = \sqrt{y^2 + h^2} - \sqrt{y^2 + (h - H)^2}. \qquad (14.144)$$

For a target with a relatively small elevation like the flagpole in this example, this may be approximated by

$$R_B - R_T \approx H \cos\theta, \qquad (14.145)$$

where the approximation is obtained from the figure by using the chord instead of the arc of constant R_T.

This difference in slant range results in the apparent cross-range position of the top of the flagpole at

$$y_T = \sqrt{R_T^2 - h^2} = \sqrt{y^2 - 2hH + H^2}, \qquad (14.146)$$

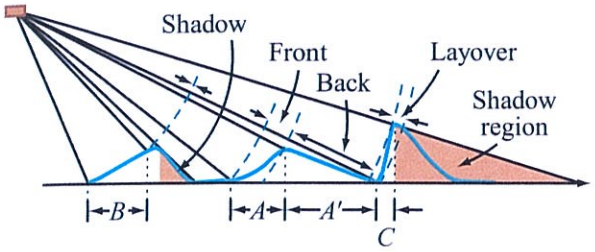


Figure 14-41: Examples of height distortion in SAR images.

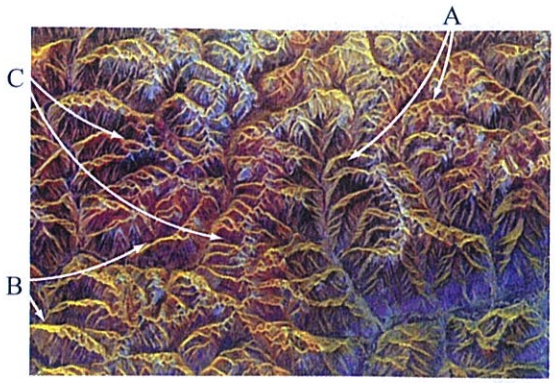


Figure 14-42: Foreshortening in mountainous terrain. The radar illumination direction is from the top of the image. [SIR-C/XSAR image courtesy of NASA/JPL.]

rather than at y. In other words, the top of the flagpole is apparent in the radar image at a cross-range distance that is closer to the radar than to the apparent location of the bottom of the image. This occurs because, in fact, the top of the flagpole *is* closer to the radar than the bottom is. This phenomenon is known as ***overlay***.

The effect of overlay can be significant in cities and steep mountainous areas. Overlay is an extreme example of the inherent ***foreshortening*** effect that occurs in radar imaging. Figure 14-41 provides several examples of different amounts of foreshortening. In this figure, the dashed arcs represent lines of constant slant range. For the center peak in Fig. 14-41, the horizontal distance A between the mountain base and its peak is much larger than the corresponding points in the slant-range image, while for the back-side positions, A' is stretched out in the slant plane compared with its horizontal displacement. Further, because the radar return from a large area ends up at the same distance in the radar image, the forward face of the mountain is often much brighter than the slant-range image of the back side. Peak A in Fig. 14-41 illustrates even more foreshortening where the entire forward slope of the mountain appears at a single slant range. The resulting SAR image has essentially no effective cross-range resolution at this point. We also note that the back side of the mountain is in the shadow of the peak and thus provides no reflection in the SAR image. Peak C in Fig. 14-41 illustrates when the foreshortening effect causes the top of the peak to appear closer to the radar than the bottom of the valley, creating overlay conditions. These effects are clearly evident in the SAR image of mountainous terrain shown in Fig. 14-42. Note that the mountains are "laid over" toward the radar and that their front sides appear brighter than the back sides. Elevation distortion is also visible in the distorted bridge under construction in Fig. 14-43.

As previously indicated, image shadows result from areas that do not provide any reflection for the radar. Shadows can lead to loss of information about the surface. However, the *shape* of the shadow can reveal information about the object creating the shadow. For example, the reflective metal of the aircraft in Fig. 14-44 reflects much of the radar signal away from the radar, providing limited information about the aircraft shape. However, the radar shadow on the flat ground clearly reveals the shape of the wings and the presence of the otherwise invisible helicopter blades. The shadows of trees in the upper right hint at their shape.

14-10.2 Range Distortion

The radar measures slant range, but to produce maps or ***plan view images***, the image must be converted to an image in ground range. The horizontal displacement y of the base of the flagpole in Fig. 14-40 is given by

$$y = \sqrt{R_B^2 - h^2} = h \tan \theta_B. \tag{14.147}$$

Figure 14-43: Multilook SAR image with image distortions due to digital elevation model (DEM) errors. The bridge at the center is distorted since it is not included in the DEM. Streaks along the freeway are due to moving vehicles. Parallel lines at the bottom are railroad tracks.

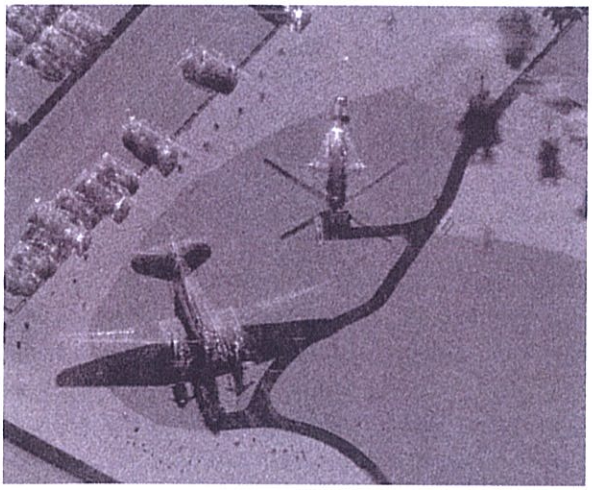


Figure 14-44: Example of how a shadow can reveal the shape of the object creating the shadow. The radar illumination direction is from the top of the image. [Image courtesy of Sandia National Laboratories.]

Since y is not linearly related to R_B, the horizontal scale in the range direction is distorted in an image. If we denote the displacement on the image by Y and a scale factor for relating displacement on the image to displacement in space by a, we find

$$Y = a(R - R_0) = a\left(\sqrt{h^2 + y^2} - h\right), \qquad (14.148)$$

where we have deleted the subscript B from R_B because we are dealing with general relations rather than simply a single location. We also replaced R_0 with the value of R at $y = 0$, namely $R_0 = h$. The rate of change of displacement on the image relative to displacement on the ground is given by

$$\frac{dY}{dy} = a\frac{y}{R} = a\sin\theta. \qquad (14.149)$$

Since the rate of change of displacement on the image relative to that on the ground is proportional to y or to $\sin\theta$, the value of the derivative is small near vertical,

where y is small, and grows larger across the swath. This means that the slant-range image compresses the horizontal variations near vertical, relative to those far from the flight line. This inherent distortion associated with the difference between slant range and ground range cannot be totally overcome in the radar processing.

14-10.3 SLAR Motion Distortion

SLAR images are generated by displaying the echo power versus time (or range) for one pulse in one dimension, and "stacking" the range lines in the other dimension to create an image. Nonideal motion and pointing can create distortions in the resulting image. These include the following:

1. **Varying speed:** The PRF is not adequately synchronized with the speed of the aircraft. This results in a nonlinear ***stretching*** or ***compression*** in the along-track direction.
2. **Lateral or vertical motions:** Deviations away from a straight-line flight path result in curvilinear distortions that make straight lines (such as roads) parallel to the flight track appear curved, often sinuous.
3. **Yaw of the aircraft:** The effect of yaw motion is to distort the direction of different points relative to others, depending upon their displacement from the flight line. Extreme yaw, such as occurs during a turn, can completely distort the image.
4. **Pitch of the aircraft:** The effect of a pitching motion is to move the intersection of the beam with the ground ahead of or back from the position directly to the side of the point beneath the aircraft.
5. **Roll of the aircraft:** The roll of the aircraft changes the antenna gain at different points in the image and therefore modulates the grayscale.

Items 1–4 are illustrated in Fig. 14-45. If these motions are adequately sensed by the aircraft navigation system, corrections can be made by including frequency shifting in the processing, by adjusting the timing or positioning of the pixels in the image, or by rectifying the image after formation.

14-10.4 SAR Motion Errors

Stripmap synthetic-aperture processing ordinarily depends upon the assumption that the radar is traveling horizontally in a straight line. Figure 14-46 illustrates the geometry for determining the effect of deviations from this straight-line path. The assumption is that the radar is traveling in the x direction at height h and is observing a ground cell at an angle θ from the vertical. The coordinates of the forward edge of the ground cell are $(r_a/2, y, 0)$. The coordinates for the radar are $(0, 0, h)$, and the range vector from the radar to the front edge of the cell is

$$\mathbf{R} = \hat{\mathbf{x}}\,\frac{r_a}{2} + \hat{\mathbf{y}} R \sin\theta - \hat{\mathbf{z}} h.$$

If the radar velocity $\mathbf{u}$ has components in the y and z directions (in addition to the x direction),

$$\mathbf{u} = \hat{\mathbf{x}} u_x + \hat{\mathbf{y}} u_y + \hat{\mathbf{z}} u_z,$$

the total Doppler shift associated with the point $(r_a/2, y, 0)$ is given by

$$\begin{aligned} f_D &= \left(\frac{2\mathbf{u}}{\lambda}\right) \cdot \left(\frac{\mathbf{R}}{R}\right) \\ &= \frac{2}{\lambda R}(\hat{\mathbf{x}} u_x + \hat{\mathbf{y}} u_y + \hat{\mathbf{z}} u_z) \cdot \left(\hat{\mathbf{x}}\,\frac{r_a}{2} + \hat{\mathbf{y}} R \sin\theta - \hat{\mathbf{z}} h\right). \end{aligned}$$

Hence,

$$f_D = \frac{2}{\lambda}\left(\frac{u_x r_a}{2R} + u_y \sin\theta - u_z \cos\theta\right). \qquad (14.150)$$

This Doppler frequency f_D can be expressed in terms of the desired Doppler frequency f_{D0} $(= B_{Df}/2)$ and an error component f_{De} $(= f_{Dy} + f_{Dz})$:

$$f_D = f_{D0} + f_{De} = \left(1 + \frac{f_{De}}{f_{D0}}\right) f_{D0}.$$

If we define a limit bound ε_r to set the per-unit error allowed in the Doppler frequency, expressed as a fraction of B_{Df}, then Eq. (14.150) may be rewritten as

$$f_D = (1 + 2\varepsilon_r) f_{D0}.$$

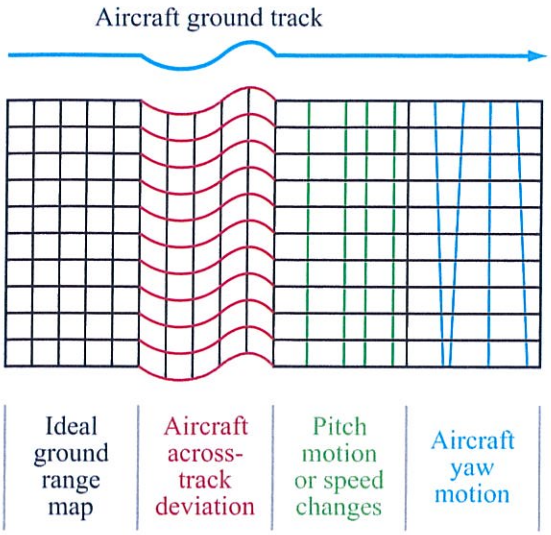


Figure 14-45: SLAR map of a square grid with uncompensated aircraft motion.

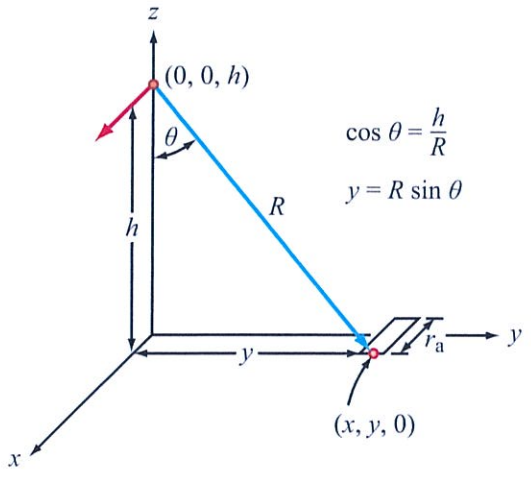


Figure 14-46: Geometry for SAR motion-error determination.

With this definition we have

$$\frac{f_{De}}{f_{D0}} \leq 2\varepsilon_r,$$

which sets the bounds on the size of the error velocities u_y and u_z as given in

$$u_y \sin\theta - u_z \cos\theta \leq \lambda f_{D0} \varepsilon_r = \frac{\lambda B_{Df} \varepsilon_r}{2}. \qquad (14.151)$$

If the velocity component errors are smaller than this, no compensation is required, but if they are larger, some sort of compensation is required to minimize adverse effects.

Consider an aircraft example with $\varepsilon_r = 0.1$, $B_{Df} = 2$ Hz, and $\lambda = 3$ cm. Substituting these values into Eq. (14.151) yields

$$u_y \sin\theta - u_z \cos\theta \leq 3 \times 10^{-3} \text{ m/s} = 3 \text{ mm/s}.$$

For a spacecraft, using the same ε_r and λ, B_{Df} is 25 Hz and the limits are

$$u_x \sin\theta - u_z \cos\theta \leq 3.75 \text{ cm/s}.$$

These are very small velocities when considered in terms of absolute values or values relative to the aircraft and spacecraft velocities.

With such tiny errors causing serious trouble for the SAR, some sort of compensation is absolutely necessary. It is unrealistic to imagine that either an aircraft or a spacecraft can be made to travel in a perfectly straight line. Although spacecraft are extremely stable, aircraft tend to bounce around in the turbulent air, and the corrections required for this nonideal movement can be very sizable.

To compensate for velocity and acceleration errors, one must first sense them. The sensing can be achieved by mounting integrating accelerometers directly on the antenna, or at least nearby. Typical aircraft navigation systems do not have the sensitivity required to sense and measure motion errors as small as required for SAR compensation. Further, motion measurement errors may be compounded by the long lever arm between the antenna mounting location and the aircraft motion sensors.

The effect of the uniform-velocity errors is to shift the center frequency for the expected Doppler spectrum away from the carrier frequency, as illustrated in Fig. 14-47(a). Here the expected spectrum is shown centered about the carrier frequency and the actual spectrum is shifted by the velocity errors. For a system in which the processing is performed at baseband, the signal is mixed down from the carrier frequency by beating the return signal with a local oscillator at the carrier frequency. One solution to the problem of velocity errors is to change the frequency of the local oscillator by just the right amount so that the LO frequency is centered in the passband shown in the solid blue lines in Fig. 14-47(a). This may be done either at the carrier frequency, at the intermediate frequency, or at baseband in the processor.

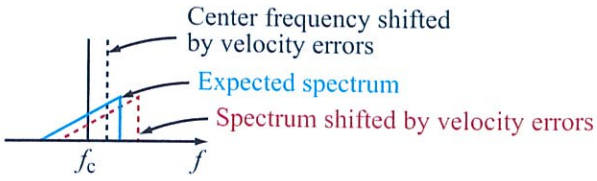
Correcting for the acceleration errors is, in effect, correcting for the resulting changes in velocity, so the correction method is similar. However, correcting for uniform-velocity shifts may be done with a system having a longer time constant than that required for the acceleration correction, so correction is easier to do.

The problem of compensating for motion errors is one of the primary limitations on the precision that can be achieved with a SAR.

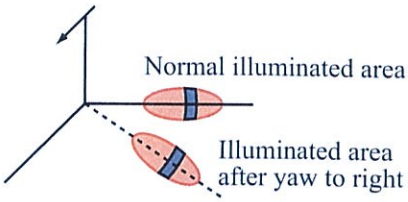
14-10.5 SAR Attitude Errors

The aircraft or spacecraft may yaw, pitch, or roll, or do all three at the same time. Each of these causes an error in the SAR. The effect of the roll error is the same as for the SLAR; that is, it simply changes the gain for a particular point on the ground. If the antenna rolls too much this can be a serious problem, but it does not affect synthetic-aperture processing and its effect is not as severe as the yaw and pitch errors, which do affect the processing.

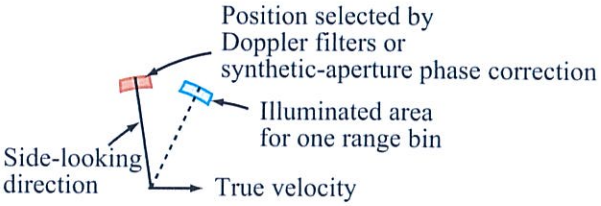
Figure 14-47(b) illustrates the geometry when the aircraft yaws to the right. This has the effect of moving the area illuminated by the real aperture away from the side-looking direction and rotating it slightly ahead of that direction. If the yaw were to the left it would result in the beam moving back of the side-looking position. This causes distortions in the image; but, more important for the SAR, it causes a shift of the Doppler frequencies



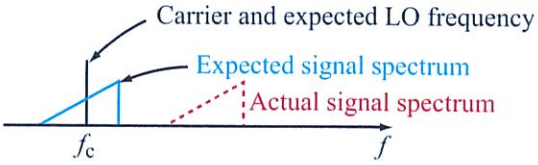
(a) Effect of velocity errors



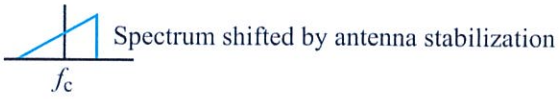
(b) Effect of yaw on SAR illuminated area



(c) Effect of yaw on beam position



(d) Effect of yaw on received spectrum



(e) Spectrum after correction of antenna position

(f) Spectrum after correction of LO frequency

Figure 14-47: Illustration of velocity and attitude errors.

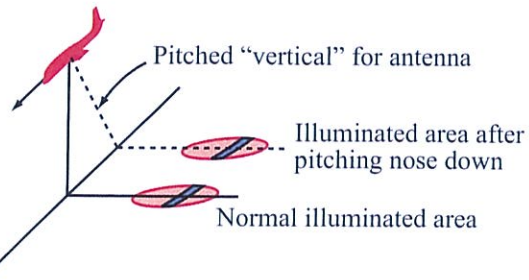


Figure 14-48: Effects of pitch on SAR illuminated area.

away from the beam. Figure 14-47(c) illustrates the situation looking down on the illuminated area for a given range position on the ground. If nothing is done about the problem, the Doppler filters select the dashed area shown, which is outside the area illuminated by the antenna beam, so the signal received cannot be correlated, as shown in Fig. 14-47(d). The signal spectrum is in the wrong place, just as it is for the velocity error.

Two types of correction are possible for this effect. Perhaps the best, if it can be done perfectly, is simply to stabilize the antenna so it always looks directly to the side. When this happens, there is no yaw of the antenna, even though the aircraft or spacecraft carrying it yaws. The result is that the spectrum is in the right place, as indicated in Fig. 14-47(e). The other solution is like that applied for velocity errors, as indicated in Figure 14-47(f); that is, one moves the local-oscillator frequency to the center of the passband for the Doppler frequencies. In practice, a combination of antenna stabilization and shifting of the LO is often required. This correction may be achieved with the same local-oscillator shifting that is done for the velocity errors. That is, the combination of velocity error and yaw error may be used to provide the error signal that shifts the local-oscillator frequency.

The effect of pitch is somewhat different, but the solution is the same. Figure 14-48 illustrates the effect of pitch on the illuminated area. Instead of rotating the beam about a vertical axis, the rotation is about a horizontal axis parallel to the side-looking direction. Thus, the two illuminated areas—the proper one and the one after pitching—are parallel to each other but displaced. The effect on the Doppler frequency, however, is the same as illustrated for yaw except that the shift may not be exactly the same at all ranges. Nevertheless, the correction can be most easily achieved by the same methods used for correcting for yaw, that is, stabilizing the antenna and shifting the local-oscillator frequency.

Note that even after the corrections have been made to the local-oscillator signal so that the SAR image can be produced, the image has errors in it. In the case of yaw, the errors are along-track displacements that are proportional to the distance from the point beneath the radar. In the case of pitch, they are approximately the same at all distances away from the radar track, but still represent along-track displacements of the pixels. When this occurs, the geometric fidelity of the image is degraded. It is possible to warp the final image to partially correct for this problem.

An inherent yaw effect is likely to occur with most spacecraft radars. This comes about because the spacecraft is aligned with the direction of travel along the orbit plane, but the zero-Doppler line is not quite perpendicular to this plane. Because the Earth is rotating beneath the spacecraft, the vector velocity relative to a point along the Earth contains a y component even if the spacecraft is traveling perfectly in the x direction in its orbit plane. The zero-Doppler line on the surface of the Earth is perpendicular to the vector sum of the velocity along the orbit direction and the velocity relative to the rotating Earth. Consequently, it is tilted from perpendicular to the orbit plane by an amount that varies up to 3.5° at the equator, where the Earth's linear velocity is the greatest. The inherent yaw built into most spacecraft systems may be corrected either by rotating the antenna about the spacecraft axis or by designing the spacecraft itself to be continuously reoriented perpendicular to the zero-Doppler line.

14-11 Elevations from SLAR and SAR

Elevations can be estimated using the interferometric techniques described in Chapter 15. Other approaches to height estimation that do not require the coherence of

interferometric techniques include exploiting shadows in the image and image parallax with the aid of stereo techniques. These are briefly considered below.

14-11.1 Shadows

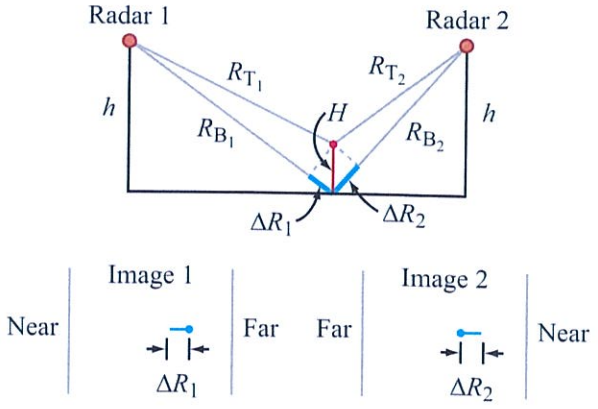
In the geometry shown in Fig. 14-40, a significant shadow exists behind a flagpole, and similar shadows may exist behind a mountain, a tree, or a building. In some cases, the shadow is long enough so that it may by used to estimate the height of the object. The length of the shadow, ΔR_S is given by

$$\Delta R_S = \frac{H}{\cos\theta} . \qquad (14.152)$$

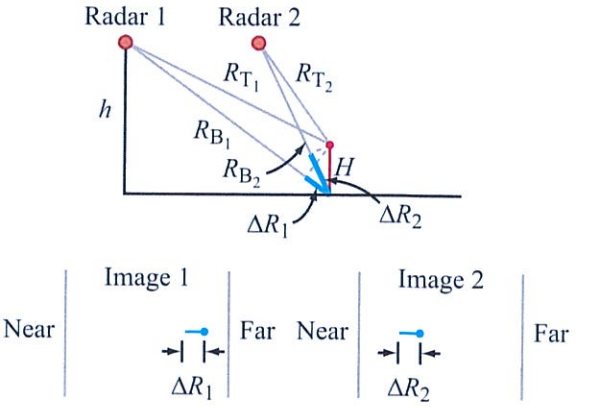
Since the denominator of the right-hand side of the equation is small for large θ, even relatively low objects may have their height determined by the shadow length. For example, consider a slant-range resolution r_r of 10 m. If the object has a height $H = 5$ m at $\theta = 80°$ then $\Delta R_S = 28.8$ m, which is almost 3 pixels. This is enough to estimate the height of the object, although the estimate is not very good, because of the small number of pixels involved. Finer range resolution enables better height estimate accuracy.

14-11.2 Stereo with Radar

The stereoscopic effect is widely used with aerial photography for determining elevations. Most topographic maps are made by stereo interpretation of overlapping pairs of aerial photographs. Similar techniques can be used with radar. Radar stereo, however, is different from optical stereo, as suggested in the discussion of Fig. 14-40, since higher objects appear further away in a photograph but closer in a radar image. With photography, the stereo effect is obtained by taking successive overlapping photographs as the aircraft travels. With radar the normal procedure is to fly different flight lines with overlapping radar swaths and use the resulting images for stereo. Targets in the radar images have different incidence angles. The flight lines may be flown on the same side or the opposite side of the area being studied. Figure 14-49(a) illustrates opposite-side stereo. An object of height H is observed from the



(a) Opposite-side stereo



(b) Same-side stereo

Figure 14-49: Two types of stereo imaging.

two indicated positions. It "leans" a distance of ΔR_1 toward the aircraft on flight line 1 and a distance of ΔR_2 toward the aircraft on flight line 2. When the images are superimposed and translated to the side until the two images of the same point on the ground coincide, the amount of translation required is the sum of ΔR_1 and ΔR_2. By knowing the location in the image and the spacing between the flight lines, one can compute the height above the mean ground level. When a pair of such images is observed through a stereo viewer, the effect of the parallax is to make the mountain appear in three dimensions, much as it would in an optical stereo view,

except that the displacements are the opposite of what they would be in the photographs.

An important difficulty with opposite-side stereo is shadows. Since the shadows from one image lie to one side of the object and from the other image to the other side of the object, objects are often difficult to recognize. In fact, a given object may be in the shadow of some taller object for one of the images and not for the other, so that it cannot even be seen on both images. The solution to this is to use same-side stereo, as indicated in Fig. 14-49(b). Two flight lines are also used for same-side stereo, but they are translated from each other, either horizontally, as illustrated, or vertically. Because in this case ΔR_2 also is different from ΔR_1, a stereo effect occurs. When the two images are translated so that a particular object is superimposed from one image to the other, the amount of translation can be used to establish the height above mean ground just as with opposite-side stereo or with photography. In this case, however, it is the difference of ΔR_2 and ΔR_1 that determines the height. Because the shadows are on the same side for both images, this method is easier to use than opposite-side stereo. It has been used rather widely in mapping remote areas where other elevation information is lacking.

14-11.3 Squint Stereo

A dual-beam SAR can produce stereo images in a single pass that may be used for a single flight line to determine altitudes. The situation is illustrated in Fig. 14-50. At a given point on the flight path, two beams are used—a forward beam and an aft beam. While a squint angle of 45° ahead of and behind the side-looking direction is used here, other squint angles can be used. Pulses are typically transmitted alternately on the two beams and two recording systems are used—one for each beam. In the figure, the direction of the scan lines is indicated first. Then the locations of the beams are shown for two positions at which a given point on the ground is imaged (position 1 on the forward beam, and position 2 on the aft beam). If the point on the ground represents a flagpole, its top is displaced from the bottom toward the radar on each image, but the direction toward the radar is different on the two images, as shown. The images may then be superimposed at the base of the flagpole, and

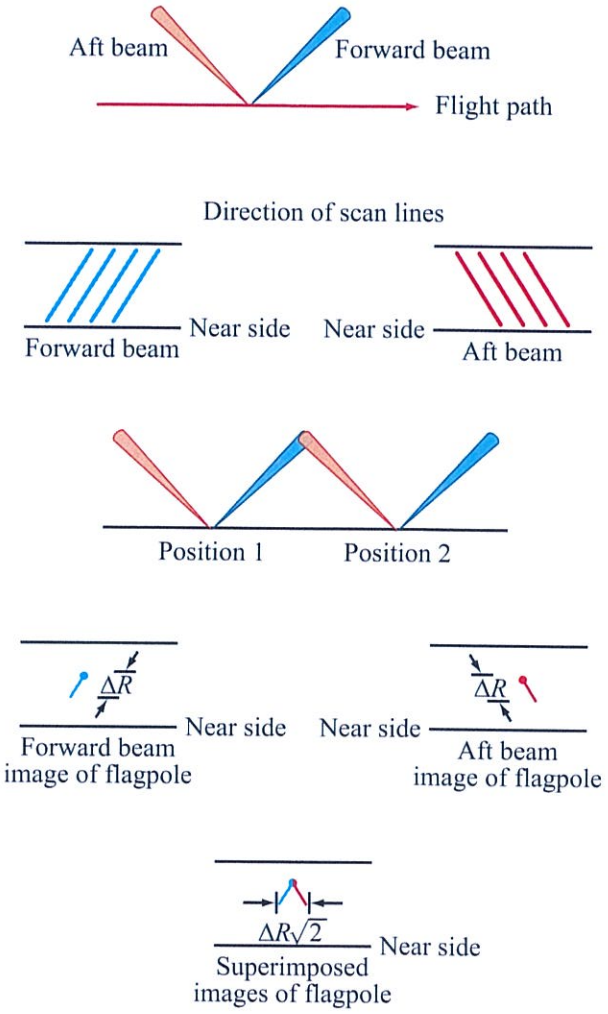


Figure 14-50: Dual squint-beam stereo radar with 45° squint angles.

the distance between the images of the tops indicates the height of the flagpole. Similar techniques can be used in which a given point at some elevation, such as the top, is superimposed on the two images, and the amount of translation required to align the relative positions of the images for a reference level determines the height. Squint-stereo systems require fan-beam antennas that are tilted relative to the axis of the aircraft.

14-11.4 Mountains and Buildings

Because of shadows behind mountains and buildings, images in mountainous terrain and cities with tall buildings necessarily fail to contain information about the areas that are in the shadow. In such regions, multiple flight lines are normally required to image all parts of the regions. Orthogonal flight lines, in which the region is observed from both sides along parallel flight lines and from both sides along parallel lines orthogonal to the first pair, are considered the best approach. When such images are superimposed, the shadows are generally filled in. Unfortunately, however, superimposing them in mountainous country can cause difficulties because the parallax displacement is different for the different images.

One way to accomplish part of this purpose with a single flight line is to use the dual squinted-beam imaging geometry illustrated in Fig. 14-50. The shadows are filled in to the degree that orthogonal looks can fill them in. Of course, under some circumstances, parts of a region may be shadowed from both positions on the flight line and a separate flight line on the other side may be required. Distortions due to deviation from a straight flight path and due to aircraft altitude changes still remain in this pair of images, and will differ between them, so that proper superposition of the images depends upon the maintenance of a straight flight line and a stable platform. This technique is viable for both real-aperture and synthetic-aperture systems.

14-12 Ionospheric Effects

The Earth's ionosphere, which extends from about 60 km above the Earth surface to as high as 1000 km depending on solar activity, consists of an ionized gas containing free electrons and positively charged ions (Kelley, 1989; Brekke, 1997; Hunsucker and Hargreaves, 2003). Under the influence of the Earth's magnetic field, a radio wave passing through this ionized medium may experience the ***Faraday effect*** that causes the wave's electric field vector to rotate, which, in turn, alters the polarization state of the radio wave. The amount of angular rotation, termed the ***Faraday rotation***, depends on the frequency of the wave, its direction of propagation relative to the direction of the Earth's magnetic field, and the state of the ionosphere (free electron density profile).

This Faraday rotation may have consequences to both radar and radiometric Earth observations if the observation platform is above 60 km, which includes all satellite orbits.

For a satellite-borne radar transmitting a v-polarized wave, the Faraday rotation alters the polarization state of the wave, so the wave illuminating the surface is no longer purely v polarized, but rather a combination of a v-polarized component and an h-polarized component. Consequently, the scattering measurements recorded by the radar no longer match the assumptions of the inversion algorithm used to convert the radar measurements into estimates of the biophysical or geophysical parameters of the observed scene. The mismatch usually translates into estimation errors. A similar scenario applies to the radiometer; the v-polarized energy captured by the radiometer antenna is a result of emission by the Earth's surface at a polarization direction other than vertical, such that after getting rotated by the ionosphere arrives as v-polarized at the radiometer antenna. Fortunately, the Faraday rotation can be ignored completely if $f \gtrsim 3$ GHz, and for frequencies in the range $0.5 \leq f \leq 3$ GHz, the rotation angle can be estimated and the measured radar observations can be corrected accordingly.

14-12.1 Rotation Angle

The Faraday rotation effect is illustrated by the simple diagram shown in Fig. 14-51, which depicts an EM wave traveling through an ionized layer (usually called a ***plasma***). Here, **B** represents the Earth's magnetic field and $B_{\|}$ is the component of **B** along the propagation direction. The rotation angle χ is given by Hunsucker and Hargreaves (2003) as

$$\chi = 2.365 \times 10^{-14} \frac{B_{\|}}{f^2} N_{\text{TECU}} \sec\theta \qquad \text{(rad)}, \quad (14.153)$$

where $B_{\|}$ is the magnetic flux density (in teslas) of the Earth magnetic field parallel to the direction

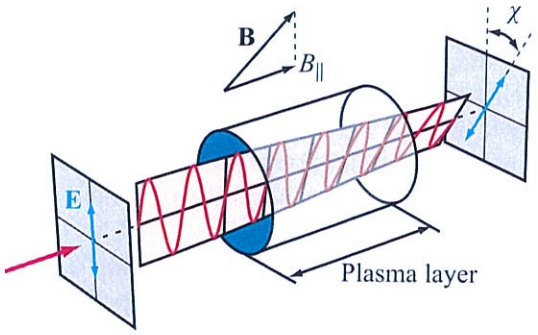


Figure 14-51: Illustration of the Faraday rotation effect resulting from EM propagation in a plasma under the influence of an external magnetic field **B**.

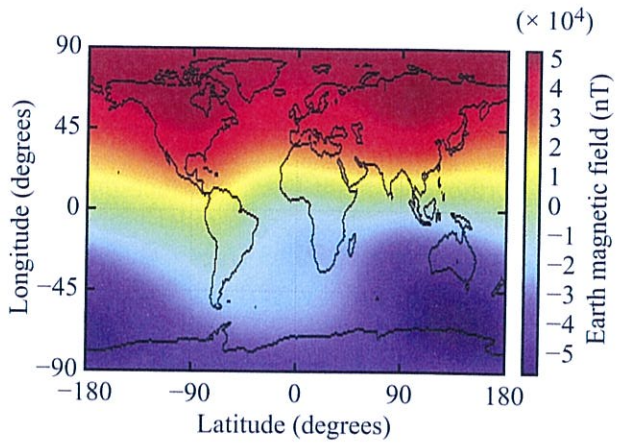


Figure 14-52: Vertical geomagnetic field as modeled by the IGRF10 for June 21, 2007 [Jehle et al., 2009].

of propagation (also called the ***line-of-sight*** (LOS) magnetic field), f is the frequency in GHz, θ is the incidence angle relative to nadir, and N_{TECU} is the integrated free-electron density for a vertical column through the entire ionosphere, measured in TECU, where TECU stands for ***total electron content unit*** and 1 TECU $= 10^{16}$ electrons/m^2. At night, N_{TECU} may be as low as 0.2 TECU, whereas in the daytime it may reach 100 TECU, and when sunspots are highly active N_{TECU} may reach 200. The magnitude of N_{TECU} also varies with season and geographic latitude. The magnetic field component $B_{\|} = B\cos\psi$, where ψ is the angle between the Earth's magnetic field and the propagation direction, and B can be computed from the International Geomagnetic Reference Field (IGRF) model (Macmillan and Maus, 2005), which varies with both latitude and longitude (Fig. 14-52). Using this model, we show in Fig. 14-53 plots for the one-way rotation angle χ as a function of f for a satellite sensor observing the Earth at an incidence angle $\theta = 40°$, at a midlatitude location. At 3 GHz, χ varies between $\approx 0.06°$ for $N_{\text{TECU}} = 1$ (typical nighttime condition) to only about 6° for $N_{\text{TECU}} = 100$. This means that the Faraday rotation effect may be ignored at $f \gtrsim 3$ GHz. At lower frequencies, the Faraday rotation may be more significant, as discussed next.

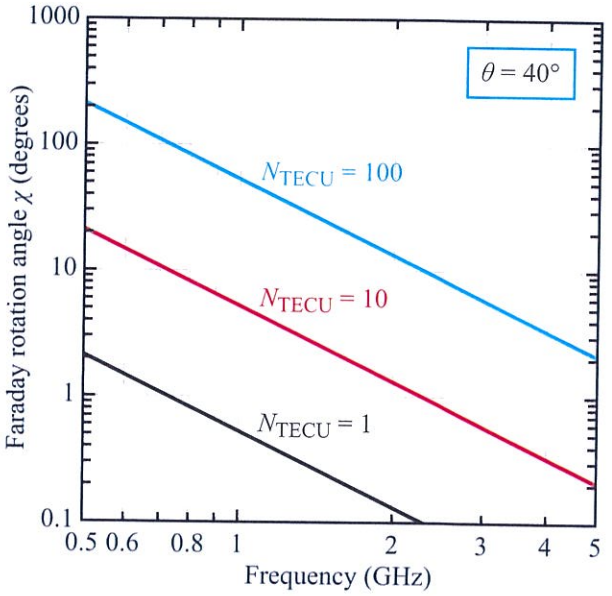
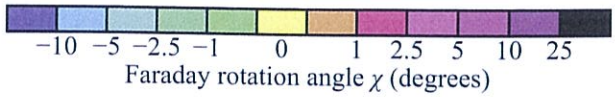


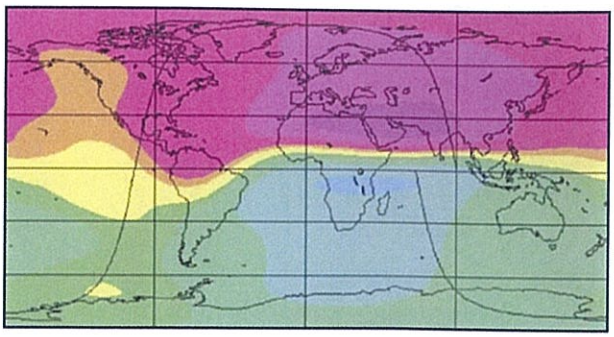
Figure 14-53: Computed variation of the one-way Faraday rotation angle with frequency for various values of N_{TECU}, assuming a midlatitude region and an incidence angle $\theta = 40°$.

14-12.2 Impact on SAR Data

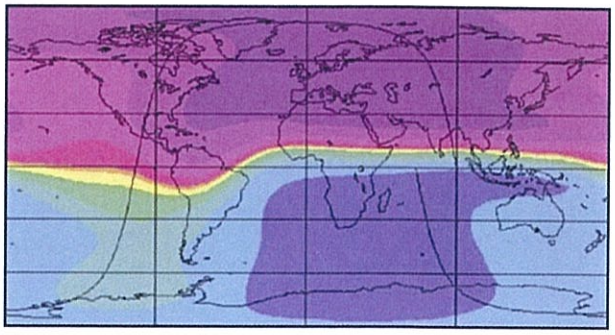
To illustrate the impact of the Faraday rotation on SAR data, we show the results of two simulation studies. In the first study (Wright et al., 2003), the one-way Faraday rotation angle was computed at L-band (1.26 GHz) as a function of latitude and longitude for both low sunspot activity with N_{TECU} varying geographically between about 1 and 80, and high sunspot activity with N_{TECU} between 10 and 120. The incidence angle was set at 35°. The results of the study are displayed in Fig. 14-54. Under the low-sunspot-activity condition, $|\chi|$ seldom exceeds 5°, but when the sunspot activity is high, $|\chi|$ sometimes exceeds 15°. The impact of the Faraday rotation on the backscattering coefficient is illustrated by the computed values given in Table 14-2 for an agricultural area. Application of two-way Faraday rotation leads to insignificant errors (for $\chi \leq 10°$) for hh and vv polarizations, but the error for hv polarization is 1.61 dB for $\chi = 10°$ and as much as 4.23 dB at $\chi = 20°$. To provide a visual illustration of what such errors mean, the authors used the multipolarization image shown in Fig. 14-55(a) to simulate the same image after incorporating the Faraday rotation twice, once to simulate the transmission from the satellite to the ground and a second time for the return trip. Whereas the multipolarized image in Fig. 14-55(a) provided significant discrimination between different crops, the simulated images in parts (b) and (c) of the figure clearly demonstrate that much of the discriminating power has vanished because of the Faraday rotation effect.

The impact of Faraday rotation on SAR data was also evaluated by Jehle et al. (2009) for three SAR configurations with operating center frequencies of 9.65 GHz, 1.27 GHz, and 0.45 GHz. Their results, summarized in Table 14-3, suggest that whereas the impact of the Faraday rotation on the SAR parameters is negligibly small at X-band (9.65 GHz) and somewhat significant at L-band (1.27 GHz), it is very serious at P-band (450 MHz). For $N_{\text{TECU}} = 100$, the two-way Faraday rotation angle is $2\chi = 470°$! Even if the 2π-ambiguity problem can be resolved and χ can be estimated using the models described earlier, the errors associated with the estimate can pose serious problems





(a) Low sunspot activity

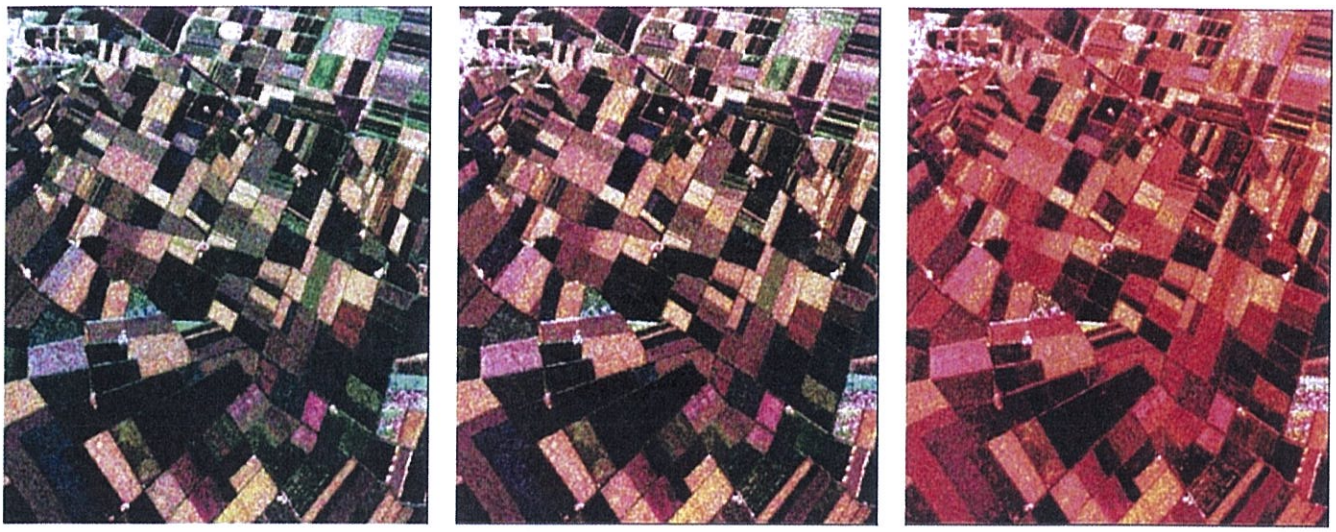


(b) High sunspot activity

Figure 14-54: Predictions of one-way Faraday rotation angle under (a) low sunspot activity and (b) high sunspot activity, for 12:00 GMT, April and $\theta = 35°$ [Wright et al., 2003].

Table 14-2: Change in backscattering-coefficient magnitude for an agricultural scene observed at L-band due to the application of two-way Faraday rotation [Wright et al., 2003].

One-way χ	Change in mean hh (dB)	Change in mean vv (dB)	Change in mean hv (dB)
5°	−0.08	−0.09	0.51
10°	−0.35	−0.36	1.61
20°	−1.39	−1.49	4.23



(a) $\chi = 0°$ (b) $\chi = 10°$ (c) $\chi = 20°$

Figure 14-55: L-band SAR imagery of Feltwell, U.K. with one-way Faraday rotation of: (a) 0° (original image), (b) 10°, and (c) 20°. Blue = hh, green = vv, red = hv. The same color balance was used for the three images. [Wright et al., 2003.]

Table 14-3: Satellite sensor details and estimates of the two-way impact of 50 and 100 TECU on the path delay, chirp length, and Faraday rotation for TerraSAR-X, ALOS PALSAR, and a possible future spaceborne P-band sensor configuration. The Earth's magnetic field is modeled for June 21, 2007, 45° North and 0° East, at a height of 300 km and a nadir-looking sensor configuration [Jehle et al., 2009].

Sensor	TerraSAR-X		PALSAR		P-band	
Frequency (f_c) [GHz]	9.65		1.27		0.45	
Bandwidth (B) [MHz]	max. 300		max. 28		6	
Chirp duration [μs]	40		27		27	
Sampling rate [MHz]	max. 300		max. 32		8	
Chirp form	up		down		down	
Orbit (altitude) [km]	514		695		695	
TECU	50	100	50	100	50	100
Path delay at f_c [m]	0.5	1	27	54	218	436
Change in chirp length for B [ns]	−0.1	−0.2	4	8	19.4	38.7
Two-way Faraday rotation at f_c [°]	0.5	1	29.5	59	235	470
Δ Faraday rotation within chirp [°]	0.03	0.06	1.3	2.6	6.3	12.5

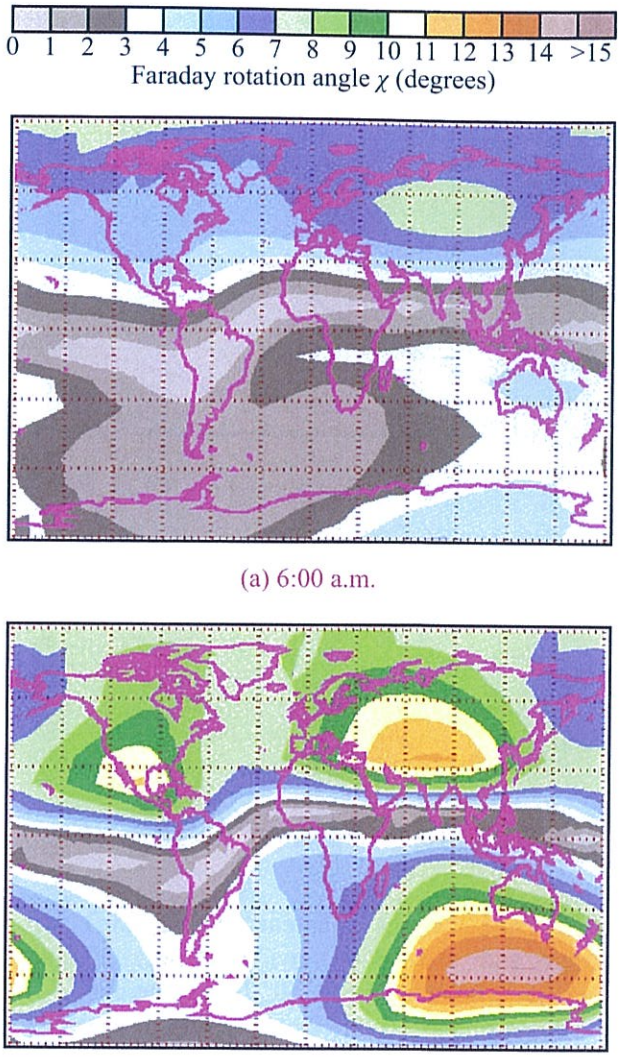
to both the interpretation of the measured data and the efficacy of correction algorithm applied to it.

Other noteworthy investigations of the impact of the Faraday rotation angle on SAR data include Rignot (2000), Gail (1998), Qi and Jin (2007), Freeman and Saatchi (2004), Lin et al. (2003), Kimura (2009), and Meyer and Nicoll (2008).

14-12.3 Impact on Radiometric Data

Using an ionospheric model similar to that discussed earlier in Section 14-12.1, Le Vine and Abraham (2002) modeled the effect of the Faraday rotation on radiometric emission from the ocean at L-band (1.4 GHz). The motivation was to determine the degree of error caused by the Faraday rotation relative to the requirement to measure sea surface salinity with an accuracy on the order of 0.1–0.2 psu (practical salinity unit), as discussed later in Section 18-2.4. An accuracy of 0.1 psu corresponds to a change of brightness temperature on the order of 0.05 K (Le Vine and Abraham).

Figure 14-56 displays simulated world maps of the Faraday rotation angle at L-band at $\theta = 30°$, for both 6 am and noon local time. The simulations, which were performed for high solar activity, show values for χ extending from 0 to greater than 15°. The change in brightness temperature, ΔT_B, due to the Faraday rotation is shown in Fig. 14-57 as function of the sunspot number R_s, which varies between 10 for a quiet sun to 150 for a highly active sun. We observe that at noon and $\theta = 50°$, ΔT_B increases from about 0.3 K at low sunspot activity ($R_s = 10$) to as much as 2.1 K at $R_s = 150$. These brightness-temperature errors are certainly much larger than the required precision of 0.05 K specified for measuring ocean salinity, which calls for the application of highly accurate correction algorithms (Le Vine and Abraham, 2002; Yueh, 2000).



(a) 6:00 a.m.

(b) Noon

Figure 14-56: Global distribution of Faraday rotation for local time of 6:00 a.m. and noon. The data are for high solar activity (June 1989), an altitude of 675 km, and looking perpendicular to the satellite heading to the right at an incidence angle of 30°. [Le Vine and Abraham, 2002].

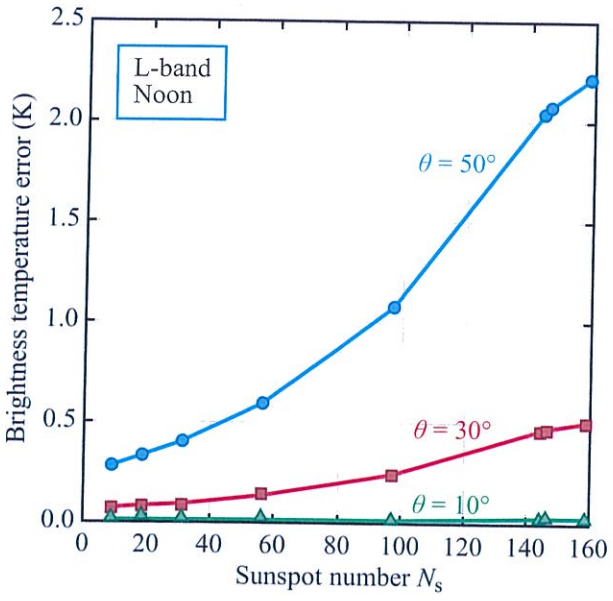


Figure 14-57: Error in brightness temperature due to neglecting Faraday rotation at noon at 30° N, 220° E for high solar activity (June 1989). The data are for a sensor at altitude of 675 km and looking perpendicular to the satellite heading to the right. The surface is ocean with $S = 35$ psu and $T_0 = 20$ °C. [Le Vine and Abraham, 2002.]

PROBLEMS

14.1 A SLAR uses an antenna with a $[(\sin x)/x]^2$ gain pattern with a 3 dB beamwidth of 30°. The antenna boresite is pointed at $\theta = 60°$ from a height of 100 m. For this problem assume that σ^0 over the ocean is given by

$$\sigma^0 = K \exp -\theta/6.5°,$$

and over land σ^0 is given by

$$\sigma^0 = K \exp -\theta/25°.$$

Calculate the relative signal power as a function of θ from 30° to 80° for land and ocean. Compare and discuss the effects of the antenna gain pattern and the incidence angle variation of σ^0 on the received signal power.

14.2 A SLAR operating at a frequency of 10 GHz uses 50 ns pulses and a 20 MHz receiver bandwidth. It has a 4 m long antenna with a uniform aperture distribution. Assuming a flight altitude of 5 km, determine the following as a function of across-track distance at 5 km intervals from 5 to 40 km:

(a) Azimuth resolution.

(b) Range resolution.

(c) Equivalent square pixel dimensions.

(d) Potential number of independent samples.

14.3 A SLAR has the following characteristics when flying at 100 m/s at a height of 2 km:

Peak power:	25 kW
Pulse duration:	100 ns
Peak antenna gain:	30 dB
Horizontal beamwidth:	0.01 mrad (two-way)
Noise figure:	10 dB
Bandwidth:	15 MHz
Frequency:	10 GHz
PRF:	2000 Hz

Calculate the following at 10 km ground range, assuming the antenna boresight is pointed at 10 km ground range:

(a) Ground-range resolution.

(b) Along-track resolution.

(c) Signal-to-noise ratio for $\sigma^0 = -20$ dB.

(d) Number of independent samples.

14.4 Separately sketch to scale the Doppler frequency versus time for a radar with 3 cm wavelength, 3 m antenna length, and speed of 200 m/s for targets at slant ranges of 5 km and 20 km. Indicate the frequency limits for an unfocused SAR in the sketch.

14.5 A SAR has a 3 m long antenna with $\lambda = 30$ cm and flies at 200 m/s. Plot the phase versus distance along a synthetic aperture for a point target at slant ranges of 5 and 20 km.

14.6 A radar operates at 10 GHz with a 3 m long antenna. Compare the azimuth resolution for SLAR, unfocused SAR, and fully focused SAR at slant ranges of 1, 5, 10, 50, 100, and 500 km.

14.7 A satellite SAR has the following characteristics:

$r_a = 25$ m	$u = 7.5$ km/s
$r_r = 25$ m	$\sigma^0 = -20$ dB
$l = 10$ m	$h = 500$ km
$G = 55$ dB	$\lambda = 3$ cm
$\theta = 30°$	range chirp $TB = 500$
$F = 5$ dB	$S_1 = 5$ dB

Calculate P_{av}, P_t and S_{N_s}.

14.8 A spaceborne radar with an aperture of $l = 10$ m at a 3 cm wavelength and elevation beamwidth of 7° is flown at 7.5 km/s at an altitude of 500 km and observes at $\theta = 20°$ directly to the side of the flight track. Plot isodops on the ground for this case. If the spacecraft is pitched forward 1°, locate the beam on the ground and plot isodops through it. Discuss frequency correction to compensate for the difference in mean Doppler shift. For simplicity, assume a flat Earth.

14.9 A dual-frequency SAR operates on an aircraft at frequencies of 1 and 10 GHz. Both frequencies use 2 m long antennas.

(a) For resolutions of 3 m in slant range and azimuth, compute and discuss the range curvature at ranges of 5 and 25 km.

(b) Repeat for 15 m resolution.

14.10 Calculate depth of focus for a SAR with $l = 6$ m, $r_a = 3$ m, $\lambda = 5$ cm, and $R = 600$ km. Repeat for $r_a = 30$ m and 60 m.

14.11 A spaceborne SAR has 3 cm wavelength and a 10 m long antenna. It operates at an orbital height of 500 km and a velocity of 7.5 km/s.

(a) Determine the swath width if the inner edge of the swath has an incidence angle of 20°. Use spherical-Earth geometry.

(b) If the resolution is to be 10 m × 10 m, determine the raw data rate in complex samples per second.

(c) Repeat part (b) for 50 m × 50 m resolution.