Chapter 3

Incoherent Synthetic Aperture Sonar Processing

3.1 Introduction

Clearly, the limitations and its related setbacks suffered by a coherent SAS system raise questions of its reliability under non-ideal operating conditions. Incoherent SAS processing is a viable and promising alternative to a fully coherent SAS because it directly addresses some of the constraints present, and at the same time offers features that are otherwise unavailable.

This chapter investigates the incoherent approach applied in synthetic aperture sonar processing. The details of implementation are shown, and the technical properties of this technique are presented. Simulations are presented along the way for purposes of illustrations and comparisons. Discussions of the potential and limitation of this approach are made. All the simulations and analytical investigations in this chapter are performed using Matlab. Unless otherwise specified, they are based on two assumptions: monostatic configurations of transducers; and transducers’ movements during pulse transmission and reception are negligible.

3.2 Phase Removal

In incoherent SAS processing, the phase information is removed before along-track integration, but only after cross-track processing. Cross-track processing is performed by a matched filter in the frequency domain for computational efficiency. A Hilbert transform is then applied, so that taking the inverse fast Fourier transform will give the analytical form of the cross-track processed signal in the time domain. Hence, phase information is usually preserved up to the stage of along-track integration.
The removal of phase information is performed by a modulus square operation \([35]\), which is basically a multiplication of the received signal which has been processed in the cross-track domain, \(\Gamma(t)\), by its complex conjugate, such that

\[
P(t) = \Gamma(t) \cdot \Gamma^*(t)
\]  

(3.1)

to give the power of the signal \(P(t)\). At this stage, the image becomes

\[
I_{\text{image}}(n_a, n_r) = P_{n_a}(t_{n_r})
\]  

(3.2)

which is a representation of the reflected energy density at locations corresponding to the pixels \((n_a, n_r)\), where \(n_a\) and \(n_r\) are discrete positions in the along-track (azimuth) direction and cross-track (range) direction respectively.

### 3.3 Time-domain Back-projection

It has been established in section 2.3.2 that time-domain algorithms are more suitable for the along-track processing of SAS data. With an incoherent SAS approach, time-domain techniques are also the most suitable because the exact reconstruction algorithm can be used. This reduces the number of approximations being made, and hence avoids any unnecessary degradation of along-track resolution. Long computing time is a rapidly diminishing issue as the ratio of processing speed to cost continues to increase exponentially.

Consider a stationary point target located at coordinate \((x_T, y_T)\) on a 2-D plane surveyed by a transducer (monostatic) towed linearly pinging at positions \((x_0, y_n)\), where \(n\) is the index of along-track positions (Fig.3.1). Back-projection here implies the calculation of the range from the target to each ping position of the transducer in the along-track direction, such that

\[
R_n = \sqrt{(x_T - x_0) + (y_T - y_n)}
\]  

(3.3)
Chapter 3: Incoherent Synthetic Aperture Sonar Processing

Thus effectively back-projecting, in range, the point target’s location to the synthesized aperture. $R_n$ here is a set of range values corresponding, and unique, to the point target concerned. These are the exact ranges from the target to the ping positions.

This set of range values can be converted into the total propagation time of a transmitted pulse to and from the point target for each ping location. As such, a set of time-delays $t_n$ is produced from the set of range values, $R_n$. The time-delays for the point target, if plotted against the along-track ping positions, will form a hyperbolic curve, and can also be called the range-curvature of the target. Based on equation (3.3), $t_n$ can be expressed as:

$$t_n = \frac{(n \cdot 2 \cdot v)^2}{(c \cdot PRF)^2 + (t_0)^2}$$  \hspace{1cm} (3.4)
where $t_0$ is the time-delay corresponding to the closest point of approach (CPA) of the target to the synthesized aperture. $t_n$ is unique to the point target located at $(x_T, y_T)$, and will be the set of time-delays used in order to integrate in the along-track direction to reconstruct the reflectivity at $(x_T, y_T)$. The process of SAS image reconstruction is undertaken by performing such an operation for all other point of interest on the $x$-$y$ plane.

As $t_n$ is a function of $t_0$, a new set of time-delays is needed for each point of interest to be reconstructed (Fig. 3.2). This is the exact reconstruction algorithm, and is also known as a “fully focussed” SAS image reconstruction. The concept of depth of focus can be used to avoid re-estimating the time-delay function for every CPA. Targets with CPAs in the same depth of focus will share the same time-delay function. The larger the depth of focus, the greater will be the degradation of overall image quality due to the errors between the assumed and actual time-delay functions within the depth of focus.
One can tell by visual observation that this error increases as CPA decreases given the same depth of focus. This is due to the increasing gradient of the hyperbola as CPA decreases. Let this focussing error be defined by $E_f$ such that

$$E_f = t_{n,CPA2} - t_{n,CPA1}$$  \hspace{1cm} (3.5)$$

where $t_{n,CPA2}$ and $t_{n,CPA1}$ are the set of return time-delays for two reflective points with CPA2 and CPA1 respectively, and $t_{n,CPA} = 2CPA/c$. Let $e_{CPA} = CPA2 - CPA1$, the focussing error can be expressed as:

$$E_f = \sqrt{\frac{2\cdot CPA1}{c}} + \frac{e_{CPA}}{c} - \sqrt{1 + \frac{2\cdot CPA1}{c}}$$  \hspace{1cm} (3.6)$$

**Figure 3.3:** Plot of $E_f$ versus CPA1, based on equation 3.6, showing that the error increases as a fixed depth of focus, $e_{CPA}$, approaches the synthesized aperture.

$e_{CPA} = 22.5m$

$e_{CPA} = 15.0m$

$e_{CPA} = 7.5m$
which dictates that, with the same depth of focus, the maximum error due to the mismatch of time-delay curves increases as the depth of focus draws nearer to the synthetic aperture (Fig.3.3).

3.3.1 Coordinate mapping

To perform image reconstruction, it is convenient to have two separate matrices, $I_{\text{data}}$ and $I_{\text{image}}$, to respectively represent the collected data and the result of along-track integration on the collected data. Based on the sampling criteria, the coordinates corresponding to any discrete samples on $I_{\text{data}}$ can be known. However, this coordinate system may not always match the coordinate system in the desired form of $I_{\text{image}}$, in which case the necessary mapping function needs to be taken into account. One example will be the case when the reconstructed image, in the cross-track direction, needs to be represented with the coordinates of the seabed plane, rather than just the ranges from the synthesized aperture. Unless specified otherwise, this thesis assumes the same coordinates in both $I_{\text{image}}$ and $I_{\text{data}}$.

3.3.2 Image reconstruction

The image reconstruction process uses the data in $I_{\text{data}}$ to produce $I_{\text{image}}$, where $I_{\text{data}}$ contains $\text{Image}(n_a,n_r)$ stated in (3.2). To reconstruct the SAS image, for each point on $I_{\text{image}}$, a sample will be chosen from each cross-track slice of $I_{\text{data}}$ based on the back-projected ranges, and then summed to be the result, such that

$$I_{\text{image}}(n_a,n_r) = \sum_{n=1}^{N} I_{\text{data}}(n,n_d(n,n_a,n_r))$$

where $(n_a,n_r)$ is the coordinate of the point of interest on $I_{\text{image}}$, $N$ the total number of pings in the synthesised aperture, $n_d$ the sample delay which determines which sample on the $n^{th}$ cross-track slice is selected for summation. The sample delay $n_d$ is based upon the set of back-projected ranges:

$$n_d(n,n_a,n_r) = \frac{2 \cdot f_s \cdot R_n(n_a,n_r)}{c}$$

(3.8)
where $f_s$ is the sampling rate and $R_n$ is the back-projected range function of the $n^{th}$ element.

The process in equations (3.7) and (3.8) can also be expressed in terms of back-projected time-delays, as they are easily interchangeable with the monostatic transmitter-receiver assumption.

### 3.3.3 Two-way propagation distance in bi-static configuration

It is worth noting that the term *bi-static* and *multistatic* used in this project stresses on the fact that the transmitter and receivers are not strictly co-located in the proposed schemes, and is *not* a proposal for a deliberately separated transmitter/receiver system with a wide bi-static angle (angle between the transmitter and receiver as seen from the target at nearest range of interest) of $>30^\circ$ as seen in fully bi-static systems. This is clarified further below.

The term bi-static here refers to two specific situations in SAS imaging:

1) A towed SAS where the pulse is received at a different location than where it was transmitted. At 50m range, a typical imaging SAS towed at 1.5ms$^{-1}$ would have moved by approximately 0.1m (in the along-track direction) from its original position when the return pulse is received. The bi-static angle is $0.1^\circ$.

2) A single transducer/multiple sparsely spaced hydrophone array is used. In the sea-trial in chapter 6, the furthest hydrophone is 11.25m away from the boomer transmitter. At 30m range (the depth of the trial environment), the return path is approximately 2m longer, compared to the transmit path. The bi-static angle is $20.5^\circ$.

The term multistatic is used to describe the proposed scheme in chapter 7, where there are 2 transmitters that correspond to other hydrophones attached onto the same laterally towed-array. With the proposed array dimension and element separation, the maximum bi-static angle is approximately $18.4^\circ$. 
Hence, because of the small bi-static angle, the scope of this work does not cover the technical properties for a fully bi-static system. It must be noted that properties of monostatic system such as along-track resolution and cross-track resolution no longer hold when the bi-static angle is large. In such cases, the interested reader is referred to [83,84] to gain an appreciation of the complexities involved in synthetic aperture systems with a wide bi-static angle ($\geq 30^\circ$ and $\leq 180^\circ$).

This work only stresses the need to take into account the difference of range between the outgoing and return signal path so that the range/time-delay curvature is matched optimally in the process of along-track integration.

The two-way path for a bi-static case, where the transmitters and receivers are not collocated, is not just $2R_n$. Rather, the total range of the two-way path $R_T$ is now the sum of the range from the transmitter to the target, $R_{n,tx}$, with the range from the receiver to the target, $R_{n,rx}$. The two-way propagation delay is expressed as

$$n_d(n,n_r) = \frac{f_s \cdot (R_{n,tx}(n_a,n_r) + R_{n,rx}(n_a,n_r))}{c}$$

(3.9)

Hence, in the sea-trial case in chapter 6, the two-way path is 62m for the furthest hydrophone, and not 60m. Equation 3.9 replaces 3.8 for the calculation of total two-way range in the summation algorithm to take this into account. If this is not taken into consideration in the along-track integration of samples, then at the sampling rate of 51.2kHz, it results in an error of 68 range samples.

### 3.3.4 Interpolation

It is obvious from equations (3.7), (3.8) and (3.9) that there needs to be interpolation in the cross-track domain due to the fact that $n_d$ is seldom a whole number, meaning that the required sample falls between two existing samples. Thus, interpolation is needed to provide a more accurate value for that sample. A linear interpolation between the two samples is believed to be sufficient. Alternatively, $n_d$ can be rounded to a whole number, avoiding the need for interpolation, and its effect on image quality should be
minimal as long as the sampling interval in the cross-track direction is much smaller than the achievable resolution in that direction.

Interpolation can be used as a tool for reducing the memory requirement for image reconstruction. In order to keep the data matrix as small as possible, it can be down-sampled so that minimal memory is required for processing, and interpolation is employed during image reconstruction to estimate the removed samples.

### 3.3.5 Computational complexity

Let $O$ be the complexity of the operation to obtain $n_d$ as in (3.9), such that there are two hypotenuse calculations and an interpolation, the number of computational calculations to implement the process in (3.7) to reconstruct $I_{image}$ with $N_a \times N_r$ pixels is approximately $O.N_r.(N_a)^2$.

Making a mono-static assumption will reduce the complexity of $O$. The same happens when a depth of focus is introduced, where the same set of back-projected ranges and its corresponding delays are used for more than one point of interest. The reduction in computational complexity directly decreases the necessary computational time.

### 3.3.6 Processing gain

Coherent and incoherent processing yields a different processing gain given that the same number of along-track samples is used for aperture synthesis. For coherent processing, where the signals are summed vectorially [72(pp.70)], the processing gain in integrating $N$ spatial samples from consecutive pings is [73(pp.25-26),74(pp.365)]:

$$G_c = 10 \cdot \log(N)$$  \hspace{1cm} (3.10)

expressed in dB. For example, for an active sonar that requires a signal-to-noise ratio (SNR) of 12dB in a single pulse (for probability of detection, PD, of 0.95 and probability of false alarm, PFA, of $10^{-3}$ [72]), integration of 100 coherent pulses will decrease the SNR requirement by 20dB to achieve the same level of PD and PFA.
If the pulses are summed incoherently (i.e. after square-law detection), the processing gain achievable is less than in coherent integration. This is treated as a loss in SNR relative to what would have been obtained in coherent integration, by Marcum [75]. This integration loss is a function of an incoherent integration efficiency that is less than 1 [73(pp.28)]. Defining this efficiency parameter as $\gamma$, the gain achievable in incoherent integration is:

$$Gi = 10 \cdot \gamma \cdot \log(N)$$ \hspace{1cm} (3.11)

The incoherent integration efficiency, $\gamma$, is dependent on $N$, $PD$ and $PFA$ (and hence SNR) [Meyer Plots in 72]. By example, with $PD$ of 0.95 and $PFA$ of $10^{-6}$, the Meyer Plots in [72] give an incoherent integration efficiency ranging from 80% for $N=10$, 72.5% for $N=100$, 68% for $N=1000$, dropping to 64% for $N=10000$. The target model used in this example is non-fluctuating.

Setting $\gamma$ to be 50% such that the incoherent processing gain is $5\log(N)$ is a conservative approximation, or for very large value of $N$. It is used in [76] and [77] as a simplistic and conservative approximation of incoherent processing gain, and with relation to SAS in [15(pp.86)]. When an incoherent processing gain of $5\log(N)$ is used in this work for analytical purposes, it is worth noting that it is a conservative approximation, and in practice, incoherent processing gain in the region of $6\log(N)$ to $7.5\log(N)$ is realistic ($N$ is usually within the region of 100 to 10000).

In cases where the trade-off in processing gain due to a lesser number of elements being integrated are discussed, a realistic processing gain (between $6\log(N)$ to $7.5\log(N)$) is applied to give a more accurate representation of the loss in processing gain.
Chapter 3: Incoherent Synthetic Aperture Sonar Processing

3.4 Pulse Design

Broadband pulses are used to obtain optimal cross-track resolution, and as shown in section 3.5, the along-track resolution as well in the case of incoherent processing. The transmitted signal used in this work is a linear frequency modulated (LFM) or ‘chirp’ signal. It can be expressed as

\[ s(t) = \cos(2\pi f_c t + \frac{\pi B t^2}{T}) \]  

The Q-factor, \(\frac{f_c}{B}\), is a good measure of the genuine degree of broadbandness, where \(f_c\) is the centre frequency and \(B\) the bandwidth of the pulse.

The use of windowing on the chirp pulses will reduce the level of range side-lobes on the matched-filtered pulses. However, because the effective bandwidth is reduced by windowing, the cross-track resolution becomes wider. This is shown in Fig 3.4.

In practical situations, unilateral weighting (where only the matched filter replica is weighted) is useful for maximising the power of transmission, such that the loss of

Figure 3.4: Matched filtered chirp pulse (a) without and (b) with Hanning window.
amplitude is offset by the much larger energy content of a non-weighted pulse transmission [36].

Where Doppler effects induced by the movements of transmitters and receivers during the process of transmission and receipt are detrimental to the gain of the matched-filter, the use of linear period-modulated (LPM) signals may be more appropriate [37],[38].

3.5 Along-track Resolution

The along-track resolution of an incoherent processing scheme can be derived from the geometry as shown in Fig.3.5. \( r_x_a \) and \( r_x_b \) are two transducer positions with beam angle \( \theta_a \), and \( r_0 \) is the closest path of approach to the target. The dotted lines are range-ambiguity curvatures with respect to the transducer positions, and they correspond to the matched-filtered responses from a point target marked ‘x’. The point target matched-filtered responses range from \(-\Delta r\) to \(\Delta r\).

**Figure 3.5:** Geometry for the derivations of along-track resolution in incoherent processing.
Chapter 3: Incoherent Synthetic Aperture Sonar Processing

The along-track resolution of the system is defined in terms of the correlation of the overlapping responses. This resolution is taken to be twice the along-track increment that causes the correlation to drop to half its peak value, considering only a cut along the along-track dimension of the azimuth-range ambiguity function [39]. Thus, the cross-track resolution, \( P_r = 2\Delta r \), and the along-track resolution, \( P_z \), can be expressed as

\[
P_z = x_h - x_a
\]  

(3.13)

With the assumptions that

\[
x_b \gg x_a \quad \text{and} \quad r_o \gg \Delta r/2
\]  

(3.14)

which implies a long synthetic aperture length and a fine cross-track resolution. It can be shown that (Appendix A):

\[
P_z = \frac{2\Delta r}{\sin \theta_a}
\]  

(3.15)

By considering the directional response of the transducers having a 3dB effective beamwidth [40],

\[
\sin \theta_a = \frac{1.39 \lambda}{\pi D_p}
\]  

(3.16)

Substituting (3.16) into (3.15) gives, with a conservative approximation,

\[
P_z = \frac{4.5 \cdot \Delta r \cdot D_p}{\lambda} \approx \frac{3P_r D_p}{\lambda}
\]  

(3.17)

and for broadband pulses
\[ P_z \equiv 1.5QD_p \]  

which is an important result in incoherent SAS processing. It shows that the use of broadband pulses is advantageous since the cross-track resolution influences the along-track resolution proportionally. It also shows that the wider is the beamwidth, the better the along-track resolution will be, and also implied that the along-track resolution improves as the synthetic aperture length covering the point of interest increases.

Simulations were used to confirm the results obtained. A point target 150m away is pinged by a chirp pulse with a \( Q \)-factor of 2.5. The along-track spacing is 3.33m.

The first simulation studies the dependence of the along-track resolution on \( \theta_a \). The along-track resolution is measured as the along-track spacing that corresponds to the half-level of the target response in the processed image, and it is computed for synthetic apertures of varying length to represent varying \( \theta_a \) in (3.14). The cross-track resolution, \( P_r \), is first established from the pulse-compressed signal to be 12.75m. Hence a plot based on (3.15) can be made and compared to what is obtained from the simulation. The results are shown in Fig. 3.6. When \( \theta_a \) increases, the \( x_b \approx x_a \) assumption becomes valid and thus the two theoretical plots converge. The simulation verifies the validity of equation (3.15) in terms of the dependence of the along-track resolution on the beam angle \( \theta_a \).

Such dependence of \( \theta_a \) implies that the omni-directionality of the physical elements in the synthetic aperture is a desirable feature for finer along-track resolution. Besides that, one can also deduce that in the case of non omni-directionality, which is true under most circumstances, there will be a limit on the synthetic aperture length beyond which there is a smaller improvement in the along-track resolution. This synthetic aperture length depends on the CPA of the point of interest. It then follows that in order to achieve uniform along-track resolution at all ranges, a varying synthetic aperture length will be needed in the image reconstruction process.
The second simulation maintains $\theta_a$ constant ($\theta_a=1.15\text{rad}$), but varies the $Q$-factor of the chirp pulse so that the cross-track resolution, $P_r$, of the matched-filtered output changes. The corresponding along-track resolutions are then measured and plotted. The results are shown in Fig. 3.7. As the cross-track resolution becomes finer, the validity of the assumption in (3.14) improves. Thus the theoretical plots at fine cross-track resolution converge. The simulation verifies the relationship between the along-track resolution and the cross-track resolution. It highlights the benefit of transmitting broadband pulses with typical $Q$-factor $<1$, where the cross-track resolution is fine, resulting in an improvement of the along-track resolution.
Figure 3.6: (a) The along-track slice shows the change in along-track resolution as $\theta_a$ changes. (b) The dependence of the along-track resolution on $\theta_a$ is shown by simulation and compared to the theoretical predictions.
Figure 3.7: (a) The along-track slice shows the change in along-track resolution as $P_r$ changes. (b) The dependence of the along-track resolution on $P_r$ is shown by simulation and compared to the theoretical predictions.
3.6 Dynamic Focussing

Having established the relationship between the along-track resolution and the beam angle $\theta_b$, it is clear that in order for the along-track resolution to be consistent over all cross-track ranges, it is necessary to have a variable synthetic aperture length for the different ranges. Such a requirement fulfils the relationship between the physical and synthetic beam, a relationship which produced equation (2.2). This equation underlines the linear dependence of the synthetic aperture length $D_s$, on the range between the target and the aperture, $R_0$. Fig. 3.8 illustrates this dependence.

Figure 3.8: Dynamic focussing by varying the length of the synthetic aperture according to the range to the aperture, hence keeping $\theta_s$ and the along-track resolution constant.

In order to implement this, the synthetic aperture length used to focus a target with smaller $R_0$ may be shorter than it could have been, thus limiting its along-track resolution. In other words, the along-track resolution is capped by the level achievable at the farthest range of interest. Thus in practice, if achieving constant along-track resolution for all ranges is not paramount, it is not uncommon to use the largest
synthetic aperture available so that an optimal along-track resolution is always achieved albeit it degrading with increasing $R_\theta$.

### 3.7 Self-noise

One characteristic of images summed incoherently, due to the removal of phase, is the presence of self-noise or background-clutter surrounding the focussed point-target. It can be defined as the energy contained beyond the half-of-peak level perimeter surrounding the incoherently focussed image (Fig. 3.9). The pattern of distribution and the level of self-noise are both influential factors affecting the quality of the processed image. A low signal to self-noise ratio will increase the likelihood of targets with low reflectivity being masked by stronger targets nearby, as well as the appearance of false-targets contributed by the self-noise of adjacent strong targets.

The self-noise pattern can be managed by the selection of elements and the use of weighting functions in SAS processing. This fact is further exploited in Chapter 4.
3.8 Grating Lobes

In coherent processing, grating lobes appear when the inter-ping spacing exceeds $\lambda/2$, and this is especially evident when the pulses are narrowband. However, with incoherent processing, under-sampling in the along-track direction does not cause the appearance of grating lobes. Rather, it causes a reduction in the signal to self-noise level due to a lesser number of elements insonifying the same target. This is also true in the coherent case. In both cases, there is no reduction in the quality of along-track resolution. Along-track slices from the simulation results of the incoherent and coherent processing of uniformly under-sampled data are shown in Fig. 3.10.

In the case of incoherent processing, such a property provides the convenience of under-sampling in the along-track direction without the presence of grating lobes. The

**Figure 3.10:** Along-track slices show that grating lobes appear when under-sampled data were processed coherently, but is absent in the incoherently processed image. The samples are uniformly spaced in both cases.
constraint in equation (2.4) governing coherent SAS processing is no longer valid, and the new lower constraint on PRF becomes

\[
PRF \geq \frac{\nu}{P_z} = \frac{\lambda \cdot \nu}{3 \cdot P_r \cdot D_p}
\]  

(3.19)

The broader synthesized beam (lower along-track resolution) via incoherent processing, has relaxed the constraint on the PRF. Because \(P_z\) is bigger than \(P_s\), \(\nu\) can be increased without violating the constraint. Thus with the same PRF, the tow speed, \(\nu\), can now be increased to achieve a higher mapping rate. This increase in mapping rate, however, is at the expense of signal to self-noise ratio since less elements are available within the maximum length of the synthesized aperture. The limit of acceptable signal to self-noise ratio degradation, which varies depending on the target applications, must not be exceeded.

### 3.9 Tolerance to Motion Errors

Incoherent SAS processing does not demand a high level of precision in the knowledge of the actual along-track path compared to coherent SAS processing. It is thus more tolerant to unknown path deviations.

It is fair to assume that in coherent processing, even though minor unknown path deviations cause blurring to the image, it, nevertheless, still has a better sharpness than an incoherently processed image. However, this is only true if the intended target has a strong reflectivity in contrast to its background, and that the path deviations have zero mean value. The first condition may be met in some applications, but the latter is rarely fulfilled. This means that the path deviations will accumulate as the synthesized aperture gets longer, and consequently, the coherently processed image becomes severely grainy.
The incoherently processed image will also be blurred, but it retains the target’s cognitive properties far more satisfactorily with little graininess. There are two main reasons for this. Firstly, because coherent processing produces high resolution images, such that it has a fine synthetic beam, it is inevitable that the image appears “grainy” when there is mismatch between the estimated and the actual back-projected range-curvature function. Secondly, for the same reason that the gain of $N$ samples added coherently is higher than the gain achievable if added incoherently, the fall in processing gain is far greater in coherent processing than in incoherent processing if there are any mismatch in range-curvature functions. Hence, the detection properties are better with incoherent processing despite a reduction in sharpness.

A simulation is run to illustrate the points made. Four equally reflective point-targets are simulated with path deviations generated from a uniformly random distribution with zero mean and a standard deviation of $0.9m$. 
Chapter 3: Incoherent Synthetic Aperture Sonar Processing

Figure 3.11: SAS image of 4 point targets processed (a) coherently and (b) incoherently with random unknown motion errors. The plots are normalised by its respective maximum, and represent the integrated energy in dB.

Figure 3.12: Along-track and cross-track slices showing the greater tolerance to random motion errors in incoherent processing. Note the difference in scale between the left and right plots.
Fig. 3.11 shows the images of the 4 point-targets processed coherently and incoherently. Grains are obvious on the coherently processed image and the point targets appear unfocussed and smeared, whereas on the incoherently processed image, the targets, although not as sharp, still have a relatively strong focus of energy with no appearance of grains. Slices of the images are taken along the along-track direction, cutting through the first two point targets, and compared to the corresponding slices from the processed images where no motion errors were introduced. These slices, in Fig. 3.12, show a more severe degradation in the processing gain of the coherently processed targets when there are random unknown motion errors. It is worth noting that the motion errors commonly experienced may not have a zero mean, in which case the unknown errors are cumulative and yield a more detrimental effect to the coherently processed image.

3.10 SAS Imaging Modes

There are two main imaging modes applicable in SAS imaging; stripmap and spotlight modes, terms originated from, and coined in, synthetic aperture radar [41].

Stripmap mode is the typical mode of imaging in survey and terrain mapping applications where large areas need to be covered. It is the imaging mode implemented throughout this work. In this mode the transducer’s real beam faces the same direction, without any angular bias, for each and every ping, thus ensuring an even level of beam exposure across the swath of interest. Transducers with a wide beam are preferred so that a higher processing gain is obtainable via a longer synthesizable aperture.

Spotlight mode, on the other hand, involves steering the beam of the transducer towards a fixed spot on the scene for every ping. This is either in the form of mechanical steering, or electronic beam-steering with an array of transducers. This mode of operation is usually applied when a high resolution image is required of a specific, well-defined area. The raw data of a point target obtained in spotlight mode is similar to the raw data in stripmap mode with omni-directional transducers. This is due
to the fact that, ignoring acoustic absorption, the returns are equally strong across the whole synthesized aperture regardless of its length.

Most proposed SAS systems work in the stripmap imaging mode, however, in practice, omni-directional transducers that satisfy both the bandwidth and sensitivity requirements of the system are rarely available. In incoherent processing, the possibility of having a large inter-ping spacing suggests that it is not unusual to have a long synthesized aperture with few elements, where the limiting factor on the processing gain is the effective (3dB) beamwidth of the transducers. In this case, it is possible to have a longer synthesized aperture to improve the processing gain by including elements whose 3dB beam does not cover the point of interest. The condition for this is that the directional response of each individual element with respect to the point of interest must be taken into account in the SAS processing. This is done by introducing a weighting function into the image reconstruction algorithm in equation (3.7).

### 3.11 Weighting Function

In order to include elements whose 3dB beam does not cover the point of interest, it is necessary to introduce a weighting function in the image reconstruction algorithm. In order to obtain a maximum gain, the weighting function applied must match the directional response of the individual elements in the synthetic aperture. The argument for this is straight forward.
Chapter 3: Incoherent Synthetic Aperture Sonar Processing

Consider the scenario in Fig. 3.13, where there are three transducers’ ping positions named as ‘A’, ‘B’ and ‘C’ surveying a target represented by the hexagonal shape. The target is within the 3dB beam of ‘B’ ping, but not for ‘A’ and ‘C’ pings. Thus, on the image registration matrix, the returns of the target detected by ‘A’ and ‘C’ is weaker than in ‘B’, as shown with a shade of grey versus black. Now, consider ‘C’ alone. For this single target, its return is recorded in pixel [C,3]. However, if there are other targets present, the energy in pixel [C,3] will represent the sum of the reflected energy from all the targets that lie on the same range-line of ambiguity. Based upon the directional response of the transducer, it can be deduced that the amount of contribution of each target along the range-line of ambiguity towards the energy in [C,3] depends on the angle of the targets with respect to the transducer. Hence, matching the weighting function with the directional response of the transducer is to acknowledge the fact that the targets along the range-line of ambiguity do not contribute equally to the total detected energy. By taking this into account, the error between the actual and the estimated energy distribution pattern is minimised, thus reducing the level of self-noise around the focussed target.

Figure 3.13: A target seen from three ping positions, with returns at different intensities recorded on the data registration matrix due to the directional response of the transducer.
Equation 3.7 is thus modified to

\[
I_{image}(n_a, n_r) = \sum_{n=1}^{N} w(n, n_a, n_r) \cdot I_{data}(n, n_d(n_a, n_r))
\] (3.20)

where \( w(n, n_a, n_r) \) is the weighting function applied across the synthetic aperture that corresponds to the point of interest at \((n_a, n_r)\). The fact that the weighting function is also a function of \(n_a\) and \(n_r\) says that it is a function of the directional angle of the point of interest with respect to the element on the synthetic aperture.

Simulation results to compare the effect of using a matched weighting function are shown in Fig. 3.14. Three point targets are imaged with an element that has non-uniform directional response. The image processed with a matched weighting function has a significantly lower level of self noise. This may not be obvious on the processed images, but the plot of slices across targets in both directions showed that the self-noise level is up to 3dB lower in the cross-track domain if a matched weighting function is applied. In the along-track direction, however, there is a very small degradation. The real improvement is the reduction of the level and spread of self-noise in areas surrounding the focussed target. If more slices are taken at different angles instead of just strictly in the along-track or cross-track direction, it can be observed that this improvement is present for most directions around the target, albeit a slight degradation in the along-track direction alone. Hence, it produces a more contained, limited and even spread of self-noise around the target. Further results of this are shown in the tank trial results in chapter 7. This is true assuming that the signals are not noise-limited. Otherwise, the benefit is not justifiable in the light of the fact that an inverse, rather than matched, weighting function is usually optimal in a noise-limited scenario.
In the case where exactly the same synthetic aperture is used to focus all the points of interest, the final image will be non-uniformly weighted in the along-track direction. It can be avoided by using a different synthetic aperture for points at different positions in the along-track direction, such that the point of interest is always centred relative to the synthetic aperture. In another words, the initial value of $n$ in the summation of equation (3.20) becomes a function $n_a$. This is practical because in many imaging situations, the synthetic aperture length used in processing is usually smaller than the overall length of the sonar towed path. Otherwise, the weighting in the final image can be removed by compensating it with a gain function. This, however, will increase the level of self-noise as well.

Figure 3.14: The image of 3 point targets processed (a) without weighting and (b) with weighting matched to the response of the transducer. For comparison, the slices in the (c) cross-track and (d) along-track direction over the targets as indicated by the dotted line over the processed image.
3.12 Point Target Energy Response

In order to study the effect of the synthetic aperture on the energy distribution pattern, it is necessary to have a scheme that enables an accurate representation of the energy distribution on and around a focused point target, or the point target energy response. This scheme needs to be able to take into account the configuration of the synthetic aperture, the processing algorithm and the application of weighting function.

In this work, a range-ambiguity back-projection scheme is used. It is based upon the plot of range-ambiguity loci crossing on a point target (Fig. 3.15(a)). The distribution of the loci on and around the point target is the point target energy response of the synthetic aperture.

This is a simple yet powerful tool in the design process of an incoherent SAS system. It allows a quick and accurate estimate of the energy distribution pattern for a particular configuration or arrangements of elements in the synthetic aperture. Fig. 3.15(b) shows the point-target response plot for a uniform, monostatic, non-weighted synthetic aperture in two locations, both of which centred on the aperture and at a range of 80m and 100m respectively. The overlap of self-noise is highlighted. It shows how distributed energies from adjacent targets can contribute to an overall higher level of self-noise in the image.
Figure 3.15: (a) The range-ambiguity loci of elements on a synthetic aperture towards a point target. (b) The point-target energy response of a uniform 2 m inter-ping spacing synthetic aperture by plotting corresponding range-ambiguities loci.
Note that the scheme of merely plotting lines in Fig. 3.15(b) relies on the density of the lines to represent the intensity of the energy present, which is thus limited especially for purposes of analytical comparisons between two different synthetic aperture arrangements, or when the intensity of the lines vary such that the elements are directional. This scheme can be easily extended to address these concerns, and take into consideration the weighting function as well as a bi-static configuration of elements. An image can be used such that the lines are plotted as pixels on the image. In this way, points on the lines that overlap will have pixels with larger values. Fig. 3.16 illustrates this with the same synthetic aperture as the one used in Fig. 3.15(b).

**Figure 3.16:** Intensity plot of range-ambiguity loci over two point targets, showing the point-target energy response of the synthetic aperture at the two locations.

This method allow one to predict, quickly and visually, the point-target energy response in the incoherent processing of a synthetic aperture with bi-static configuration of any form, and can be extended to predict the volumetric pattern of
point-target response in 3-dimension. This scheme is applied in the study of the algorithms presented in the later chapters.

### 3.13 Image Quality Benchmark

How does one compare the quality of an image processed with two different algorithms? There are some general criteria that a high quality image fulfils, like fine resolution, good contrast and strong signal to noise ratio. The first two influence the cognitive property of the image while the latter influences its detection property. The cognitive property of the image relates to its quality of recognition [60]. The detection property, though not completely distinct, relates to the probability of detecting a target in the presence of noise, reverberations or shadows of stronger nearby targets. Depending on the type of application, the property of which the quality benchmark is based upon and optimises may be different. The detection property is of priority in a mine-hunting application, whereas cognitive property is more important in the imaging of underwater archaeological sites.

In the context of incoherent SAS imaging, the resolution of the processed image is often difficult to quantify, especially if there are no available a priori knowledge of the target scene. The contrast of the image is a form of statistical measure, requiring no pre-determined knowledge of the target scene, and the simplest model being the standard deviation normalised by the mean of the image [27]. This differs from the standard definition of contrast in image processing, but rather focuses on the measure of structural content in the image. The measure of signal to noise ratio too, like resolution, requires the a priori knowledge of a target scene to give an accurate comparison.

Thus, a reliable way of comparing image quality is to take the measure of its contrast. It is worth noting that its accuracy in representing the overall image quality increases with the presence of structural content in the image. Such a comparison is usually done on the statistic of an individual, or a small group of slices, in the along-track direction.
as most of the algorithms in question deal with the along-track processing of the SAS data. Let this quality benchmark be $q$, and measuring the contrast over individual along-track slices, $n_r$, will give

$$q_{n_r} = \frac{\sigma_{n_r}}{\mu_{n_r}} = \frac{\sqrt{\frac{1}{N}\sum_{n=1}^{N} (I_{\text{image}}(n,n_r) - \mu_{n_r})^2}}{\sum_{n=1}^{N} I_{\text{image}}(n,n_r) / N}$$

(3.22)

where $\sigma$ and $\mu$ are the corresponding standard deviation and mean respectively, and $N$ is the total number of along-track samples in the slice.

As an example of its implementation, this quality benchmark is tested on two incoherently processed point targets, one of which is (a) deliberately simulated perfectly and another which is (b) processed with a mismatched range-curvature function (Fig.3.17).

**Figure 3.17**: (a) Point target with processed with perfect range-curvature, and (b) point target processed with a deliberately mismatched range-curvature.

The dotted lines over the targets indicate the position of the along-track slices that are chosen for the calculation of $q$, and the three slices correspond to $n_r$ (cross-track range) of 40m, 41m and 42m. They are chosen because they are within an area influenced by
the presence of a target. The standard deviation is calculated for values along the dotted lines and divided by its mean, and the results are as such:

<table>
<thead>
<tr>
<th></th>
<th>( n_r = 40 \text{m} )</th>
<th>( n_r = 41 \text{m} )</th>
<th>( n_r = 42 \text{m} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image (a)</td>
<td>( q_{nr} = 1.59 )</td>
<td>( q_{nr} = 1.75 )</td>
<td>( q_{nr} = 1.77 )</td>
</tr>
<tr>
<td>Image (b)</td>
<td>( q_{nr} = 1.34 )</td>
<td>( q_{nr} = 1.35 )</td>
<td>( q_{nr} = 1.30 )</td>
</tr>
</tbody>
</table>

The higher value of \( q_{nr} \) indicates a stronger structural presence in the along-track slice at \( n_r \). The mismatch in range curvature had caused \( q \) to drop in image (b).

### 3.14 Design Considerations in Incoherent SAS Imaging

It is obvious from the discussions in this chapter that there are many parameters that influence the outcome of the incoherent imaging process. The consideration for the design parameters should start with the definition of the quality that is most desirable, one of which the design of the system would seek to optimise. This is then offset by limited resources related to the cost and availability of transducers, survey time, processing time and the system’s complexity.

In order to have an incoherent imaging system that optimises the image quality, one important factor is the consideration of how many pings per point of interest are needed first, which is based upon the required signal to self-noise ratio. Then, by setting a requirement for the along-track resolution and cross-track resolution, the bandwidth of the transmitted pulse can be decided. The along-track resolution can be easily constrained by the use of highly directional elements, especially if bandwidth is scarce. The inter-ping spacing, which determines the coverage rate of the system, is often limited by the needed number of effective pings per point of interest when given the beamwidth of the physical elements. The coverage rate can be increased at the expense of signal to self-noise ratio and along-track resolution.

An alternative approach would be to set the needed coverage or mapping rate first. Given the beamwidth or directionality of the element, the effective pings per point of
interest can be known. Setting the bandwidth of the transmitted pulse, and thus the cross-track resolution, will give the achievable along-track resolution. The signal to self-noise ratio can be estimated from the available effective pings per point of interest. It can be increased at the expense of mapping rate.

Hence, the overall design paradigm encapsulates the balance between the achievable signals to self-noise ratio plus along track resolution, and the achievable mapping rate, where their mutual effect reduces as the beamwidth of the individual elements increases. Having fixed the point of balance between these two factors, the along-track resolution can be further increased via the increase in cross-track resolution. Weightings need to be applied in cases where the elements are highly directional and yet many elements are needed in the process of incoherent summation.

As a design example, a 50kHz imaging system with bandwidth of 20kHz ($Q=2.5$) is considered. Given that 100 pings are needed per point of interest (to achieve gain of approximately 13dB), the range of interest is 20m-150m, and that 3 transducers with beamwidth of 1.05rads (60degrees) are available, the transducers can be built as an array and their maximum separation can be calculated (Fig. 3.18).
The maximum synthetic aperture length is found to be 11.5m, and hence to meet the requirement of 100 pings per point of interest (such that $N=100$ in along-track integration), the maximum *inter-element* spacing of the transducer is $11.5/100=0.12$m. Because there are 3 transducers the maximum *inter-ping* spacing is $3 \times 0.12=0.36$m. The upper limit of PRF (as set by the cross-track range of interest) is $1500 \text{ms}^{-1}/(2 \times 130 \text{m}) = 5.8$Hz. For the lower PRF limit, choosing a PRF of 5Hz gives a maximum achievable tow speed of $1.13 \text{ms}^{-1}$. Setting a tow-speed of $1 \text{ms}^{-1}$ will then give an inter-ping spacing of 0.23m, and this is acceptable because it does not exceed the maximum inter-ping spacing of 0.36m. The 3 transducers should then be attached 0.08m apart onto the tow platform. The mapping rate in this case is $130 \text{m}^2 \text{s}^{-1}$. The cross-track resolution and along-track resolution is 0.08m and 0.24m respectively.

Designing a coherent system with the same available hardware and complexity, while satisfying the $\lambda/2$ constraint (a bigger constraint than the required $N$ in coherent cases), and pushing the PRF to 5.5Hz, will only provide a maximum tow speed of 0.08ms$^{-1}$, and a mapping rate of $10.4 \text{m}^2 \text{s}^{-1}$. The along-track resolution, however, is 0.02m.

**Figure 3.18:** Maximum synthetic aperture length limited by the beamwidth of the transducer and the minimum range of interest.
Chapter 3: Incoherent Synthetic Aperture Sonar Processing

The computational speed for incoherent processing is at least two times faster than for coherent processing, given the use of the same time-domain algorithm. However, where the system needs to process information close to real-time, the computational requirements still prove to be too demanding. Where shorter processing time is a primary requirement, along-track resolution and signal to self-noise ratio can be traded off by using a smaller number of elements in processing, as well as by the implementation of a depth of focus such that the same range back-projection function is used for points of interests in more than one range.

3.15 Core Advantages and Disadvantages of Incoherent SAS Processing

Having established the key properties of incoherent SAS processing, it is clear that there are some key advantages in incoherent processing in terms of overcoming the limitations present in coherent SAS imaging, and these are summarised against the background of the necessary trade-offs.

The first advantage is the allowance of sparse sampling in the along-track without the occurrence of grating lobes. One sparse sampling scenario is to keep the synthetic aperture length constant while reducing the number of elements, and the other is to keep the number of elements constant while increasing the synthetic aperture length. Sparse sampling can be implemented to the convenience of the surveying process, such as to have a higher mapping rate, shorter survey time, and smaller hardware requirement. It can also be implemented just before the process of incoherent summation on readily available dataset meant for coherent processing by using only selected samples, thus reducing the processing time needed especially if a quick and crude image is needed. The price paid because of the lesser number of elements is a loss of SNR by approximately $6\log(N)$ to $7\log(N)$. This trade-off is illustrated in Fig. 3.19, based on typical $N$ (total number of integration samples) of 100, 1000 and 10000.
The absence of grating lobes means that the arrangement of elements in the synthetic aperture does not have to be linear, nor do they have to be uniformly spaced or symmetrically shaped. This provides a greater degree of freedom in the surveying process, as well as the prospect for a more flexible imaging approach. The synthetic aperture can be practically extended to be 2-dimensional, or even 3-dimensional, which potentially produces images with better cognitive properties. This helps overcome the limitation in traditional imaging with linear array in the “side-scan” manner which suffers from shadows and the lack of accurate object-height representation.

The second highlight of this technique is its tolerance to motion errors. Given the same set of motion errors, the net effect of image degradation is much less in an incoherently processed image. In cases where motion compensation is needed, the motion estimations need only be accurate in the order of half the compressed pulse length, whereas an accuracy of the order of $\lambda/2$ is required for coherent dataset. This greatly increases the robustness of an incoherent SAS system.

![Figure 3.19: Loss in SNR due to sparse sampling, as a function of an undersampling factor. This undersampling factor is just a dividing factor for $N$, such that 1 means that there are no undersampling, and 2 means that the number of samples is reduced by half.](image)
By example, for a 50kHz system (assuming $Q=2.5$), the motion error tolerance is of the order of ±0.015m for a coherent system, whereas for an incoherent system, the tolerance is of the order of ±0.04m. It must be noted here that as the bandwidth of the signals are increased to improve the cross-track/range resolution of the system, the compressed pulse length becomes smaller and hence the tolerance also decreases. If the $Q$ factor is now 1.5 instead of 2.5, the tolerance will drop to ±0.023m. A simulation showing its tolerance is shown in section 3.9.

One crucial disadvantage of using an incoherent SAS system is the inferior along-track resolution. This can be evaluated by comparing the expression of the along-track resolution of a coherent SAS (equation 2.3) and an incoherent SAS (equation 3.17). The comparison of along-track resolution for typical SAS imaging systems is shown in Fig. 3.20.

\[
P_s = \frac{D_p}{2} \quad P_z \approx \frac{3P_rD_p}{\lambda}
\]

<table>
<thead>
<tr>
<th>Coherent SAS</th>
<th>Incoherent SAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_s = \frac{D_p}{2}$</td>
<td>$P_z \equiv \frac{3P_rD_p}{\lambda}$</td>
</tr>
<tr>
<td>25kHz system, $Q=5$, $\theta_b=1.05\text{rad}$</td>
<td>$P_s = 0.030\text{m}$</td>
</tr>
<tr>
<td>50kHz system, $Q=5$, $\theta_b=1.05\text{rad}$</td>
<td>$P_s = 0.014\text{m}$</td>
</tr>
<tr>
<td>50kHz system, $Q=2.5$, $\theta_b=1.05\text{rad}$</td>
<td>$P_s = 0.014\text{m}$</td>
</tr>
</tbody>
</table>

**Figure 3.20:** Comparison of along-track resolution between coherent and incoherent SAS, where $P_s$ is for coherent SAS, $P_z$ is for incoherent SAS, and $\theta_b$ is the transducer’s beamwidth.

With incoherent SAS system, remote SAS imaging is more feasible with autonomous unmanned vehicles (AUV) without a great deal of increase in system and processing complexity. A recent experiment of the viability of coherent SAS imaging by an AUV done in [42] has indicated that the top challenge lies in the obtaining of accurate trajectory information, hence requiring the overlapping of looks between pings and the introduction of targets of opportunity. With incoherent processing, remote SAS imaging is far more feasible, and in good sea conditions (sea-state-0 to sea-state-1),
motion compensation is not needed at all as long as the AUV’s trajectory is inherently straight.

### 3.16 Conclusion

A time domain incoherent processing technique is presented, which encourages the use of broadband pulses, in order to gain improvement in along-track resolution. The technical properties of incoherent SAS imaging have been established with comparisons to coherent processing. The key differences are the achievable along-track resolution, the absence of grating lobes, the tolerance to motion errors and the characteristics of the processed image.

The concept of signal to self-noise ratio is introduced to enable the analysis of the effect of the self-noise level, and the self-noise energy distribution pattern, on the image quality. Two image quality benchmarks, based upon the image’s detection properties and cognitive properties, are defined as the signal to self-noise level and the contrast of the image respectively.

The main driving parameters in the design of incoherent SAS are identified. The optimisation of mapping rate stands against the optimisation of signal to self-noise ratio and along-track resolution. Wide bandwidth and beamwidth of the transducers are factors that allow one to improve the achievable along-track resolution.

Finally, the core advantages of incoherent processing are compared against the trade-offs being made, and that forms the basis and foundation on which the research works presented in the following chapters are based upon.