Synthetic Aperture Radar Fundamentals and Image Processing

Robert O. Harger

Electrical Engineering Department - University of Maryland College Park, MD 20742

ABSTRACT

The basic concept of the generation of a synthetic aperture and its incorporation into an imaging radar are reviewed. Several methods approximately describing electromagnetic scattering are discussed, along with an oceanographic application. Recent studies of image processing against clutter, for multiple images, and for image segmentation are presented, along with applications to polarimetric SAR imagery.

INTRODUCTION

The synthetic aperture radar (SAR) system employs data processing to estimate the along-track, or azimuth, position of a scatterer using the Doppler modulation due to relative motion. The manner in which SAR achieves azimuth resolution is reviewed for a CW radar: the basic idea, the formation of the complex azimuth history, and the processing, especially by an optical processor. The SAR employs periodic pulse modulation to measure time delay, hence range, to a scatterer: this introduces range and, due to sampling, azimuth ambiguities. Data formatting, compression and signal-to-noise ratio (SNR) are reviewed.

Some topics in electromagnetic (EM) theory are reviewed. The Kirchoff-Huyghens diffraction theory adequately characterizes the free space propagation used by the SAR system, in particular the optical processor. The approximation of the EM scattering phenomenon by ray theory, surface waves, the method of physical optics and the two-scale theory are discussed. An application to oceanography is mentioned.

Topics, some advanced, are discussed in SAR image processing. Background clutter is the serious limiter of performance in some applications: modelling and processing are reviewed. Constantly increasing performance requirements and advances in SAR technology have led to processing multiple images: a specific near-optimal processing algorithm is reviewed. A quite recent and advanced method of segmenting SAR images into classes, or homogeneous regions, employing hidden Markov random field models, is reviewed.

1. SAR PRINCIPLES

The way in which the SAR system (Cutrona, et. al., 1961) achieves fine azimuth resolution may be explained from several points of view: the construction, through data manipulation, of a virtual antenna (synthetic aperture) of relatively fine beamwidth (Heimiller, 1962), as a Doppler filter (Sherwin, et. al., 1962), as a compression, or matched filter, processing (McCord, 1962), and as an hybrid holographic system (Cutrona, et. al., 1960; Cutrona, et. al., 1966). The attendant analytic descriptions are all equivalent (McCord, 1962; Cutrona and Hall, 1962): the matched filter processing viewpoint most easily leads into the broader aspects of signal and image processing and will be taken here. The hologram viewpoint (Leith and Upatnieks, 1962) is of great power, pointing out that the compression filter is realized by free space propagation!

In order to get a second coordinate and thereby locate a scatterer on a terrain plane, pulses are emitted periodically, allowing measurement of time delay and hence range. Unavoidably attendant are "range ambiguities" due to "multiple sweep returns" and "azimuth ambiguities" due to unavoidable aliasing caused by the sampling operation. The concept of a SAR range-Doppler coordinate system enables a simple description of basic properties of the mapping from scene to image.

Synthetic aperture concept

1. *SAR basic idea.* - The SAR achieves an along-track, or azimuth, resolution much finer than that afforded by its antenna beamwidth by data processing that exploits the Doppler modulation due to the relative motion of the SAR and the scene. The basic idea is seen most simply by considering a continuous-wave (CW) radar, transmitting and receiving at wavelength λ_0 , borne by a vehicle in uniform translation $x = v\,t$, past a "point scatterer" at an abeam range R_0 : the transition from "up Doppler" to "down Doppler" locates the scatterer abeam. This "azimuth history" B is analytically described by a complex function whose phase is given, in radians, by $2\,\pi$ times the round-trip ray path measured in wavelengths and whose amplitude $|B|$ is determined by the farfield antenna pattern. In the usual case of relatively narrow beamwidths the phase has a quadratic approximation. Thus

$$B(x) = |B(x)| \exp(i K x^2/2)$$

where the "linear FM rate" $K = 4\,\pi/\lambda_0 R_0$.

The nominal extent of B is $X = (\lambda_0/D_h) R_0$, the product of the antenna beamwidth and the nominal range; D_h is the antenna aperture's x-dimension. X is the resolution of the unprocessed data. The "local wavenumber", Kx, the derivative of the quadratic phase, varies nominally over a wavenumber bandwidth $\Omega = K X = 4\,\pi/D_h$, affording, as will be seen, a nominal SAR azimuth resolution $\rho_a = 2\,\pi/\Omega = D_h/2$. As $X/\rho_a = 2\,\lambda_0 R_0/D_h^2 >> 1$, the SAR data processing affords a much smaller resolution- which, unlike X, is independent of wavelength and range. Therefore, the azimuth modulation is said to be "dispersed" and the SAR processor is said to "compress" it.

Since the complex signal locally has a wavenumber Kx, a "delay line" which assigns a "local delay" to a wavenumber accomplishes this compression operation (Klauder, et. al., 1960). The idea is sketched in Fig. 1.

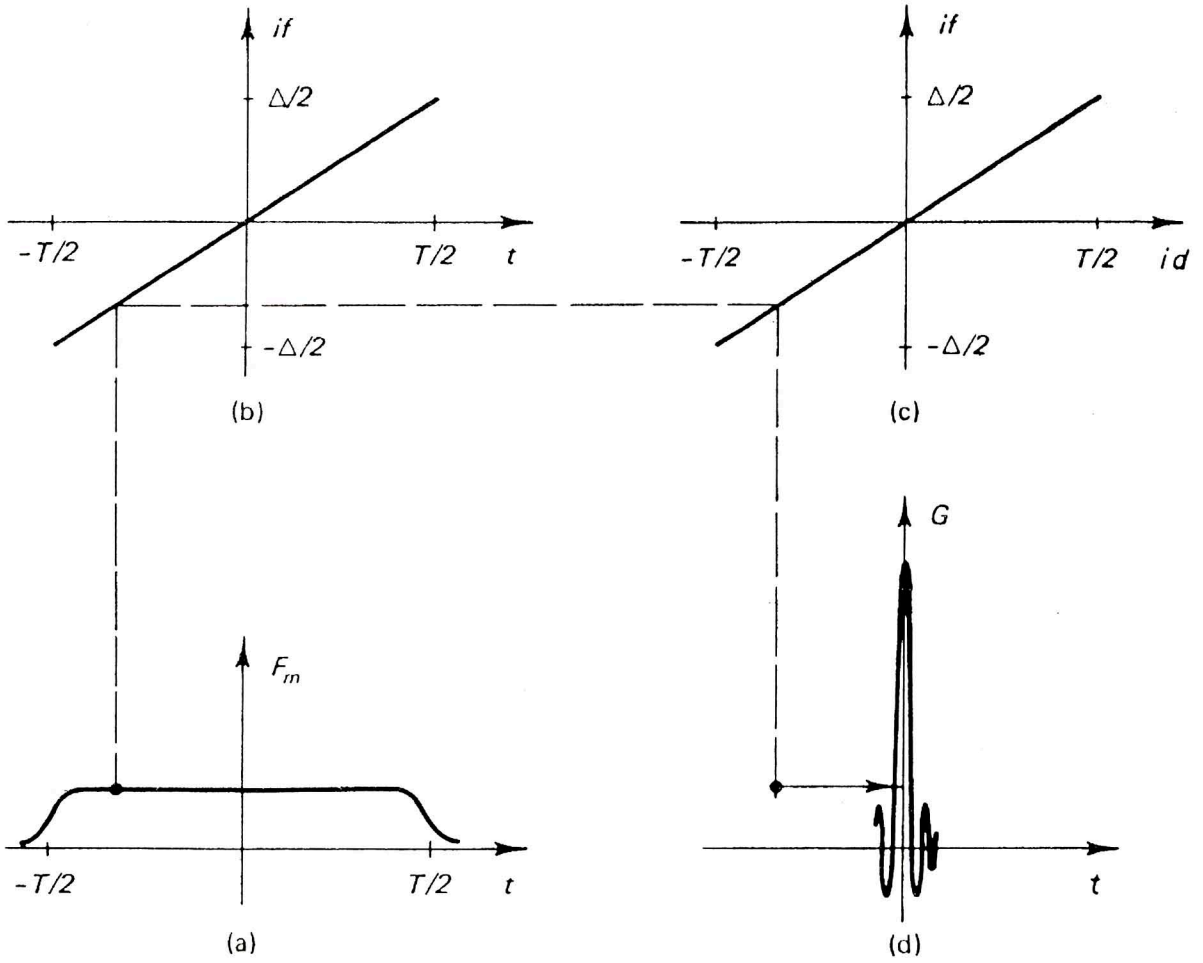

Fig. 1 - A graphical method of computing features of the compression of a linear FM signal. (a) Envelope F_m (t) and (b) instantaneous (local) frequency of input signal. (c) Compression filter's delay versus instantaneous (local) frequency characteristic. (d) Output signal G.

To find the detail of the compressed image one computes the filter's output: it is informative to do this "in the transform domain". Always in SAR systems the "space-bandwidth product" $X\Omega = 4\pi\lambda_0 R_0/D_h^2 >> 1$: then the method of stationary phase gives a close approximation to the spectrum (Fourier transform) $\tilde{B}$ of B as proportional to $B(k/K)$:

$$\tilde{B}(k) = \tilde{B}_0 \left| B\left(\frac{k}{K}\right)\right| \exp\left(\frac{ik^2}{2K}\right)$$

The linear delay line then has a wavenumber response $\tilde{H}$ that must have the phase $(-k^2/2K)$ over a bandwidth selected by the "window" $|B(k/K)|$: thus the compression filter has a wavenumber response

$$\tilde{H}(k) = |B(k/K)| \exp(-ik^2/2K)$$

The filter's output is then the inverse transform of $(\tilde{B}\tilde{H})$, or

$$q_a(x) = (q_{a\,o}/2\pi)\int dk \exp(ixk)\,|\tilde{B}(k)|^2,$$

the azimuth impulse response of the SAR. Note that the processing filter depends on the range.

Thus, as $|\tilde{B}(k)|^2 = |B(k/K)|^2$, the far-field antenna pattern determines the nature of q_a . As $|B|$ has nominal extent $X = (\lambda_0 R_o/D_h)$, $|B(k/K)|$ has nominal extent $\Omega = KX = 4\pi/D_h$: as it has no phase, q_a has extent inversely related as $2\pi/\Omega = D_h/2$ as simple examples illustrate. Alternatively, this impulse response has the convolution form

$$q_a(x) = q_{a\,o}\int dx_1 B[-(x-x_1)] * \int dx_2 B(x_1 - x_2)\,\delta(x_2):$$

the scene is described by (i) the Dirac delta function modelling a "point" scatterer $\delta(x)$, (ii) the interior integration modelling the SAR "scanning operation" producing the complex SAR signal, and (iii) the exterior integration modelling the SAR compression processing. This "matched filter processing" can be arrived at from other points of view, as mentioned above, and also under various statistical criteria, given a model with additive thermal noise- an early derivation is due to North (1943).

Experience and analysis show that this signal compression operation is quite sensitive to spurious phase and motion effects. For example, if the scatterer has a radially directed velocity component of magnitude v_R , its image appears at a false zero-Doppler location with a position error of

$$\Delta_X = 2\left(\frac{V_R}{V}\right)\left(\frac{R_0}{D_h}\right)$$

where v is the speed of the SAR-bearing vehicle. Since $R_0/D_h >> 1$, one must have $v_R/v < 1$. For example, suppose v = 25,000 ft/sec and R_0 = 150 n.mi.: then the position error is 36 v_R ft. This phenomenon has the often observed, amusing effect of vehicles appearing midst fields adjoining their thoroughfares! If the scatterer has an azimuth-directed velocity component of magnitude v_p then its linear FM rate is not matched to the processor and the compression will be imperfect. In order to preserve a well-designed SAR impulse response q_a, the maximum quadratic phase error, at the nominal extremes $\pm X/2$, should be small relative to $\pi/2$, the Rayleigh criterion. This implies the requirement that

$$\frac{v_p}{v} << \frac{2\pi}{\Omega X}.$$

E.g., if $\frac{2\pi}{\Omega X} \approx 10^{-3}$ and $v \approx$ 18,000 mph, the condition is v_p << 18 mph, which is easily violated.

The extent of the azimuth histories, and hence their bandwidth and intrinsic resolution, was set by the beamwidth of the physical antenna. If the latter has no directivity at all in azimuth then the infinite duration azimuth history, no longer linear FM in its extremities, has a limiting bandwidth set by the forward and aft Doppler frequencies, affording a limiting resolution of

$$\rho_\infty = \lambda_0/4 .$$

The physical antenna beam may alternatively be dwelled upon a fixed area, thereby generating a longer azimuth history, as in the SPOTLITE SAR system. The linear FM approximation still obtaining, an extent X_* yields a bandwidth $\Omega_* = KX_* = \theta_*(4\pi/\lambda_0)$ where $\theta_* \equiv X_*/R_0$ is the aspect angle over which the history is gathered. The commensurate resolution is

$$p_* = \left(\frac{\theta_*}{\pi}\right)\left(\frac{\lambda_0}{4}\right).$$

2. *Elementary SAR optical processor.*- If one were able to record the complex azimuth modulation on photographic film and illuminate it with a monochromatic plane wave, the emerging light would be observed to come to a focus. Thus propagation in free space performs the compression! Thus an optical computer, highly parallel and essentially instantaneous, aside from a fixed delay for film exposure and development, is a natural and very attractive analogue computer in this application.

The surface of focus will be a plane inclined to the optical axis normal to the plane of the film, analogous to the

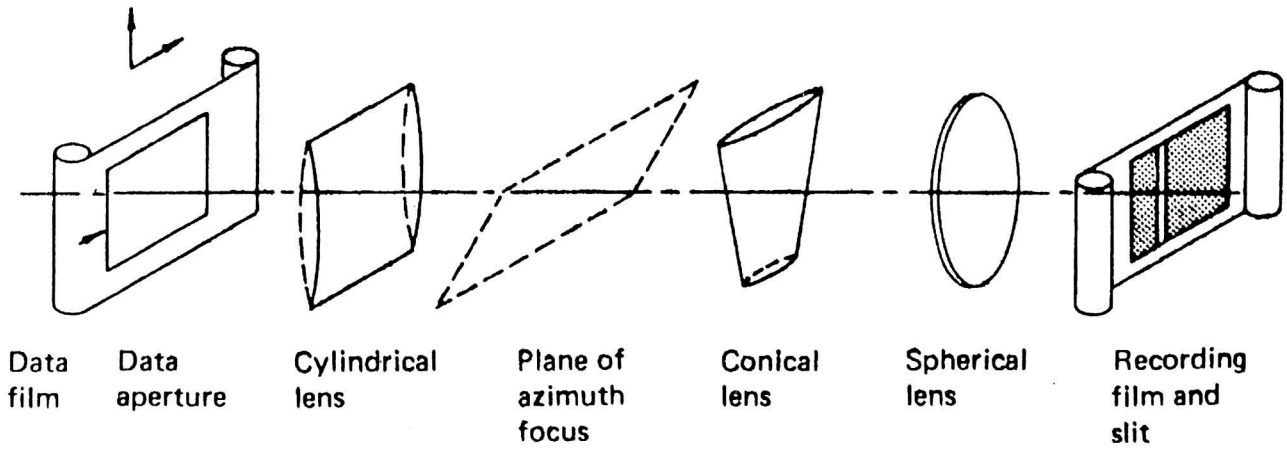

Fig. 2 - A coherent optical system to compress the SAR signal film.

inclination of the mean terrain plane of the scene to the antenna's boresight. A "conical" lens, whose focal length depends on transaxial position, will erect this focal plane and a spherical lens will refocus it to a convenient location.

Since film responds to intensity, a clever trick is required. The complex signal is kept on a suitable "spatial carrier" $\exp(i k_c x)$ of wavenumber k_c and added to a suitable "bias" b_0. Recorded on the film will then be

$$|B(x)\exp(i k_c x) + b_0|^2 = [\,|B(x)|^2 + b_0^2\,] + 2\,Re\,[\,B(x)\exp(i k_c x)\,].$$

The first term emanates along the optical axis while the second term creates divergent and convergent wavefronts at angles to the optical axis, corresponding to, respectively, virtual and real images of the scene, all thus separable.

The offset wavenumber k_c may be created using an offset frequency in the SAR, derivable by offsetting the radar receiver's local oscillator ("range offset mode") or by pointing the antenna off the zero Doppler direction ("azimuth offset mode"). The critical role of the offset frequency/ wavenumber was recognized by Leith and Palermo (Cutrona, et. al., 1960) in the SAR context and contemporaneously by Leith in the holography context (Leith and Upatnieks, 1962).

3. *Digital processing: a remark.*- The basic operation required by the azimuth compression operation is a complex convolution, equivalent to four real convolutions. If these are performed directly, for each range resolution interval in the unambiguous range interval- discussed below, and the unprocessed data is not allowed to accumulate, then, aside from any range processing that might be required, the allowed time, T_c, for a real multiply-add is about

$$\left(\frac{D_h^2}{\lambda_0 R_0}\right)\left(\frac{D_h}{8C}\right) << \frac{D_h}{8C} \approx 4 * 10^{-10} D_h \, sec$$

(D_h in meter units). This is a formidable requirement and, for emphasis, may be put in an amusing form,

$$4\,c\,T_c << \frac{D_h}{2}:$$

in the time allowed for a real multiply-add, a light wave travels a distance very much smaller than the SAR azimuth resolution.

While techniques such as the "fast convolution"- i.e., computing the convolution with fast Fourier transforms and using methods devised in digital signal processing such as "overlap-add" and "overlap-save" tricks- the success of digital processing in all but special systems has awaited the development of special computing architectures.

Synthetic Aperture Radar

In order to locate an object on the terrain the azimuth position estimate must be supplemented with another cordinate estimate. Hence, modulated pulses are periodically transmitted, rather than a CW, and the round-trip delay measured. The azimuth histories are now sampled: because it is impossible to confine the far-field azimuth

antenna pattern to the desired small angle and because a useful range interval must be imaged, in practice there is aliasing due to inadequate sampling. Depending upon the offset mode of SAR operation- mentioned above- this causes a dispersed background and/or false images at locations $n\,X$, $n = \pm 1, \pm 2, \ldots$. Evidently the design of the far-field antenna pattern and choice of sampling frequency, or PRF, is a crucial SAR system design consideration.

The employment of periodic pulses also introduces the problem of range ambiguities, known from the earliest days of radar as "second (third,...) sweep returns". In SAR systems there is an inherent suppression of such returns because their azimuth histories' linear FM rate is mismatched to that of the unambiguous range return and because the scene is viewed at an intermediate incidence angle that allows the design of the "vertical" antenna pattern to assist in the their suppression.

The data available for processing are now basically a collection of "range traces". With the intent of producing a two-dimensional image and, in early SAR systems, of using an optical processor, these range traces are recorded "side-by-side" on film, sample smoothing occurring en passant with appropriate design. When the azimuth histories are collected over sufficiently small angles, the subsequent processing, and the SAR system impulse response, are separable into an "azimuth channel" and a "range channel". The already-discussed optical processor may be used, preceded by range processing as required.

If a linear FM is used for the pulse modulation, then, again, propagation in free space provides the needed compression operation. In an optical processor a cylindrical lens, effective in the range coordinate, brings the plane of focus into coincidence with the erected plane of focus of the azimuth histories (Cutrona, et. al., 1960; Cutrona, et. al., 1966).

The analytic description of the SAR complex signal is now

$$S(x,y) = B(x)\,F(y)$$

where $F(y)$ is the pulse modulation, expressed as a function of slant range, and ignoring the effects of aliasing and ambiguous range returns. The compression filter is now of two dimensions and a motivation for its choice remains as discussed above: the matched filter is now

$$H(x,y) = S(-x,-y)^*$$

and so the corresponding output is given by

$$q(x,y) = \int d\,x_1\,B\,[-(x-x_1)\,]^* \int d\,y_1\,F\,[-(y\,H\,x_1)\,]^* \int d\,x_2\,B\,(x_1 - x_2) \int d\,y_2\,F\,(y_1 - y_2)\,\delta\,(x_2, y_2)\,.$$

The integrations over x_1, y_1 may be carried out, yielding

$$q(x,y) = q_B(x)\,q_F(y)$$

the central response of the SAR ambiguity function or impulse response, a product of the central response in azimuth and the central response in range.

The magnitude of the peak of the impulse response relative to the processed thermal noise level is a measure of the SAR system's ability to detect and locate objects. A simple, approximate derivation of the "signal-to-noise ratio, SNR" begins by recalling the radar transmission equation giving the received, instantaneous power $P_{r\,1}$ for a transmitted pulse of instantaneous power $P_{t1} = |F(t)|^2/2$,F the complex pulse modulation, scattered by a "point-like" object of radar scattering cross-section (RCS) σ, namely

$$P_{r\,1} = \frac{A^2\,\sigma\,P_{t\,1}}{4\,\pi\,\lambda_0^2\,R_0^4}$$

where **A** is the area of the physical antenna aperture. The matched-filter processor produces the maximum possible SNR, the ratio of signal energy to the thermal noise spectral density level $\mathbf{N_0}$. If f_s pulses per second are transmitted, then, in the duration of an azimuth history, a total of $N = (X/v)\,f_s$ pulses are summed, of total received energy

$$\mathbf{E} = \frac{\mathbf{A}^2\,\sigma\,P_{ave}}{4\,\pi\,\lambda_0\,R_0^3\,v\,D_h}$$

where the average transmitted power

$$P_{ave} = f_s\,\frac{1}{2}\int\,|F(t)|^2\,d\,t$$

Thus $SNR = \dfrac{\mathbf{E}}{\mathbf{N_0}}$; $\mathbf{N_0} = k_B\,T_e$, a product of Boltzmann's constant and an "effective system noise temperature", in SAR normally determined by the temperature (in degrees Kelvin) of the viewed scene.

The design of the ambiguity function (see Fig. 3) of the SAR is a much-discussed topic, as it was for the classical range-Doppler radar (Woodward, 1953). If the azimuth history is sampled at least once in a resolution interval, a time interval of ρ_a/v sec, and as the time required to scan a slant range interval Δ_R is $2\,\Delta_R/c$, one has the fundamental SAR restraint of

$$2\,\frac{\Delta_R}{\rho_a} = \frac{c}{v}$$

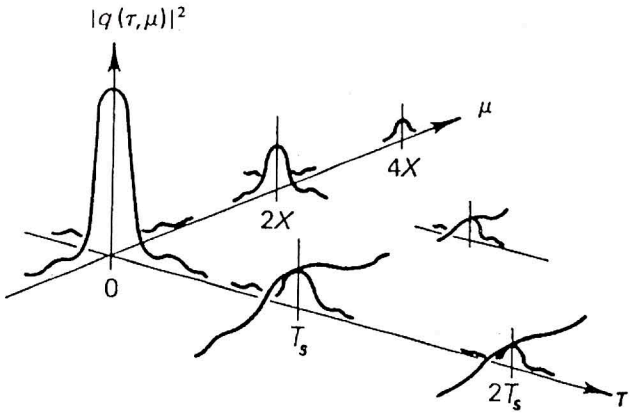

Fig. 3 - A typical SAR ambiguity function.

Since the unambiguous range interval can be quite small in, e.g., fine resolution satellite SAR systems, much thought was given to obviating this restraint, the general conclusion being that, unless one builds more complicated systems - clearly two SAR sytems can map twice the range interval - the ambiguity design problem is controlling, rather than avoiding, the ambiguity reponse. It may be argued that a certain "maximum energy in an interval" criterion is particularly appropriate for antenna pattern and signal modulation design, leading to aperture weighting functions (Slepian and Pollock, 1961) and nonlinear FM pulse modulations (Fowle, 1964): extensive design procedures for major SAR system parameters have been described (Harger, 1965).

While the "delta function" provides a description of the reflection by a point scatterrer, one may hope to describe a scene by some function $G(x,y)$ that could be used in its place in the convolution characterization of the SAR image. Finding useful such descriptions requires approximation methods from electromagnetic scattering theory in general, a matter taken up below.

When the duration of an azimuth history is sufficiently long, given a fixed range resolution, the range to a scatterer will observably change during the azimuth history. The received complex signal is now better modelled as

$$S(x,y) = S_0 B(x) F\left(y - \frac{x^2}{2} R_0 \right)$$

Now the range and azimuth dependencies do not factor, as will not the dependencies in the matched filter and SAR system impulse responses. In order to use an optical processor a clever re-examination of the two-dimensional formatting of the range sweeps led to their being arrayed in a fan-like pattern (Walker, 1980).

This technique was employed in SPOTLITE SAR along with another clever technique, STRETCH, to reduce the processor's range channel bandwidth for commensurate range resolution: a delayed FM wave was used as a local oscillator, thereby converting range delay uniquely into frequency difference so that a Fourier analysis estimated range- albeit over a restricted range interval.

SAR range-doppler space

A simple but powerful concept is the range-Doppler coordinate system of a SAR whose construction requires a knowledge only of the varying ray path from transmitter to scatterer to receiver. This idea is discussed here in the more inclusive context of a bistatic SAR, a system whose transmitter and receiver can be on separate vehicles. Such an arrangement can be useful in several ways. For example, the scattering characteristics of an object can be quite different in a bistatic configuration; also, expendable receivers can be used that approach the imaged scene directly.

The physical configuration of a bistatic SAR is determined by the assignment to the transmitter and receiver a position, respectively $\vec{P_T}, \vec{P_R}$, at the reference time $t = 0$ and a velocity, respectively $\vec{v_T}, \vec{v_R}$. It is convenient to use a coordinate system centered in the scene whose locations are denoted $\vec{p}$; it is assumed that, for any location in the scene, $p << P_T, p << P_R$ to simplify discussion here. It is also assumed that, over the time required to generate the azimuth histories, the vehicle motions are rectilinear: then the transmitter and receiver loci are, resp.,
$(\vec{P_T} + \vec{v_T} t)$, $(\vec{P_R} + \vec{v_R} t)$.

The bistatic SAR, as the conventional, monostatic SAR, will radiate pulses to measure location in slant range, thereby sampling the Doppler modulation due to relative motion which provides another position coordinate. Thus, the total ray path from transmitter to receiver, as a function of time, contains the information that allows position estimation in the scene: merely by examining it, much is revealed about the bistatic SAR's capability.

The range to a scatterer at location $\vec{p}$ is

$$R(\vec{p}) = <\vec{P_{TN}} + \vec{P_{RN}}, \vec{p}>$$

where <.,.> denotes inner product and $\vec{P_{TN}}, \vec{P_{RN}}$ are unit vectors in the direction of, resp., $\vec{P_T}, \vec{P_R}$. Therefore, the surfaces of constant total range are approximately planes whose intersection with the mean plane of the scene are straight lines of constant range that are orthogonal to $\left(\vec{P_{TN}} + \vec{P_{RN}}\right)$.

The range rate to a (stationary) scatterer at location $\vec{p}$ is

$$v(\vec{p}) = <\vec{V}, \vec{p}>$$

where

$$\vec{V} = -\left(\frac{v_T}{P_T}\right)\left(\vec{v}_{TN} - <\vec{v}_{TN}, \vec{P}_{TN}>\vec{P}_{TN}\right) - \left(\frac{v_R}{P_R}\right)\left(\vec{v}_{RN}, - <\vec{v}_{RN}, \vec{P}_{RN}>\vec{P}_{RN}\right)$$

where $\vec{v}_{TN}$, $\vec{v}_{RN}$ are unit vectors in the direction of, respectively, $\vec{v}_T$, $\vec{v}_R$. Therefore, the surfaces of constant range rate are planes whose intersections with the mean terrain plane give straight lines of constant range rate that are orthogonal to $\vec{V}$. Note that its first component is orthogonal to $\vec{P}_{TN}$ and its second component is orthogonal to $\vec{P}_{RN}$.

The rate of change of range rate is

$$K = \left(\frac{v_T^2}{P_T}\right)\left(1 - <\vec{v}_{TN}, \vec{P}_{TN}>^2\right) + \left(\frac{v_R^2}{P_R}\right)\left(1 - <\vec{v}_{RN}, \vec{P}_{RN}>^2\right)$$

independently of $\vec{p}$.
In the conventional, monostatic SAR in the abeam mode,

$$\vec{v}_T = \vec{v}_R = \vec{v}, \vec{P}_T = \vec{P}_R = \vec{P}, <\vec{v}, \vec{P}> = 0$$

Thus

$$\left(\vec{P}_{TN} + \vec{P}_{RN}\right) = 2\vec{P}, \vec{V} = -2\left(\frac{V}{P}\right)\vec{v}, <\vec{P}_{TN} + \vec{P}_{RN}, \vec{V}> = 0:$$

this configuration affords an orthogonal coordinate system. Also, as already noted, $K = 2v^2/R$, $R = P_T = P_R$.

In the conventional, monostatic SAR in the "squint mode" it is not true that $<\vec{v}, \vec{P}> = 0$. Since $\vec{V}$ is orthogonal to $\vec{P} = \vec{P}_T = \vec{P}_R$, the coordinate system remains orthogonal but now

$$K = 2\left(\frac{v^2}{R}\right)\left(1 - <\vec{P}_N, \vec{V}_N>^2\right)$$

Thus, for the same range rate resolution, a longer azimuth history must be gathered, rendering the processing more complex.

Note that if the "squint" is dead ahead, $<\vec{V}_N, \vec{P}_N> = 1$ and so the range rate $\vec{V}$ and its rate of change K are zero: only range can be measured.

But now consider a bistatic SAR in which the receiver-bearing vehicle is closing on the scene: $\vec{v}_{RN} = -\vec{P}_{RN}$. Then

$$\vec{V} = -\left(\frac{v_T}{P_T}\right)\left(\vec{v}_{TN} - <\vec{P}_{TN}, \vec{v}_{TN}>\vec{P}_{TN}\right)$$

a vector orthogonal to $\vec{P}_{TN}$. As the constant range lines are normal to $\vec{P}_{TN} + \vec{P}_{RN}$, if the receiver's position is chosen to be in the same vertical plane as the transmitter's position, the lines of constant range rate will be orthogonal to the lines of constant range! The rate of change of range rate is now

$$K = \left(\frac{v_T^2}{P_T}\right)\left(1 - <\vec{v}_{TN}, \vec{P}_{TN}>^2\right)$$

one-half that of the conventional SAR.

2. EM PROPAGATION AND SCATTERING

In order to usefully describe SAR systems it is necessary to accurately model elecromagnetic (EM) propagation and scattering. Understanding the SAR image in applications can require a quite sophisticated understanding of the scene's nature and the manner in which it reflects EM waves. With the concept of an "effective reflectivity density" distributed over the mean plane of the terrain, the "EM problem" divides into a propagation part, more easily described and made part of the SAR "system model", and leaves the determination of the effective reflectivity density for specific scenes for treatment by appropriate approximation methods from EM scattering theory.

Propagation theory

In order to characterize fully the SAR system's image of a point object, it is necessary to describe propagation in space. The Kirchoff-Huyghens diffraction theory (Born and Wolf, 1965) is appropriate to the task: it describes the relation between an antenna aperture illumination and its far-field pattern and also the Fourier transform relation of the fields in the front and back focal planes of a quadratic lens. The ability of an optical system to perform a Fourier transform gives it the ability to perform linear filtering operations- in particular, the compression, or matched filtering required in SAR. Here the first topic is overviewed; for a discussion of the second topic, see, e.g., Harger (1970).

Given a physical antenna aperture illumination function E_0, an emergent coherent, scalar EM wave, of wavenumber k, produces a far field E_f that, over sufficiently wide angles, is the two-dimensional, scaled Fourier transform, $\tilde{E}_0$, times a factor accounting for a quadratic approximation to the (essentially) spherical wavefront and the round-

trip ray path. The scattered field is described as $E_s = g\,E_f$ over a suitable surface Σ. The propagation back to the antenna aperture of the field E_s is similarly described by the Kirchoff-Huyghens diffraction theory. The received field E_r is weighted by E_0 and summed over the aperture. Given that an ensemble of such coherent waves are propagated, with frequency weighting described by the spectrum, or Fourier transform, of the pulse modulation F_0, centered on the RF carrier frequency, a final summation over frequency is performed.

Thus a "range sweep" may be so described analytically as

$$S(t) \approx \left(\frac{i\,k_0}{4\,\pi\,R_0}\right) e^{i\,2\,k_0 R_0} \iint d\,x_1\,d\,y_1\,F_0 \left[t - \frac{2\,(R_0 + y_1 \sin\delta_0)}{c}\right] B(x_1 - v\,t\,;k_0)\; A(y\,;k_0)\; e^{-i\,2\,k_0 y_1 \sin\delta_0}\, g(x_1\,,y_1\,;k_0)$$

The SAR incidence angle to the mean terrain plane is δ_0. Several approximations have been made.

(i) The wavelength dependencies of the antenna patterns are approximated by their dependencies at λ_0.

(ii) Any roughness of the actual terrain is not resolved by the range pulse F_0; e.g., range inversion is not modelled.

(iii) All range dependencies are approximated by that at the nominal range R_0, except in the rapidly varying phase $-2\,k_0\,y \sin\delta_0$ and in the phases of A, B.

(iv) The reflectivity density g is defined over the mean terrain plane, of coordinates (x, y) located in the scene, assumed independent of the small variation in aspect angle, and with a wavelength dependence approximated by that at λ_0.

(v) The physical antenna's aperture illumination function is factorable as $E_0(x,y) = E_{oh}(x)\,E_{ov}(y)$.

(vi) for the usual narrow antenna beamwidths

$$B(x;y,k_0) = \tilde{E}_{oh}\left(-\frac{k_0\,x}{R_0}\right) \exp\left[\frac{-\,i\,k_0\,x^2}{(R_0 + y \sin\delta_0)}\right]$$

and

$$A(y;k_0) = \tilde{E}^2_{ov}\left(-\frac{k_0\,y\cot\delta_0}{R_0}\right) \exp\left[-\frac{i\,k_0\,y^2\cos^2\delta_0}{(R_0 + y \sin\delta_0)}\right]$$

These range sweeps are recorded side-by-side, creating the two-dimensional, complex "signal film": the "fast time" $t \rightarrow 2y \sin\delta_0/C$ in S and F and the "slow time" $t \rightarrow x/v$ in B. Thus

$$S(x,y) \approx S_0 \int dx_1 \int dy_1\, B(x - x_1;y_1,k_0)\; F(y - y_1)\, A(y_1;k_0)\; e^{i\,2\,k_0\,y \sin\delta_0}\, g(x_1,y_1)$$

models the two-dimensional complex signal, again ignoring range and azimuth ambiguities. Here $F(y) = F_0(2y \sin\delta_0)/c$ where c is the speed of light.

This form defines more precisely the complex signal as a convolution of a scanning beam with an effective reflectivity density. Note in particular that

$$G(x,y) = A(y;k_0)\; e^{i\,k_0\,y \sin\delta_0}\; g(x,y)$$

Therefore it is only the wavenumber content of the reflectivity density g about $(0,-2\,k_0 \sin\delta_0)$ that can influence the SAR image.

EM scattering models

The reflectivity density, describing the EM scattering mechanism at the scene, conceptually isolates this phenomenon in the SAR system description. It is closely related to, and often would be extracted from, the commonly calculated "scattering matrix", which usually characterizes the received scattered field in terms of a ficticious, isotropic scatterer. Two broadly applicable approximations will be discussed: the method of physical optics which often applies to "cultural objects" and the two-scale model which often applies to "natural scenes".

Not to be overlooked is the method of ray optics which can easily predict phenomena described only with great difficulty otherwise- e.g., multiple scattering, shadowing and obscuration. Software systems are commonly used along with "solid modelling" software to compute the scattered field near the object- in effect computing the reflectivity density. Thus, given a suitable physical description of an object, combined with a suitable SAR simulator (processor), one can predict its SAR image.

There are other special methods that can be useful in specific instances. For example, to model accurately the scattering by an object near the terrain (within a few wavelengths) it is necessary to resort to an old theory due to Norton (1937) which includes a "surface wave", discussed earlier by Sommerfeld and Weyl. The theory can be summarized succinctly by a "planar Earth propagation factor" $\mathcal{F}$ which directly multiplies the radar transmission equation and hence is easily used. In conventional radar,

$$\mathcal{F} = |\,\mathcal{F}(\lambda_0)\,|^4$$

where

$$F(\lambda_0) \approx 1 + \Gamma\,e^{i\,\Delta} + (1 - \Gamma)\,A\,e^{i\,\Delta} + \Psi$$

In order, the terms account for the direct ray, a ground-reflected ray with phase shift Δ due to the longer ray path, the surface wave present with a finite dielectric constant which determines and higher order surface effects; Γ is the polarization-dependent plane wave reflection coeficient.

In harmonic SAR (Harger, 1976), which receives scattered waves on the transmitted frequency's third harmonic, generated by nonlinear scattering from, likely, metal-oxide-metal junctions on cultural objects,

$$\mathcal{F} \approx | \mathcal{F}(\lambda_0) |^{2\alpha} \; | \mathcal{F}(\lambda_0/3) |^2$$

where experimentally α is found to be typically . Thus the surface wave correction is essential to accurately modelling these systems: e.g., system detection performance depends dramatically upon polarization choice.

1. Method of physical optics. - For scenes or objects described by a surface height elevation and of good conductivity - or, more generally, of large dielectric constant magnitude, a powerful but tractable description of the EM scattering mechanism is provided by the method of physical optics (Kerr, 1951; Mentzer, 1955). If the incident tangential magnetic field at a point on the scattering surface is $\vec{H}_{it}$, then the total field at that point is assumed to be $2\vec{H}_{it}$. This approximation is accurate provided the surface is smoothly varying, with radii of curvature large relative to λ_0 . The Kirchoff-Huyghens diffraction integral, in vector form, describes the incident and scattered fields, the latter involving an integral over the surface.

If $\vec{k_0}$ is the direction of the incident field and $\vec{n}(\vec{r})$ is the surface normal at the point $\vec{r}$ on the surface, the reflectivity density is well-known to be

$$g(\vec{r}) = 2 < \vec{k_0}, \vec{n}(\vec{r}) >$$

Further, there is no depolarization. The surface integral describing the scattered field may be expressed as an integral over the mean terrain plane: then, in effect, there is a reflectivity density distributed over this plane of

$$g(x,y) = g_0 [\cos\delta_0 + (\partial h/\partial y) \sin\delta_0]$$

Thus the SAR image of such a scene is the convolution of the SAR system impulse response with the overall scene description

$$G(x,y) = B_v(y;k_0)\, e^{i2k_0 y \sin\delta_0 - i2k_0 h(x,y)\cos\delta_0}\, g_0 \cos\delta_0 \; [1 + (\partial h/\partial y)\tan\delta_0)]$$

At typical intermediate incidence angles, $(0,2k_0 \sin\delta_0)$ is a wavenumber very large relative to the SAR bandwidth: then the only way $G(x,y)$ can have wavenumbers sufficiently large to be hetorodyned down to the SAR bandwidth by the factor

$$\exp(i2k_0 y \sin\delta_0)$$

is via the factor $\exp[i2k_0 h(x,y)\cos\delta_0]$. An evaluation by the method of stationary phase yields the "specular points" at which a ray is backscattered directly to the antenna.

2. Two-scale model.- An improvement upon the method of physical optics in describing the scattering mechanism is obtained for surfaces which additionally have a roughness, $\xi(x,y)$, small relative to the RF wavelength and not too rapidly varying spatially (Valenzuela, 1968). Often random models of the surface are useful. If at some wavenumber $k_c << k_c$, the surface's spectral density has sufficiently small integral over, then the surface may be decomposed into a "large scale", h, and an additive "small scale", ξ . This approximation has wide application and appears to be robust in that it serves well even beyond its apparent range of applicability- e.g., in sea surface imaging (Wright, 1968; Bass, et. al., 1968).

This approximation is a perturbation of the physical optics approximation. Hueristically one may imagine that the small scale roughness results in a small perturbation to the round-trip ray path, resulting in the phase factor

$$e^{i2k_0 \xi \cos\delta_0} \approx 1 - i\,2k_0 \xi \cos\delta_0$$

Thus one expects the reflectivity density for the perturbed field to be proportional to $(2k_0 \xi \cos\delta_0)$. In fact, one finds (Bass, et. al.,1968) that to first order in the large-scale slope and for a dielectric constant of large magnitude, this reflectivity density is, for horizontal transmit and receive polarizations,

$$g_{HH}(x,y) = g_{HH_0} \cos\delta_0 [\cos\delta_0 + 2(\partial h/\partial y)\sin\delta_0] 2k_0 \xi(x,y)$$

For vertical transmit and receive polarizations

$$g_{VV} = g_{VV_0}\, 2k_0 \xi(x,y)$$

and for unlike transmit and receive polarizations

$$g_{HV} = g_{VH} = g_{HV_0} \sin\delta_0 (\partial h/\partial x)\, 2k_0 \xi(x,y)$$

which describes a depolarization effect.

Similar, but more complicated forms of the reflectivity density are available (Bass and Fuks, 1979) for arbitrary surface dielectric constant. The overall scene description is now of the form

$$G(x,y) = B_v(y;k_0) \exp[i2k_0 y \sin\delta_0 - i2k_0 h(x,y)\cos\delta_0] * g_0 [\cos\delta_0 + (\partial h/\partial y)\sin\delta_0]\; [1 - i2k_0 \xi(x,y)\cos\delta_0]$$

Now ξ itself, unlike h, has broad wavenumber content: ignoring the phase factor due to h for the moment, the heterodyning due to the phase factor $\exp(-i2k_0 y \sin\delta_0)$ means that wavenumbers of ξ centered on $(0, -2k_0 \sin\delta_0)$ are actually imaged by the SAR system. The presence of the phase factor $\exp(i2k_0 h \cos\delta_0)$ means that an even broader bandwidth of such wavenumbers influence the SAR image.

This is a complicated description of the scene sensed by the SAR and makes clear that generally one should regard the SAR image as containing artifacts of the large and small scale elevations.

3. Oceanographic applications.- EM scattering from the sea surface has been described quite successfully by the two-scale model. The SAR responds to the small scale structure with wavenumbers near $(0, -2k_0 \sin\delta_0)$. These waves, in turn, interact, in an often subtle and complicated, nonlinear manner (Phillips, 1981a), with the wind, large-scale waves, currents, wakes, bathymmetry and other phenomena, all of interest to oceanography and its application areas. Such a scene may vary with "slow time" $t \rightarrow x/v$ that can effect its SAR image: then such a scene is no longer described as a convolution of the SAR system impulse response with a (stationary) scene description, greatly complicating image understanding and simulation- entire azimuth histories now must be generated and processed.

Two-scale (WKB) methods are also useful in describing the aforementioned hydrodynamic interactions- for example, the nature of short waves propagating on a long wave: in effect they experience a current and an altered acceleration of gravity due to the orbital motion and inclination of the long wave (Phillips, 1981b). Suppose the long wave is

$$h(\vec{x},t) = A_{1/W} \cos\Phi(\vec{x},t)$$

$$\Phi(\vec{x},t) = \langle \vec{K}_{1/W}, \vec{x} \rangle - \sigma(K_{1/W})\, t$$

where σ is the gravity wave dispersion relation. Then, to first order in the long wave slope $A_{1/W} K_{1/W}$, a short wave of wavenumber $\vec{K}$ experiences an amplitude variation factor of $(1 + A_{1/W}\, K_{1/W} \cos\Phi)$ and a phase containing the contribution $K_x (x + A_{1/W} \sin\Phi)$, x the direction of the long wave, accounting for the "to-and-fro" translation by the orbital motion of the long wave.

This two-scale hydrodynamic model of small waves on long waves is combined with the two-scale EM scattering model, the SAR image of such a sea, with conventional matched-filter processing, is

$$I(x,r) = \int dx_1\, B[-(x - x_1)] \int dx_2\, B(x_1 - x_2)$$

$$\int dr_1\, q_F(r - r_1)\, G(x_2, r_2, t = x_1/v)$$

where the "slow-time" variation of the sea surface does not effect the range compression. G is composed of three factors,

$$G = G_1\, G_2\, G_3 :$$

$$G_1 = G_{10}[1 + A_{1/W} K_{1/W} \cos\Phi - A_{1/W} K_{1/W} \sin\Theta \sin\Phi \tan\delta_0]$$

where θ is the angle between the long wave and the SAR directions, containing the amplitude "straining" and "tilting" due, respectively, to the nonlinear hydrodynamics and EM scattering from the long wave; giving the phase due to the round-trip ray path to the large wave;

$$G_2 = \exp(i2k_0 r_2 - i2k_0 h \cos\delta_0)$$

giving the phase due to the round-trip ray path to the large wave;

$$G_3 = \xi_{(-)}[x_2 + A_{1/W} K_{1/W} \cos\theta \sin\Phi\, ,\, r_2 / / \sin\delta_0 + A_{1/W} K_{1/W} \sin\theta \sin\Phi\, ,\, x_1/v]$$

the ensemble of influential small waves tranlated to-and-fro by the long wave.

This model has been simulated and compared to actual SAR imagery (Harger and Korman, 1988) and has been found to accurately predict the image and important features of it. The data required for the simulation, acquired roughly contemporaneously, was the long wave directional frequency spectrum. A typical result is Fig. 4, showing the long wave, the simulated SAR image and the actual SAR image.

3. SAR IMAGE PROCESSING TOPICS

The SAR technique has been brought to a high level of reliable, specified performance. The development of digital signal processing, much more flexible than optical processing, has led to a rapid development of the field of image processing which is now influencing SAR image processing. Three representative topics are selected for discussion. First, clutter, rather than thermal noise, will determine SAR performance in some applications and it can obviate costly efforts devoted to system design against thermal noise. Second, well-registered multiple images in multiple frequencies and/or multiple aspect angles and/or multiple polarizations are now available and the question of their appropriate processing arises. The methods discussed in these two topics are quite standard from a

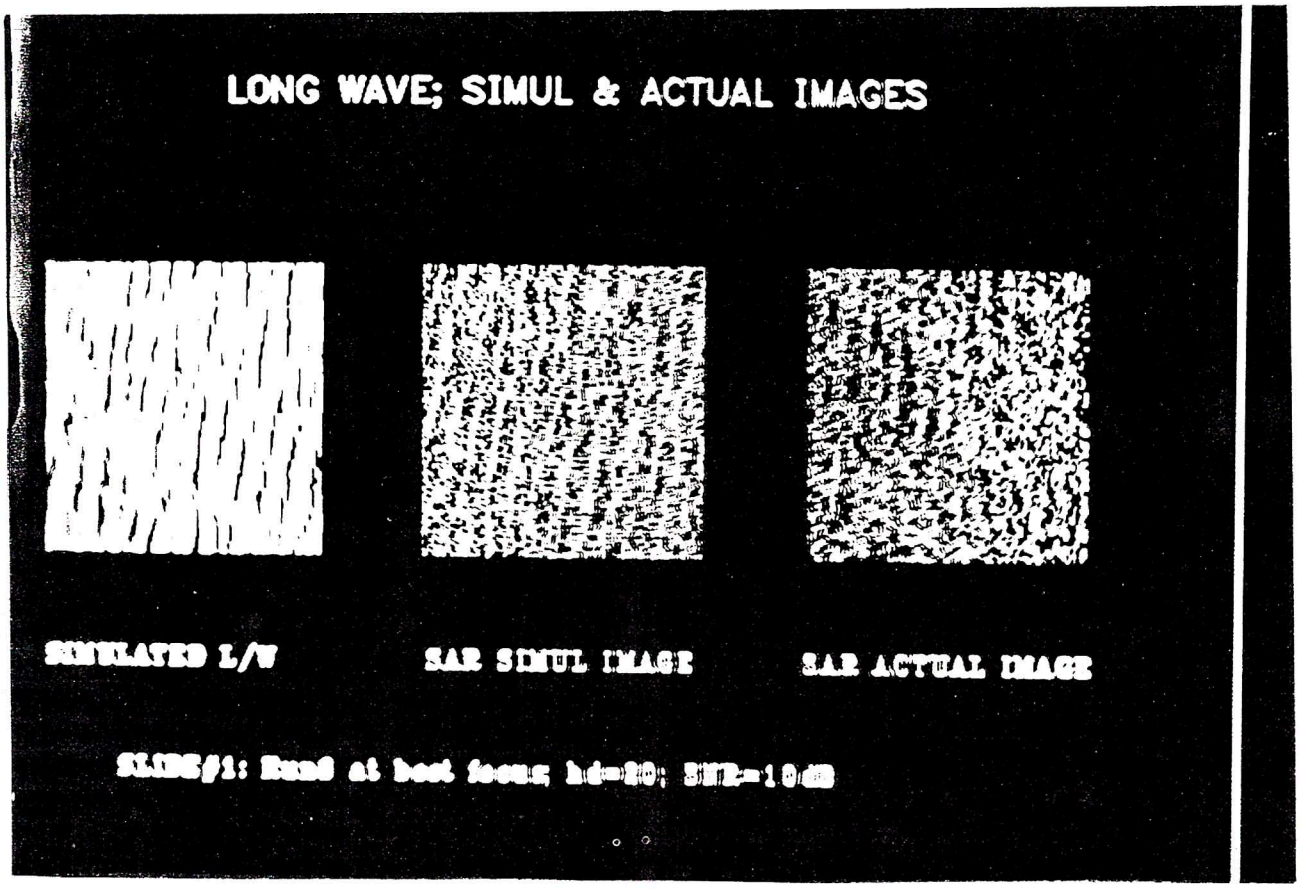

Fig. 4 - A comparison of an actual SAR image (right) of ocean gravity waves and a simulated SAR image (center) employing a simulated long wave ensemble (left, at an instant) based on a contemporary measurement of the long wave directional frequency spectrum.

statistical decision theory point of view. Third, segmentation of SAR images into classes or regions can be the object in some SAR applications and is a useful stage of image processing in other applications. The model and method discussed is quite advanced, using complicated statistical models with processing algorithms using random sampling, stochastic relaxation and simulated annealing.

Processing Against Clutter

In a remote sensing application, the object of interest often resides on terrain, whose artifacts in the SAR image are referred to as "clutter". The terrain can sometimes be modelled as of two scales, as discussed above. Usually the clutter is taken as a random field, often Gaussian, and additive to the object's image. Unlike thermal noise, the clutter, being sensed by the SAR, has a nature dependent on the SAR system.

For example, if the clutter is modelled as an homogeneous random field with a spectral density $P_{c1}(k_x, k_y)$, then the complex signal generated by the SAR is (locally) homogeneous with spectral density $P_{c1}(k_x,k_y) \, | \tilde{B}(k_x) \tilde{F}(k_y) |^2$. The thermal noise is usually taken to be of constant spectral density N_0, say. If the SAR system is designed to optimize the SNR at the object's location, then the optimal filter has a wavenumber response

$$\tilde{H}(k_x,k_y) = \frac{[\tilde{B}(K_x)\tilde{F}(k_y)\tilde{G}(k_x,k_y)]^*}{N_0 + |\tilde{B}(k_x)\tilde{F}(k_y)|^2 \rho_{c1}(k_x,k_y)}$$

The filter with this wavenumber response may be regarded as a concatenation of three filters:

(i) $[\tilde{A}(k_x)\tilde{F}(k_y)]^*$, the conventional SAR processing;
(ii) $\tilde{G}(k_x,k_y)^*$, a processing matched to a specific object;
(iii) $[N_0 + |\tilde{A}\tilde{F}|^2 P_{c1}]^{-1}$, a processing to reduce clutter effects.

Notice that, when "clutter dominates thermal noise", as is often the case in SAR systems, this optimal filter is approximately

$$\frac{\tilde{G}\,(k_x,k_y)^*}{\tilde{A}\,(k_x)\,\tilde{F}\,(k_y)\,P_{c1}\,(k_x,k_y)}$$

The appropriate system design for a desired SAR system impulse response is now quite different from that appropriate for solely additive noise.

If (i) the clutter random field is white and the background is described by a radar scattering cross-section density σ_0 and (ii) the object is a point scatterer of RCS σ, then the "signal-to-(clutter plus noise) ratio" can be expressed approximately as

$$SCNR \cong SNR\left[1+\left(\frac{\rho\,\sigma_0}{\sigma}\right)SNR\right]^{-1}$$

for a SAR area resolution ρ ; $(\rho\,\sigma_0)$ is the RCS of a resolution cell. Thus, as system resources are devoted to increasing SNR,

$$SCNR \dashrightarrow \left(\frac{\sigma}{\rho\,\sigma_0}\right)$$

so that, then, for fixed background and object, only decreasing the SAR area resolution ρ can improve matters.

The above clutter model, widely used, is clearly only a first order model in that:
(i) mutual shadowing of object and background is not modelled;
(ii) the object sits on and obscures the terrain so that clutter and object signals are strictly not additive;
(iii) the "white" assumption does not model the long-range correlations and scale similarities empirically observed in SAR imagery.

One way to model clutter with extended correlation is as a "fractional Brownian random field" (Mandlebrodt and Van Ness, 1968). Such a structure is a fractal structure in that it is similar at various scales, can have extended spatial correlation, and processing based on such a model has successfully classified natural and cultural scenes (Keller, et. al.,1987). This random field is an homogeneous, isotropic, zero mean normal random field with spectral density of the form $\beta\,k^{-\alpha}$ over an appropriate band of wavenumbers (k_1,k_2) . As α varies from 0 to 1, the random field varies from "white" to a two-dimensional version of "1/f noise". A limiting $(k_1 \to 0, k_2 \to \infty)$ form of the covariance is shown in Fig. 5.

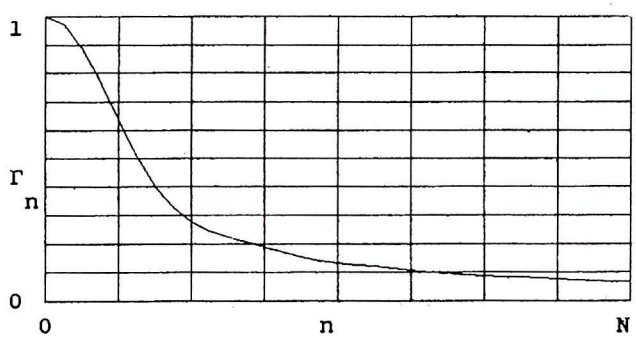

Fig. 5 - Normalized covariance function of a fractional Brownian motion random field ($\alpha = 1$ and typical SAR wavenumber response). (Abscissal units are SAR resolution units.)

The optimal processor (Harger, 1990a) is the linear filter with wavenumber response given above. It is of interest to let the scattering object be a point and then to examine this SAR system's impulse response as α varies. In Fig. 6a is shown the situation for $\beta = 0$ - no clutter- and a well-designed SAR impulse response with about -35 db first sidelobe level. With clutter present and strong relative to the thermal noise and "white", the response in Fig. 6b shows the predicted "stripping" of the designed weightings, the first sidelobe level rising to about -13 db. for $\alpha = 1$.

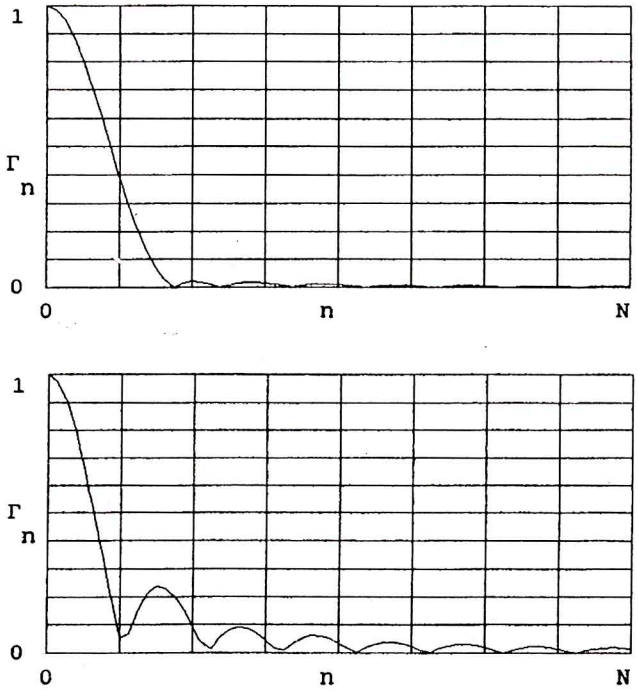

Fig. 6 - (a) The SAR system impulse response for no clutter. (b) The SAR system impulse response for strong clutter ($\alpha = 1$).

An object sitting on terrain of relatively small amplitude variation obscures the background where it resides and also casts a shadow which, in a SAR of sufficiently fine resolution, results in an image area free of clutter, containing only thermal noise effects, that can enable easy object detection. A simple model for this phenomenon leads to a "standard" statistical decision problem where the means and the variances depend on the hypotheses, object present or not, and only the image pixels where the object or shadow reside are used. The optimal processor may be regarded as an apropriate combiner of (i) a linear processing matched to the object and (ii) a quadratic processing that measures energy. The latter provides increasing detection performance for increasing clutter level: the shadow becomes increasingly obvious.

Processing Multiple Images

In the presence of strong clutter, the detection performance of even the optimal processor, assuming it is known and constructible, may be unsatisfactory in specific applications. In this event, one can consider schemes to gather additional data. Recent advances in SAR technology can offer multiple images in a single pass using combinations of multiple frequencies, multiple aspect angles and multiple polarizations.

Consider the case of two images in which corresponding pixel values are highly correlated in the case of clutter and weakly correlated in the case of signal: then differencing the images should reduce the clutter and enable enhanced detection performance (Harger, 1975). This theory has been more recently developed (Margalit, et. al., 1985; Chen and Reed, 1987) and applied to optical imagery, resulting in impressive performance gains.

Discussed here is the case of two SAR images with clutters that are uncorrelated within and between each image except that corresponding pixel values have a correlation coeficient ρ that may be nonzero. The joint distributions are assumed normal, perhaps after an appropriate preprocessing. In Chen and Reed (1987) a slowly varying mean intensity was removed; often the empirical data shows the SAR image intensity to be lognormal and so a simple logarithmic transformation renders the (univariate) distribution normal.

The data of each $\sqrt{N} * \sqrt{N}$ image are arrayed into an N dimensional vector and together arrayed into a partitioned $2N$ dimensional vector $Z = (Z_1 \mid Z_2)^t$. Because of the correlation model assumed, the $2N * 2N$ correlation matrix C has a 2 by 2 partition of $N * N$ diagonal correlation matrices: such a matrix is easily inverted analytically and hence the normal joint distribution of this data vector is known explicitly. In practice there may be unknown parameters of the distributions. When the object is present, additive to the clutter and thermal noise, there is a 2N dimensional mean value vector, partitioned into two N dimensional vectors describing the object in the two images. This mean value vector is assumed known: e.g., given a physical description, its SAR image could be generated by a simulator as mentioned above. In practice it is likely that the clutter and noise variances, the clutter correlation and an overall gain are unknown and have to be adaptively estimated. Under the usual statistical decision criteria, such as Neyman-Pearson, the optimal processor computes and compares the log-likelihood ratio to a threshold.

Known parameters.- It is useful to review the case when parameters of the distributions are known. Then the familiar theory states that a sufficient statistic has the form

$$t(Z) = Z^t C^{-1} S$$

with detectability parameter

$$d^2 = S^t C^{-1} S$$

Working out the details here, one finds that, for example,

$$Z^t C^{-1} = K\left[Z_1^t - \left(\frac{\rho}{a}\right) Z_2^t \mid Z_2^t - \left(\frac{\rho}{a}\right) Z_1^t \right]$$

which gives precisely the "generalized differences" of the images that are processed. Here $a \equiv 1 + \sigma_n^2/\sigma_{cl}^2$ where the thermal noise and clutter variances are, respectively, σ_n^2, σ_{cl}^2.

The detectability parameter is

$$d^2 = f * d_0^2$$

where

$$d_0^2 = \frac{T}{\sigma_n^2 + \sigma_{cl}^2},$$

$$E_T = S_1^t S_1 + S_2^t S_2,$$

and

$$f = \frac{1 - \left(\frac{\rho}{a}\right)\rho_s}{1 - \left(\frac{\rho}{a}\right)^2}$$

where

$$\rho_s = \frac{2S_1^t S_2}{T}$$

is the "signal correlation coeficient".

Thus the form f shows the dependence of detection performance upon the interimage clutter and signal correlation coeficients. In particular, when $\rho_s \neq \pm 1$, as $|\rho|/a$ approaches unity, f and hence d^2 can become quite large, relative to 1 and d_0^2, respectively.

Unknown parameters.- When the variance parameters and correlation coeficient are unknown, the likelihood ratio is maximized over them prior to comparison with a threshold: that is to say, the data is used to adaptively make estimates of these unknown parameters which are then use in the likelihood ratio as if known. The so-maximized likelihood ratio can be found concisely and precisely and is well-known in "analysis of variance" discussions (Anderson, 1975).
An unknown gain can enter this resultant likelihood ratio to the fourth power. While in practice it could be numerically computed, an analytical approximate solution can be found (Harger, 1990b), giving the approximate sufficient statistic

$$t(Z) = [\,(1 - R_+)\;(1 - R_-)\,]^{-1}$$

where

$$R_\pm \equiv \frac{[\,(Z_1 \pm Z_2)^t\;(S_1 \pm S_2)\,]^2}{\|\,Z_1 \pm Z_2\,\|^2\;\|\,S_1 \pm S_2\,\|^2}$$

This is an appealing processor whose implemenation mainly involves inner-product computations that can be done in parallel.

When the scatterer location is also to be estimated by a maximum likelihood estimate, $t(Z)$ is computed for all possible locations: these values may be conveniently arrayed into a "processed image". All inner products are now convolutions which may be implemented efficiently with two-dimensional fast Fourier transforms.

Under reasonable approximations the performance can be shown to be characterized as above and the approximations were supported by simulations. The cost, in increased signal-to-noise ratio, of having to estimate the unknown parameters, is about 4 db.

This processing scheme was applied to a pair of HH and VV polarimetric SAR images- Fig. 7, supplied by the Environmental Research Institute of Michigan- in order to show the promise of the idea. A particular object's images were excised from the HH and VV images, Fig. 8, and used as the apriori known signal images. The processor then estimated position and computed the "processed image" as just discussed: it is displayed in Fig. 9. As can be seen, the object response- the "bright point"- is very strong.

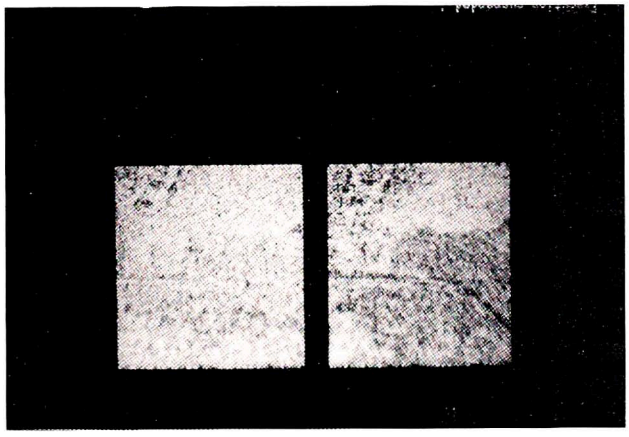

Fig. 7 - Polarimetric SAR images: (a) HH and (b) VV. Parameters: 2 m resolution, L-band.

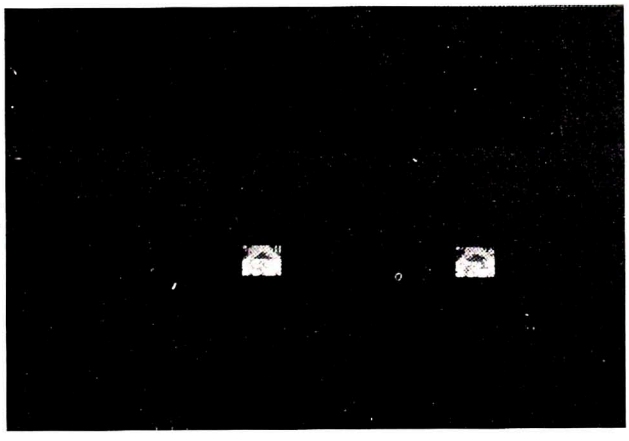

Fig. 8 - Excised SAR object images: (a) HH and (b) VV.

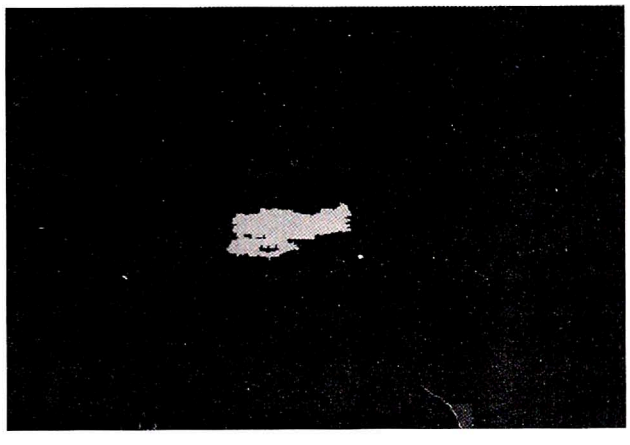

Fig. 9 - Near-optimally processed polarimetric SAR image pair.

Segmenting SAR Images

In some remote sensing applications of SAR systems, the segmentation of the image into labelled regions of statistically similar characteristics is an important aim: e.g., classification of agricultural fields or of ice fields. The resultant boundaries,e.g., tree lines, can also be of interest. As practical segmentation algorithms will adaptively estimate unknown parameters, a "texture analysis" would be performed en passant. Observing the sequence of segmentations produced by an iterative algorithm is itself informative. More generally, a segmentation algorithm would be useful as one step in a sequence of image processing operations.

An heirarchical hidden Markov random field model (Geman and Geman, 1984; Kelly et. al., 1988) may be constructed as follows. An underlying Markov random field, "MRF", realization X defines the regions or classes: that is, each pixel is assigned to one of K classes. Random field realizations $[\, Y^{(k)}, k = 1, \ldots K \,]$, distinct and independent, are available: they are called "hiding" random fields. The observed image Y is constructed as follows: if a pixel belongs to class k, it is assigned assigned the value of the corresponding pixel in realization $Y^{(k)}$.

Given such an image, one wishes to recover the MRF realization, thereby segmenting the image into K classes. The decision theory criterion can be to maximize the aposteriori probability of correct classification: using Bayes rule, this requires knowledge of the probabilities $P(Y \mid X), P(X)$. In practice there will be unknown parameters to be estimated. There are severe numerical problems in carrying out this estimation algorithm.

The appeal of MRFs is that their global behavior is determined by relatively simple local specifications, given by conditional probabilities with respect to "local neighborhoods" (Derin and Kelly, 1989). The explanation of ferromagnetism was perhaps the first, and a famous, application- the Ising model. From an image modelling point of view, extensive class regions can be modelled; further, if the hiding random fields are only locally correlated, then there is the possibility that optimal, or near optimal, processing algorithms will also be "local" in nature and hence amenable to highly parallel implementation.

The enormous number of possible realizations renders optimization by exhaustive search infeasible: a random, "Gibbs sampler" can be used, iteratively in stages of scanning over the image, to locally improve the "fit" to an MRF. Frequently in this sequence of stages, the unknown parameters must be estimated: these are maximum likelihood estimators, usually approximated, especially those of the hidden MRF. In order to avoid local maxima, one can use a simulated annealing technique.

This type of segmentation algorithm has been successfully applied to SAR images, both intensity and complex (Kelly, et. al., 1988; Derin, et. al.,1990). Here we illustrate the segmentation of a polarized pair, (HH,VV), of SAR images (Harger, 1991). Shown in Figs. 10 and 11 are 128 by 128, 2 m resolution L-band intensity images. A preliminary analysis showed the intensity to have a lognormal univariate density and so a logarithmic transformation of the intensity was presumed to yield a normal random field. The means and variances had to be estimated for each class. The images of all classes were assumed to be uncorrelated except that corresponding pixel values in the two images were assumed to be correlated. A multi-level logistic MRF with respect to a {N,S,E,W} local neighborhood was assumed and its parameter estimated also.

The segmentations are shown in Figs. 12,13,14 and it is clear that the results are reasonable. Such results may be useful in studies of the effect of polarization in SAR imagery. The parameters estimated are tabulated below.

Class	1	2
mean (HH)	4.191	4.762
mean (VV)	4.456	5.055
variance	0.178	0.198
correlation coeficient	0.226	0.319
population	11,332	5,052

ACKNOWLEDGEMENTS

Studies in some of the topics discussed here have been supported by the Environmental Institute of Michigan, under the direction of R.E.Mitchel.

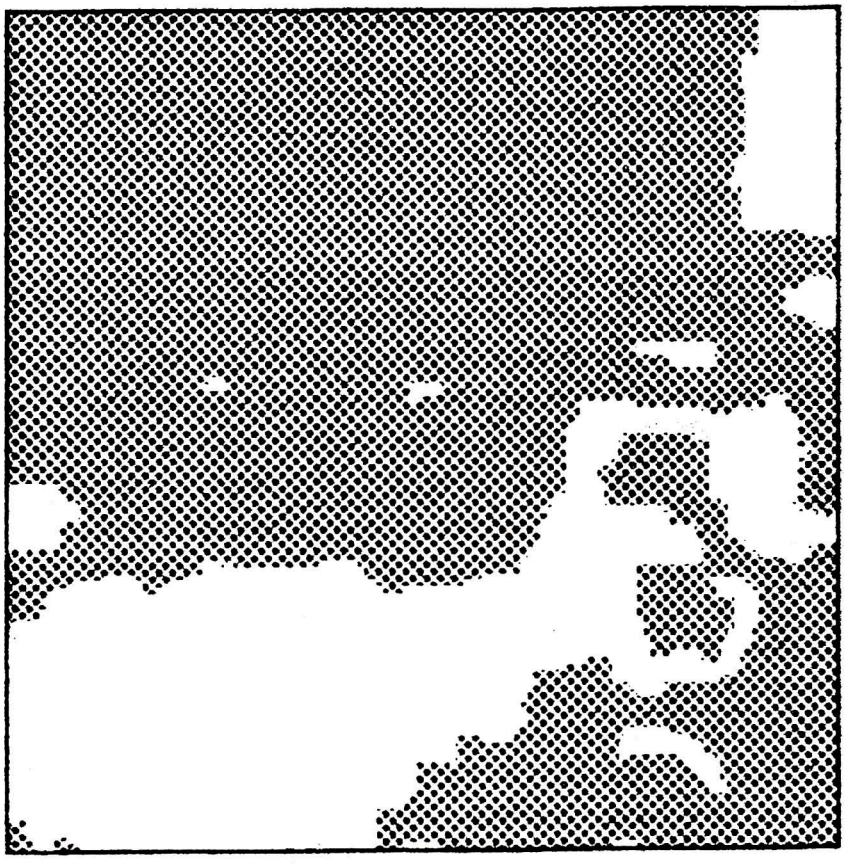

Fig. 10 - Segmentation of HH image into 2 classes.

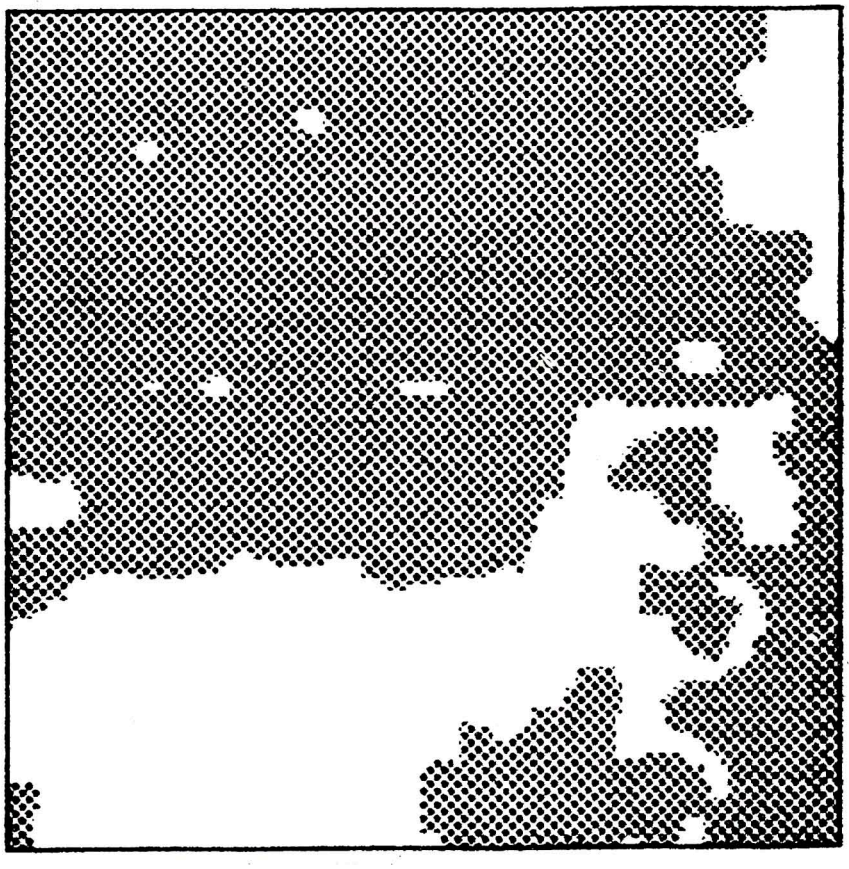

Fig. 11 - Segmentation of VV image into 2 classes.

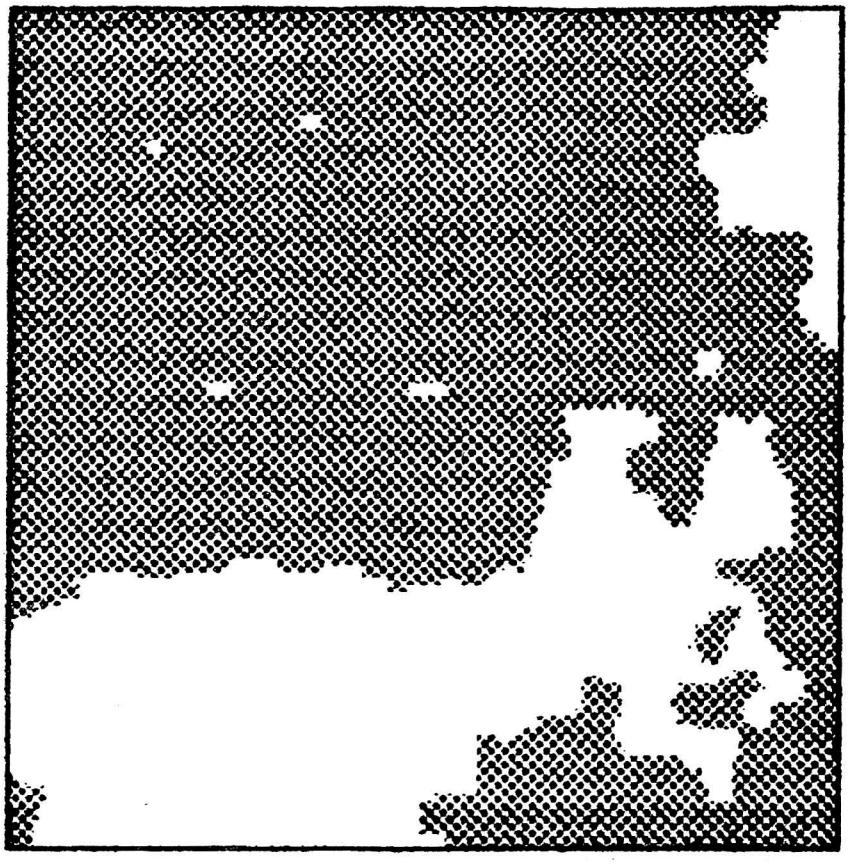

Fig. 12 - Segmentation jointly of HH and VV images into 2 classes.

REFERENCES

Anderson T.S., 1975, *Multivariate Statistics* (New York: Wiley)

Bass F.G. and Fuks, I.M., 1979, *Wave Scattering from Statistically Rough Surfaces* (London: Pergamon)

Bass F.G., Fuks I.M., Kalmykov A.I., Ostrovsky I.E. and A.D. Rosenberg, 1968, Very high frequency radiowave scattering by a disturbed sea surface. IEEE Trans. Ant. Prop., **16**, 554-568.

Born M. and Wolf E., 1965, *Principles of Optics*, 3rd Ed. (London: Pergamon)

Chen J.Y. and Reed I.S., 1987, A detection algorithm for optical targets in clutter. IEEE Trans. Aerosp. Electr. Systs., **23**, 46-59.

Cutrona L.J. and Hall G.O., 1962, A comparison of techniques for achieving fine azimuth resolution. IRE Trans. Mil. Electr., **6**, 119-122.

Cutrona L.J., Leith E.N., Palermo C.J. and Hall G.O., 1961, A high resolution radar combat surveillance system. IRE Trans. Mil. Electr., **5**, 127-131.

Cutrona L.J., Leith E.N., Palermo C.J., and Porcello L.J., 1960, Optical data processing and filtering systems. IRE Trans. Inform. Theo., **6**, 386-400.

Cutrona L.J., Leith E.N., Porcello L.J., and Vivian W.E., 1966, On the application of coherent optical processing techniques to synthetic aperture radar. Proc. IEEE, **57**, 1026-1032.

Derin H. and Kelly P., 1989, Discrete-index Markov type random processes. Proc. IEEE, **77**, 1485-1510.

Derin H., Kelly P.A., Vezina G., and Labitt S.G., 1990, Modelling and segmentation of speckled SAR complex images. IEEE Trans. Geosci. Rem. Sens., **28**, 76-87.

Fowle E.N., 1964, A method of designing FM pulse compression signals. IRE Trans. Inform. Theory **IT-10**, 61-67.

Geman S., and Geman D., 1984, Stochastic relaxation, Gibbs distributions and the restoration of images. IEEE Trans. Pattern Anal. Mach. Intell., **6**, 721-741.

Harger R.O., 1965, An optimum design of ambiguity function, antenna and signal for side-looking radars. Trans. Mil. Electr., **9**, 264-278.

Harger R.O., 1970, *Synthetic Aperture Radar Systems*. (New York: Academic)

Harger R.O., 1975, Signal sequence detection given noisy, common background image sets. IEEE Trans. Aerosp. Electr. Syst., **11**, 629-635.

Harger R.O., 1976, Harmonic detection and imaging radar systems for nonlinear, near-ground, in-foliage scatterers. IEEE Trans. Aerosp. Electr. Syst., **12**, 230-245.

Harger R.O. and Korman C.E., 1988, Comparisons of simulated and actual synthetic aperture radar gravity wave images". Jo. Geophys. Res., **93**, 13867-13882.

Harger R.O., 1990a, SAR object detection with fractal terrain models". Proc. IGARRS'90, 313-316.

Harger R.O., 1990b, A near-optimal processor for multiple SAR images. Proc. IGARSS'90, 1329-1331.

Harger R.O., 1991, Polarimetric SAR image segmentation with hidden Markov random field models. Proc. IEEE SEASCON'91, 640-644.

Heimiller R.C., 1962, Theory and evaluation of gain patterns of synthetic arrays. IRE Trans. Mil. Electr., **6**, 122-130.

Keller J.M., Crownover R.M., and Chen R.Y., 1987, Characterization of natural scenes related to fractal dimension. IEEE Trans. Pattern Anal. Mach. Intell., **9**, 621-627.

Kelly P.A., Derin H., and Hartt K.D., 1988, Adaptive segmentation of speckled images using an hierarchical random field model. IEEE Trans. ASSP, **36**, 1628-1641.

Kerr D.E., 1951, *Propagation of Short Radio Waves*. New York: McGraw-Hill.

Klauder J.R., Price A.C., Darlington S., and Alberscheim, W.J., 1960, The theory and design of chirp radars. Bell Syst. Tech. Jo., **39**, 745-808.

Leith E.N., and Upatnieks J., 1962, Reconstructed wavefronts and communication theory. Jo.Opt.Soc.Am., **52**, 1123-1130.

Mandlebrodt B.B., and Van Ness, J.W., 1968, Fractional Brownian motion, fractional noises and applications. SIAM Rev., **10**, 427-437.

Marglit A., Reed I.S., and Gagliardi R.M., 1985, Adaptive optical target detection using correlated images. IEEE Trans. Aerosp. Electr. Syst., **21**, 394-405.

McCord H.L., 1962, The equivalence among three approaches to deriving synthetic array patterns and analyzing processing techniques. IRE Trans. Mil.Electr., **6**, 116-119.

J.R.Mentzer J.R., 1955, *Scattering and Diffraction of Radio Waves*. (London: Pergamon)

North, D.O., 1943, An analysis of the factors which determine signal-to-noise discrimination in pulsed carrier systems. Reprinted (1963) Proc. IEEE, **51**, 1016-1027.

Norton K.A., 1937, The physical reality of space and surface waves in the radiation field of antennas. Proc. IRE, **25**, 1192-1212.

Phillips O.M., 1981a, The structure of short gravity waves on the ocean surface, in R.C.Beal, P.S.deLeonibus and I.Katz, Eds., *Spaceborne Synthetic Aperture Radar for Oceanography* (Baltimore: Johns Hopkins Univ. Press)

Phillips O.M., 1981b, The dispersion of short wavelets in the presence of a dominant long wave. Jo.Fluid Mech., **107**, 465-485.

Sherwin C.W., Ruina J.P., and Ratcliff, R.D., 1962, Some early developments in synthetic aperture radar systems. IRE Trans. Mil.Electr., **6**, 111-116.

Slepian D. and Pollock H.O., 1961, Prolate spheroidal wave functions, Fourier analysis and uncertainty-I. Bell Systems Tech. Jo. **40**, 43-63.

Walker J.L., 1980, Range-Doppler imaging of rotating objects. IEEE Trans. Aerosp. Electr. Syst.,**16**,23-52.

P.M. Woodward, P.M., 1953, *Probability and Information Theory with Applications to Radar* (New York: McGraw-Hill)*

Wright J.R., 1968, A new model for sea clutter. IEEE Trans. Ant. Prop., **16**, 217-223.

Valenzuela G.R., 1968, Scattering of electromagnetic waves from a tilted, slightly rough surface. Radio Sci., **3**, 1057-1066.