Imaging Algorithms for a Strip-Map Synthetic Aperture Sonar: Minimizing the Effects of Aperture Errors and Aperture Undersampling

Peter T. Gough and David W. Hawkins

***Abstract*—Imaging the sea floor using high-precision synthetic aperture sonar (SAS) techniques is now at the stage where the efficiency and the robustness of the various imaging algorithms are of concern. There have been several block processing algorithms developed for relatively narrow-band-, narrow swath-, and narrow beamwidth synthetic aperture systems mainly for use by the synthetic aperture radar (SAR) community. These algorithms are summarized and their relevance to the slower speed of propagation, broad-band, broad swath-, and broad beam-width sonar equivalents are noted.**

Additional algorithms intended to ameliorate distortions injected by motion errors of the tow fish and medium turbulence are also assessed. One of the significant differences between the sonar and radar systems is that most synthetic aperture sonars travel faster than that required to meet the spatial sampling criterion and so the aperture is under- or insufficiently sampled. The digital spotlighting approach can be shown to reduce the grating-lobe images generated by this undersampling to a significant degree.

The operational effectiveness of these various algorithms are shown on real data as collected by an ocean-going, boat-towed, rather than a rail or otherwise guided, sonar. What is important is that these algorithms in various combinations can ultimately produce near diffraction-limited imaging on real data. Typical results are shown when using the Kiwi-SAS to image point retro-reflectors (either as isolated targets or deployed in pairs) on a sea floor of bland silt. To date, no unclassified SAR or SAS uses the range or along-track spatial bandwidths employed by the Kiwi-SAS. The final SAS image resolution of 16 cm $\times$ 5 cm is a considerablely finer resolution than achieved by any SAR of equivalent carrier wavelength. The fine resolution is due to the correspondingly high spatial bandwidths covered by the system; that of range due to the chirp bandwidth coupled with the slow speed of sound in water and that of along-track due to the small real apertures employed. Access to this wide spatial bandwidth makes the applicability of normal SAR algorithms uncertain and we explore some of the trade-offs.

***Index Terms*— Sonar, sonar array, sonar imaging/mapping, sonar signal processing, synthetic aperture imaging, synthetic aperture sonar.**

I. Introduction

THE sequence of echoes detected by an active synthetic aperture radar (SAR) or synthetic aperture sonar (SAS) are coherently added in an appropriate way to produce an image with greatly enhanced resolution in the azimuth, or along-track, direction when compared with the image obtained from a standard side-looking radar or sonar. The received echoes are essentially a mapping of the two-dimensional (2-D) terrain reflectivity modulated by the transmitted signal and delayed in time by the appropriate amounts. The received echoes are also modulated by the movement of the platform between transmitted pulses. The movement of the platform along the track gives the echoes a hyperbolic phase history in azimuth which is similar to a linear frequency chirp in spatial coordinates. As a consequence, this spatial modulation is often called a Doppler modulation. It is this spatial Doppler modulation of the received waveform that enables the various SAR and SAS imaging algorithms to produce images with such enhanced azimuth resolution.

The first real-time synthetic aperture processor was for airborne radar and was based on Wiley's 1952 unfocused Doppler beam sharpening (DBS) concept [31]. This idea exploits the fact that a repeated transmission of an FM pulse gives a signal with discrete spectral lines spaced by the pulse repetition frequency (PRF). The discrete nature of the spectrum means that temporal Doppler shifts that lie within the PRF/2 of a spectral line can be measured quite accurately [3]. The Doppler shifts exploited by DBS are due to Doppler effects *between* pulses and represent relative platform-target speeds, Doppler shifts *within* pulses due to movement during pulse transmission are ignored. The inversion of the DBS data to produce high resolution images can be termed a *velocity-induced* SAR [27]. The original DBS concept was based on a plane-wave assumption similar to that now used in the tomographic formulation of spotlight SAR, and because of this assumption, DBS suffers from an impairment known as *motion through resolution cells* [1], [5]. The approximation-free FM–CW SAR inversion scheme recently developed [27] does not suffer from this problem and is a complete generalization of the original DBS scheme. (This work also shows the equivalence of the FM–CW formulation and of the range-Doppler and wavenumber SAR inversion schemes used in this paper.)

The first fully focused *spatially induced* SAR system was proposed in 1953 [25]. Spatially induced synthetic aperture processing is based on the idea that a sequence of pulses recorded from a single moving real aperture can, with suitable computation, be treated as the output of a much longer array. When this synthetic array is focused, it can achieve an along-

Manuscript received April 2, 1996. The work of D. W. Hawkins was supported by the R. H. T. Bates Scholarship from the Royal Society of New Zealand.

The authors are with the Electrical and Electronic Department, University of Canterbury, Christchurch, New Zealand.

Publisher Item Identifier S 0364-9059(97)01264-8.

track resolution that is independent of range and wavelength and dependent only on the physical size of the real aperture. However, at that time, the computational load exceeded existing capabilities so the pulse echo data was recorded coherently and the inversion to high-resolution imagery used a ground-based processor. Other spatially induced focused SAR systems were developed independently [12] and it was these publications that first coined the term *synthetic aperture radar*; now universally known as SAR. The most well-known inversion algorithm for the early spatially induced SAR was "computed" optically using a sophisticated system of spherical, cylindrical, and conical sectioned lenses [13], nicely reviewed in [28]. As the phase histories of the demodulated echoes were recorded on a continuously scrolling film during the fly-past, the film, after it was developed, could be coherently illuminated and used as the input to the optical system. Aside from the cost of the optical system, disadvantages include a low signal-to-noise ratio (SNR) in the recorded data along with all the noise problems inherent in any coherent optical system. The first digital-based processor marked the start of the modern SAR era [11], [19], [20]. Most of these schemes assume the data is recorded in a 2-D matrix of pulse echoes arranged into the following format.

Assuming only a single channel detector, the signal received is a continuous one-dimensional (1-D) stream of coherently summed pulse echoes, each echo being a scaled replica of the transmitted waveform. Unless the transmitted signal is already a short pulse waveform, synthetic aperture algorithms usually time- or range-compress the incoming echoes by correlating them against the transmitted waveform. They then reorder the 1-D time-compressed echoes into 2-D coordinates of along-track distance u and delay time t where discrete increments in along-track correspond to the distance moved along the track between pulse repetition periods τ_{rep}. The time or range compression is sometimes done after the 1-D to 2-D reordering process and provided care is taken to ensure that reflecting objects at the extreme edge of the maximum unambiguous range are not truncated, the actual sequence of compression and reordering is immaterial. The 1-D to 2-D reordering operation produces what is frequently referred to as the *data space* in the (t, u) domain (or sometimes referred to as the *compressed range/azimuth domain*). With the temporal coordinate delay time t scaled to the spatial coordinate of range (or cross-track) x, the data space is converted to a raw image in the coordinates of range x and along-track y which is the normal output of side-looking/side-scan sonars. The along-track coordinate is always discrete whereas the range coordinate may be continuous or discrete. The raw image is at best a poor representation of the sea floor reflectivity and is characterized by the almost parabolic appearance of point targets, the parabola getting more and more extended in azimuth as the reflecting point targets are located further and further away from the aperture.

The significant part of most synthetic aperture algorithms starts from the function as recorded in data space. The most accurate and inefficient of these algorithms being a point by point, brute force correlation of the function recorded in data space with the appropriate point spread function for the point under consideration. This has been called the *exact* reconstruction algorithm. As the correlation is a point-spread variant computation, very little can be gained in the way of array or bulk (usually Fourier) processing without some approximations. Depending on the type of approximation, there are a host of reconstruction algorithms with a variety of names; the two most well-known are the *range-Doppler* and the *wavenumber* reconstruction algorithms [2]. A fundamental problem with many of the algorithms, and their derivatives, is the necessity to change the variables in a nonlinear way which, when using sampled data, means interpolation from one sampling grid to another. Unfortunately, the interpolation itself is costly in computing resources and injects phase and amplitude errors into the process. Some, such as the *chirp-scaling* algorithm [22], have been written to do the interpolation in an alternative form (usually by multiplication in an equivalent frequency domain) to avoid brute force grid-to-grid interpolation.

Regardless of the method of grid-to-grid interpolation, there are still two major problems limiting the widespread use of SAS. These are the image blurring caused by imperfect motion compensation resulting in aperture errors and the generation of artifacts caused by undersampling along the aperture.

A. Aperture Errors

When the sonar moves down the track, the locus is seldom perfectly straight. Many SAS and SAR have on-board inertial navigation systems to compensate for gross departures from the straight locus, but still, minor residual errors corrupt the detected echoes. In the SAR community, there are three main autofocusing algorithms that are routinely used to remove these residual errors. *Contrast optimization* algorithms [4], [21] determine the final focused image on the basis that the image with highest contrast is the optimal image. The algorithms perform a trial-and-error search on the along-track chirp rate and Doppler centroid, parameters which are related to the along-track instantaneous velocity and squint angle of the synthetic aperture system. *Map-drift* or *multilook registration* algorithms [11], [21] are based on the correlation of multilook images. Usually, the raw data is filtered in along-track into a number of "looks" with each look corresponding to a certain bandwidth of the along-track Doppler wavenumber. Each of these Doppler sections corresponds to a different look at the object field. When each look is focused, we get a set of images with each of these images having a reduced along-track resolution. The low-resolution image generated from each look will overlay exactly if the optimal spatial chirp rate and Doppler centroid have been chosen correctly in the along-track or azimuth focusing section of the processing algorithm. These low-resolution images may be added coherently (provided the sampling of each image is increased before the coherent addition) to produce a maximum-resolution image. The more usual alternative is to add them noncoherently to produce a lower resolution image with greatly reduced speckle. However, if the spatial chirp rate is incorrect or other parameters used by the algorithm were in error, the resulting low-resolution images are displaced relative to each other. The amount of image

displacement can be measured by cross-correlating each of the low-resolution images, and these displacements used to obtain an estimate of the parameter error. *Phase gradient autofocus* (PGA) [14], [18], [29] is typically used in the processing of spotlight-SAR images but has recently been adapted for strip-map systems [30]. PGA models the along-track error as a phase-only function of the Doppler wavenumber in the range-Doppler domain. By using redundancy over range in the image, an estimate of this phase error can be obtained and removed from the image. All of these algorithms are iterative in nature with the fastest algorithm, straight PGA with no enhancements, typically converging in 2–3 iterations.

B. Aperture Undersampling

A major advantage of airborne radar systems over equivalent sonar systems is that the high speed of light relative to platform velocity allows the along-track dimension of SAR recordings of data space to be well sampled. (This is not always the case in space-borne SAR.) In sonar, the low propagation speed of sound causes this sampling requirement to be in conflict with the need for the sonar to travel as fast as possible to maximize the mapping rate. This leads to the second of the two problems limiting the widespread use of SAS: aperture undersampling.

When the sonar moves down the track and, for this particular section, we assume without motion error, the aperture is always sampled by the sonar. A strip-map SAS (i.e., a sonar without transducer beam-steering) has a simple formula that specifies the appropriate distance between samples that will result in an image virtually free of artifacts, and this is related to the physical size of the radiating and receiving transducers, *viz.*,

$$\delta y_{3\,\mathrm{dB}} = \frac{D}{2} \tag{1}$$

$$\Delta_u = \frac{D}{4} \tag{2}$$

where Δ_u is the distance between samples along the aperture, $\delta y_{3\,\mathrm{dB}}$ is the along-track resolution in the ground plane reflectivity function, and where D is the size of the largest transducer. This means that the sonar can travel no more than $D/4$ between pulses and, for anything but a very short range sonar, this implies a very low platform speed and consequential mapping rate. If the sonar travels at

$$v > \frac{\delta y_{3\,\mathrm{dB}}}{2\tau_{\mathrm{rep}}} \tag{3}$$

where v is the velocity, then the aperture is no longer appropriately sampled (i.e., since $\Delta_u > D/4$) and the reconstructed beam pattern of the synthetic aperture has grating lobes. This sampling rate is correctly stated as $D/4$ whereas many SAR references incorrectly state $D/2$. Appendix A explains the origin of this confusion and a complete mathematical analysis of the system model also results in a $D/4$ aperture sampling requirement [27, p. 293]. Airborne SAR's do not have a problem satisfying this aperture sampling criterion; however, it is often relaxed for space-borne SAR in a tradeoff of dynamic range for an increased mapping rate [11, p. 298]. The tradeoff is parameterized using what is known as the along-track-ambiguity-to-signal ratio (AASR). Systems sampled at $D/2$ that only process 70% of the Doppler bandwidth available typically have an AASR of -20 dB. This figure is considered adequate for most space-borne SAR imaging applications. As a consequence, a single strong reflecting target often reconstructs to multiple images in along-track and this has been noticed in SAR images of strong targets surrounded by low-contrast backgrounds [11, p. 299] and also in similar contrasted SAS images [23]. We call the extra ambiguous images grating-lobe images. There are some circumstances where the grating-lobe images can be minimized by suitable computation; however, it should be realized that once a strip-map SAS recording has been undersampled along the aperture, there are very few options available; the missing data cannot be retrieved without errors, approximations, or *a priori* information.

II. Basic SAS System Model

Consider a continuous 2-D distribution of omnidirectional reflecting targets on the sea floor described by a complex object reflectivity function $ff(x, y)$ (where the utility of the double notation will eventually become clear). The sea floor is insonified by a side-looking sonar traveling along a straight locus u having a velocity v, flying h meters directly above and parallel to the y axis located in the plane of the sea floor. For simplicity, let $h \ll x$, which implies that the sonar platform is flying very nearly in the plane of the object. The sonar transmits a waveform $p_l(t)$ of duration τ_c which repeats every τ_{rep}. Obviously, $\tau_c \leq \tau_{\mathrm{rep}}$. It is usual to arrange the recorded echoes into a 2-D representation of distance along the aperture u and echo delay time from the start of the current pulse t to give

$$ee(t, u) = \int_y \int_x ff(x,y) \cdot p_l\left[t - \frac{2}{c}\sqrt{x^2 + (y-u)^2}\right] dx\, dy \qquad t \in [0, \tau_{\mathrm{rep}} + \tau_c]. \tag{4}$$

For simplicity, we have absorbed all amplitude spreading losses into our definition of $ff(x, y)$ and we deliberately ignore the convolutions necessary to account for the finite size of the real transmitting and receiving apertures as this has been dealt with at some length elsewhere [15]. Since almost all sonars use a waveform with some form of coding or frequency chirp, $ee(t, u)$ is still not readily interpretable and it is normal to pulse- or range-compress this signal by correlating it with the unit waveform $p_l(t)$ to produce a 2-D range-compressed presentation. This is the *data space* or the *compressed range/azimuth domain*, given by

$$\begin{aligned} ss(t, u) &\equiv ee(t, u) \star p_l(t) \qquad t \in [0, \tau_{\mathrm{rep}}] \\ &= \int_y \int_x ff(x,y) \\ &\qquad \cdot p\left[t - \frac{2}{c}\sqrt{x^2 + (y-u)^2}\right] dx\, dy \end{aligned} \tag{5}$$

where $\star$ represents 1-D correlation and $p(t) \equiv p_l(t) \star p_l(t)$ which represents the time- or range-compressed pulse. For most modern sonars, $p(t)$ is often no more than a few cycles of the nominal center frequency ω_0 in extent.

The appearance of $ss(t, u)$ is of a collection of parabolas, one parabola for each reflecting point in $ff(x, y)$, that become more extended as the time delay (i.e., the range) increases. The extent to which a single point in $ff(x, y)$ maps into an almost parabolic locus which covers several range cells is called *range curvature* or *range migration*. Unfortunately, each range has a different amount of range curvature which makes the reconstruction of $ff(x, y)$ from $ss(t, u)$ far from trivial. With one notable exception (the chirp-scaling algorithm), synthetic aperture imaging algorithms start with (5) and proceed to estimate $ff(x, y)$ to the diffraction limit of the system.

Before we discuss the various imaging algorithms, it is helpful to see the relationship of the 2-D Fourier transform of $ff(x, y)$ to $ss(t, u)$. First a 1-D transform of $ss(t, u)$ from time t to temporal frequency ω gives

$$\begin{aligned} Ss(\omega, u) &= \mathcal{F}_t\{ss(t, u)\} \\ &= P(\omega) \int_y \int_x ff(x,y) \\ &\quad \cdot \exp\left[-j2k\sqrt{x^2 + (y-u)^2}\right] dx\, dy \end{aligned} \tag{6}$$

where $k \equiv \omega/c$ and $P(\omega)$ is the Fourier transform of the compressed pulse $p[t]$. Using the principle of stationary phase, $Ss(\omega, u)$ given by (6) is followed by a further 1-D Fourier transform from space u to spatial frequency or Doppler wavenumber k_u to give the transfer function of the data recording [27] (see Appendix B) as

$$\begin{aligned} SS(\omega, k_u) &= \mathcal{F}_u\{Ss(\omega, u)\} \\ &= P(\omega) \int_y \int_x ff(x, y) \exp\left[-j\sqrt{4k^2 - k_u^2} \cdot x\right] \\ &\quad \cdot \exp(-jk_u y)\, dx\, dy. \end{aligned} \tag{7}$$

Now with a change of coordinates inside the integrals given by

$$k_x \equiv \sqrt{4k^2 - k_u^2}, \quad k_y \equiv k_u \tag{8}$$

$$\begin{aligned} SS(\omega, k_u) &= P(\omega) \int_y \int_x ff(x, y) \exp(-jk_x x) \\ &\quad \cdot \exp(-jk_y y)\, dx\, dy \\ &= P(\omega) FF(k_x, k_y). \end{aligned} \tag{9}$$

Note that now the coordinates of $SS(\omega, k_u)$ and $P(\omega)$ and those of $FF(k_x, k_y)$ are different. The implications of the coordinate warping between the (ω, k_u) and (k_x, k_y) domains will be referred to later in this paper.

III. Imaging Algorithms for Strip-Map SAR's

The object of any SA image reconstruction algorithm is to estimate the magnitude of the object's complex reflectivity function $ff(x, y)$ from $ss(t, u)$ to the diffraction limit imposed by the real aperture(s) used and the bandwidth of the transmitted pulse.

Many of these algorithms have been developed by the SAR community and so are optimized for the various SAR systems and geometries. These SAR systems are usually characterized by narrow bandwidth transmissions (cf. most SAR having megahertz bandwidths at gighertz carriers with the SAS having kilohertz of bandwidth at kHz center frequencies). A second significant difference are the speeds of propagation (cf., 3×10^8 m/s with 1500 m/s) and the platform velocities (cf. spaceborn SAR up to 7500 m/s with SAS up to 10 m/s but usually around 1 m/s). Also, the ratio of swath width to average slant range is very different (cf., spaceborn SAR 50–100 km at 900 km with SAS 500 m at 500 m). Finally, the beamwidths and squint angles are often very different (cf., SAR having a 1° beamwidth at up to 8° squint with SAS 10° at broadside). Even with such narrow beamwidths, the dwell time of a SAR over a target is typically several thousand pulse repetitions, whereas a SAS is lucky to have a few hundred. All of these practical details need to be considered when evaluating the usefulness of the following algorithms to SAS. A comparison of the operational parameters of a satellite SAR (Seasat) with two ocean-going, cable-towed SAS systems is given in Table I.

TABLE I
Representative Parameters of SEASAT, Acoustical Imaging Development (ACID) Sonar and KIWI-SAS

	Seasat	ACID	Kiwi-SAS
c (m/s)	$3.0{\cdot}10^8$	$1.5{\cdot}10^3$	$1.5{\cdot}10^3$
Platform V_s(m/s)	7400	1.5–2.5	0.25–0.5
Carrier f_0	1.3GHz	8kHz	30kHz
Bandwidth B_c	20MHz	6kHz	20kHz
Pulse Rep. Period	650μs	0.4–0.8s	0.15–0.3s
Antenna D (m)	10.7	2.0	0.30
Beamwidth	1.5°	5.4°	10°
Depression Angle	75°	30°–50°	5°
Stand-off r_0	700km	300m	100m
Swathwidth	100km	400m	200m
Ground Resolution	25m × 25m	0.13m × 1.0m	0.05m × 0.15m
No. Looks	4	1	1

A. The Exact and Exact Transfer Function Algorithms [2], [8]

The simplest algorithm is known as the *exact* image reconstruction algorithm and it is the only purely time-domain algorithm. Since each point reflector in $ff(x, y)$ produces a unique hyperbolic locus in the $ss(t, u)$ domain, it is possible to do brute force correlation of $ss(t, u)$ against a continuous point reflector function $ss(t, u)|_{x_0, y_0}$ calculated for the particular point under consideration x_0, y_0. Since the data is sampled and the point reflector function describing the hyperbolic locus is continuous, interpolating functions are needed, which means this algorithm is very slow and seldom used for anything but calibration. The simplest of the Fourier-based algorithms is the *exact transfer function* algorithm. The recorded data is first Fourier-transformed in 2-D to the $SS(\omega, k_u)$ domain and then multiplied by the complex conjugate of the hypothetical $SS(\omega, k_u)|_{x_0, y_0}$ for a single point at x_0, y_0. The inverse Fourier transform gives the reflectivity of the x_0, y_0 point in the object domain.

In effect, the exact and exact transfer function algorithms are identical to a beam-forming process for each pixel in the final image. This individual pixel beam-forming technique has been the basis for most of the previous SAS prototypes reported in the unclassified literature.

A common modification of the exact transfer function algorithm is to use a common $SS(\omega, k_u)|_{x_0, y_0}$ for pixels within the same depth of focus. Unfortunately, this injects phase errors into the reconstructed image, and while phase preservation has not been an issue with previous SAS imaging systems, interferometric and bathymetric SAS, as well as those that use autofocus to correct aperture errors, all require deterministic phase errors to be eliminated from the image. So despite the efficiency gains of the nonphase preserving version of the exact transfer function algorithm, a new type of algorithm is required when a full-phase preserving strip-map image reconstruction is needed. For that, the complete global point variant nature of the recorded data must be accommodated.

B. The Range-Doppler Algorithm [2], [8], [19]

The first of these to be developed was the *range-Doppler* algorithm. This algorithm starts with a 1-D Fourier transform of $ss(t, u)$ from a coordinate of along-track distance u to along-track/Doppler wavenumber k_u producing what is known as the range-Doppler domain $sS(t, k_u)$. In this domain, the variation in delay time (range) migration with k_u is removed by a simple coordinate remapping of the sampled data (sometimes called reformatting) before azimuth compression by a phase-only inverse of the spatial Doppler chirp. The major advantage is that since each range pixel has a similar phase history (but displaced in position along the aperture), a single compression in the range-Doppler domain applies to all pixels at the same range. Thus the processing has progressed from a point-by-point computation in the exact and exact transfer function algorithms to a [range] line-by-line computation.

In beam-forming parlance, the convolutional nature of the SA model in the along-track direction implies that beam-forming parameters are the same for any pixel at the same range and so the focusing can be performed simultaneously for a range line of pixels. The beam-forming equivalent of coordinate reformatting is implemented by phase shifting and summing networks and if the phase information is to be preserved, each image pixel must have its own unique set of beam–forming parameters. In an effort to improve efficiency, the same beam-forming parameters are often used to focus all the pixels within a depth of focus. Unfortunately, this commonly used technique is not phase preserving and so cannot be used when subsequent processing such as autofocus is required.

In systems where the range cell migration is less than 1/3 of a range resolution cell, we can use the *Fresnel approximation* algorithm. The range-Doppler algorithm normally requires precise interpolation as part of the coordinate remapping. If the range migration is small, as is typical for an airborne SAR system, this step can be dispensed with and so the Fresnel approximation algorithm becomes no more than two 1-D Fourier transforms with an intermediate phase multiplication for every range line of pixels.

Additional complications arise when the squint angle or the spatial bandwidths in both range and Doppler of the physical apertures are sufficiently large that the range chirp is no longer constant, with k_u the Doppler wavenumber. In this situation, there is an extra k_u-dependent LFM component in the range chirp, and what is known as secondary range compression (SRC) is needed [2].

To our knowledge, no true wide-bandwidth range-Doppler algorithm has been used on SAS data and this Fourier-based algorithm represents a far more efficient method for producing SAS images than the "exact" algorithms.

C. The Wavenumber Algorithm [2], [6], [7], [22], [26]

SAR processing algorithms based on the *wavenumber* or range migration or seismic wave equation formulation have formed the basis of the next generation of highly efficient SAR processors. Basically the raw or range-compressed data space in (t, u) is 2-D Fourier transformed into the temporal frequency/wavenumber domain in the coordinates of (ω, k_u). Given that the transfer function of the data collection process is given by (9), the matched filter wavenumber image-reconstruction algorithm can be neatly summarized by

$$\widehat{FF}(k_x, k_y) = \mathbf{S}\{SS(\omega, k_u)\} \tag{10}$$

where the coordinate reformatting (the Stolt mapping) $\mathbf{S}\{\cdot\}$ is given by (8). Thus, we have progressed from a point-by-point algorithm (the exact) to a line-by-line (the range-Doppler) to a block-by-block (the wavenumber). As a matter of practical detail, the Stolt mapping from the (ω, k_u) to the (k_x, k_y) domain involves precise interpolation with sampled data and, as the center of the illuminated footprint is well offset from the origin of coordinates, $SS(\omega, k_u)$ varies rapidly. A more practical algorithm to implement with sampled data is

$$\widehat{FF}(k_x, k_y) = \exp(-jk_x r_0) \cdot \mathbf{S}\left\{\exp\left(j\sqrt{4k^2 - k_u^2}\cdot r_0\right) SS(\omega, k_u)\right\} \tag{11}$$

where r_0 is the distance from the center of the aperture to the center of the illuminated object domain and would be considered the average cross-track of the illuminated object. It should be noted that the various phase functions in (11) are appropriate for coordinate systems that have time and range ordinates defined from the aperture. Different phase functions would be necessary for systems with coordinates defined from the center of the object, for instance [2], [26].

The exponential inside the Stolt map in (11) represents a spatial shift that performs an analogous process to that of carrier removal in the demodulation of a spectrum. The phase multiplication moves the entire swath down to the origin of spatial coordinates in the same way demodulation moves a spectrum to baseband. Interpolation can now be performed over slowly varying quantities and so the interpolation is more likely to be successful. Using the beam-forming analogy, the exponential inside the operator achieves the focusing for pixels at r_0 while the Stolt mapping achieves focusing for ranges other than r_0.

Even with this modification, the algorithm is still relatively inefficient due to the need for interpolation, so in that regard the wavenumber algorithm is no better than the range-Doppler algorithm.

An interesting variant of the wavenumber algorithm is the FM–CW algorithm. Both end up by calculating the 2-D Fourier transform $SS(\omega, k_u)$ of the measured pulse echoes. The wavenumber algorithm segments the incoming pulse echoes into the 2-D matrix we have described as $ss(t, u)$ and then does a 2-D Fourier transform on it to get $SS(\omega, k_u)$. The FM–CW algorithm takes one long 1-D Fourier transform of all the incoming pulse echoes and then segments the vector result of the 1-D transformation into a 2-D matrix. Either way, we end up with the same $SS(\omega, k_u)$ and in all other respects the remainder of the reconstruction algorithms are the same [27].

Efficient derivatives of the wavenumber algorithm can be used if the swath width is small compared to the average cross-track r_0 and if the bandwidth is small compared with the center frequency [2]. For these derivative algorithms, the interpolation process in the (ω, k_u) domain is replaced by phase multiplication in the (t, k_u) domain, thus avoiding interpolation. Although the restrictions of swath- and bandwidth are frequently true in SAR systems, they are seldom true in SAS systems, which renders these derivative algorithms of lesser interest to the sonar community.

D. The Chirp-Scaling Algorithm [7], [10], [22], [24]

Rather than start with the range-compressed data, $ss(t, u)$, as most SAR algorithms typically do, the *chirp-scaling* algorithm starts with the raw echo data, $ee(t, u)$ as defined in (4). The chirp-scaling algorithm exploits the linear FM (LFM) signal structure of the transmitted pulse. If coded or pseudorandom sequences have been employed for the transmitted pulse, the raw data will need to first be range-compressed with the appropriate deconvolving function, then respread by a convolution with a LFM pulse.

Now recall that the range-Doppler algorithm removes a range migration time delay in the (t, k_u) domain to bring all the Doppler loci spread in k_u and t to 1-D lines spread in k_u only. As this time delay is not constant with t, the k_u-dependent time-delay subtraction is done by time-domain interpolation, which can be expensive and error inducing if not done accurately. The chirp-scaling algorithm does the equivalent of this operation in the uncompressed range or range *signal*/Doppler wavenumber (t, k_u) domain. How it does so is quite clever.

The chirp-scaling algorithm starts with a 1-D Fourier transform in along-track of the raw echo data, $ee(t, u)$, the resultant $eE(t, k_u)$ is then multiplied by a mild phase-only frequency chirp which *scales* the range chirp for each k_u and t just enough so that the Doppler loci of all reflecting targets end up with a common range migration in phase (but not amplitude). The scaled data is then range-Fourier-transformed into the (ω, k_u) domain where bulk range migration correction, pulse compression, and SRC is performed. These compressions and corrections leave a residual phase function which is then removed in the (t, k_u) domain along with the overall azimuth compression, followed by a transformation back into the (t, u) domain. Finally, coordinate rescaling gives our estimate of the diffraction limited image $\widehat{ff}(x, y)$.

E. Secondary Range Compression [11], [19]

Although the wavenumber algorithm is "exact" in that it contains no approximations and so is good for any wavelength, bandwidth, beamwidth, or swath width; the range-Doppler and the chirp-scaling algorithms both have inherent approximations and when the span of spatial bandwidths gets large enough or the system is highly squinted, the effect of these approximations needs to be corrected. The most well known of these correction factors is the secondary range compression (SRC).

Some extensive simulations using our system parameters have shown us that there is little visual difference in the images produced by the range-Doppler and chirp-scaling algorithms with or without SRC. The peak-signal-to-side-lobe ratio is almost unchanged at -12 dB, and if the spatial wavenumber domain is calculated using a Hamming weighting function $WW(k_x, k_y)$, the side lobes go down to -40 dB. So the question needs to be asked, "why use SRC at all as it would appear to add an unnecessary computational load?"

The answer lies in the *phase* of the diffraction-limited image. Normally this is of little interest as only the modulus or intensity of the image is displayed, but if *any* form of autofocus such as PGA is needed to correct for aperture errors or medium turbulence, this phase is critical whether we can "see" it in the uncorrected diffraction-limited image or not. It is important in autofocus algorithms to have the residual phase errors due to the image-forming algorithm as small as possible, and as SRC is a deterministic phase correction, little is lost by computing it for every image and a lot is gained if some form of autofocus is shown to be necessary.

F. Algorithm Summary

The wavenumber algorithm represents the exact inversion of the SA system model. If high-order interpolators are used, then the final image is the "best" possible of all possible images. The range-Doppler and chirp-scaling algorithm represent an approximate inversion of the system model and with the inclusion of SRC, the approximation has only a minimal impact on the final image quality. The range-Doppler and chirp-scaling algorithms produce images that are indistinguishable. In terms of efficiency, the range-Doppler and wavenumber algorithms perform comparably. This is because both algorithms require an interpolator to operate on the same volume of data; however, the chirp-scaling algorithm out performs both of these algorithms by a small margin. When a Hamming weighting function $WW(k_x, k_y)$ is applied to the windowed wavenumber data $\widehat{FF}(k_x, k_y)$, all the algorithms produce identical images (even without SRC the loss in peak height and the rise in side lobe levels is minor) for the Kiwi-SAS parameters. Recent analysis and simulations have shown us that a modification to the chirp-scaling algorithm improves its efficiency dramatically, and this will be the subject of a forthcoming publication [16].

IV. Image Recovery with Residual Motion Errors

As the sonar "flies" down the track u, motion errors and turbulence insert timing errors into the recorded echoes. With suitable tow fish design and perhaps some inertial navigation equipment, these errors can be reduced to a fraction of a period at the mean frequency. However, this still may not be good enough for image reconstruction of $\widetilde{ff}(x, y)$ from $\widetilde{ss}(t, u)$ (which is no more than $ss(t, u)$ with the residual positional/timing errors included). So, given that there are residual timing errors and these are to be estimated and removed, there appear to be two approaches. The first is to make an initial estimate of the image $\widetilde{ff}(x, y)$ which includes all the errors and iterate toward an "error-free" image based on some convergence criterion. Let us call this the image-compensation approach. The second is to estimate the actual timing error as a function of distance along the aperture u and remove these errors before the synthetic aperture processing begins. Let us call this the aperture-compensation approach.

A. Image-Compensation Algorithms

The image-compensation approach begins by calculating $\widetilde{ff}(x, y)$ from $\widetilde{ss}(t, u)$ [or $\widetilde{ce}(t, u)$] using whatever synthetic aperture algorithm is deemed appropriate and then iterates around $\widetilde{ff}(x, y)$ using some criterion to force it to converge to $\widehat{ff}(x, y)$. The most successful of these algorithms is called *phase gradient auto focus* (PGA). The algorithm's name is due to the fact that it averages the phase difference (gradient) of the along-track Fourier transform of selected targets to determine the aperture error [14], [18], [29].

1) PGA for Spotlight Systems: With the notation we've chosen, the PGA model of the corrupted scene is given by,

$$\widetilde{ff}(x, y) = \widehat{ff}(x, y) \odot_y \mathcal{F}_{k_y}^{-1}\{\exp[j\phi(k_y)]\} \tag{12}$$

where the effect of the timing error is modeled as a convolution in the along-track y direction [14], [18], [29]. The use of a convolution implies that the scene is distorted by the same point spread function.

The following development is appropriate for spotlight-SAR, which is what the PGA algorithm was initially designed to focus; after explaining the spotlight operation of this algorithm we will indicate the modifications required for strip-map operation. Fourier transforming (12) in along-track y gives

$$\widetilde{fF}(x, k_y) = \widehat{fF}(x, k_y) \exp[j\phi(k_y)] \tag{13}$$

if the scene consisted of a noisefree distribution of point reflectors where there was only one target along any x constant line, then the phase of $\widetilde{fF}(x, k_y)$ along any constant x will consist of a linear component due to a target's offset from $y = 0$, plus the phase error. The PGA algorithm determines this phase error by proceeding as follows: in the target scene, $\widetilde{ff}(x, y)$, the strongest targets are shifted to the aperture center, $y = 0$ m, the weaker targets are removed, and the resulting matrix is called $gg(x, y)$. A Fourier transform with respect to y gives $gG(x, k_y)$, the phase of which (in the noise-free case) consists of only the phase error as the shifting operation has removed the linear components due to target location. Using $gG(x, k_y)$, the gradient of the phase error $\dot{\phi}(k_y)$ is determined. This gradient is then integrated and the complex conjugate of the phase error, $\exp[-j\hat{\phi}(k_y)]$, is applied to $\widetilde{fF}(x, k_y)$ which is then inverse-Fourier-transformed with respect to y to give a new estimate of the image. The algorithm iterates until some form of convergence criteria is reached, e.g., this criteria can be based on window size or the size of the phase error detected.

2) PGA for Strip-Map Systems: The convolution assumption in (12) is unwarranted for strip-map systems. Basically, (12) assumes that all parts of the image have suffered the same distortion, i.e., they have been created using exactly the same aperture. In a strip-map system, targets at different along-track positions subtend different portions of the aperture traversed by the sonar and so they are subject to different distorting functions. As well, for broad swath-width systems, targets at minimum range have only a small portion of their synthetic aperture in common with targets at the maximum range and so they too show different realizations of the aperture-distorting function even though the targets may be at the same along-track position.

Recently, modifications to the PGA algorithm have shown that it is possible to do image compensation for strip-map SAR [29]. Although the modifications were promulgated for a narrow swath-width system where the subapertures for all the reflecting targets are of the same extent even if not covering the same part of the full along-track aperture, the technique can be applied to wide swath-width SAS systems where the subapertures increase in length for targets at greater ranges.

The starting point for the strip-map PGA is also $\widetilde{ff}(x, y)$ and again the strongest targets in the image are selected. However, the spotlight-PGA step of shifting the strong targets to the center of the along-track aperture is replaced by a convolution in along-track with the spatial chirp appropriate to the range of the target currently under consideration. Now the phase differential of this new function cannot be averaged with that of the other targets directly as they cover different parts of the full along-track aperture, and as well there is an unknown phase shift between each subaperture. Consequently, it is the *second* differential of the phase function $\widehat{\ddot{\phi}}(u)$ that is averaged over all the strong targets and then finally integrated twice, by whatever method is deemed appropriate, to get our distorting phase function $\widehat{\phi}(u)$.

B. Aperture Compensation Algorithms

Looking at the aperture-compensation approach in a little more detail, we first make the assumption that the synthetic aperture (along-track distance) is sampled at $D/4$ as required by the spatial sampling criterion. Let us also assume that the timing error is constant for the duration of a single pulse so that the timing error can be described as a function of u alone. Consequently, instead of measuring $ss(t, u)$, we measure $\widetilde{ss}(t, u)$ where we model the error as

$$\widetilde{ss}(t, u) \approx ss[t + \tau(u), u]. \tag{14}$$

Note that this assumption is quite different from that used by PGA [as defined in (12)] and has the added advantage that it

is applicable to strip-map and broad swath-width systems as well as spotlight and/or narrow swath-width systems.

To perform the inversion given by (10) or (11), we require

$$\begin{aligned}\widetilde{SS}(\omega, k_u) &= \mathcal{F}_u\{\exp[j\omega\tau(u)] \cdot Ss(\omega, u)\} \\ &= ZZ(\omega, k_u) \odot_{k_u} SS(\omega, k_u) \end{aligned} \tag{15}$$

where $ZZ(\omega, k_u)$ is the along-track Fourier transform of the error function. Because convolution is not distributive, the matched filtering and Stolt mapping given by (10) or (11) when applied to (15) does not result in simply a bandlimited version of $FF(k_x, k_y)$ multiplied with a blurring function. Any motion-compensation algorithm that attempts to estimate the timing errors will be limited by how closely the inversion of (15) can be modeled as a convolution. The only algorithm that successfully does this is the PGA algorithm. The other autofocusing algorithms mentioned in Section I-A can only determine the equivalent of a low-order polynomial fit to the aperture error, whereas PGA can determine errors up to the along-track folding frequency.

V. Image Recovery with Undersampled Apertures: Digital Spotlighting

Up till now, we have tacitly implied that $ss(t, u)$ is known continuously along both axes. Although this could conceivably be true along the delay time axis t [for instance, we could have used analog delay lines to produce $ss(t, u)$ from $ce(t, u)$], it is not true and never can be true along the aperture axis u. We can make the sampling nature of the u axis more specific by stating

$$ss(t, u) = 0, \qquad u \neq m\Delta_u \tag{16}$$

where m is an integer. Provided Δ_u is small enough (i.e., $\Delta_u \leq D/4$), the sampling along the aperture u is adequate to meet the appropriate sampling criterion. Thus, $sS(t, k_u)$ and $SS(\omega, k_u)$ are not aliased in Doppler wavenumber k_u and we can follow any one of the usual algorithmic paths to calculate $\widehat{FF}(k_x, k_x)$ and so reconstruct an estimate of the object reflectivity $\widehat{ff}(x, y)$.

However, when we travel too fast (i.e., $\Delta_u > D/4$), the appropriate sampling criterion is not met and spatial undersampling in the u direction occurs. Although $ss(t, u) = 0$ for $u \neq m\Delta_u$ is still a correct statement, $sS(t, k_u)$ and $SS(\omega, k_u)$ are aliased in Doppler wavenumber k_u and so $\widehat{ff}(x, y)$ is corrupted by grating-lobe images if we use any of the normal reconstruction algorithms.

Unfortunately and for the reasons noted in the introduction, there is usually a commercial or operational necessity to travel much faster than the maximum speed allowed to meet the spatial sampling criterion. In this case, the aperture is undersampled and we lose valuable information. Unless that information can be recovered using some other *a priori* information, there is little that can be done. However, the *a priori* information can often impose quite powerful constraints. Let us presuppose that the sonar is being used in the mine-hunting or a harbor-clearance environment where we know that objects of interest are limited to 4 m in extent. Under these conditions, we can perform what is known from the SAR literature as digital spotlighting and, using it, we can recover an image of limited extent as though the aperture was properly sampled.

To see how digital spotlighting works, consider a target of limited extent in the (x, y) plane which is centered about (x_0, y_0) and, by limited, we mean considerably smaller than the extent of the synthetic aperture and lying within one depth of focus. This target of limited extent gives rise to a compressed signal data space that can be appropriately described by $ss'(t, u)$ which is undersampled in u. Consequently, its 1-D transform, $sS'(t, k_u)$, is aliased in the range-Doppler domain.

As we know approximately where the target of interest is located, we can center the aperture to be synthesized about y_0, the position of closest approach. (We can think of this process as the equivalent of arranging the target of interest to lie close to the boresight of the synthesized aperture.) Consider now a single temporal frequency ω in the (ω, u) domain and recall that u is really a discrete set of samples $m\Delta_u$. Now $Ss'(\omega, u)$ has an approximately linear spatial frequency dependence in u and so, at some value of u, this linear FM waveform is spatially undersampled. For the small extended target on boresight, the linear spatial frequency dependence of u is approximately known and may be removed to produced a spatially compressed signal $Ss'_c(\omega, u)$ given by [27]

$$Ss'_c(\omega, u) = Ss'(\omega, u) \exp\left(j2k\sqrt{r_0^2 + u^2}\right) \tag{17}$$

where it is important to note that both $Ss'_c(\omega, u)$ and $Ss'(\omega, u)$ are both sampled at the same rate Δ_u. This operation is identical to the along-track deramping operation used by spotlight SAR. In spotlight SAR, the phase factor in (17) is equivalent to gating the signal about the scene center where the distance from the current aperture position to the scene center is given by $R_c = \sqrt{r_0^2 + u^2}$.

Now, since $Ss'_c(\omega, u)$ has lost most of its high spatial frequencies, it is usually well sampled at Δ_u and so its Fourier transform $SS'_c(\omega, k_u)$ is not aliased. This now has enormous benefits as we can zero pad $SS'_c(\omega, k_u)$ well out beyond its normally accepted limit of $|k_u| \leq \pi/\Delta_u$, i.e., we can define a new Doppler wavenumber/temporal frequency domain, *viz.*,

$$SS'_{cd}(\omega, k_u) \equiv \begin{cases} SS'_c(\omega, k_u), & \text{for } |k_u| \leq \dfrac{\pi}{\Delta_u} \\ 0, & \dfrac{\pi}{\Delta_u} < |k_u| < \dfrac{4\pi}{D}. \end{cases} \tag{18}$$

It then follows that we can calculate $Ss'_{cd}(\omega, u)$ from $SS'_{cd}(\omega, k_u)$. Finally, the decompressed signal that we would have recorded at a higher spatial sampling rate can be calculated by reinserting the linear FM waveform we originally removed from $Ss'(\omega, k_u)$ in (17) to obtain

$$Ss'_d(\omega, u) = Ss'_{cd}(\omega, u) \exp\left(-j2k\sqrt{r_0^2 + u^2}\right). \tag{19}$$

Note now that $Ss'_d(\omega, u)$ and $ss'_d(t, u)$ are spatially upsampled versions of $Ss'(\omega, u)$ and $ss'(t, u)$ without the effects of the original undersampling, and we can proceed to use $SS'_d(\omega, k_u)$ and $ss'_d(t, u)$ in the reconstruction algorithms

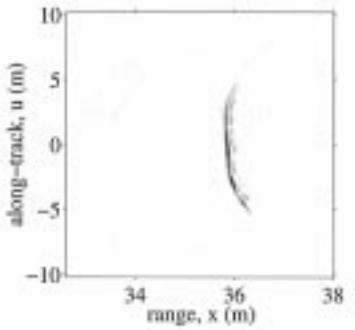


Fig. 1. Range-compressed echoes $ss(t = 2r/c,\ u)$ for a single retroreflector deployed 35.8 m from the towfish path. Sway effects are observable as an indent in what should be a parabolic locus.

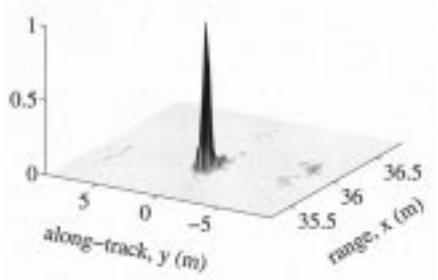


Fig. 3. Focused image generated from the sway affected echoes shown in Fig. 1 after autofocus has removed the sway error. The point target is resolved to $\delta x_{3\,\mathrm{dB}} = 6$ cm in slant-range and $\delta y_{3\,\mathrm{dB}} = 24$ cm in along-track. The highest range side lobe level is at -23 dB and the highest along-track grating lobe is at -22 dB.

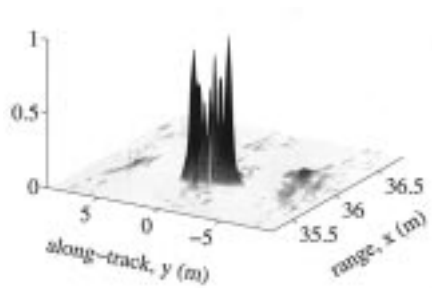


Fig. 2. Focused image generated from the sway affected echoes shown in Fig. 1. The target blur covers 12 cm in slant-range and 3 m in along-track. This image and those following has had its magnitude normalized to 1.0 and has not been corrected into ground plane coordinates.

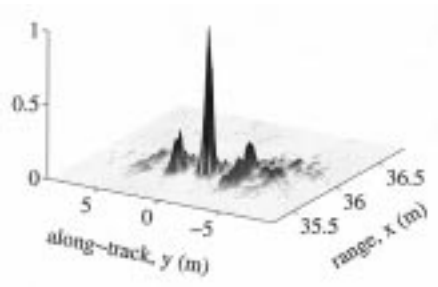


Fig. 4. Focused image of along-track spatially undersampled data with motion errors removed. The main peak is resolved to $\delta x_{3\,\mathrm{dB}} = 6$ cm in slant-range and $\delta y_{3\,\mathrm{dB}} = 24$ cm in along-track. The highest range side lobe is at -23 dB and the highest along-track grating lobe at -9 dB.

as if they have no aliased spatial frequencies at all and had been adequately sampled in the original $(t,\ u)$ domain.

VI. Results of Sea Trials

In this section, we present the results of a series of sea trials performed in Lyttelton Harbour, near Christchurch, New Zealand, on November 25, 1994 and on May 2, 1995. The sea state during each trial was not ideal (dead calm) with swells of around 3 ft occurring (sea state 1–2). Fig. 1 shows a 2-D view of the intensity of the reflections from a single retroreflector deployed 35.8 m from the tow fish path (note that this is just a small region of the sea-floor data collected). The effects of sway can clearly be seen in Fig. 1: the returns from the retroreflector should show a smooth parabolic locus, whereas the actual locus is indented where the tow fish has moved away (across-track) from the target during echo reception. The data in Fig. 1 has been sampled such that our along-track aperture sampling $\Delta_u = D/4$. With this in mind, the 3-D display of the focused image shown in Fig. 2 has not resolved into a single peak due to the corruption caused by the sway over the target, as opposed to any corruption due to undersampling effects. The effect of sway on the image is to spread the point target predominantly in along-track. The blurred target is contained within an area 12 cm in range and 3 m in along-track and so is still an improvement over the raw data shown in Fig. 1, which is spread 50 cm in range and 10 m in along-track. Fig. 3 shows the same focused target after the corrupting phase error has been removed using autofocusing. The target has resolved into a single peak with a resolution of $\delta x_{3\,\mathrm{dB}} \times \delta y_{3\,\mathrm{dB}} = 6$ cm $\times$ 24 cm, and maximum range sidelobes and along-track grating lobes of -23 and -22 dB, respectively.

As a Hamming weighting function was applied to the range dimension of the wavenumber domain before forming the image estimate, the theoretical range resolution is given by $1.30c/(2B) = 5$ cm, and with a transducer length $D = 32.5$ cm and with very little weighting (just that caused by the selection of elements and the damping effects of the housing), the theoretical along-track resolution is 16 cm. Thus, a comparison of the theoretical results to those observed shows that the autofocusing procedure has resulted in an image with resolution near the diffraction limit of the imaging system.

Fig. 4 is the result of focusing only every third along-track pulse of the data in Fig. 1 after the effects of sway had been removed. That is, Fig. 4 is the result of focusing data that has been collected along a rectilinear path, but has been *spatially undersampled* in along-track. The data is displayed as a 3-D plot to emphasize the existence of ambiguous or grating-lobe targets either side of the main peak. The distance of these ambiguous targets from the main peak is frequency-dependent, so they have been smoothed in along-track by the transmission of the low-Q waveform employed by the Kiwi-SAS system (20 kHz bandwidth pulse around a 30-kHz carrier). The main peak in Fig. 4 retains the resolution of the peak as shown in Fig. 3, but the along-track grating-lobe level has risen 13 to -9 dB. The result of applying digital spotlighting compression and upsampling, followed by focusing of the then upsampled data, produces an image which is identical to Fig. 3 and

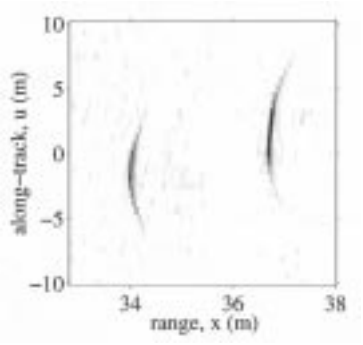


Fig. 5. Range-compressed echoes $ss(t = 2r/c, u)$ for two retroreflectors deployed 3 m apart, 35.5 m from the tow fish path. This raw data was undersampled four times in along-track; the effects of sway can be seen in the target at around 37 m.

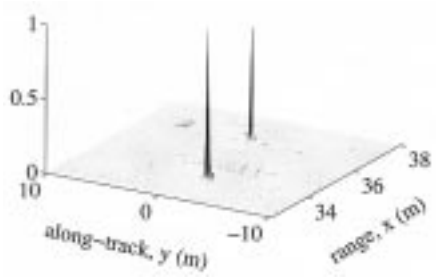


Fig. 6. The focused image of the twin retroreflectors after spotlighting and autofocusing. The target at 34 m is resolved to $\delta x_{3\,\mathrm{dB}} = 5$ cm in slant-range and $\delta y_{3\,\mathrm{dB}} = 19$ cm in along-track. The target at 37 m is resolved to $\delta x_{3\,\mathrm{dB}} = 7$ cm in slant-range and $\delta y_{3\,\mathrm{dB}} = 17$ cm in along-track.

shows no evidence at all of the ambiguous targets apparent in Fig. 4. This shows that even in an undersampled situation the system can achieve near diffraction-limited imaging of compact targets. In this particular case, a mine-hunting vessel could have traversed the aperture at three times the speed normally specified for synthetic aperture systems, yet still resolved the returns from compact mines.

Fig. 5 shows the range-compressed returns from two retroreflectors deployed 3 m apart, 35.5 m from the tow fish path. This raw data was undersampled four times in along-track, that is, $\Delta_u = D$, and as well the effects of tow fish sway can be seen quite clearly in the target at around 37 m. Fig. 6 is the result of focusing the data after spotlighting and autofocusing the data, again near diffraction-limited performance has been achieved with the target at 34 m being imaged to a resolution of $\delta x_{3\,\mathrm{dB}} \times \delta y_{3\,\mathrm{dB}} = 5$ cm $\times$ 19 cm and the target at 37 m being imaged to 7 cm $\times$ 17 cm.

The focused images Figs. 2–4 and 6 were processed with the exact algorithm for the target range, the range-Doppler algorithm with SRC, the wavenumber algorithm, and the chirp-scaling algorithm with SRC. All these algorithms produced target responses that were visually identical which is why we show only one example. A similar result has previously been shown by Bamler [2] but only for narrow-band SAR images.

VII. Conclusion

We can reach four significant conclusions.

All of the major SAR algorithms work on synthetic aperture sonar raw data; however, care should be taken to ensure that the appropriate versions of the algorithms are used since several of the narrow-band, beam-, and swath-width approximations used by the SAR community cannot be used for SAS. This is especially true of some of the more computationally efficient derivative algorithms. If there is any doubt, the safest algorithm to use is the wavenumber as it is inherently wide band-, beam-, and swath-width and, provided care is taken with the grid-to-grid interpolation, produces close to diffraction-limited images.

Given that there is a choice of waveforms allowed, the most efficient SA imaging algorithm to use is the chirp-scaling algorithm since it avoids the nonlinear grid-to-grid interpolation needed by the range-Doppler and the wavenumber algorithms. Although the chirp-scaling algorithm can be applied to waveforms other than a linear swept FM, it is optimized for that type of waveform. However, it is still only an approximation to the wavenumber algorithm, so its limitations must be considered.

If the targets of interest are relatively isolated and of limited extent, then digital spotlighting is a valuable adjunct to the imaging algorithm as it enables us to minimize the effects of imperfect motion compensation as well as extensive aperture undersampling. Here, the term "of limited extent" means that the object needs to be smaller the synthetic aperture appropriate for that range. For example, the mean range to our target in Fig. 5 was 35.5 m. The synthetic aperture at that range is approximately 5 m long, so our two-point reflector target of 3 m between reflectors is slightly shorter than its subtended synthetic aperture even oriented in the along-track direction, i.e., parallel to the flight path—which is the worse-case orientation. In fact, looking at Fig. 6, it is clear that the orientation is rather more fortuitous and lies almost in the cross-track direction.

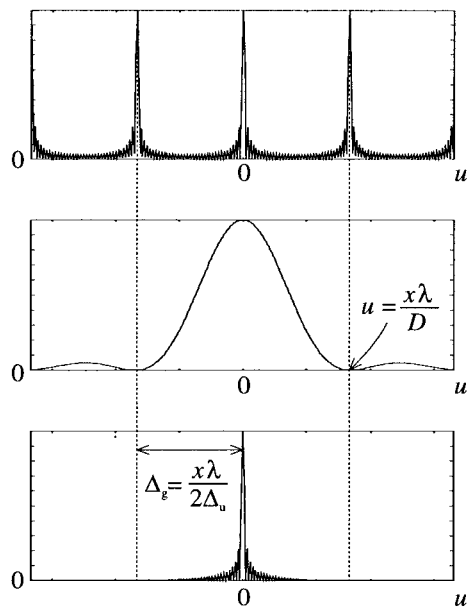
The results of the imaging algorithms with digital spotlighting on real sonar data collected in a series of unconstrained (i.e., ocean-going, boat-towed as opposed to rail-guided) sea trials proves that it is possible to get very close to the diffraction limit of the imaging system for isolated targets.

Appendix A
Aperture Sample Spacing of $D/4$

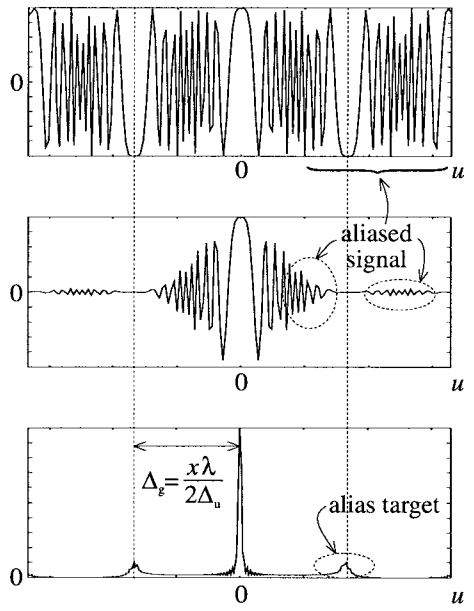
A. Aperture Sample Spacing of $D/4$

In 1978, Tomiyasu published a tutorial paper on strip-map SAR [28]. In his paper, he described several methods from which the required along-track sample spacing of $D/2$ could be determined. To date, no publications appear to refute his analysis, even though it has been obvious to a number of investigators that sample spacings of $D/2$ are barely adequate in many situations [11], [23].

The spatial sampling requirement is classically referred to as $D/2$ in the SAR literature due in part to an incorrect interpretation of the grating-lobe suppressing effect of the real aperture radiation pattern [28]. Fig. 7 clearly explains the origin of this confusion; Fig. 7(a) can be interpreted as the array factor of the synthetic aperture at frequency ω; this figure is formed by the along-track compression (phase matching) of



(a)



(b)

Fig. 7. The effect of the real aperture pattern on the along-track undersampling (a) classical grating-lobe suppression and (b) correct interpretation.

the along-track signal at the top of Fig. 7(b). This along-track compression forms a main lobe at $u = 0$ and grating lobes spaced at $n\Delta_g$ where $|n| = 1, 2, \cdots$ and $\Delta_g = x\lambda/(2\Delta_u)$. The first null of the real aperture radiation pattern occurs at $u = x\lambda/D$ (Fig. 7 assumes both transmitter and receiver have the same length, D). The sample spacing required for this null to suppress the grating lobe is $\Delta_u = D/2$.

This interpretation assumes an *incorrect sequence* of events. The correct sequence is shown in Fig. 7(b): the along-track phase function at the top of Fig. 7(b) is weighted by the real aperture radiation pattern, giving the signal shown in the center. To determine the point at which aliasing occurs requires the instantaneous position of the phase function in the (ω, k_u) domain (9), $u_i = \partial\phi(k_u)/\partial k_u \approx k_u x/(2k)$. At the folding Doppler wavenumber, $k_u = \pi/\Delta_u$, the instantaneous signal will be at spatial location $u_i = x\lambda/(4\Delta_u) = \Delta_g/2$; after this point the signal aliases. The along-track signal passes through $k_u = 0$ again at multiples of $u = \Delta_g$. With an along-track sample spacing of $D/2$, the effect of the real aperture pattern on the aliased signal is to suppress the terms at $|u| = n\Delta_g$ and "attenuate" the aliased signals near the grating lobes. When this signal is along-track compressed, the aliased energy near the first grating lobe is compressed into an alias or grating-lobe target that is 10% of the main lobe height. To reduce this alias level, the along-track sample spacing should ideally be $\Delta_u = D/4$. The classical interpretation incorrectly assumes that *all* the aliased energy is compressed into the grating-lobe peak *before* the null of the real aperture radiation pattern suppresses it. This interpretation is correct for a real aperture array because all the elements are energized simultaneously, however, for a synthetic aperture system, this is not the case.

Note that a sample spacing of $D/4$ still results in a small amount of spatial aliasing. However, this aliasing can be minimized still further by weighting the illumination patterns of the real apertures.

Appendix B
The Principle of Stationary Phase

Briefly for integrands having a wide phase variation and with an envelope of $g(u)$ [9]:

$$\int g(u)\,\exp\left[j\vartheta(u)\right]du = \left\{\left[\frac{j2\pi}{\vartheta''(u^*)}\right]^{1/2}\right\} \cdot g(u^*)\,\exp\left[j\vartheta(u^*)\right] \quad (20)$$

where the "stationary point" u^* is given by the solution of the expression

$$\vartheta'(u^*) = 0$$

assuming that the derivative of the phase with respect to u is single valued and has only one value of physical significance.

Appendix C
The Spatial Fourier Transform Pair

Repeating (6) for convenience:

$$Ss(\omega, u) = P(\omega) \iint ff(x, y) \cdot \exp\left[-j\,2k\sqrt{x^2 + (y-u)^2}\right]dx\,dy. \quad (21)$$

So let us apply a 1-D Fourier transform over u to the phase part of (21):

$$\mathcal{F}_u\left\{\exp\left[-j2k\sqrt{x^2 + (y-u)^2}\right]\right\} = \int \exp\left[-j2k\sqrt{x^2 + (y-u)^2}\right] \cdot \exp\left(-jk_u u\right)du. \quad (22)$$

Now the phase and derivatives of the RHS of (22) are given by

$$\vartheta(u) = -2k\sqrt{x^2+(y-u)^2} - k_u u$$
$$\vartheta'(u) = \frac{2k(y-u)}{\sqrt{x^2+(y-u)^2}} - k_u$$
$$\vartheta''(u) = -\frac{2kx^2}{[x^2+(y-u)^2]^{3/2}}$$

and solving $\vartheta'(u)$ for the stationary point gives

$$u^* = y - \frac{k_u x}{\sqrt{4k^2-k_u^2}}$$

and when substituted back gives

$$\vartheta(k_u) = \vartheta(u^*) = -\sqrt{4k^2-k_u^2}\cdot x - k_u y.$$

So, as a result

$$\mathcal{F}_u\left\{\exp\left[-j2k\sqrt{x^2+(y-u)^2}\right]\right\} = B\exp\left(-j\sqrt{4k^2-k_u^2}\cdot x\right)\exp(-jk_u y) \quad (23)$$

where there is complex amplitude factor B given by

$$B = \left[\frac{j2\pi}{\vartheta''(u^*)}\right]^{1/2} = \sqrt{\frac{8\pi k^2 x}{j(4k^2-k_u^2)^{3/2}}} \approx \sqrt{\frac{\pi x}{jk}}$$

and where the approximation for B is valid for wide bandwidth and low-Q system parameters.

Note that this amplitude factor B is mildly k dependent, so that in a wide bandwidth system, it represents an uncompensated amplitude modulation. Ideally, any processor should include a *deconvolution* filter to remove this deterministic amplitude function; however, the amplitude factor is insignificant for most realistic windows of the Fourier domain. Should the need arise, for instance, in a wide bandwidth system, appropriate compensation can be applied as part of the inversion processing.

Finally then, the Fourier transform of $Ss(\omega, u)$ over u is

$$SS(\omega, k_u) = P(\omega)\iint ff(x, y)\exp\left(-j\sqrt{4k^2-k_u^2}\cdot x\right)\cdot\exp(-jk_u y)\,dx\,dy \quad (24)$$

where the complex amplitude B resulting from the integration has been ignored.

ACKNOWLEDGMENT

The authors are grateful for the assistance given by M. Cusdin and A. Vernon for electronics and transducer manufacture.

REFERENCES

[1] D. A. Ausherman, A. Kozma, J. L. Walker, H. M. Jones, and E. C. Poggio, "Developments in radar imaging," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-20, pp. 363–399, July 1984.

[2] R. Bamler, "A comparison of range-Doppler and wave number domain SAR focusing algorithms," *IEEE Trans. Geosci. Remote Sensing*, vol. 30, pp. 706–713, July 1992.

[3] B. C. Barber, "Theory of digital imaging from orbital synthetic aperture data," *Int. J. Remote Sensing*, vol. 6, no. 7, pp. 1009–1057, 1985.

[4] D. Blacknell, A. Freeman, S. Quegan, I. A. Ward, I. P. Finley, C. J. Oliver, R. G. White, and J. W. Wood, "Geometric accuracy in airborne SAR images," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 25, pp. 241–258, Mar. 1989.

[5] W. M. Brown and R. J. Fredricks, "Range-Doppler imaging with motion through resolution cells," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-5, pp. 98–102, Jan. 1969.

[6] C. Cafforio, C. Pratti, and F. Rocca, "SAR data focusing using seismic migration techniques," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 27, pp. 194–207, Mar. 1991.

[7] W. G. Carrara, R. N. Goodman, and R. M. Majewski, *Spotlight Synthetic Aperture Radar: Signal Processing Algorithms.* Boston, MA: Artech House, 1995.

[8] C. Y. Chang, M. Y. Lin, and J. C. Curlander, "SAR processing based on the exact two-dimensional transfer function," in *Int. Geosci. Remote Sensing Symp.*, 1992, vol. 1, pp. 355–359.

[9] C. E. Cook and M. Bernfield, *Radar Signals: An Introduction to Theory and Application.* New York: Academic, 1967, Electrical Science Series.

[10] I. Cumming, F. Wong, and K. Raney, "A SAR processing algorithm with no interpolation," in *Int. Geosci. Remote Sensing Symp.*, 1992, vol. 1, pp. 376–379.

[11] J. C. Curlander and R. N. McDonough, *Synthetic Aperture Radar: Systems and Signal Processing.* New York: Wiley, 1991.

[12] L. J. Cutrona, W. E. Vivian, E. N. Lieth, and G. O. Hall, "A high-resolution radar combat-surveillance system," *IRE Trans. Military Electron.*, vol. 5, pp. 127–131, Apr. 1961.

[13] L. J. Cutrona, E. N. Leith, L. J. Porcello, and W. E. Vivian, "On the application of coherent optical processing techniques to synthetic aperture radar," *Proc. IEEE*, vol. PROC-54, pp. 1026–1032, Aug. 1966.

[14] P. H. Eichel, D. C. Ghiglia, and C. V. Jakowatz, "Speckle processing method for synthetic-aperture-radar phase correction," *Opt. Lett.*, vol. 14, no. 1, pp. 1–3, Jan. 1989.

[15] P. T. Gough and D. W. Hawkins, "A unified framework for modern synthetic aperture imaging," presented at the Int. Conf. Imaging and Image Processing Technology, Dec. 9–11, Santa Barbara, CA.

[16] D. W. Hawkins and P. T. Gough, "An accelerated chirp scaling algorithm for SAR and SAS image reconstruction procedures," to be published.

[17] M. P. Hayes and P. T. Gough, "Broad-band synthetic aperture sonar," *IEEE J. Oceanic Eng.*, vol. 17, pp. 80–94, Jan. 1992.

[18] C. V. Jakowatz, D. E. Wahl, P. H. Eichel, D. C. Ghiglia, and P. A. Thompson, *Spotlight-Mode Synthetic Aperture Radar: A Signal Processing Approach.* Boston, MA: Kluwer Academic, 1996.

[19] M. Y. Jin and C. Wu, "A SAR correlation algorithm which accommodates large range migration," *IEEE Trans. Geosci. Remote Sensing*, vol. GRS-22, pp. 592–597, Nov. 1984.

[20] J. C. Kirk, "A discussion of digital processing in synthetic aperture radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-11, pp. 338–348, May 1975.

[21] C. J. Oliver, "Synthetic-aperture radar imaging," *J. Phys. D, Appl. Phys.*, vol. 22, pp. 871–890, 1989.

[22] R. K. Raney, H. Runge, R. Bamler, I. G. Cumming, and F. H. Wong, " Precision SAR processing using chirp scaling," *IEEE Trans Geosci. Remote Sensing*, vol. 32, no. 4, pp. 786–799, July 1994.

[23] K. D. Rolt and H. Schmidt, "Azimuthal ambiguities in synthetic aperture sonar and synthetic aperture radar imagery," *IEEE J. Oceanic Eng.*, vol. 17, pp. 73–79, Jan. 1992.

[24] H. Runge and R. Bamler, "A novel high precision SAR focusing algorithm based on chirp scaling," in *Int. Geosci. Remote Sensing Symp.*, 1992, vol. 1, pp. 372–375.

[25] C. W. Sherwin, J. P. Ruina, and R. D. Rawcliffe, "Some early developments in synthetic aperture radar systems," *IRE Trans. Military Electron.*, vol. 6, pp. 111–115, Apr. 1962.

[26] M. Soumekh, "A system model and inversion for synthetic aperture radar imaging," *IEEE Trans. Image Processing*, vol. 1, pp. 64–76, Jan. 1992.

[27] ———, *Fourier Array Imaging.* Englewood Cliffs, NJ: Prentice-Hall, 1994.

[28] K. Tomiyasu, "Tutorial review of synthetic aperture radar (SAR) with applications to imaging the ocean surface," *Proc. IEEE*, vol. PROC-66, pp. 563–583, May 1978.

[29] D. E. Wahl, P. H. Eichel, D. C. Ghiglia, and C. V. Jakowatz, "Phase gradient auto focus—A robust tool for high resolution SAR phase

correction," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 30, pp. 827–835, July 1994.
[30] D. E. Wahl, C. V. Jakowatz, and P. A. Thompson, "New approach to strip-map SAR auto focus," in *Sixth IEEE Digital Signal Processing Workshop*, Oct. 1994, pp. 53–56.
[31] C. A. Wiley, "Synthetic aperture radars," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-21, pp. 440–443, May 1985.

Peter T. Gough was born in Dunedin, New Zealand, in 1947. He received the B.E. (1st Class Honors) and Ph.D. degrees from the University of Canterbury, Christchurc, New Zealand, in 1970 and 1974, respectively.

After completing his Ph.D. work, he continued at the University of Canterbury for some five years as a Post-Doctoral Fellow and then as an Assistant Professor at the Institute of Optics, University of Rochester, Rochester, NY, and at the University of Manitoba, Winnipeg, Canada. After working as a consultant in California, he returned to the Electrical and Electronic Engineering Department, University of Canterbury, in 1980, where he was, until recently, the Head of Department.

David W. Hawkins was born in Waipukurau, New Zealand, in 1970. He received the B.S. degree in physics from Victoria University, Wellington, New Zealand, in 1991. He is currently working towards the Ph.D. degree in the Electrical and Electronic Engineering Department of the University of Canterbury, Christchurch, New Zealand.