The Development and Analysis of Scanned Pattern Interferometric Radar

by

Karen Marais

B.Eng., Electrical and Electronical Engineering
University of Stellenbosch, South Africa, 1994
B.Sc., Mathematics
University of South Africa, 1997

SUBMITTED TO THE DEPARTMENT OF
AERONAUTICAL AND ASTRONAUTICAL ENGINEERING
IN PARTIAL FULFILLMENT OF THE DEGREE OF

MASTER OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2001

Signature of Author .....
Department of Aeronautics and Astronautics
August 24, 2001

Certified by .....
Raymond J. Sedwick
Research Scientist in the Department of Aeronautics and Astronautics

Certified by .....
David W. Miller
Director of the MIT Space Systems Laboratory

Accepted by .....
Wallace E. Vander Velde
Professor of Aeronautics and Astronautics
Chair, Committee on Graduate Students

AERO

The Development and Analysis of Scanned Pattern Interferometric Radar

by

Karen Marais

Submitted to the Department of Aeronautics and Astronautics
on August 24, 2001 in Partial Fulfillment of the
Requirements for the Degree of Master of Science
at the Massachusetts Institute of Technology

ABSTRACT

A method of performing space based GMTI using radar interferometric processing is presented. The algorithm, referred to as Scanned Pattern Interferometric Radar (SPIR), uses the high angular variability of a sparse array Point Spread Function (PSF) to collect sufficient data from the signal return that the clutter and targets can be separated without an a priori assumption of the clutter statistics. By deconvolving the PSF from the received signals the true ground scene is revealed. Performance of the algorithm is highly dependent on cluster design. Aperture placement in the cluster determines the (PSF). We show that minimum redundancy arrays are not appropriate for SPIR systems. An alternative design algorithm to select aperture placements that result in good SPIR performance is developed. Numerical solution of SPIR systems by conventional linear algebra techniques is not feasible since the system conditioning is poor. The CLEAN algorithm, developed in the astronomical interferometry field, delivers promising results. Although clutter amplitude is random in nature, its position and doppler shift are geometrically related. An adapted version of CLEAN that uses this information shows improved target recovery. Prior target information can be used to further improve detection of both existing and new targets.

Thesis Supervisor:
Dr. Raymond J. Sedwick
Department of Aeronautics and Astronautics

ACKNOWLEDGMENTS

The opportunity to study at MIT has been the greatest of my life. A world that I did not know existed has been revealed to me. I would like to thank the following people: Raymond Sedwick, Tienie van Schoor, Rade Klaric, Elisabeth Lamassoure, John Enright, Joseph Saleh, and my parents, Sarel and Joan Marais.

This research was carried out under the supervision of Dr. Raymond Sedwick of the Space Systems Laboratory, MIT. I thank him for his guidance and support.

Tienie, you opened the door to MIT to me. Thank you for your unquestioning willingness to help.

Rade, you were unselfish enough to encourage me to embark on this voyage. My new world was your gift and I thank you for it.

Elisabeth, you became my best friend in your quiet and dignified way. I thank you for your support in all my endeavours, both academic and personal. You are wiser than you know and have taught me so much. I wish you all the best in your new life.

John, thank you for always being willing to help. Your insight has been invaluable in the creation of this thesis.

Joe, my debt to you is incalculable. If I have grown at all it is because of you.

Mamma en Pappa, dankie vir alles. Wat ek ook al bereik kom van julle af.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE

AMTI	Airborne moving target indication
A	Area [m^2]
c	Speed of light [m/s]
CFAR	Constant False Alarm Rate
CW	Continuous Wave
d	Diameter [m]
DFT	Discrete Fourier Transform
DPCA	Displaced Phase Centre Antenna
E	Electric field intensity [V]
f	Frequency [Hz]
f_c	Carrier frequency [Hz]
GMTI	Ground moving target indication
g	Gravitational acceleration [m/s^2]
h	Height [m]
i	Current [A]
λ	Carrier wavelength [m]
l	Length [m]
MTI	Moving target indication
PRF	Pulse Repetition Frequency [Hz]
PRI	Pulse Repetition Interval [s]
r	Radius [m]
R_{earth}	Radius of the earth
SPIR	Scanned Pattern Interferometric Radar
sr	Steradian
STAP	Space Time Adaptive Processing
τ	Pulse width [s]

Chapter 1

INTRODUCTION

The recent demise of MIT's Building 20 brought to mind a remarkable era of discovery, much of which took place here at the Massachusetts Institute of Technology. Designed in an afternoon, and erected in just six months in 1943, Building 20 was the heart of the Radiation Laboratory, affectionately known as RadLab. Here, during the frenzied war years, a remarkable collaboration between the British and the Americans, the military and civilians, resulted in discoveries that would eventually impact all aspects of modern life.

Radar, an acronym for RAdio Detection And Ranging, is based on the deceptively simple concept of locating objects using transmitted and reflected high-frequency radio waves. Waiting in the wings for technology to advance, it was seized on by eager governments at its first appearance. The ensuing fifty-odd years have seen an explosion in radar applications, both in the civilian and military worlds.

Ground Moving Target Indication (GMTI) using separated spacecraft interferometry at radio wavelengths promises to be a powerful new all-weather surveillance tool. GMTI radar is frequently implemented on both ground-based and airborne systems, but spaceborne systems are less common. This thesis develops and analyses a novel processing algorithm by which radar data from separated spacecraft can be combined to yield high resolution radar maps.

1.1 Historical Background

In 1873, James Clark Maxwell published his landmark work unifying the equations governing electromagnetic waves, *Treatise on Electricity and Magnetism.* Nearly fifteen years later, Heinrich Hertz conducted a classic series of experiments in which he generated and detected radio waves. He showed that the waves travelled at the speed of light and could be reflected, refracted, diffracted and polarized like visible light, as Maxwell had predicted. The discovery inspired scientists across the world. When Guglielmo Marconi successfully sent radio signals across the Atlantic in 1901 the modern communications industry was born.

Credit for the invention of radar cannot be given to a single person. Hertz's experiments had showed that electromagnetic waves could be bounced off objects. Both Nikolai Tesla and Marconi realised that objects could be detected by bouncing radio signals off them. By patenting a radio echo device for locating objects at sea, a German engineer named Christian Hülsmeyer proclaimed himself the inventor of radar in 1904.

The idea was not practicable at the time however. The technology to generate high frequency signals, with wavelengths shorter than 50 m did not exist. Low frequency signals could not maintain sufficient focus over long distances to be detectable. It took years of technological advancement before radar could come into its own.

Rightly or wrongly known as the father of radar, Sir Robert Watson-Watt and his junior scientific officer Arnold F. "Skip" Wilkins proposed the concept of using radio waves to detect aircraft to the British government in 1935. The Air Ministry had offered a £1000 reward for anyone who could design a "death" ray that could kill a sheep at a hundred yards. In his capacity as superintendent of the Radio Department of the National Physics Laboratory, Watson-Watt quickly determined that the idea was not feasible at the time. He did come up with an alternative proposal however:

> "Meanwhile attention is being turned to the still difficult but less unpromising problem of radio-detection as opposed to radio-destruction, and

> numerical considerations of the method of detection by radar waves will be submitted when required."

Nervous about the growing threat from Hitler's Germany, and with no means of protecting Britain against an air attack, the British government pounced on Watson-Watt's idea, requesting the promised facts and figures. Watson-Watt turned to Wilkins to perform the detailed calculations. The results were so inspiring that an excited Watson-Watt immediately drafted a memo that would become a landmark in the history of radar, "Detection of Aircraft by Radio Methods" on February 12, 1935.

Following Watson-Watt's proposals, the Chain Home network was created along Britain's east and south coasts. It was a system of antennas that could detect aircraft up to 150 miles away. During the London Blitz the system provided early warning of incoming bombing raids, saving countless lives.

British scientists aggressively experimented with shorter wavelengths, narrower beams, more compact equipment, and greater power generation to improve their radar capability. In 1940 Watson-Watt, together with John Randall and Henry Boot, invented the cavity magnetron. The magnetron was a compact source of high frequency radio waves that could be used over distances on the order of tens of kilometres to detect small, airborne targets.

In August 1940, the British government dispatched the top-secret Tizard Mission to the United States to exchange information on radar. In a pivotal meeting the British revealed their resonant cavity magnetron. The magnetron, an efficient, high-power (10-kilowatt) pulsed oscillator that operated at 10 centimetre wavelengths, was the seed for the British-American collaboration, centered at the newly created MIT Radiation Laboratory.

The laboratory's first three projects focussed on protection of Britain. Project I created a 10 cm airborne intercept radar that would enable bombers to detect enemy aircraft at night. Project II developed the highly successful SCR-584 gun-laying radar, which is credited with leading to the safe destruction of 85% of the V-1 buzz bombs that were dropped

on London. Project III developed a long range navigation system for ships and aircraft, known as LORAN (LOng RAnge Navigation), which is still in use today.

Building 20 began as a collaboration between Britain and the United States to make microwave radars. It rapidly developed into a centralised laboratory, which laid much of the theoretical radar groundwork while simultaneously creating working systems for use in the Allied war effort. In the years from 1940 to 1945 RadLab designed over one hundred different radar systems, which comprised close to half of the radar systems deployed in World War II, and had a monetary value of $1.5 billion.

Radar has not failed to deliver on its promise. Since the heady days of RadLab radar has been adapted for many purposes. These include high-resolution mapping of ocean levels, freeway speed enforcement, air-traffic control and many military applications.

Spaceborne radar was initially used as an aid for space rendezvous, on the Gemini spacecraft. It was not long before radar antennas were turned towards Earth and in present day the primary application of space-based radar has become observation. Spaceborne radar permits high resolution and high fidelity observation of the entire surface of Earth regardless of weather conditions. It is not surprising that a wide range of applications have been found in both the military and civilian worlds. Applications include altimetry, ocean level observation, weather observation, terrain mapping, and navigation.

The interested reader is invited to consult Robert Buderi's fascinating account of the development of radar at the Radiation Laboratory, *The Invention that Changed the World* [Buderi, 1996].

1.2 Spaceborne Radar

GMTI is typically used to locate large vehicles such as tanks and trucks in military applications. Current systems, such as JSTARS (Joint Surveillance Target Attack Radar System, a US Air Force Research Project), use airborne platforms to provide temporary

coverage in regions of interest. These systems provide good performance but require large support infrastructures and significant deployment time and require pilots to fly in active theatres. Spaceborne radar can offer nearly continuous global coverage without appreciable operator risk.

As discussed by Shaw [Shaw, 1998], there are three fundamental design considerations for spaceborne radar. First, the radar must have sufficient power-aperture product to detect targets at the required search rate. Second, the angular and range resolution must be high enough to locate the target with the required degree of accuracy. Third, the radar must reject clutter and noise sufficiently to provide the specified probability of detection and false alarm rate. Thus, development of an effective spaceborne radar system presents many challenges.

In this thesis we discuss the design and implementation of a spaceborne GMTI interferometric radar system. We address high-level system design aspects and describe a novel processing algorithm by which images can be constructed.

1.3 Multiple Aperture Systems

Single aperture systems are a compromise between high resolution and broad coverage. Broad coverage, or large field of view, requires a wide beamwidth, while high resolution requires a narrow beamwidth.

Unwanted return from the ground, known as clutter, presents an additional challenge. Clutter suppression is computationally intensive and requires prior knowledge of the target as well as an appropriate clutter model. Performance is therefore limited by the quality of prior target information and the accuracy of the clutter model.

In contrast, multiple aperture systems can provide coverage corresponding to the diameter of a single aperture and resolution proportional to the maximum separation between individual apertures.

Space-Time Adaptive Processing (STAP) has been proposed for use in multiple aperture systems [Ward, 1994]. STAP takes advantage of the high resolution properties of the interferometer, but does not use all the information available in the return signals. Like single aperture methods, STAP is limited by the accuracy of the clutter model. Good prior estimates of the position and velocity of any targets are required. Careful maintenance of aperture positions in the cluster is necessary to achieve an adequate level of clutter suppression.

TECHSAT21 is a technology demonstrator program with an experimental multiple aperture radar payload. It consists of a cluster of microsatellites (less than 100 kg) that orbit in close proximity (on the order of hundreds of metres). Each microsatellite is capable of coherently detecting not only return signals resulting from its own transmission, but also return signals resulting from the transmissions of other satellites in the cluster.

Scanned Pattern Interferometric Radar (SPIR) combines the individual aperture signals in a way that allows ground clutter to be characterized completely and removed. This method depends not on the statistics of the clutter, but solely on the clutter position and doppler shift. While the clutter amplitude and phase are random in nature, the doppler shift is entirely predictable due to the known angular location of the clutter footprint. The algorithm relies on "deconvolving" the gain pattern of the synthesized aperture from the received signals to reveal the true ground scene.

In conventional interferometry, deconvolution is usually performed entirely in the spatial domain, and removes the artefacts that appear in the image as a result of incomplete (u, v) plane filling. SPIR processing uses the distinctive features of the interferometer response, and their effect on both spatial and frequency information, to separate ground clutter return from target return. The statistical nature of the clutter is irrelevant and the probability of detection is based solely on the signal-to-noise ratio of the received signals.

1.4 Outline

This thesis develops and analyses the Scanned Pattern Interferometric Radar method, beginning with the initial intuitive concept and concluding with a simulated implementation.

A description of the process begins with some fundamentals of radar and interferometry in Chapter 2. Chapter 3 describes SPIR and gives a mathematical derivation. Optimum placement of apertures within the cluster is discussed in Chapter 4. Solution of SPIR systems is not trivial, Chapter 5 examines this problem. Chapter 6 gives conclusions and recommends future work.

Chapter 2

RADAR INTERFEROMETRY

Scanned Pattern Interferometric Radar (SPIR) borrows techniques from the field of interferometry to create high-resolution radar systems. Understanding radar interferometry therefore requires an appreciation for both radar and interferometric imaging theory. To that end, this chapter discusses background material. We begin with a short discussion of some radar concepts that are necessary to understand Scanned Pattern Interferometric Radar (SPIR). Fourier analysis is a powerful tool of which we make much extensive use in SPIR. A summary of useful Fourier techniques is provided. Finally, the last two sections discuss interferometry techniques, with specific reference to radar applications.

2.1 Radar Concepts

In this section we introduce some concepts that will aid in the understanding of SPIR. Readers unfamiliar with radar are referred to [Stimson, 1998] for an excellent introduction.

2.1.1 Pulse-Doppler Radar

Radar has many applications, from ground mapping to altimetry to moving target indication. Ground mapping radar maps stationary objects and may have very high resolution on the order of millimetres. Altimetry measures the altitude of air- or spaceborne objects. Moving Target Indication (MTI) radar searches for targets that may be on the ground or air

(space) borne. MTI radar has an additional source of information on targets, namely their closing velocity. Most MTI radars are of a specific type referred to as pulse-doppler radar.

There are two general types of radar transmitter, Continuous Wave (CW) and pulsed. CW radar continuously transmits and receives. The main problem with air- and spaceborne CW radars is electrical noise, or interference. The limited available space makes it difficult to physically isolate the transmitter from the receiver. As a result noise arising from generation of the high energy transmitted signal corrupts the received echoes, preventing accurate radar mapping. Pulsed radar eliminates this problem by transmitting short pulses and listening for the reflected echoes in the intervals between transmissions. Pulsed radars that sense doppler frequency are called pulse-doppler radars, those which only sense range are simply called pulse radars.

The Pulsed Waveform

The pulsed waveform has four basic characteristics: carrier frequency, f_c [Hz], pulse width, τ [s], interpulse modulation, and pulse repetition frequency, PRF [Hz].

The carrier frequency is usually chosen based on physical size limitations, available transmitted power, and atmospheric attenuation characteristics. In the case of air- and spaceborne systems the choice of carrier frequency is dominated by the size limitations inherent on a satellite. As frequency is increased, RF analog processing hardware dimensions decrease and the antenna diameter required to provide a given beamwidth decreases. The maximum beamwidth should be comparable with the expected region of interest; this ensures that transmission energy is not "wasted" outside the region of interest. The X-band ($10GHz$) is popular for space-based radar applications since radar hardware dimensions are compatible with satellite size limitations while atmospheric attenuation is still reasonably low.

The pulse width determines the highest possible resolution of the radar in the range direction. Range resolution is defined as the minimum range between two objects at which they can be resolved.

The extent of a pulse in space is

$$L = c\tau \tag{2.1}$$

Objects that are more closely spaced than $L/2$ cannot be resolved. Increasing the resolution therefore requires decreasing the pulse length. Decreasing the pulse length decreases the energy contained in the pulse. Since the peak power that a radar can transmit is limited by size and mass constraints, very narrow pulses will not contain enough energy to allow target detection.

Pulse width therefore seems to place a limit on the range resolution. However, if the effective pulse width is decreased while the actual pulse is kept sufficiently wide to ensure target detection, good range resolution can be obtained. This can be achieved by encoding the transmitted pulse in frequency and/or phase so that the received pulse can be compressed before target detection is performed. Two popular techniques are chirp compression, where the frequency of the carrier is linearly increased over the length of the pulse, and binary phase modulation, where successive *chips* of the pulse have their phase changed by $0°$ or $180°$.

The pulse repetition frequency (PRF) is the rate at which the pulses are transmitted. The inverse of the PRF is the pulse repetition interval (PRI). The choice of PRF determines to what extent the ranges and doppler frequencies sensed by the radar will be ambiguous, as discussed in the next two sections.

The probability of detection of a target is directly related to the amount of signal energy that illuminates it. In addition to using pulse compression to increases the individual pulse energy, the received echoes from several successive pulses are usually summed. This process is referred to as *pulse integration.* The number of pulses integrated is selected to be sufficiently low that there is no appreciable change in the targets' positions and velocities during the integration time.

Pulse Integration

There are two ways of integrating the signal returns from multiple pulses: coherent and non-coherent integration.

Coherent integration is performed on the signal before it has been stripped of its carrier, hence phase is preserved. Non-coherent integration is performed on the modulating signal (envelope), which does not carry phase information.

If n pulse are coherently integrated, in the ideal case the resulting signal to noise power ratio would be n times that of a single pulse. Non-coherent integration always has a gain less than n, due to information loss in the envelope detection process.

However, since non-coherent integration is easier to implement, it is more popular.

Thresholding

The process of deciding whether a given signal return corresponds to a target or noise is called thresholding. In essence, signals that are greater than a certain level, or threshold, are declared to be targets, whereas smaller signals are assumed to be noise. The selection of the threshold is therefore crucial. If the threshold is too high, small targets may be missed. If the threshold is too low, noise spikes may be mistaken for targets.

The threshold directly affects the probability of detection (P_d) as well as the probability of false alarm (PFA). An ideal system would have a high P_d and a low PFA. Unfortunately these requirements conflict with each other. Current systems settle for a constant false alarm rate (CFAR), which depends on the required P_d. CFAR processing is complex and the literature abounds with references. [Skolnik, 1980] is a good starting point.

Measuring Range

Range measurement with pulsed radar is simple. The range to the target can be determined by measuring the time interval t between transmission of the pulse and its reception

$$range = \frac{ct}{2} \tag{2.2}$$

where c is the speed of light.

This method of ranging works as long as the round trip time for the pulse is less than the interpulse period. If the radar detects a target whose round-trip time is larger than the interpulse period, the echoed pulse will appear to come from a subsequent pulse, as shown in Figure 2.1. The resulting calculated range is ambiguous since it is impossible to determine which of the transmitted pulses it corresponds to.

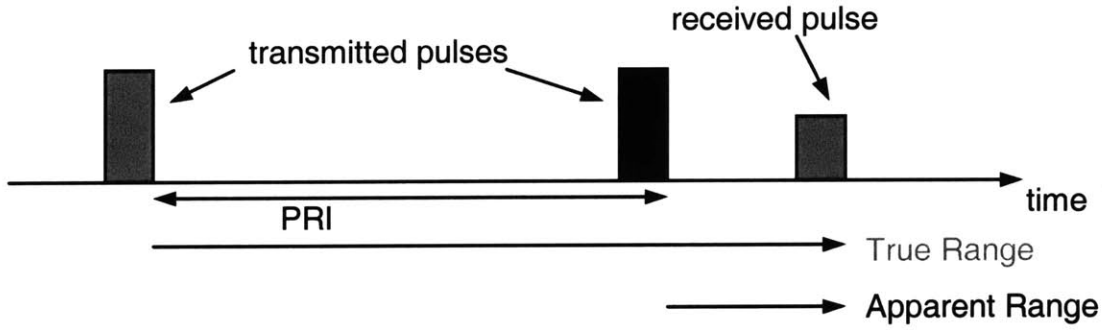

Figure 2.1 Range Ambiguity

The maximum range from which return may be received without the observed range being ambiguous is called the unambiguous range R_u

$$R_u = \frac{c}{2 \cdot PRF} \tag{2.3}$$

At high PRF's measured ranges are likely to be ambiguous, as many targets may fall outside R_u. Range ambiguities can be avoided by setting R_u greater than the maximum expected target range. However, at low PRF's doppler ambiguities appear, as discussed in the next section.

The Doppler Effect

The doppler effect is the shift in frequency experienced by any type of wave that is radiated, reflected, or received by a moving object. It is commonly observed when standing on the side of a road listening to a car passing at high speed. The pitch of the sound is higher when the car is approaching than when it is receding. When objects approach one another the doppler shift is positive, and when they are receding from each other the doppler shift is negative.

In radar the doppler effect is used to determine the relative velocity between the radar antenna and the object(s) it is tracking. The doppler frequency shift is related to the closing velocity as follows

$$f_{doppler} = 2\frac{v_{closing}}{\lambda} \tag{2.4}$$

where λ is the wavelength of the carrier. Given the radar platform's velocity, the target's velocity along the line of sight (LOS) to the radar can be determined. In spaceborne radar three relative motions contribute to the observed doppler shift: the Earth's rotation, the satellites' orbit and the object's motion.

Doppler ambiguities occur when the PRF is low. Since the radar pulses are the observed scene at the PRF, the frequency spectrum of the radar return is repeated at the PRF. Targets with doppler frequencies greater than the PRF will be undersampled and appear to have lower doppler frequencies, as shown in Figure 2.2.

At low PRFs measured doppler frequencies are likely to be ambiguous, since target doppler shifts may be greater than the PRF. Doppler ambiguities can be avoided by setting the PRF higher than the expected doppler shift bandwidth. However, as we have seen in the previous section, this introduces range ambiguities.

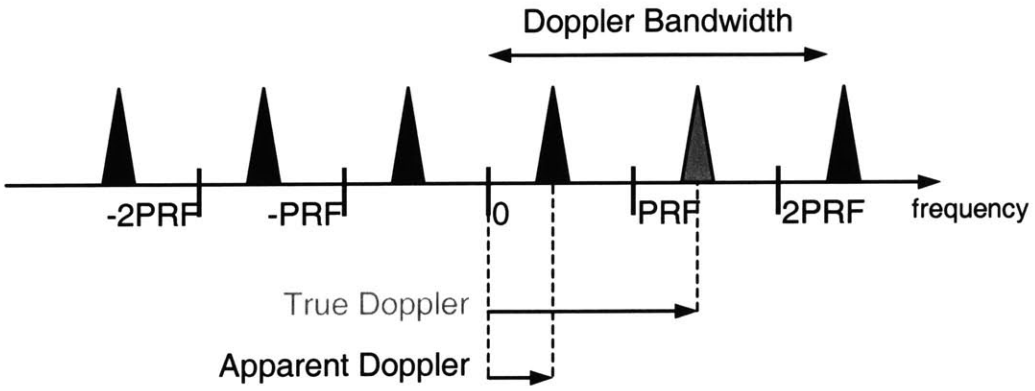

Figure 2.2 Doppler Ambiguity

Range and doppler ambiguities can be alleviated by varying the pulse repetition frequency (PRF), this is known as PRF switching or jittering. Refer to [Stimson, 1998] for a detailed description of this technique.

Range Gating and Doppler Filtering

Digitally sampling the received radar signal allows the use of sophisticated processing algorithms. Therefore, in modern radars most processing is performed on a digitally sampled version of the signal. Although analog versions of some processes described here were used in the past, this discussion is restricted to the digital processes.

The first steps in processing are range gating and doppler filtering.

Range gating is the process of assigning (ambiguous) ranges to the received signal samples. Ideally, each signal sample corresponds to a particular ambiguous range, or range gate. However, individual signal samples are likely to be corrupted by noise. Therefore, a range gate will usually straddle several samples and average their contributions in some manner. By overlapping the range gates, as shown in Figure 2.3, range resolution can be maintained.

The next step is to determine the doppler frequency content of each range gate. The process is referred to as the *corner-turn FFT* (fast Fourier transform) and is illustrated in

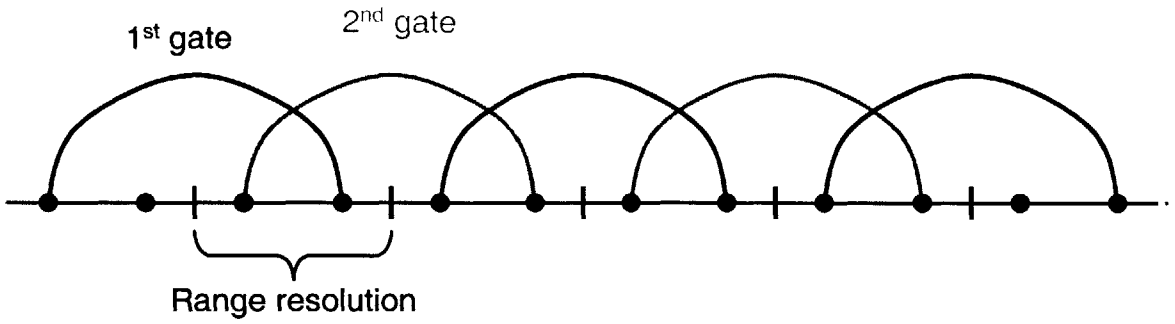

Figure 2.3 Range Gating

Figure 2.4. The FFT is a computationally efficient implementation of the Discrete Fourier Transform (DFT). Returns from successive pulses are placed below each other and an FFT is taken for each range gate. The sampling rate for the FFT is therefore equal to the PRF.

2.1.2 Radar Cross Section

The radar cross section (RCS) of an object is an indication of the amount of reflected energy the radar will receive from that object. It is commonly indicated by σ and is a function of the geometric cross section, reflectivity and directivity of the object

$$\sigma = \left(\begin{array}{c} Geometric \\ Cross \\ Section \end{array} \right) \times (Reflectivity) \times (Directivity) \tag{2.5}$$

The *geometric cross section*, A_{object}, is the area of the object as seen from the radar. This area determines how much signal power the object will intercept

$$P_{intercepted} = P_{transmitted} \times A_{object} \tag{2.6}$$

The *reflectivity* indicates how much of the intercepted signal power the object reflects, or scatters

$$Reflectivity = \frac{P_{scatter}}{P_{intercepted}} \tag{2.7}$$

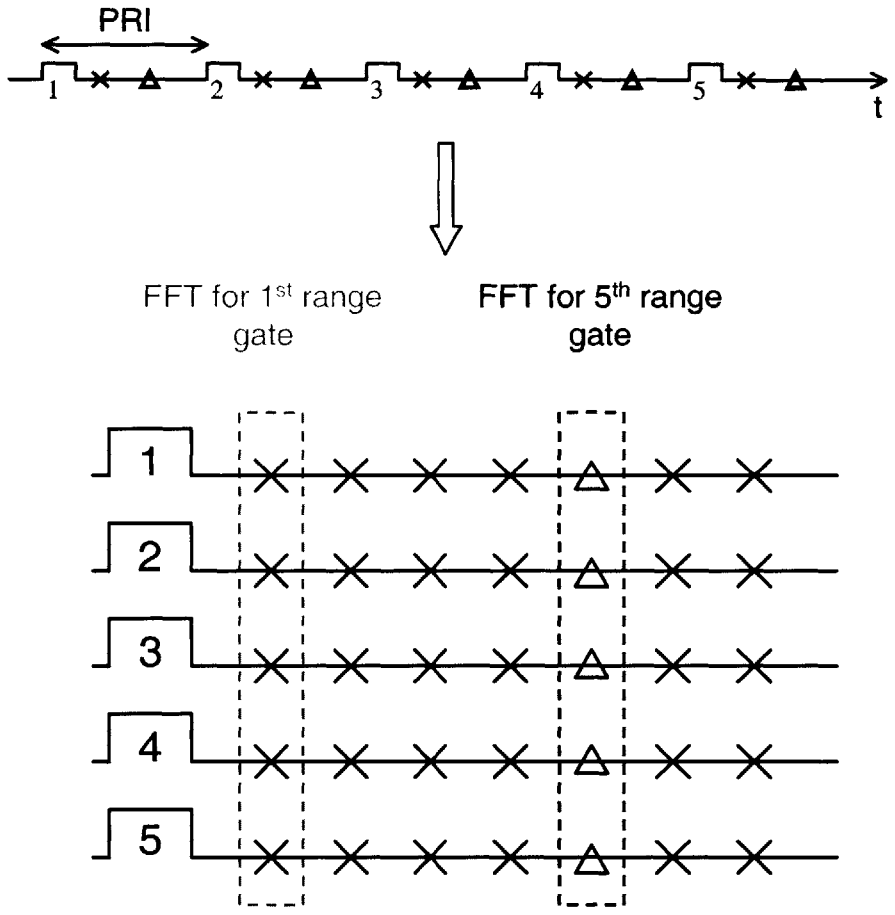

Figure 2.4 The Corner-Turn FFT

The *directivity* is the ratio of the power scattered back in the radar's direction to the power that would have been scattered isotropically

$$Directivity = \frac{P_{backscatter}}{P_{isotropic}} \tag{2.8}$$

Normally $P_{backscatter}$ and $P_{isotropic}$ are expressed as power per unit solid angle (steradian). $P_{isotropic}$ is then equal to the total scattered power divided by the number of steradians in a sphere

$$P_{isotropic} = \frac{P_{scatter}}{4\pi} \tag{2.9}$$

The directivity can therefore be expressed as

$$Directivity = \frac{P_{backscatter}}{(1/4\pi)P_{scatter}} \tag{2.10}$$

In the case of reflection from the ground, the ground is characterised in terms of the incremental backscattering coefficient, σ_0. It is the radar cross section of a unit area of ground.

2.1.3 The Radar Equation

The radar equation predicts the detection range of a radar based on the characteristics of the radar, target and environment. One form of the equation is

$$P_{received} = \frac{P_{average} G \sigma A_e t_{int}}{(4\pi)^2 R^4 L} \tag{2.11}$$

where

$$\begin{aligned} P_{received} &= \text{Received power} \\ P_{average} &= \text{Average transmitted power} \\ G &= \text{Antenna gain} \\ \sigma &= \text{Target RCS} \\ A_e &= \text{Effective antenna area} \\ t_{int} &= \text{Integration time} \\ R &= \text{Range} \\ L &= \text{Loss factor} \end{aligned} \tag{2.12}$$

The received power decreases as range to the fourth. This means that at long ranges the received power will be significantly lower than the transmitted power. The probability that a target will be detected is directly related to the received power. For a given target and range the probability of detection can be improved by increasing the transmitted power, integration time or effective antenna area, or by decreasing the system losses L.

The integration time is the time that the radar illuminates the footprint. As the integration time is increased, the signal to noise ratio of the integrated signal is increased. Increased dwell time also results in higher frequency resolution, which aids in clutter rejection. However, the integration time must be sufficiently short that the relative velocity between the cluster and the footprint does not change appreciably, as this decreases the accuracy of the doppler measurements. The integration time directly affects probability of detection, false alarm rate and doppler frequency measurement accuracy and should be carefully chosen to obtain the desired system performance.

The design of spaceborne Earth-observing radars presents additional challenges since the available transmission power is limited and the received signal is severely attenuated by the large range. Excellent clutter and noise rejection characteristics are essential for successful systems.

2.1.4 Radar Clutter

Any unwanted radar return is referred to as clutter. In a ground-mapping radar, return from the ground is desired, and is used to construct the ground map. However, in the case of moving target indication (MTI) radar, return from the ground obscures, or *clutters*, the desired target return.

Ground clutter is commonly classified into three categories: main lobe return, side lobe return, and altitude return. Main lobe and side lobe return is collected by the radar antenna's main and side lobes respectively. Altitude return is side lobe return received from directly beneath the antenna and appears at a range corresponding to the altitude of the antenna above the ground. Figure 2.5 shows a typical ground clutter profile seen by an airborne or spaceborne radar, as a function of doppler frequency. Main lobe clutter has a higher average amplitude than side lobe clutter, since the main lobe antenna gain is higher. Altitude return has a high average amplitude since it experiences the least scattering.The doppler shift experienced by clutter from a specific patch on the ground can be predicted based on the geometric relationship between the ground patch and the antenna.

$$f_{doppler} = \frac{2v_{radar}\cos\theta}{\lambda} \tag{2.13}$$

where θ is the angle between the antenna velocity vector v_{radar} and the ground patch, referred to as the *look angle*. Altitude return is centered about $\theta = 90°$ and hence has close to zero doppler shift. The maximum theoretical doppler shift occurs at $\theta = 0°$, which corresponds to looking straight ahead at infinite range. Similarly, the maximum theoretical doppler shift in the negative direction occurs at $\theta = 180°$, which corresponds to looking straight back at infinite range. The loci of constant doppler shift, or isodopplers, are formed by the intersections of cones centered about the antenna velocity vector with the ground.

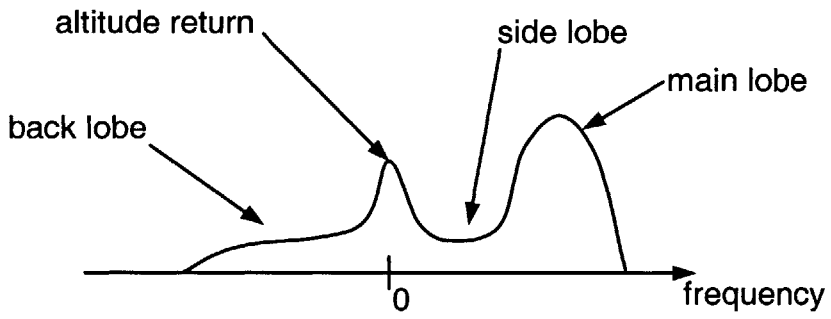

Figure 2.5 Ground Clutter Observed by a Forward-Looking Air- or Spaceborne Radar Platform

The amplitude of the clutter at any one time is more difficult to predict. It depends on the angle of incidence, frequency, polarisation, electrical characteristics of the ground, roughness of the terrain, and the nature of the objects on it [Stimson, 1998]. The creation of clutter models for different types of ground terrain is an entire field of study. Many references exist, (for example [Skolnik, 1980]).

Some radar interferometric methods, such as STAP, use clutter models when recovering targets. Such methods are limited by the applicability and accuracy of the particular clutter model used.

2.2 Fourier Analysis

The key to understanding interferometry is a solid background in Fourier theory. This section presents the Fourier transform and some useful transform pairs.

2.2.1 The Time Domain Fourier Transform

The time domain Fourier transform is a common tool in signal processing. It is used to convert from time to its inverse, frequency. Certain signal characteristics are more readily discernible in the frequency domain than the time domain and vice versa. Operations that are computationally intensive in the time domain, such as convolution, are often less demanding in the frequency domain.

A time domain signal $s(t)$ and corresponding frequency domain representation $S(f)$ are related by the one-dimensional Fourier transform and its inverse as follows:

$$S(f) = F[s(t)] = \int_{-\infty}^{\infty} s(t)e^{-j2\pi ft}dt \tag{2.14}$$

$$s(t) = F^{-1}[S(f)] = \int_{-\infty}^{\infty} S(f)e^{j2\pi ft}df \tag{2.15}$$

We may also take the Fourier transform of sampled signals, this is known as the Discrete Fourier Transform (DFT)

$$S[k] = \sum_{n=0}^{N-1} s[n]e^{-j2\pi\frac{kn}{N}} \qquad k = 0, 1, \ldots, N-1 \tag{2.16}$$

$$s[n] = \frac{1}{N}\sum_{n=0}^{N-1} S[k]e^{-j2\pi\frac{kn}{N}} \qquad n = 0, 1, \ldots, N-1 \tag{2.17}$$

For convenience the following notation will be used to denote the Fourier transform.

$$\begin{aligned} s(t) &\leftrightarrow S(f) \\ s[n] &\leftrightarrow S[k] \end{aligned} \tag{2.18}$$

Time and frequency are not the only one-dimensional Fourier pair; other examples are aperture dimension and radiation pattern in radiation theory, and brightness and visibility in image processing.

2.2.2 The Spatial Fourier Transform

The definition of the Fourier transform is easily extended to an arbitrary number of dimensions. If $\mathbf{p}$ and $\mathbf{q}$ are vectors in n dimensions

$$\begin{aligned} \mathbf{p} &= [p_1, p_2, \ldots, p_n] \\ \mathbf{q} &= [q_1, q_2, \ldots, p_n] \end{aligned} \tag{2.19}$$

and $s(\mathbf{p})$ is a scalar function of $\mathbf{p}$, then $S(\mathbf{q})$ is the n-dimensional Fourier transform of $s(\mathbf{p})$

$$S(\mathbf{q}) = F[s(\mathbf{p})] = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} s(\mathbf{p}) e^{-j2\pi(\mathbf{p} \cdot \mathbf{q})} d\mathbf{p} \tag{2.20}$$

The inverse transform is

$$s(\mathbf{p}) = F^{-1}[S(\mathbf{q})] = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} S(\mathbf{q}) e^{j2\pi(\mathbf{p} \cdot \mathbf{q})} d\mathbf{q} \tag{2.21}$$

Of particular use is the two-dimensional Fourier transform as applied to image processing. Given spatial variables x and y, the Fourier transforms are

$$S(u, v) = F[s(x, y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(x, y) e^{-j2\pi(ux + vy)} dx dy \tag{2.22}$$

and

$$s(x, y) = F^{-1}[S(u, v)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(u, v) e^{j2\pi(ux + vy)} du dv \tag{2.23}$$

What are the quantities (u, v)? When applied to time domain signals the Fourier transform relates time to temporal frequency, measured in cycles per second. Similarly, in the

spatial domain the Fourier transform relates length in a particular direction to spatial frequency in that direction, measured in cycles per metre.

2.2.3 Aperture Response

Another useful application of the Fourier transform is in the field of antenna theory. Given a two-dimensional aperture S' with source illumination $E(x', y')$ and carrier wavelength λ, the electric field away from the aperture can be approximated as [Ramo et al., 1984]

$$E(x, y, z) = \frac{je^{-j\frac{2\pi}{\lambda}r}}{\lambda r}\int_{S'} E(x', y')e^{j\frac{2\pi}{\lambda r}(xx' + yy')}dx'dy' \tag{2.24}$$

This solution is valid for small angles from boresight in the Fraunhofer region. The first term accounts for the variation in electric field with range. The integral can be recognised as the two-dimensional Fourier transform of the aperture geometry. In this case we transform from *source* space to *radiation* space. The Fourier pair in this case is

$$\frac{x}{\lambda} \leftrightarrow \sin\theta \tag{2.25}$$

When describing aperture response it is usual to disregard the variation with range and evaluate only the integral portion of (2.24). A change of coordinates may be useful; spherical coordinates are often used. Two commonly used aperture distributions are given below [Ramo et al., 1984]:

For a uniformly illuminated rectangular aperture of width D, $E(\theta)$ is

$$E_{rectangular}(\theta) = \frac{\pi D}{\lambda} \cdot \frac{\sin\left(\frac{\pi D \sin\theta}{\lambda}\right)}{\frac{\pi D \sin\theta}{\lambda}} \tag{2.26}$$

For a uniformly illuminated circular aperture of diameter D, $E(\theta)$ is

$$E_{circular}(\theta) = \frac{\pi D(1+\cos\theta)}{\lambda} \cdot \frac{J_1\left(\frac{\pi D \sin\theta}{\lambda}\right)}{\frac{\pi D \sin\theta}{\lambda}} \tag{2.27}$$

The aperture response can also be interpreted as the response of an aperture to a point source, that is, the impulse response of the aperture. In image-processing the absolute value of the aperture response is often referred to as the *Point Spread Function* (PSF). Points that are closer together than half the angular separation of the first nulls of the PSF cannot be identified as separate points. This defines the resolution of the aperture. Energy that is located in the side lobes of the PSF overlays energy in the main lobe, so that the reconstructed image is smeared.

2.2.4 Brightness and Intensity

The *intensity,* I_v, of a radiation source is the radiated energy flow per unit area, per unit time, per unit frequency bandwidth, and per unit solid angle.

$$[I_v] = \frac{W}{m^2 \cdot Hz \cdot sr} \tag{2.28}$$

In radio astronomy the term *brightness* is usually used instead of intensity. Brightness (B) is the intensity over a certain field of view and is commonly measured as a function of angle. A brightness map is a standard image.

The visibility, V, of a radiation source is the spatial frequency content of the image and is measured as a function of spatial frequency.

Brightness and visibility are a Fourier transform pair. In the one-dimensional case where θ is small the quantities are related as follows

$$V(u) = F^{-1}[B(\theta)] = \int_{-\infty}^{\infty} B(\theta) e^{-j2\pi u\theta} d\theta$$
$$B(\theta) = F[V(u)] = \int_{-\infty}^{\infty} V(u) e^{j2\pi u \sin\theta} du \tag{2.29}$$

In the two-dimensional case the quantities are related by the two-dimensional Fourier transform

$$B(\xi, \eta) \leftrightarrow V(u, v) \tag{2.30}$$

where $\xi = \sin\theta$ and $\eta = \sin\phi$.

In general, measurements of brightness and visibility are discrete, in this case the DFT and inverse DFT (equations 2.16 and 2.17) are used.

2.3 Fundamentals of Interferometry

The resolution that can be obtained using a single aperture is closely related to its size. Points that are simultaneously covered by the main lobe cannot be differentiated. Resolution is therefore dependent on the width of the main lobe, or the beamwidth. For example, the 3-dB beamwidth of a uniformly illuminated circular aperture is approximately [Stimson, 1998]

$$\theta_{3dB} = 1.02\frac{\lambda}{D} \tag{2.31}$$

where λ is the carrier wavelength and D is the aperture diameter. The larger the aperture the higher the resolution.

The field-of-view of an aperture indicates how large the area observed by the aperture is. It is equivalent to the individual aperture beamwidth. In ground-mapping applications the area where the field-of-view and the ground intersect is usually referred to as the footprint. As the field-of-view is increased the footprint size is also increased.

Broad coverage, or large field-of-view, requires a wide beamwidth, while high resolution requires a narrow beamwidth. Single aperture systems are therefore a compromise between high resolution and broad coverage.

In contrast, multiple aperture systems can provide coverage corresponding to the diameter of a single aperture and resolution proportional to the maximum separation between individual apertures.

We begin by examining the response of a single slit aperture and then comparing it to the response of the simplest interferometer, consisting of two point apertures. This illustrates the key concepts. Next we introduce the point spread function of an interferometer and examine its properties.

2.3.1 Single Slit Aperture

Consider a uniformly illuminated one-dimensional aperture, formed by a slit of width d, as shown in Figure 2.6.

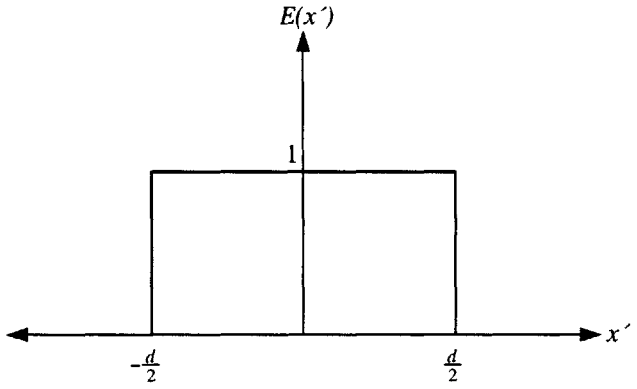

Figure 2.6 Single Slit Amplitude Distribution

The amplitude distribution across the one-dimensional slit is

$$E(x') = \begin{cases} 1 & |x'| \le \frac{d}{2} \\ 0 & |x'| > \frac{d}{2} \end{cases} \tag{2.32}$$

Applying the integral part of (2.24) to the source distribution, the point spread function (PSF) is

$$(x) = \left| d \frac{\sin \pi \frac{dx}{\lambda r}}{\pi \frac{dx}{\lambda r}} \right| \tag{2.33}$$

The PSF is more conveniently expressed as a function of angle from boresight, θ

$$\sin\theta = \frac{x}{r} \tag{2.34}$$

so that

$$(\theta) = \left| d \frac{\sin\left(\pi \frac{d}{\lambda} \sin\theta\right)}{\pi \frac{d}{\lambda} \sin\theta} \right| \tag{2.35}$$

Figure 2.7 shows the PSF for $d = 0.1\,m$ and $\lambda = 0.03\,m$

The PSF has a clearly discernible main lobe and several side lobes that are much smaller in amplitude. The null-to-null beamwidth of the main lobe is

$$\theta_{nn} = 2\,\text{asin}\frac{\lambda}{d} \tag{2.36}$$

With the given parameters, $\theta_{nn} \approx 35°$.

Points that are angularly separated by $\theta_{nn}/2$ or more can be discerned from one another. The resolution of the slit is therefore

$$\theta_{resolution} = \text{asin}\frac{\lambda}{d} \tag{2.37}$$

The beamwidth decreases as the aperture size is increased relative to the carrier wavelength. As mentioned earlier, the width of the main lobe determines the maximum achievable resolution. The side lobes bring ghost images in from outside the main lobe. These ghosts corrupt the constructed image; the lower the relative side lobe level the less effect they will have on the image. Clear, high-resolution images can be obtained if the PSF has a narrow main lobe and low side lobes. Unfortunately, one attribute generally has to be traded with the other.

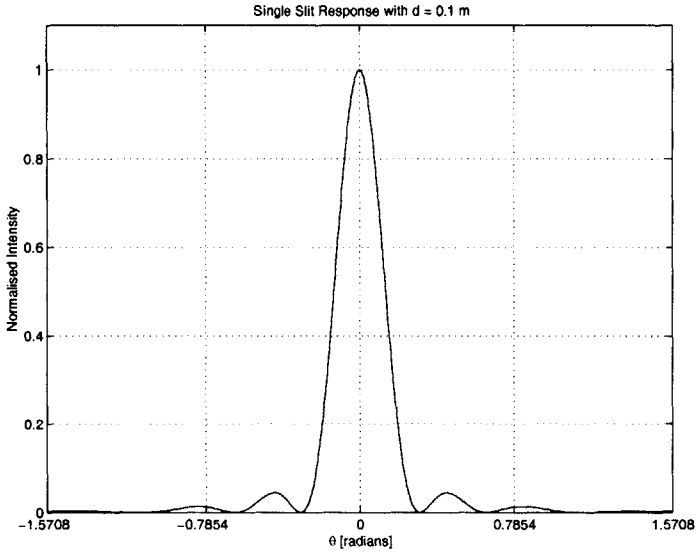

Figure 2.7 Single Slit Aperture Point Spread Function

2.3.2 Single Baseline Sparse Array

The single slit aperture can be thought of as an array of an infinite number of pinhole apertures. The simplest interferometer consists of two pinhole apertures. Let us examine how well an array of two pinholes located a distance d apart, as shown in Figure 2.8, approximates the single slit aperture.

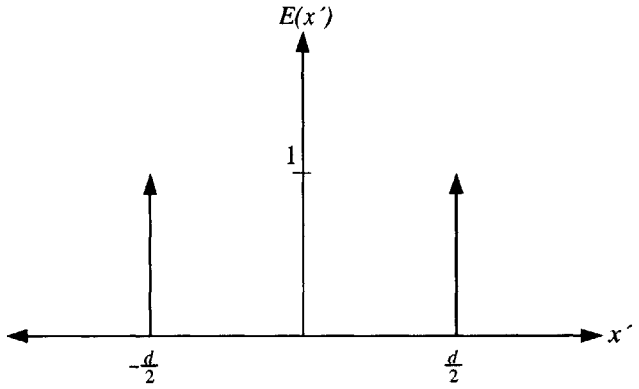

Figure 2.8 Two Pinhole Amplitude Distribution

The amplitude distribution is given by

$$E(x') = \delta(x' + \frac{d}{2}) + \delta(x' - \frac{d}{2}) \tag{2.38}$$

The distance between the two pinholes is referred to as the *baseline*.

Applying the integral part of (2.24) to the source distribution, the PSF is

$$(x) = \left|2\cos\left(\pi\frac{dx}{\lambda r}\right)\right| \tag{2.39}$$

Expressed as a function of angle from boresight

$$(\theta) = \left|2\cos\left(\pi\frac{d}{\lambda}\sin\theta\right)\right| \tag{2.40}$$

Figure 2.9 shows the PSF for $d = 0.1\,m$ and $\lambda = 0.03\,m$.

This PSF does not have a main lobe; all the lobes have equal amplitude. The null-to-null width of the centre lobe is

$$\theta_{nn} = 2\,\text{asin}\frac{\lambda}{d} \tag{2.41}$$

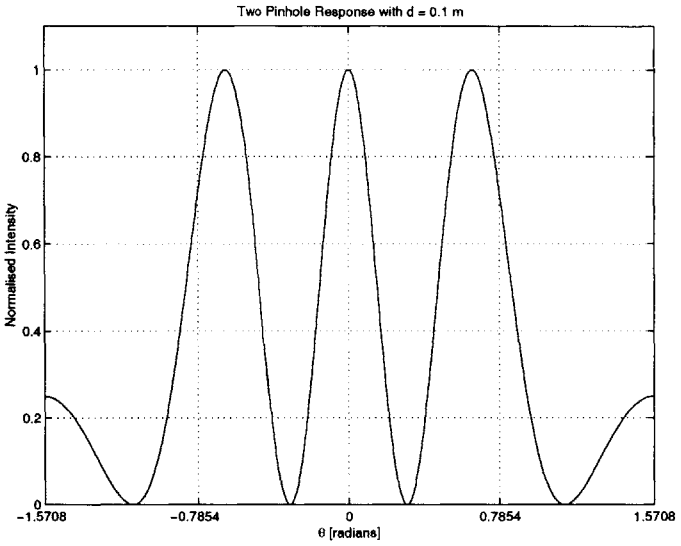

Figure 2.9 Two Pinhole Point Spread Function

This is the same as the null-to-null width obtained with the single slit (2.36). However, the side lobes of the response pattern all have the same amplitude as the main lobe. Since all the lobes have equal amplitude, energy originating from the centre lobe is indistinguishable from energy originating from the side lobes. The resulting image will contain aliased copies, or ghosts, from the off-centre lobes. Despite having the same resolution as the single slit aperture, the image quality is significantly worse.

For the general case of $2n$ apertures distributed evenly across a distance d, the amplitude distribution is

$$E_{even}(x') = \sum_{i=1}^{n} \left(\delta(x' + \frac{d}{ni}) + \delta(x' - \frac{d}{ni}) \right) \tag{2.42}$$

The associated PSF is

$$E_{even}(\theta) = 2 \sum_{i=1}^{n} \cos\left(\frac{2\pi d}{\lambda ni} \sin\theta \right) \tag{2.43}$$

For the general case of $(2n+1)$ pinholes distributed evenly across a distance d, the amplitude distribution is

$$E_{odd}(x') = \delta(x') + \sum_{i=1}^{n} \left(\delta(x' + \frac{d}{ni}) + \delta(x' - \frac{d}{ni}) \right) \tag{2.44}$$

The associated PSF is

$$E_{odd}(\theta) = 1 + 2 \sum_{i=1}^{n} \cos\left(\frac{2\pi d}{\lambda ni} \sin\theta \right) \tag{2.45}$$

Taking the limit as n tends to infinity in either (2.43) or (2.45), we obtain the PSF of the slit (2.35).

By adding additional pinholes to the sparse array, the synthesized aperture response will more closely approach that of the slit, and the quality of the image will improve.

2.3.3 Image Construction using Interferometry

Now that we have seen how the two point-aperture interferometer works we can generalise the discussion to an interferometer consisting of an arbitrary number of apertures, which are distributed in two dimensions. In the case where the apertures are distributed across three dimensions, the discussions can be applied by projecting the apertures onto a plane. We turn to Fourier theory for aid in understanding the interferometer.

The discrete Fourier transform acts as a bank of narrowband filters. Given the DFT of a band-limited time-domain signal, and assuming that the sampling frequency is high enough to prevent aliasing, we can reconstruct the time-domain signal with increasing accuracy as the spacing between the filters is decreased. This corresponds to increasing the sampling interval

$$f_{\Delta} = \frac{1}{T} \tag{2.46}$$

An interferometer can be thought of as a DFT. In this case the centre frequencies of the bank of filters correspond to the spatial frequencies, (u_i, v_i) sampled by the interferometer. The sampled spatial frequency in a particular dimension is the component of the aperture spacing projected normal to the line-of-sight (LOS) to the footprint. It is measured in carrier wavelengths, λ

$$u_i = \frac{x_{i+1} - x_i}{\lambda} \cdot \cos\theta \tag{2.47}$$

where x_i is the spatial coordinate of the i^{th} aperture and θ is the projection angle.

A two-dimensional aperture configuration samples spatial frequencies in two dimensions

$$\begin{aligned} u_i &= \frac{x_{i+1} - x_i}{\lambda} \cdot \cos\theta \\ v_i &= \frac{y_{i+1} - y_i}{\lambda} \cdot \cos\theta \end{aligned} \tag{2.48}$$

where (x_i, y_i) are the spatial coordinates of the i^{th} aperture.

The resolution of the interferometer is determined by the greatest spacing (baseline) between apertures

$$\theta_{resolution} = \operatorname{asin}\frac{\lambda}{B} \approx \frac{\lambda}{B} \tag{2.49}$$

An image of the scene being mapped is constructed by interpolating between the discrete spatial frequency measurements. As more unique spatial frequencies are introduced the constructed image more closely approaches the true scene.

2.3.4 The Point Spread Function of an Interferometer

An individual aperture's behaviour is conveniently characterised by its antenna pattern, or point spread function. The idea carries through to an interferometer, for which the PSF is

particularly useful, as we shall see in subsequent chapters. In this section we discuss the construction and characteristics of the PSF of an interferometer.

Construction

The PSF of an interferometer is a separable function of the individual aperture response as well as the distribution of the apertures in space. If we assume that the aperture responses are identical, as is the case for most interferometers, the interferometer PSF can be conveniently calculated using Fourier transform techniques, as shown for a one-dimensional interferometer in Figure 2.10.

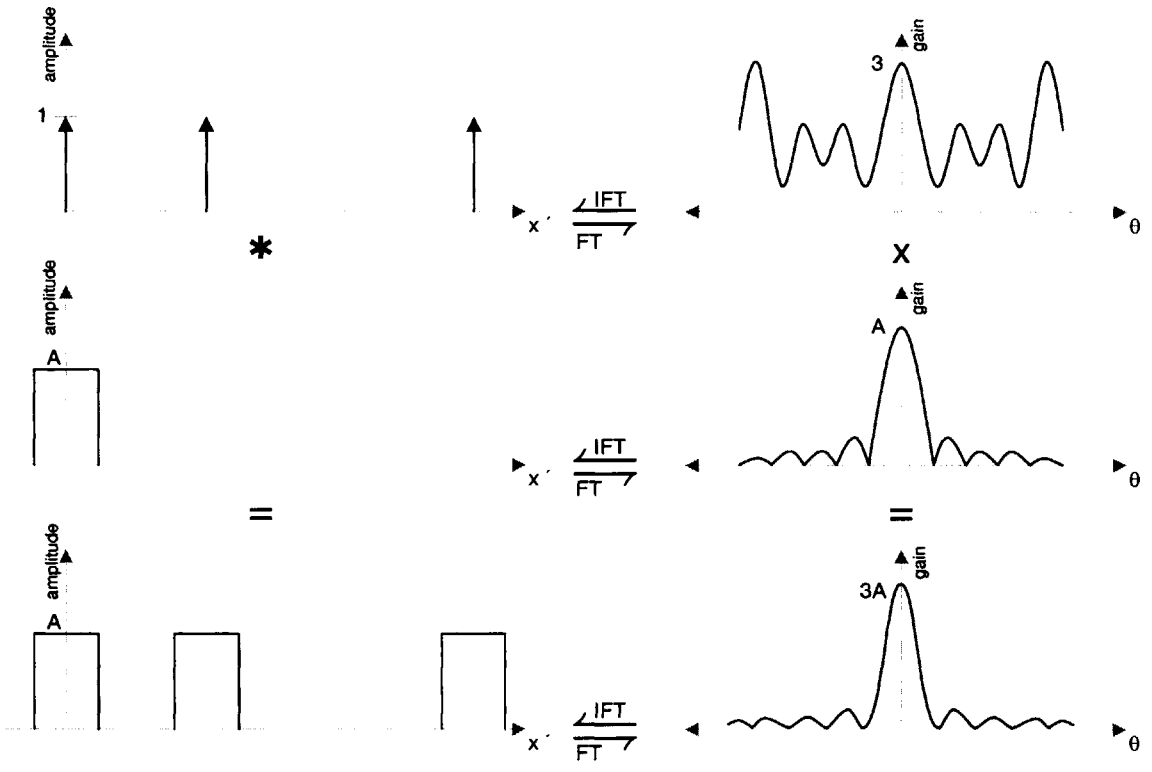

Figure 2.10 Calculation of the PSF using the Fourier Transform

An array of identically illuminated apertures is formed by convolving the individual aperture illumination with impulses placed at the aperture positions. By first taking the Fourier

transforms (2.22) of the aperture positions and illumination, the convolution operation is transformed into multiplication. The PSF can therefore be written as the product of two functions

$$PSF = array\_factor \times aperture\_response \tag{2.50}$$

The Fourier transform of the aperture positions is referred to as the *array factor*, while the Fourier transform of the aperture illumination is just the individual aperture response.

The array factor for a two-dimensional interferometer is the two-dimensional Fourier transform of the aperture positions

$$\begin{aligned} array\_factor(x, y, z) &= \left| \int_{S'} \sum_{n=1}^{N_{sats}} \delta(x - x_n', y - y_n') e^{-j\frac{2\pi}{\lambda r}(xx_n' + yy_n')} dx'dy' \right| \\ &= \left| \sum_{n=1}^{N_{sats}} e^{-j\frac{2\pi}{\lambda r}(xx_n + yy_n)} \right| \end{aligned} \tag{2.51}$$

where the individual aperture positions are given by (x_n, y_n). After transformation to azimuth-elevation coordinates (ψ_{az}, ψ_{el}), the array factor can be expressed as [Kong, 1998]

$$array\_factor(\psi_{az}, \psi_{el}) = \left| \sum_{n=1}^{N_{sats}} e^{-\frac{2\pi j}{\lambda}(x_n \sin\psi_{az} + y_n \sin\psi_{el})} \right| \tag{2.52}$$

Since the array factor is the sum of complex exponentials, it is periodic in ψ_{az} and ψ_{el}.

The PSF for a two-dimensional interferometer is therefore given by

$$I(\psi_{az}, \psi_{el}) = \left[E(\psi_{az}, \psi_{el}) \cdot \left| \sum_{n=1}^{N_{sats}} \exp\left(-\frac{2\pi j}{\lambda}(x_n \sin\psi_{az} + y_n \sin\psi_{el}) \right) \right| \right] \tag{2.53}$$

where $E(\psi_{az}, \psi_{el})$ is the individual aperture response. Figure 2.11 shows the construction of the PSF for an unrestricted minimum redundancy (see Section 4.4) array of five uniformly illuminated rectangular apertures.

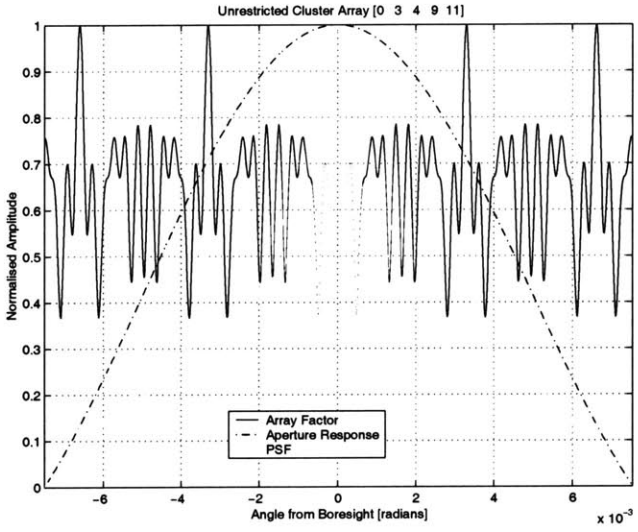

Figure 2.11 Construction of the PSF

PSF Design Parameters

For a given carrier wavelength, the aperture response is a function of the aperture geometry and illumination, while the array factor is a function of the array configuration and baseline. We can therefore isolate the effects of changing the array from the effects of changing the aperture response.

The carrier wavelength determines the scale with which the system is sampled, and is usually determined on the basis of physical size and atmospheric attenuation constraints. Therefore we assume that it is fixed and the PSF must be designed for best performance at the chosen wavelength.

Since the aperture diameter is much smaller than the array baseline, the resolution of the interferometer is determined by the baseline. The the field of view is determined by the individual aperture response.

The width of the aperture response mainbeam is inversely proportional to the aperture diameter. Increasing the aperture diameter therefore decreases the field of view and vice versa. For a uniformly illuminated rectangular aperture, the null-to-null field-of-view is

$$FOV_{rectangular} = [-\text{asin}\left(\frac{\lambda}{D}\right), \text{asin}\left(\frac{\lambda}{D}\right)] \tag{2.54}$$

The array configuration determines the shape of the PSF, and is examined in detail in Chapter 4. The array baseline determines the resolution and period of the PSF. As the baseline is increased the resolution is increased and the period is decreased. Figure 2.12 shows the array factor for two identical array and aperture configurations, one with a baseline of 500 m, the other with a baseline of 1000 m. Doubling the baseline halves the period of the array factor, and halves the widths of the main lobes.

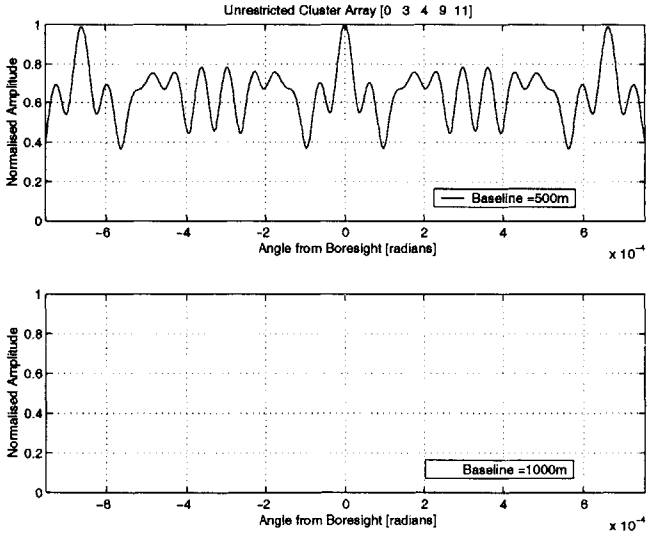

Figure 2.12 Effect of Array Baseline on the PSF

Sampling the PSF

In many applications it is necessary to sample the PSF. We define the sampling frequency as the number of samples per radian. Figure 2.13 shows a PSF that has been sampled at two different rates, one ten times faster than the other. Sampling at too low rate (under-sampling) clearly results in an inaccurate reconstruction of the PSF.

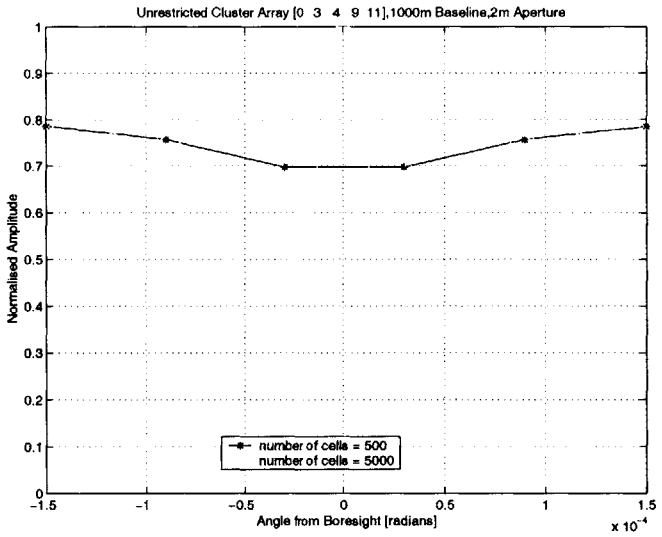

Figure 2.13 PSF Sampled at Two Different Rates

What is the minimum sampling rate that will allow accurate reconstruction of the PSF? To answer this question we refer to time domain signal processing and the Nyquist theorem. The theorem states that a band-limited signal can be accurately reconstructed provided that the sampling frequency is at least twice as high as the signal bandwidth.

$$f_s \geq 2 \cdot f_{bandwidth} \tag{2.55}$$

The signal can then be reconstructed from the samples by means of the interpolation equation [Oppenheim et al., 1999]

$$x_r(t) = \sum_{n=-\infty}^{\infty} x[n] \frac{\sin[\pi(t - nT_s)/T_s]}{\pi(t - nT)/T_s} \tag{2.56}$$

where $T_s = \frac{1}{f_s}$ is the sampling period.

The highest spatial frequency sampled by an interferometer in a particular dimension is equal to the maximum spacing (baseline) divided by the carrier wavelength

$$(u_{max}, v_{max}) = (\frac{B_x}{\lambda}, \frac{B_y}{\lambda}) \tag{2.57}$$

Therefore the Nyquist sampling frequencies in the two dimensions are

$$(u_{Nyquist}, v_{Nyquist}) = (\frac{2B_x}{\lambda}, \frac{2B_y}{\lambda}) \tag{2.58}$$

Accordingly, the required spacing between adjacent samples is

$$(\theta_x, \theta_y) = (\frac{\lambda}{2B_x}, \frac{\lambda}{2B_y}) \tag{2.59}$$

Figure 2.14 shows an example of a PSF, and the Nyquist sampled points.

Given the field of view of the antenna, the required number of sampling points in the PSF can be determined.

$$n_{points} = \left\lfloor \frac{2B}{\lambda} \cdot FOV \right\rfloor \tag{2.60}$$

The approximate resolution that can be obtained by a PSF sampled with n_{points} not necessarily corresponding to the Nyquist rate is

$$\theta_{res} \approx \frac{\lambda}{B} \cdot \frac{\frac{2B}{\lambda} \cdot FOV}{n_{points}} = \frac{2 \cdot FOV}{n_{points}} \tag{2.61}$$

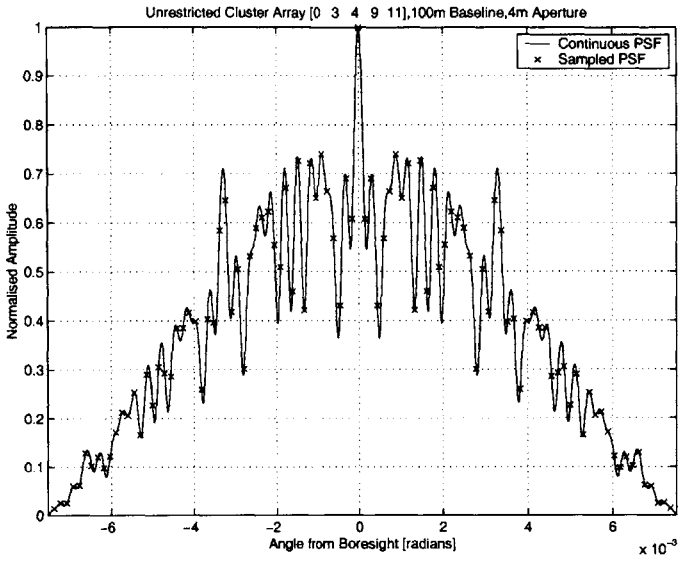

Figure 2.14 Sampling the PSF

We shall see in the next section that undersampling the PSF corresponds to underutilising the system.

2.3.5 Interferometer System Design

By definition, an interferometer consists of two or more apertures. These apertures may be physically connected, or separated in space. When the desired baseline is small, physically connecting the apertures is feasible and is often the simplest solution. However, for large baselines, the apertures must be physically separated in space. In this case the system will consist of at least two spacecraft. Accurate knowledge, and in many cases, control, of the spacecraft positions and velocities is essential for good quality image generation.

Resolution on the order of tens of metres using an X-band radar system requires baselines on the order of kilometres. MTI interferometric radar systems therefore require at least some physically disconnected apertures. A group of spacecraft performing a common purpose is referred to as a spacecraft *cluster.*

A cluster can be described in terms of its orbit, the number of spacecraft, and its distribution in space. This section provides a high-level discussion of these characteristics as applied to interferometric systems.

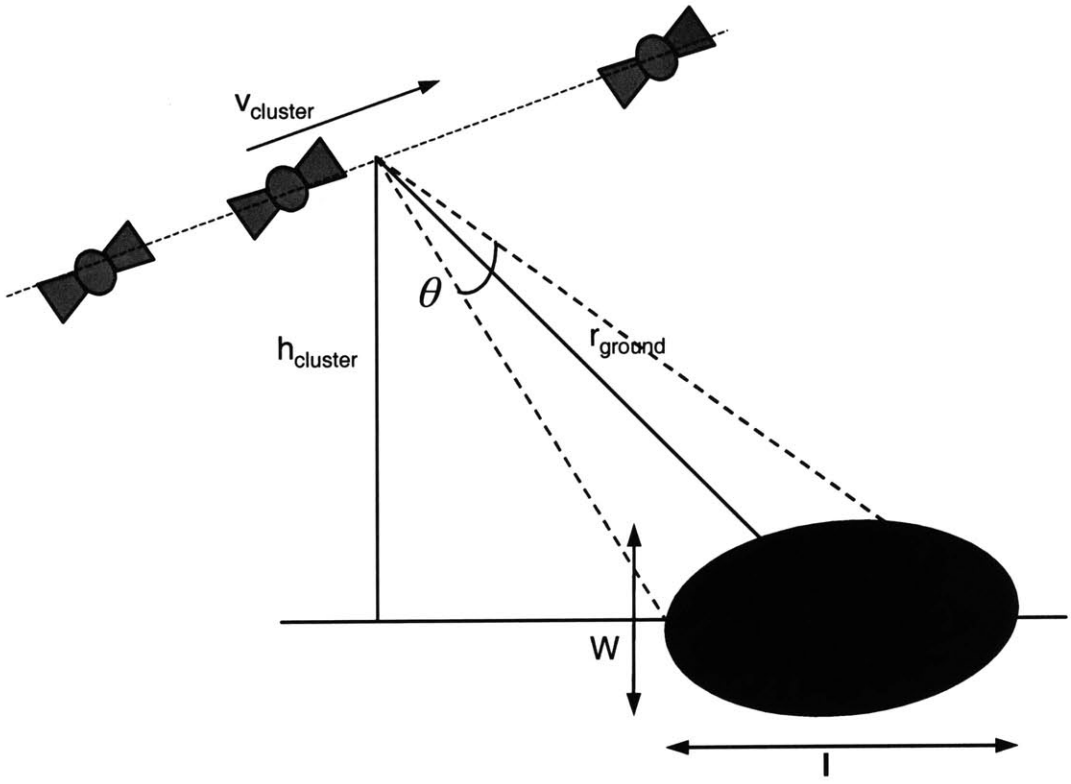

Figure 2.15 Side-looking Linear Cluster

Cluster Orbit

Selection of the cluster orbit to meet the mission requirements is complex and involves trading off a number of different parameters. The mission lifetime, costs for each phase of the mission, space environment and viewing geometry are dependent on the orbit. In the case of an observation system, orbit selection directly affects payload performance. Discussion of this task is not within the scope of this thesis. [Wertz and Larson, 1999] provides an excellent introduction to orbit selection.

Number of Spacecraft

Interferometric systems with physically isolated apertures are inherently distributed in nature. Selection of the number of spacecraft is driven by performance requirements, reliability, and cost constraints.

From an image quality viewpoint, a larger number of apertures provides greater freedom in the design of the PSF. Although the resolution of the PSF is determined by the baseline(s), the shape of the PSF depends on the placement of any additional apertures within the baseline. When all other parameters are held constant, better quality images can be obtained by interferometers with a larger number of apertures.

From a reliability viewpoint, a system consisting of multiple satellites is more reliable and can be upgraded, or degrades, gracefully, as opposed to a monolithic system design. Readers interested in the reliability of distributed space systems are referred to [Jilla et al., 1997].

The number of spacecraft is driven by the desired image quality and the required overall system reliability. In the absence of other considerations, systems with the largest number of spacecraft are preferred. However, cost constraints drive designers to trade the size and complexity of spacecraft versus the number of spacecraft in a cluster. Stated differently, a few big, complex satellites cost as much as many small simple satellites.

Cluster Shape

An interferometer can be spread across one, two, or three dimensions. We restrict the discussion to one- or two-dimensional interferemeters. The first shape characteristic of the interferometer is its baseline in each dimension. Once the baseline is determined, the distribution of the apertures must be determined. This is not a trivial process, and some approaches for SPIR systems are suggested in Chapter 4. This section provides some geometric relations that are useful in determining interferometer baselines.

As we have seen, the resolution of an interferometer in a particular direction is determined by its baseline B in that direction

$$\theta_{resolution} \approx \frac{\lambda}{B} \tag{2.62}$$

If we define r_{ground} as the range to the centre of the footprint from the centre of the cluster as shown in Figure 2.15, the cross-range resolution on the ground is

$$\Delta_{cross} = r_{ground}\frac{\lambda}{B} \tag{2.63}$$

For the side-looking linear cluster shown in Figure 2.15, there is only one baseline, which determines the resolution in the cross-range direction. In this case the range resolution is determined by the compressed pulse width, as discussed in Section 2.1.1. The number of resolution cells on the ground in the cross-range direction is determined by dividing the angular resolution into the field of view

$$n\_cells_{cross} = \left\lfloor \frac{FOV}{\lambda/B} \right\rfloor = \left\lfloor \frac{B}{\lambda} \cdot FOV \right\rfloor \tag{2.64}$$

If we refer back to (2.60) we see that this corresponds to half the required number of sampled points for the PSF. The minimum sampling rate for the PSF ensures that the maximum possible resolution is obtained. The PSF must be sampled at twice the desired ground resolution to ensure accurate image reconstruction. Undersampling the PSF results in a loss of resolution. The same loss in resolution results from decreasing the baseline.

The range to the footprint, r_{ground}, depends on the grazing angle, ψ and the cluster altitude, assuming that the Earth's curvature can be ignored.

$$r_{ground} = \frac{h_{cluster}}{\sin\psi} \tag{2.65}$$

The width of the footprint in the cross-range direction is

$$w = r_{ground} \cdot FOV \tag{2.66}$$

where FOV is the field of view of the individual apertures.

The length of the footprint in the range direction ignoring the Earth's curvature, is

$$l = \frac{w}{\sin\psi} \tag{2.67}$$

2.4 Interferometric Techniques

Radar interferometry is not new. Several techniques exist and have been successfully implemented. In this section we discuss two techniques that are used in MTI radar. Displaced Phase Centre Antenna (DPCA) was implemented on analog systems long before the advent of digital systems. Space-time adaptive processing (STAP) is a more sophisticated technique, which became popular in the 1980s.

2.4.1 Classical Displaced Phase Centre Antenna (DPCA)

DPCA is a technique used to eliminate ground clutter observed by air- or spaceborne pulse-doppler radars. The doppler shift of ground clutter is entirely determined by the relative velocity between the radar platform and the ground. Instead of appearing as a narrow peak at zero doppler shift, clutter is spread around a frequency corresponding to the relative platform motion. If the radar platform appears to be stationary, the clutter doppler shift will be reduced to a narrow band around zero frequency. The clutter doppler shift is observed as a stretching or compression of the interpulse period. For any two pulses, shifting the antenna phase centre so that the second pulse appears to come from the same location as the first pulse reduces the observed clutter doppler shift.

The required shift in antenna phase centre is typically achieved by using a two segment electronically steered antenna. The platform's speed is matched to the PRI so that the distance the aircraft travels during a PRI is equal to the distance between the phase centres of the two antenna segments.

$$d_{centres} = v_{platform} \cdot PRI \tag{2.68}$$

The distance between antenna segments is constrained by the platform's size and is fixed. The PRF is therefore limited to a range related to the platform's velocity. Precise control of the antenna's motion relative to the ground is required. [Stimson, 1998]

2.4.2 Space-Time Adaptive Processing (STAP)

Like DPCA, Space-Time Adaptive Processing (STAP) performs clutter rejection by compensating for the platform's motion. It uses the spectral and time-domain characteristics of the radar signal to filter out clutter and noise jamming.

When viewed as a function of cross-range and doppler, clutter observed by a side-looking air- or spaceborne platform appears as a narrow diagonal ridge. Noise jamming appears as a ridge along the doppler dimension. STAP applies narrow notch filters across the cross-range and doppler dimensions to attenuate clutter and noise jamming ridges without cancelling moving targets.

The filter is implemented on the range gated signals from the individual channels of an electronically steered array. If the interference (clutter, noise jamming) between channels is highly correlated, some linear combination of the channels will cancel the interference. Accurate knowledge of the clutter statistics, noise jamming frequencies, and target positions is required to construct a filter that will reveal the targets.

Although STAP can yield good results, its performance is severely degraded when the actual clutter differs from the assumed model. Suddenly applied noise jamming and new targets will not be recognised immediately and several processing cycles may be required before the filter is sufficiently adapted. [Klemm, 1999]

2.5 Summary

Scanned Pattern Interferometric Radar uses techniques from both the radar and interferometry fields. This chapter has introduced background material on signal processing, radar theory and interferometry that will be used in the remainder of this thesis. Several useful results applicable to optical and radio frequency interferometers have been derived.

The interested reader is encouraged to consult the references, especially [Stimson, 1998] (Radar), [Oppenheim et al., 1999] (Fourier Transforms), and [Thompson et al., 1994] (Interferometry) for a more in depth treatment of each area.

Chapter 3

SCANNED PATTERN INTERFEROMETRIC RADAR

Chapter 2 has laid the groundwork for our study of Scanned Pattern Interferometric Radar (SPIR). Although interferometry is most often thought of as a passive observation technique, it can also be applied profitably to active observation systems. Our concern is the observation of earth from space, specifically moving targets. By applying the theory of interferometry to radar, we develop an intuitively simple method that delivers high resolution results and is clutter-independent.

This chapter begins with an introductory discussion of the SPIR method. Next, we present some high level system design considerations. Individual aperture placement for good SPIR performance is discussed in Chapter 4.

3.1 The SPIR Method

Interferometry can be implemented in a number of ways, depending on the application. Imaging interferometers cross-correlate the time domain signals from multiple apertures to measure the magnitude and phase of the spatial Fourier components. The quality of the constructed image depends on the sampled spatial frequency resolution. Missing (u, v) components cause grating lobes, which introduce artefacts, resulting in a "smudged" image.

Cross-correlation of signals is not the only way to synthesize the sparse array. If the incoming signals are summed, the resulting signal will have contributions that were spatially weighted by the Point Spread Function (PSF). By introducing appropriate time delays in the aperture outputs before summing, the direction in which the main lobe of the PSF is located can be shifted left or right, changing array sensitivity to different regions of the scene. If the signals are at radio frequencies (X-band and below) the information from each aperture can be recorded digitally and the scanning of the PSF can be performed during post-processing. This is the concept behind Scanned Pattern Interferometric Radar (SPIR).

3.1.1 A Simplified Model of SPIR

A simple model of the SPIR concept is shown in Figure 3.1.

The three boxes labelled R, G, and B represent three footprint cells, which contain targets of unknown size. The array of arrows is the discretised PSF with values $\begin{bmatrix}1 & 5 & 3 & 2 & 4\end{bmatrix}$. As the PSF is placed on the ground in a particular position, each of the gains multiplies the strength of the signal (radar cross section) from the corresponding cell. Only the summed signal is detected by the spacecraft. As shown, with the gains $\begin{bmatrix}3 & 2 & 4\end{bmatrix}$ covering the cells, the summed signal is 27, whereas with $\begin{bmatrix}5 & 3 & 2\end{bmatrix}$ and $\begin{bmatrix}1 & 5 & 3\end{bmatrix}$ covering the cells, the summed signals are 26 and 20 respectively.

With three samples taken (corresponding to the three unknowns) the problem can be formulated as a matrix equation.

$$\begin{bmatrix}3 & 2 & 4\\ 5 & 3 & 2\\ 1 & 5 & 3\end{bmatrix}\begin{bmatrix}R\\ G\\ B\end{bmatrix} = \begin{bmatrix}27\\ 26\\ 20\end{bmatrix} \tag{3.1}$$

Provided the matrix is invertible, the values within R, G, and B can be determined.

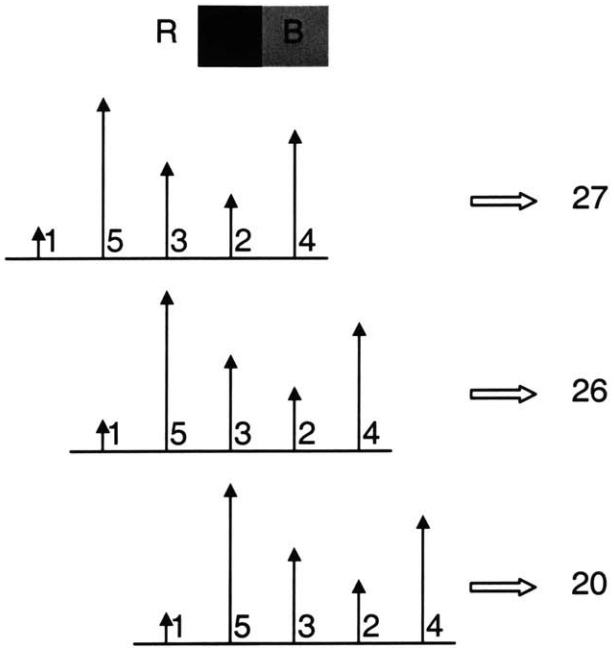

Figure 3.1 The SPIR Concept

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} -0.02 & 0.24 & -0.14 \\ -0.22 & 0.08 & 0.24 \\ 0.37 & -0.22 & -0.02 \end{bmatrix} \begin{bmatrix} 27 \\ 26 \\ 20 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix} \tag{3.2}$$

SPIR performs this same process by effectively scanning the PSF across the footprint. The "scanning" is performed during post-processing, by time shifting the outputs of the apertures in a coordinated fashion.

Unlike STAP, which must maintain precise aperture position to place nulls over competing clutter, SPIR's only requirement is knowledge of the PSF matrix, so that knowledge of the satellites' positions is more important than fine control.

In the next section we present a mathematical derivation of the SPIR method.

3.1.2 Mathematical Derivation of SPIR

Having established qualitatively the concept of SPIR, we now present its mathematical foundation. We perform the analysis for one range strip as viewed from a linear interferometer array. In the case of an interferometer with apertures distributed in two dimensions, the analysis is easily extended.

We assume that one satellite transmits and all the satellites receive the reflected signal. Multiple satellite transmits can be thought of as a series of single transmitter results. Such results may be combined to increase the probability of detection while decreasing the false alarm rate.

The transmitting satellite transmits a signal of the form:

$$s_{transmit}(t) = A(t)e^{j\omega_c t} \tag{3.3}$$

where $A(t)$ is an amplitude factor (modulation), which varies much more slowly than the carrier $e^{j\omega_c t}$. To avoid unnecessarily complicating this analysis we set $A(t) = 1$.

The received signal at satellite i is a scaled, time and doppler spread version of the transmitted signal.

The scale factor is given by the inverse of range squared (for signal amplitudes). We assume that the variation in scaling factor between satellites is negligible.

The time delay from a particular footprint cell to the i^{th} satellite is the sum of the transmission and reflection delays. Time delays are measured relative to the time delay to the transmitting satellite.

In terms of wavelengths, the received signal is delayed by

$$\Delta wavelength_i = \frac{x_i}{\lambda}\sin\theta_p = u_i \sin\theta_p \tag{3.4}$$

as shown in Figure 3.2.

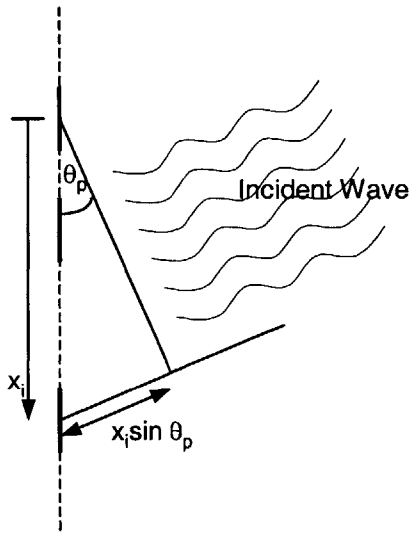

Figure 3.2 Relative Signal Delay

where θ_p is the angle to the ground and u_i is the spatial wave number corresponding to the i^{th} satellite's position, x_i, relative to the transmitting satellite.

The doppler shift of a footprint cell as observed by a given satellite is a function of the relative velocity between the transmitting and receiving satellite and the footprint cell. Each cell will have a range of doppler frequencies associated with it, representing the variation in line-of-sight velocity across a cell. In the case where a cell contains a target, the target's line-of-sight velocity is represented as an additional doppler frequency component for that cell.

In pulse-doppler radar systems the doppler frequency is obtained by digitally sampling N points of the time signal at the pulse repletion frequency (PRF) and taking the discrete Fourier transform. This gives a discrete set of frequencies for each cell

$$\omega_{Doppler} = \left[\omega_1 \ \omega_2 \ \ldots\right] \tag{3.5}$$

The radar cross-section (RCS) of a footprint cell or target represents how much of the incoming radar signal is reflected back to the transmitting aperture(s). As a first order approximation, we assume that the RCS of any given footprint cell or target is constant with respect to time and viewing angle. For short integration periods (<50 ms) this is a good approximation [Stimson, 1998].

The RCS is represented by the amplitude factor σ. For convenience we define σ to be a function of position and frequency such that it is zero at frequencies where there is no doppler return.

$$\sigma = \sigma(position,Doppler) \tag{3.6}$$

Noting that a digitally sampled frequency can be represented as

$$\omega_r = 2\pi\frac{r}{N} \tag{3.7}$$

we replace the sum of the carrier and doppler frequencies with

$$\omega_c + \omega_{Doppler} = 2\pi\frac{k}{N} \tag{3.8}$$

The signal received by satellite i from the entire footprint in the absence of noise is therefore given by

$$s_i(t) = \frac{1}{R^2}e^{j\omega t}\int_{\theta_p} E(\theta_p)\sum_{k=0}^{N-1}\sigma(\theta_p,k)e^{j2\pi\frac{k}{N}t}e^{-j2\pi u_i \sin\theta_p}d\theta_p \tag{3.9}$$

The single aperture response is a function of angle and is given by $E(\theta)$.

In order to focus on a particular cell θ_q in the footprint, the received signals at each satellite are time shifted and added

$$s(\theta_q,t) = \frac{1}{R^2}\sum_{i=1}^{N_{sats}}\left[\int E(\theta_p)\sum_{k}^{N-1}\sigma(\theta_p,k)e^{j2\pi\frac{k}{N}t}e^{-j2\pi u_i\sin\theta_p}d\theta_p\right]e^{j2\pi u_i\sin\theta_q} \tag{3.10}$$

Discretising the footprint into cells gives

$$s(\theta_q,t) = \frac{1}{R^2}\sum_{i=1}^{N_{sats}}\left[\sum_{p=0}^{N_{cells}-1}E(\theta_p)\sum_{k}^{N-1}\sigma(\theta_p,k)e^{j2\pi\frac{k}{N}t}e^{-j2\pi u_i(\sin\theta_p-\sin\theta_q)}\right] \tag{3.11}$$

Discretising the time variable and assuming that we have N samples in total

$$s(\theta_q,m) = \frac{1}{R^2}\sum_{i=1}^{N_{sats}}\sum_{p=0}^{N_{cells}-1}E(\theta_p)\sum_{k}^{N-1}\sigma(\theta_p,k)e^{j2\pi\frac{km}{N}}e^{-j2\pi u_i(\sin\theta_p-\sin\theta_q)} \tag{3.12}$$

Taking the discrete Fourier transform (DFT) to obtain the frequency response

$$\begin{aligned} s(\theta_q,l) &= \frac{1}{NR^2}\sum_{m=0}^{N-1}\left[\sum_{i=1}^{N_{sats}}\sum_{p=0}^{N_{cells}-1}E(\theta_p)\sum_{k=0}^{N-1}\sigma(\theta_p,k)e^{j2\pi\frac{km}{N}}e^{-j2\pi u_i(\sin\theta_p-\sin\theta_q)}\right]e^{-j2\pi\frac{lm}{N}} \\ &= \frac{1}{NR^2}\sum_{i=1}^{N_{sats}}\sum_{p=0}^{N_{cells}-1}\sum_{k=0}^{N-1}E(\theta_p)\sigma(\theta_p,k)e^{-j2\pi u_i(\sin\theta_p-\sin\theta_q)}\sum_{m=0}^{N-1}e^{j2\pi\frac{m}{N}(k-l)} \end{aligned} \tag{3.13}$$

Evaluating the summation over m

$$\frac{1}{N}\sum_{m=0}^{N-1}e^{j2\pi\frac{m}{N}(k-l)} = \begin{cases} 0 & k\neq l \\ 1 & elsewhere \end{cases} \tag{3.14}$$

Therefore

$$s(\theta_q,l) = \frac{1}{R^2}\sum_{p=0}^{N_{cells}-1}\sigma(\theta_p,l)E(\theta_p)\sum_{i=1}^{N_{sats}}e^{-j2\pi u_i(\sin\theta_p-\sin\theta_q)} \tag{3.15}$$

Set

$$G(\theta_p,\theta_q) = E(\theta_p)\sum_{i=1}^{N_{sats}} e^{-j2\pi u_i(\sin\theta_p-\sin\theta_q)} \tag{3.16}$$

Note that $|G(\theta_p,\theta_q)|$ is the PSF for a one-dimensional interferometer, focused on a point at angle θ_q.

Then

$$s(\theta_q,l) = \frac{1}{R^2}\sum_{p=0}^{N_{cells}-1} G(\theta_p,\theta_q)\sigma(\theta_p,l) \tag{3.17}$$

In matrix notation:

$$\mathbf{S} = \frac{1}{R^2}\mathbf{G}\Sigma \tag{3.18}$$

where

$$\Sigma_{pl} = \sigma(\theta_p,l) \tag{3.19}$$

is the RCS matrix as a function of cross-range and doppler frequency and

$$\mathbf{G}_{qp} = G(\theta_p,\theta_q) \tag{3.20}$$

is the PSF evaluated at (θ_p, θ_q) and is referred to as the *PSF matrix.*

Therefore the RCS-doppler profile of the one-dimensional strip can be recovered using the constructed signal $s(\theta_q, t)$ and the PSF of the aperture, $G(\theta_p, \theta_q)$.

3.2 SPIR Implementation

The previous section has shown that it is indeed possible to recover the true ground scene by deconvolving the PSF from the appropriately time-shifted and summed received signal.

In this section we discuss how to implement the SPIR method as part of a GMTI radar system.

3.2.1 Basic Processing Steps

The SPIR method is combined with some of the standard radar techniques introduced in Chapter 2 to generate ground maps. Figure 3.3 illustrates the basic steps that are required.

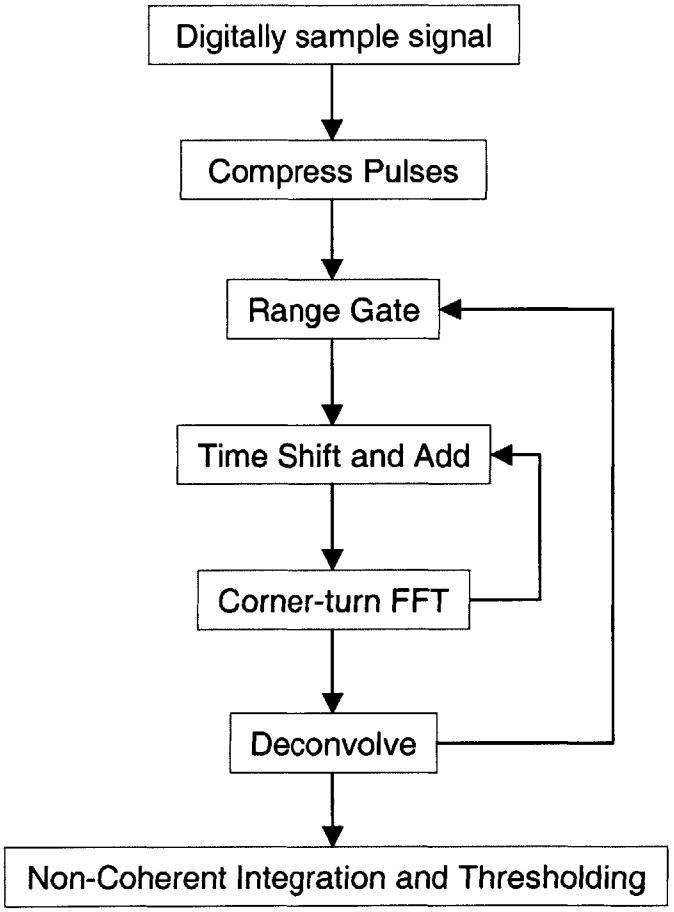

Figure 3.3 Basic SPIR Implementation

3.2.2 Transmit/Receive Configurations

To maximise the reconfigurability of the system, each aperture should be equipped with both a receiver and a transmitter. Many different transmit/receive configurations are then possible, but in this section we consider the two extreme configurations.

The simplest SPIR system has one aperture transmitting to the ground, while all the apertures receive the reflected signal, as shown in Figure 3.4. In this case a single image is constructed during each "look". The same aperture may transmit all the time, or apertures can take turns.

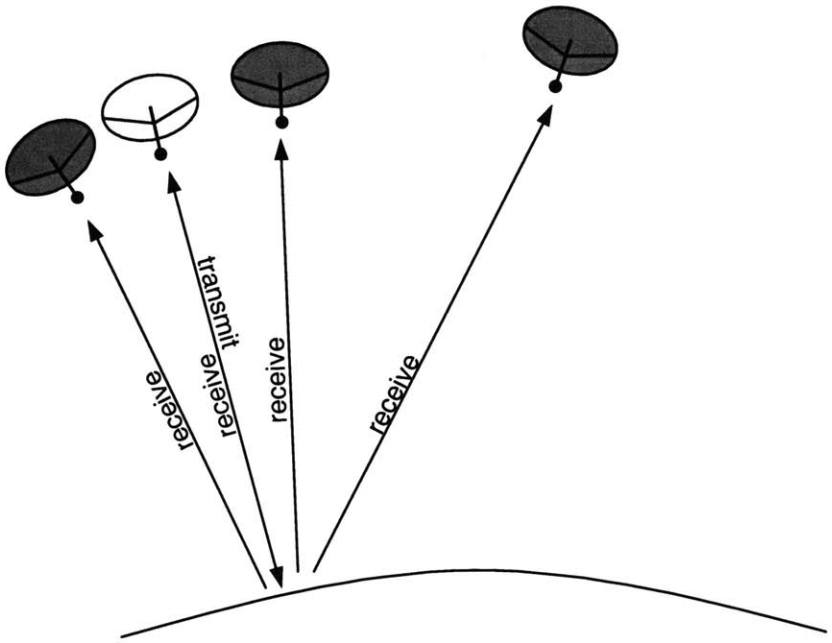

Figure 3.4 Single Transmitting Aperture

At the other extreme each aperture transmits, and then receives its own reflected signal as well as the reflected signals from all the other apertures, as shown in Figure 3.5. This configuration can be seen as a superposition of the single transmitting aperture configuration. Consider each aperture's transmission individually. An image is generated corresponding to each transmission in the "look". These images can then be non-coherently combined to

increase the signal-to-noise ratio. Unfortunately the processing load increases as the number of transmitting apertures is increased.

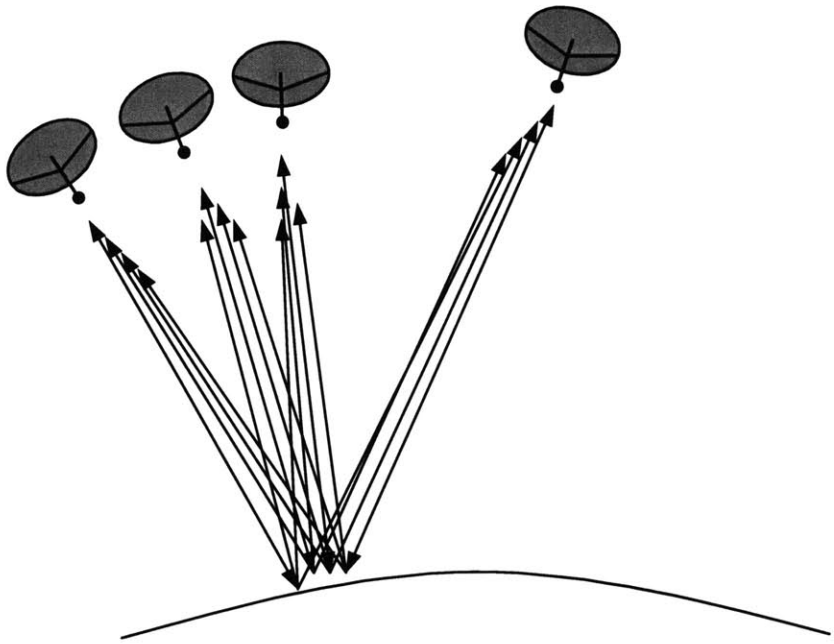

Figure 3.5 All Apertures Transmit and Receive

3.2.3 GMTI and the SPIR Method

The output of the SPIR process is a "data cube", which specifies the position in range and cross-range, and the doppler frequency of the targets, as shown in Figure 3.6. Ground clutter appears as a plane in the cube, as a result of the geometric relationship between clutter position and doppler frequency. Moving targets (circles on the figure) have different doppler shifts from the clutter and appear off the plane.

The degree to which targets can be discerned from one another and ground clutter depends on the resolution of the system in range, cross-range and doppler frequency.

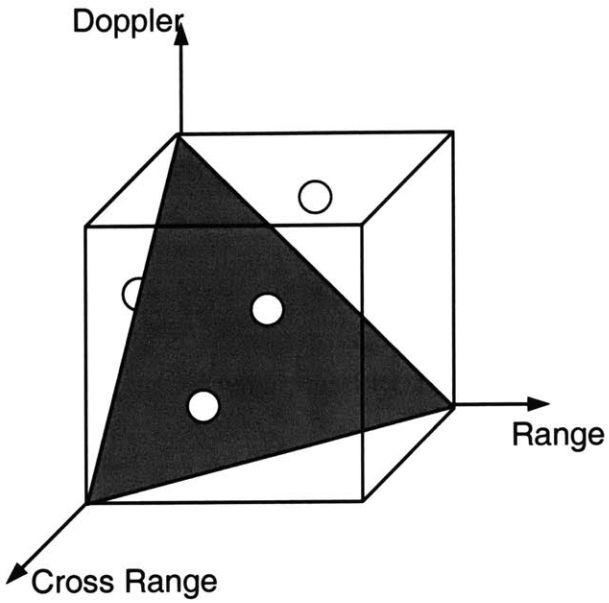

Figure 3.6 SPIR Data Cube

Processing a Two-Dimensional Footprint

The discussion till now has assumed that the PSF is scanned in the cross-range direction, but what about the range direction? There are two ways of dealing with this:

The first is to range gate the radar return before applying the SPIR method. The returns from each range gate can then be SPIR processed as one-dimensional strips (iso-range) across the footprint.

Alternatively, two-dimensional mapping can be performed as an extension of the one-dimensional case, where instead of only scanning the point spread function in the cross-range direction, it is scanned in both range and cross-range. This of course requires that the cluster have apertures distributed in two dimensions.

Range Ambiguities

The two-dimensional processing capability allows for the elimination of the problem of range ambiguities. Consider the case where the PRF is high enough that the length of the

footprint contains two range ambiguities, as shown in Figure 3.7. If range gating is performed, the data from the two ambiguous ranges will overlap. However, if two-dimensional processing is performed only on the cells that comprise the ambiguous ranges then sufficient data can be collected to separate the information coming from each range.

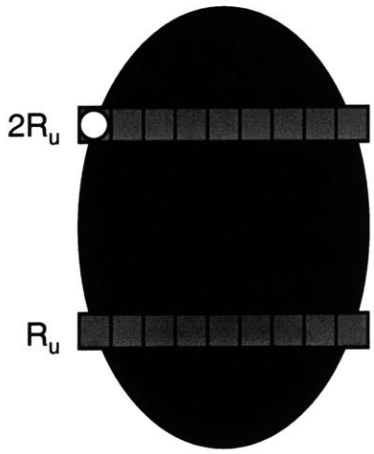

Figure 3.7 Footprint with a Single Range Ambiguity

The processing load required to remove range ambiguities is a function of the number of cross range cells and the number of ambiguities. Since the system (3.18) must be solved for each range ambiguity, the processing load scales with the square of the number of range ambiguities (see Section 3.2.4). Therefore this approach quickly becomes computationally expensive. Of course, range ambiguities can be eliminated by making the PRF sufficiently low. This introduces doppler ambiguities.

Doppler Ambiguities

In a similar manner to standard radar processing (see Section 2.1.1), doppler ambiguities can be removed by *jittering* the PRF.

Doppler ambiguities occur when the PRF is lower than the target doppler frequency. It is manifested by "ghost" targets in the data cube, which result from the repetition of the dop-

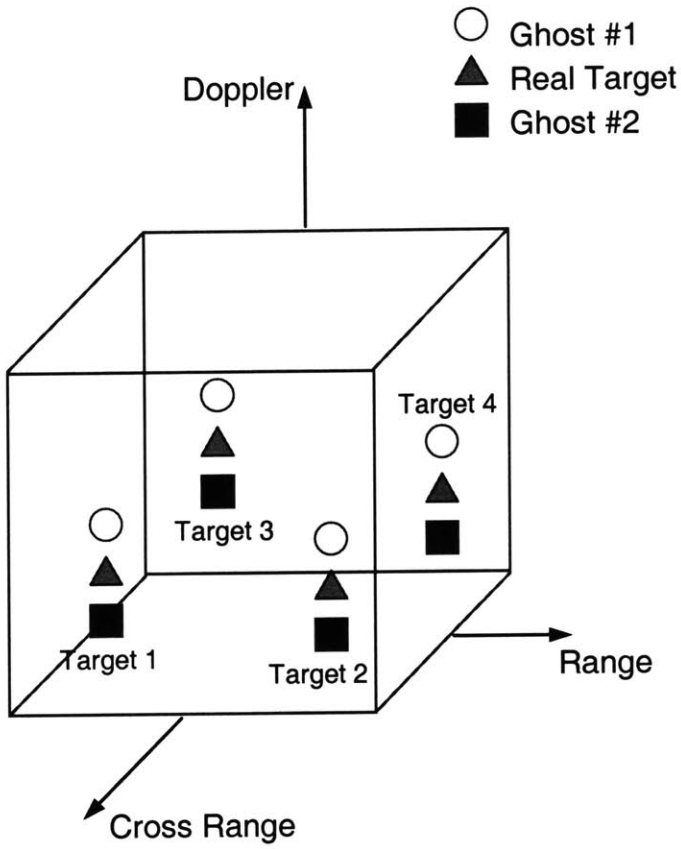

Figure 3.8 Using PRF Jitter to Eliminate Doppler Ambiguities

pler frequency spectrum at the PRF, as shown in Figure 3.8. The position of the ghosts along the doppler axis is dependent on the PRF, whereas the doppler position of the real target does not. By using multiple PRFs, ghosts can be identified by their tendency to move around. By non-coherently summing successive data cubes, entries corresponding to true targets will add up constructively, while entries corresponding to ghosts will not. The summed true targets will therefore have a higher amplitude than the ghosts.

The processing load using PRF jitter scales linearly with the number of selected PRFs (see Section 3.2.4). Therefore doppler ambiguity resolution using PRF jitter is less computationally intensive than range ambiguity resolution.

3.2.4 SPIR Processing Load

The expected processing load can be calculated based on the equations in Section 3.1.2.

To provide a feel for the numbers involved, we calculate the loads for a representative system, with the following values for the parameters defined in the processing load calculation:

TABLE 3.1 Representative System for Processing Load Calculation

Parameter	Description	Value
M	Samples per range gate	4
$N_{apertures}$	Number of apertures	5
N_{cross}	Number of cross-range cells	2000
N_{FFT}	Number of FFT points	1024
N_{PRF}	Number of PRFs	10
N_{pulses}	Number of pulses	1000
$N_{samples}$	Number of samples	4000
N_{range}	Number of range gates	2000

As in the derivation in Section 3.1.2, it is assumed that one satellite transmits and all the satellites receive. Computational loading for other schemes can be calculated in a similar manner to that shown below.

Processing Load for a Single Ambiguous Solution

We begin by calculating the load associated with solving the system (3.18) once.

The first step is to range gate the signals received by each aperture. Each range gate consists of a number of samples, while successive range gates may overlap. There are N_{range} range gates with M samples each, therefore range gating requires approximately $N_{range}M$ complex additions and N_{range} complex multiplications.

The next step is to delay each satellite's signal by the appropriate amount to focus on a particular cross-range cell. This is done by multiplying the range gated signal by a com-

plex exponential, as shown in equation (3.10). The values for the complex exponentials can be calculated beforehand. This step requires N_{range} complex multiplications per satellite. The total number of complex multiplications for this step is therefore $N_{apertures}N_{range}$.

The next step combines these signals by adding them. This requires $N_{samples}(N_{apertures}-1)$ complex additions.

For each range gate we take an N_{FFT} point discrete Fourier transform. The number of pulses taken per image is N_{pulses} and we can use the Fast Fourier Transform (FFT) by using zero-padding and setting

$$N_{FFT} = 2^{\lceil \log_2 N_{pulses} \rceil} \tag{3.21}$$

The FFTs across all the range gates require $N_{gates} \cdot \frac{N_{FFT}}{2}\log_2 N_{FFT}$ complex multiplications and $N_{range} \cdot N_{FFT}\log_2 N_{FFT}$ complex additions. Taking the absolute value squared for each FFT point in each range gate requires an additional $N_{FFT}N_{range}$ complex multiplications. These are the steps required to generate one cross-range entry across the range and doppler dimensions. Generating the composite signal $\mathbf{S}$ as given by (3.18) for each range gate results in a total of $N_{cross}N_{range}N_{FFT}$ points requiring a total of

$$\left[\left[N_{apertures} + \frac{N_{FFT}}{2}\log_2 N_{FFT} + N_{FFT}\right] \cdot N_{cross} + 1\right] \cdot N_{range} \tag{3.22}$$

complex multiplications and a total of

$$[[(N_{apertures}-1) + N_{FFT}\log_2 N_{FFT}] \cdot N_{cross} + M] \cdot N_{range} \tag{3.23}$$

complex additions.

Each complex multiplication requires four real multiplications and two real additions, while each complex addition requires two real additions. The total number of real multiplications to generate one composite signal is

$$4 \cdot \left[\left[N_{apertures} + \left(\frac{1}{2}\log_2 N_{FFT} + 1\right) \cdot N_{FFT}\right] \cdot N_{cross} + 1\right] \cdot N_{range} \tag{3.24}$$

and the total number of real additions is

$$2 \cdot \left[\left[2N_{apertures} - 1 + \left(\frac{3}{2}\log_2 N_{FFT} + 1\right) \cdot N_{FFT}\right] \cdot N_{cross} + M + 1\right] \cdot N_{range} \tag{3.25}$$

The PSF matrix can be inverted and stored before processing begins. Solving the system (3.18) for each range gate requires approximately $N_{range}N_{FFT}N^2_{cross}$ real multiplications and $N_{range}N_{FFT}N_{cross}(N_{cross} - 1)$ real additions.

Therefore the total processing load to solve the system once, without ambiguity resolution, is

$$\begin{aligned} RM &= 4\left[\left[N_{apertures} + \left(\frac{1}{2}\log_2 N_{FFT} + \frac{N_{cross}}{4} + 1\right)N_{FFT}\right]N_{cross} + 1\right]N_{range} \\ &\approx N^2_{cross}N_{range}N_{FFT} \qquad \text{for } N_{apertures} \text{ small} \end{aligned} \tag{3.26}$$

real multiplications, and

$$\begin{aligned} RA &= 2\left[\left[2N_{apertures} - 1 + \left(\frac{3}{2}\log_2 N_{FFT} + \frac{N_{cross} - 1}{4} + 1\right)N_{FFT}\right]N_{cross} + M + 1\right]N_{range} \\ &\approx \frac{1}{2}N^2_{cross}N_{range}N_{FFT} \qquad \text{for } N_{apertures} \text{ small} \end{aligned} \tag{3.27}$$

real additions. Increasing the number of range and/or cross-range cells has a dramatic effect on the processing load, which for an equal number of cells in each dimension, increases as the number of cells cubed. Large footprints with high resolution require a lot of computation. For the example system given in Table 3.1, where the number of cells in each dimension is 2000, approximately 8.1920×10^{12} real multiplications, and 4.0960×10^{12} real additions are required.

It is assumed that the system is designed to have either range or doppler ambiguities, but not both.

Doppler ambiguities are automatically resolved during non-coherent integration. Since non-coherent integration is required to increase the SNR in any case, Doppler ambiguity resolution comes at no extra computational cost. Less computationally demanding systems result if the PRF is chosen to avoid range ambiguities.

Computation Cost of Doppler Ambiguity Resolution

If doppler ambiguity resolution is used, and assuming that N_{PRF} PRFs are used to remove the ambiguities, the system (3.18) must be solved N_{PRF} times. Therefore we set

$$N_{solves} = N_{PRF} \tag{3.28}$$

In the case of doppler ambiguity resolution, non-coherent post-integration of N_{solves} solutions requires an additional $N_{cross}N_{range}N_{FFT}(N_{solves}-1)$ real additions. The total processing load is

$$N_{solves} \cdot RM \tag{3.29}$$

real multiplications and

$$N_{solves} \cdot RA + N_{cross}N_{range}N_{FFT}(N_{solves}-1) \tag{3.30}$$

real additions.

For the example system (Table 3.1), 8.1920×10^{13} real multiplications and 4.0997×10^{13} real additions are required.

3.2.5 Distributed SPIR Processing

Large amounts of data must be processed in order to generate high-resolution images using radar interferometry. Processing can be performed on-board or at ground stations. Processing the raw data at ground station(s) requires low error and high bandwidth communication to Earth. Furthermore, the cluster can only transmit data when it is in view of a ground station. Ground station processing places difficult requirements on the satellites' communication systems, and the image rate is constrained by the cluster visibility to

ground stations. If processing is performed on-board, only the reconstructed images need be transmitted to the ground. The communication link to the ground can be lower bandwidth, and images can be constructed even when the cluster is not in view of a ground station.

In the case of on-board processing, efficient distribution of the processing load across the satellites in the cluster is required to ensure minimal image generation delay as well as effective use of on-board resources. In this section we examine candidate communication architectures that enable distributed processing.

This discussion considers only homogenous clusters in which the processing capabilities of each satellite are identical. Centralised processing architectures are possible but are susceptible to single-point failures.

Three hardware capabilities are required for image generation: processing, memory, and communication. Processing capability indicates how "quickly" the image is generated and is usually measured in FLOPS (floating point operations per second). Memory capacity indicates how much data can be stored. Satellites typically have a combination of standard computer memory and mass-storage devices. Communication ability measures how much data can be transmitted (capacity) and by how many routes (connectivity). The requirement for each capability depends on the processing distribution scheme.

There is a fundamental inefficiency in distributing computation between multiple processors. Ideally, a task that takes a time T_s on a single processor would take time T_s/N when distributed between N processors. Unfortunately, there is always some overhead incurred in distribution. This can take the form of either extra time spent doing computation, or extra communication. Thus, parallel algorithms should only be employed when memory or timing requirements preclude the use of serial algorithms.

We begin by examining how SPIR processing can be distributed between satellites. Next, we propose two system architectures that illustrate different approaches to on-board processing.

SPIR Digital Processing

At a minimum, each satellite receives, performs any required analog conditioning, and digitizes the reflected signal. Once the received signals have been digitized radar processing can begin.

The first part of the digital radar processing is range gating. In general, range gating results in a reduction in the number of samples (Section 2.1.1). Furthermore, range gating can be performed independently by each satellite, it does not require knowledge of the signals received at other signals. Therefore range gating of a signal should be performed by the receiving satellite.

SPIR processing lends itself to distribution. Assuming there are N satellites, each satellite can either perform $\frac{1}{N}$ of the processing required to generate an image, or it can generate every N^{th} image. The first option promises a shorter lag between the initial reception of raw data and completion of the reconstructed image, but has a possible loss of efficiency due to the distribution overhead. The second option allows use of the most efficient serial algorithm, but there is a long lag between the initial reception of raw data and completion of the reconstructed image.

An independent cross-range/doppler data set is generated for each range gate. The processing load for a single image is conveniently distributed by assigning a set of range gates to each satellite.

The final processing step is thresholding (Section 2.1.1). CFAR thresholding schemes typically use information from adjacent cells to determine whether a particular cell contains a target or noise. If load distribution is by range gate, assigning a "block" of contiguous range gates to each satellite allows distributed thresholding of most range gates. Range

gates inside each block can be thresholded using only local data. An exchange of boundary data is required to permit thresholding of edge range gates. Figure 3.9 shows an example where the processing load to generate a single image is distributed among four satellites by range gate.

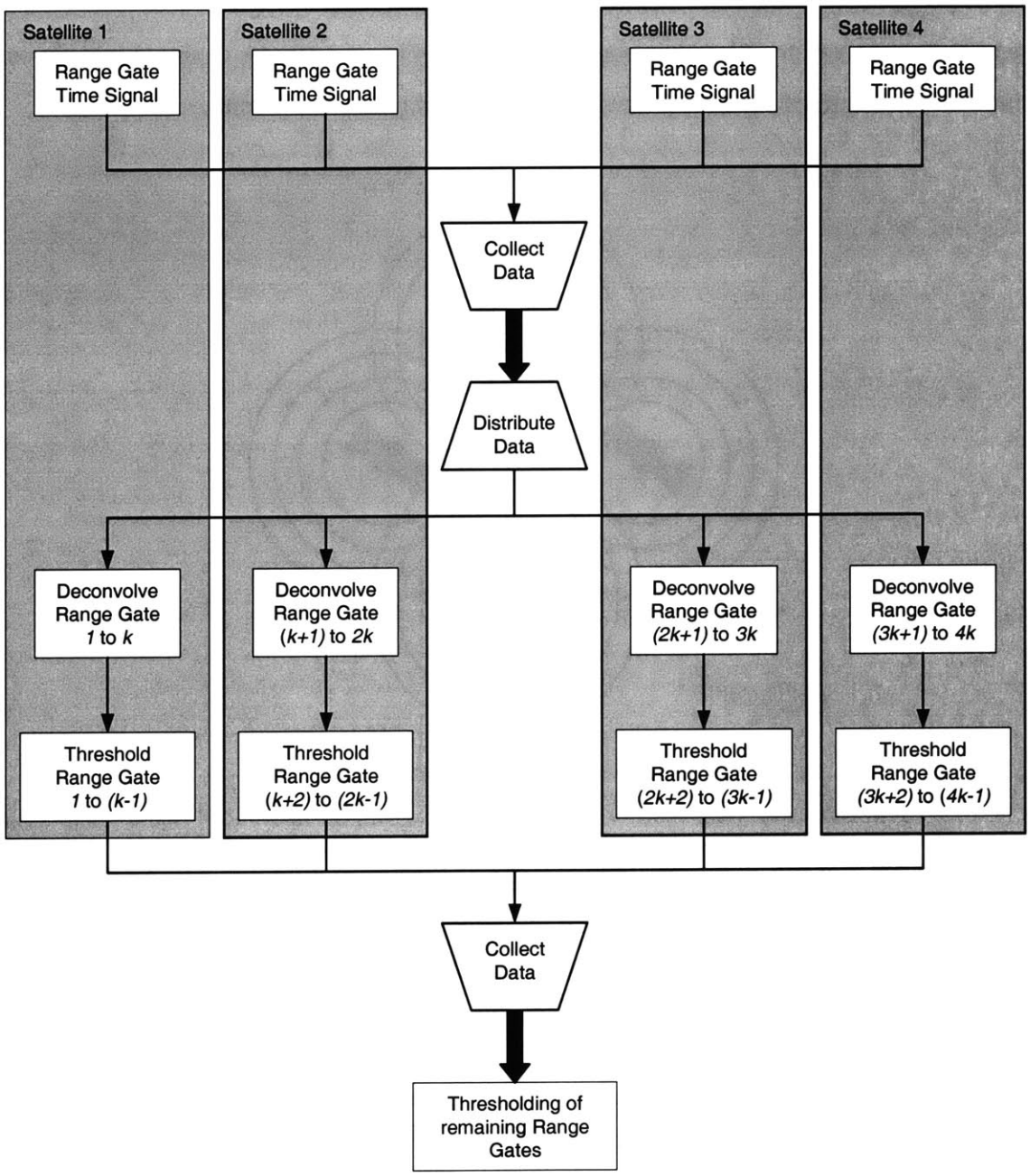

Figure 3.9 Processing Load Distribution by Range Gate

Distribution of the processing requires some kind of communication network between satellites. In the next two sections we discuss two popular architectures, broadcast and point-to-point.

Broadcast Communication

In a broadcast communication architecture, transmission from a single source node (satellite) can be received by all other nodes, as shown in Figure 3.10. An example is a cluster where all the satellites communicate using omnidirectional RF antennas.

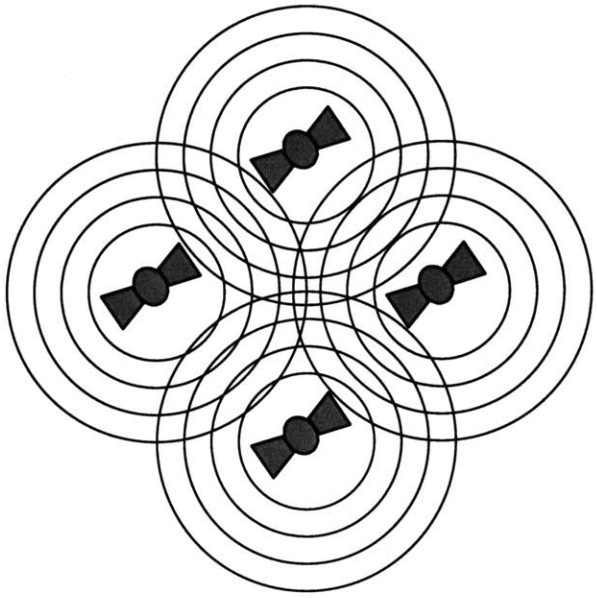

Figure 3.10 Broadcast Communication

Each node has a single high-rate transceiver, which handles in-going and out-going data traffic. Time, code, or frequency division multiple access must be used to create multiple data channels. Unless satellites are equipped with redundant transceivers, a transceiver failure will completely isolate a satellite from the cluster.

Since by the nature of the system each satellite receives data from all the other satellites, $\frac{1}{N}$ division of the processing is preferred.

Point-to-Point Communication

In a point-to-point communication scheme, fixed links connect pairs of processing nodes (satellites). This connectivity may be implemented with laser cross-links or directional RF antennas. Data travelling from one satellite to another may need to be relayed through several links. To eliminate single-point failures, each satellite should have more than one link. Nodes can be connected in simple, linear strings, or more complex network topologies. More sophisticated schemes introduce multiple links on each satellite to create redundant data pathways. For example, hypercube networks are easy to analyse and provide a good balance between the required number of links and robustness. Examples of four and eight element hypercubes are shown in Figure 3.11. A hypercube of N nodes has $N\log_2 N$ links. Each node has $\log_2 N$ links; non-adjacent nodes can be reached in a maximum of $\log_2 N$ hops.

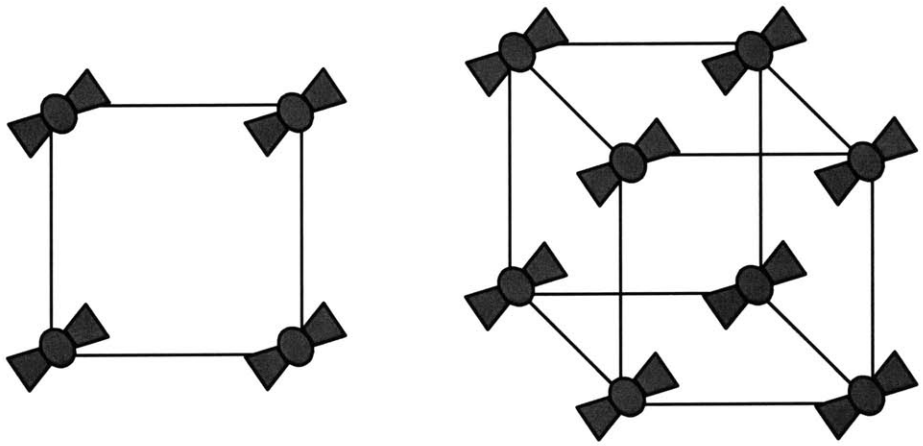

Figure 3.11 Four and Eight Element Hypercubes

Hypercube satellite communication networks are robust to failure of a transceiver, since each satellite has more than one transceiver. Data traffic to and from a satellite can be divided between the links, resulting in lower per-link traffic.

Although broadcast communication can be achieved in a point-to-point network, it is not recommended, as it requires data traffic duplication.

The choice between $\frac{1}{N}$ and N^{th} processing is not as clear as in broadcast communication, and requires more detailed study of system requirements.

3.3 Summary

This chapter has introduced a new method for processing multiple aperture GMTI radar data. Unlike other systems it does not require prior knowledge of the targets, and is independent of the clutter characteristics.

Commonly used radar processing techniques complement the SPIR method; their integration with SPIR was discussed.

High-resolution systems are computationally expensive. The processing load associated with SPIR in different configurations was determined, and ways of alleviating the system requirement by distributing the load between satellites were discussed.

In the next chapter we discuss how to design clusters to provide optimal SPIR performance.

Chapter 4

POINT SPREAD FUNCTION DESIGN

Section 3.1.2 has shown that it is indeed possible to recover the true ground scene by deconvolving the PSF of the array from the constructed return signal. We introduced the concept of a PSF matrix, in this chapter we discuss how to design the PSF for good SPIR performance.

This chapter begins by discussing characteristics of the PSF that can be used to predict SPIR performance. To gain an appreciation of the relationship between the PSF and the accompanying PSF matrix, we then discuss construction of the PSF matrix. An understanding of the construction process leads to desirable characteristics for the PSF. We use these to derive design parameters.

Minimum redundancy arrays are popular for interferometric applications. The mathematical background is discussed from an interferometry point of view. Unfortunately, MRAs do not result in invertible PSF matrices. We propose an alternative design procedure, based on the characteristics of the PSF.

4.1 PSF Performance Measures

Performance of a SPIR system is highly dependent on the PSF of the aperture array. Some performance measures of the PSF are resolution, field-of-view, and peak side lobe level.

The resolution is the smallest distance between two points at which they can be resolved as being separate. Points that are closer together than the angular separation of the first nulls of the PSF cannot be identified as separate points. The angular null separation defines the resolution of the aperture.

The requirement for high-resolution can be met by using a large baseline. Unfortunately, this decreases the spatial period of the array-factor as qualitatively discussed in Section 2.3.4. We shall see that array factors with short spatial periods may result in PSF matrices that are not full rank.

The field-of-view is determined by the width of the main lobe of the aperture response. The intersection of the field-of-view with the ground is defined as the interferometer footprint. A wide field-of-view permits coverage of a large ground area, but has higher processing requirements. In general, the field-of-view of an aperture is inversely proportional to its diameter.

Fading occurs when targets at the edge of the footprint are sampled with the outer parts of the main lobe, which have low relative gain. As a result, the targets may appear artificially small. Therefore, an aperture response with a sharply defined main lobe is preferred.

The peak side lobe level predicts the clarity of the image. Information that is located in the side lobes of the PSF overlays information in the main lobe, so that the reconstructed image is obscured. Low PSF side lobes result in clearer images.

Our goal is thus to obtain high-resolution PSFs with long spatial periods and low side lobes.

4.2 Construction of the PSF Matrix

The PSF matrix is defined by (3.20). It is a square matrix with the number of rows and columns equal to the number of cross-range cells.

The PSF matrix is constructed by placing the main lobe of the array factor on successive cross-range cells in the footprint. The footprint is defined by the intersection of the aperture response main lobe with the ground. The aperture response remains at a constant position on the ground. Figure 4.1 illustrates the construction of a 4×4 PSF matrix.

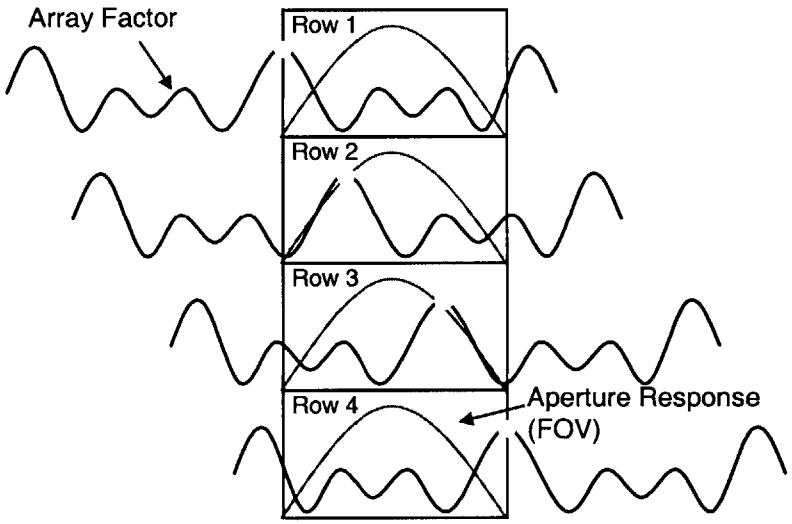

Figure 4.1 Construction of a 4×4 PSF Matrix

In this case the period of the array factor is greater than the field-of-view, therefore each row of the resulting PSF matrix is unique. Each point on the ground will be sampled once, and only once, by the main lobe of the PSF.

Figure 4.2 illustrates the effect of an array factor that repeats within the field-of-view. The last row of the PSF matrix is identical to the first row. Shifting the PSF to sample the point at very the right of the footprint also samples the point at the very left of the footprint. The result is ambiguity in the cross-range direction.

In Section 4.3.1 we determine the relationship between the array construction and the PSF period.

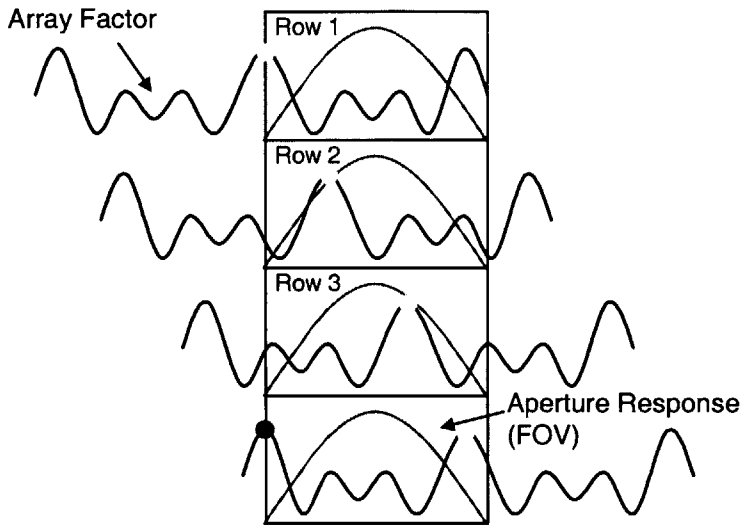

Figure 4.2 Construction of a 4×4 Ambiguous PSF Matrix

4.3 PSF Design Parameters

4.3.1 PSF Spatial Period

As we have seen the spatial period of the PSF determines whether or not the PSF matrix will be invertible. In this section we develop an expression for the spatial period, based on the aperture positions in a one-dimensional array.

A linear aperture array can be uniquely defined by specifying the relative aperture spacing (configuration) and some characteristic length l, such as baseline or minimum spacing. At a given wavelength λ, a fully specified array samples a particular set of spatial frequencies. These spatial frequencies scale inversely with l and proportionally with λ. Therefore the spatial period ψ_p of the array factor scales according to

$$\psi_p \propto \frac{\lambda}{l} \tag{4.1}$$

In particular,

$$\psi_p \propto \frac{\lambda}{B} \tag{4.2}$$

where B is the array baseline.

The constant of proportionality can be derived from the array configuration and is determined as follows:

The array factor (2.52) of a linear array is the sum of a sequence of complex exponential terms, which each have a maximum projection of one onto the real axis. Therefore the array factor is bounded above by the total number of apertures, N

$$array\_factor(\psi) = \left| \sum_{n=1}^{N} e^{-2\pi j \frac{x_n}{\lambda} \sin\psi} \right| \leq 1 \cdot N \tag{4.3}$$

By taking the squared norm of the array factor, the interaction between the complex exponential terms can be seen

$$\begin{aligned} |array\_factor(\psi)|^2 &= \left[\sum_{n=1}^{N} e^{-2\pi j \frac{x_n}{\lambda} \sin\psi} \right] \left[\sum_{n=1}^{N} e^{2\pi j \frac{x_n}{\lambda} \sin\psi} \right] \\ &= \left[\sum_{m=1}^{N} \sum_{n=1}^{N} \cos\left(2\pi j \frac{|x_m - x_n|}{\lambda} \sin\psi \right) \right] \leq N^2 \end{aligned} \tag{4.4}$$

Each aperture pair spacing $|x_m - x_n|$ samples a spatial frequency. The maxima ψ_{mn} for a particular spatial frequency occur when

$$2\pi \frac{|x_m - x_n|}{\lambda} \sin\psi_{mn} = 2\pi k_{mn} \qquad (m \neq n) \in \{1 \ldots N\} \tag{4.5}$$

where k_{mn} is a positive integer referred to as the spatial frequency number.

Solving for $\sin\psi_{mn}$

$$\sin\psi_{mn} = \lambda \frac{k_{mn}}{|x_m - x_n|} \qquad (m \neq n) \in \{1 \ldots N\} \tag{4.6}$$

Successive values for k_{mn} give the positions of successive maxima corresponding to the spatial frequency sampled by the aperture pair (x_m, x_n).

The array factor global maxima occur when the maxima of all the sampled spatial frequencies coincide, that is, whenever all the cosines evaluate to one. The first global maximum occurs at $\psi = 0$; successive global maxima occur at multiples l of the period ψ_p. Setting $\psi_{mn} = l\psi_p$ and eliminating $\sin(l\psi_p)$ from (4.6) gives a set of relations between the aperture positions and the k_{mn}

$$\frac{k_{12}}{|x_1 - x_2|} = \frac{k_{23}}{|x_2 - x_3|} = \ldots = \frac{k_{1N}}{|x_1 - x_N|} \tag{4.7}$$

which can be rewritten as

$$\frac{k_{mn}}{k_{pq}} = \frac{|x_m - x_n|}{|x_p - x_q|} \qquad (m \neq n) \in \{1 \ldots N\} \tag{4.8}$$

The set $\{\hat{k}_{mn}\}$ of smallest k_{mn} that will satisfy equations 4.6 with $\psi_{mn} = \psi_p$ is determined by simplifying each $\frac{|x_m - x_n|}{|x_p - x_q|}$ to irreducible rational form, such that the greatest common divisor for each pair (k_{mn}, k_{pq}) is minimised. If at least one k_{mn} does not have a GCD of one with all the other k_{pq}, then the entire set is reducible and the period of the array is reduced according to (4.6).

ψ_p is found by solving equations 4.6 with the set $\{\hat{k}_{mn}\}$

$$\sin\psi_p = \lambda\frac{\hat{k}_{12}}{|x_1 - x_2|} = \ldots = \lambda\frac{\hat{k}_{1N}}{|x_1 - x_N|} \tag{4.9}$$

Recognising that $|x_N - x_1|$ is the array baseline B, and setting

$$k_{max} = \hat{k}_{1N} \tag{4.10}$$

the array factor period can be found from

$$\sin\psi_p = k_{max} \cdot \frac{\lambda}{B} \tag{4.11}$$

For small ψ_p

$$\psi_p \approx k_{max} \cdot \frac{\lambda}{B} \tag{4.12}$$

Alternatively, defining the minimum spacing as x_{min}, and setting

$$k_{min} = \min[\{\hat{k}_{mn}\}] \tag{4.13}$$

the array factor period can also be found from

$$\sin\psi_p = k_{min} \cdot \frac{\lambda}{x_{min}} \tag{4.14}$$

If k_{min} is greater than one, the behaviour of the array factor is less well defined. Interaction between different spatial frequencies may result in low frequency terms that decrease the period of the array factor from the design value ψ_p. Therefore arrays with $k_{min} = 1$ are preferred.

4.4 Minimum Redundancy Arrays

Minimum Redundancy Arrays (MRAs) are a type of linear array popular for interferometric applications. In this section we discuss the theory of minimum redundancy and its application to interferometry. We shall see that, contrary to expectation, MRAs do not result in invertible PSF matrices for large-baseline arrays.

Since an interferometer constructs an image by sampling spatial frequencies, it is to be expected that image quality can be improved by sampling a large number of spatial frequencies in each dimension. Performance in different dimensions can be decoupled, therefore we consider the problem of designing a linear aperture array that samples all the

desired spatial frequencies while minimising duplication. Such an array is deemed to display *minimum redundancy.*

The simplest type of linear array has identically spaced apertures. We refer to these arrays as *regular* linear arrays. By analogy with optical diffraction gratings, the term *grating array* is also used. The (u, v) sample space of a linear aperture array is defined by the set of all differences between the elements in the array. Regular linear arrays will therefore duplicate many spatial frequencies. This section discusses the design and evaluation of arrays that display minimum redundancy.

4.4.1 Difference Bases

Leech [Leech, 1956] investigated this problem from a purely number theoretic viewpoint; this section summarises the relevant parts of his work.

A set of integers $(a_1, a_2, \ldots, a_k)$ is called a *difference basis* with respect to a difference base generator n if every positive integer v such that $0 < v \leq n$ can be represented in the form $v = a_i - a_j$. Let $k(n)$ denote the minimum value of k for given n. It can be shown that

$$2.424\ldots \leq \lim_{n \to \infty} \frac{k^2(n)}{n} \leq 2.664\ldots \tag{4.15}$$

A *restricted* difference basis $(a_1, a_2, \ldots, a_l)$ with respect to n satisfies

$$0 = a_1 < a_2 < \ldots < a_l = n \tag{4.16}$$

In this case $l(n)$ denotes the minimum value of l for given n. It can be shown that

$$2.434\ldots \leq \lim_{n \to \infty} \frac{l^2(n)}{n} \leq 3.348\ldots \tag{4.17}$$

The limits given in equations 4.15 and 4.17 place bounds on the efficiency with which large sparse arrays can sample spatial frequencies, as we shall see in the next section.

Construction of difference bases is not trivial, and becomes increasingly difficult as n increases. Table 4.1 and Table 4.2 show general (unrestricted) and restricted difference

TABLE 4.1 Unrestricted Difference Bases, up to 11 Apertures

Apertures	n	Configuration	Redundancy
1	0	{✱}	0
2	1	{✱1✱}	0
3	3	{✱1✱2✱}	0
4	6	{✱1✱3✱2✱}	0
5	9	{✱3✱1✱5✱2✱} {✱4✱1✱2✱6✱}	0.1111
6	13	{✱4✱1✱1✱7✱3✱} {✱6✱1✱2✱2✱8✱} {✱1✱3✱6✱2✱5✱} {✱1✱7✱3✱2✱4✱}	0.1538
7	18	{✱6✱3✱1✱7✱5✱2✱} {✱8✱1✱3✱6✱5✱2✱} {✱14✱1✱3✱6✱2✱5✱} {✱13✱1✱2✱5✱4✱6✱}	0.1667
8	24	{✱8✱10✱1✱3✱2✱7✱8✱}	0.1667
9		*none*	
10	37	{✱16✱1✱11✱8✱6✱4✱3✱2✱22✱} {✱7✱15✱5✱1✱3✱8✱2✱16✱7✱}	0.2162
11	45	{✱18✱1✱3✱9✱11✱6✱8✱2✱5✱28✱}	0.2222

TABLE 4.2 Restricted Difference Bases, up to 11 Apertures

Apertures	n	Configuration	Redundancy
1	0	{✱}	0
2	1	{✱1✱}	0
3	3	{✱1✱2✱}	0
4	6	{✱1✱3✱2✱}	0
5	9	{✱1✱3✱3✱2✱} {✱1✱1✱4✱3✱}	0.1111
6	13	{✱1✱1✱4✱4✱3✱} {✱1✱5✱3✱2✱2✱} {✱1✱3✱1✱6✱2✱}	0.1538
7	17	{✱1✱1✱4✱4✱4✱3✱} {✱1✱1✱1✱5✱5✱4✱} {✱1✱1✱6✱4✱2✱3✱} {✱1✱1✱6✱4✱3✱2✱}	0.2353
8	23	{✱1✱1✱9✱4✱3✱3✱2✱} {✱1✱3✱6✱6✱2✱3✱2✱}	0.2174
9	29	{✱1✱1✱12✱4✱3✱3✱3✱2✱} {✱1✱3✱6✱6✱6✱2✱3✱2✱} {✱1✱2✱3✱7✱7✱4✱4✱1✱}	0.2414
10	36	{✱1✱2✱3✱7✱7✱7✱4✱4✱1✱}	0.2500
11	43	{✱1✱2✱3✱7✱7✱7✱7✱4✱4✱1✱}	0.2791

bases for $n = 3, 4, \ldots, 11$. The *numbers* (aperture positions) in each basis are represented by ✱ and the *differences* (aperture spacings) between them by the numbers in the table entries.

4.4.2 Application to Linear Antenna Arrays

We refer to a linear aperture array corresponding to an unrestricted difference basis as an *unrestricted minimum redundancy* array. Such an array maximises the *total* number of sampled spatial frequencies. For example, given the six-element unrestricted difference base with elementary differences $\{3;1;5;2\}$, the complete set of difference pairs is $\{1 \cdot 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8 \cdot 9 \cdot \cdot 11\}$. Since 9 is the largest contiguous number, this difference base is defined with respect to 9. However, an additional spatial frequency, corresponding to 11, is also sampled.

Similarly, we refer to a linear aperture array corresponding to a restricted difference basis as a *restricted minimum redundancy array.* Such an array maximises the total number of contiguous sampled spatial frequencies, but does not guarantee uniqueness of the sampled spatial frequencies. For example, given the five-element restricted difference base with elementary differences $\{1;3;3;2\}$, the complete set of difference pairs is $\{1 \cdot 2 \cdot 3 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8 \cdot 9\}$. Like the five-element unrestricted difference base above, this base is defined with respect to 9. In this case, however, 9 is the largest difference arising from the base, and the spatial frequency corresponding to 3 is sampled twice.

In order to compare different minimum redundancy arrays, we define two metrics, *spread* and *redundancy.*

Spread

We define the *spread* of an aperture array as the normalised largest contiguous spatial frequency sampled by the array. The normalisation factor is the smallest difference pair, this permits comparison between arrays with different baselines.

In the case of minimum-redundancy arrays the spread corresponds to n; for regular linear arrays the spread is equal to the number of array elements minus one. Figure 4.3 shows the spread for both regular linear and minimum-redundancy arrays of up to eleven elements. Minimum-redundancy arrays achieve a larger spread than regular arrays for the same number of elements, as expected, since the spacing between elements is not restricted to be uniform.

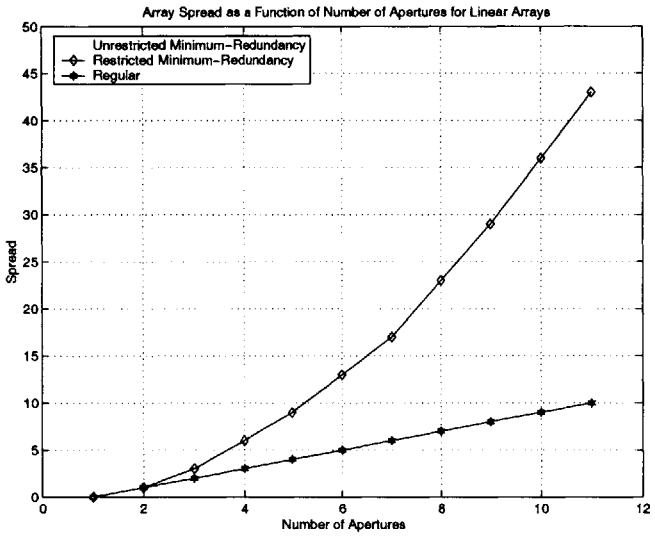

Figure 4.3 Array Spread

Redundancy

We define the array *redundancy*, R, of an array as follows

$$R = \frac{n_{pairs}}{n} - 1 \tag{4.18}$$

where n_{pairs} is the number of difference pairs (not necessarily distinct) and n is the associated difference base generator.

Zero redundancy indicates that each difference pair is unique, i.e. no spatial frequencies are sampled more than once. Bracewell [Bracewell, 1962] provides an elegant proof that there are only four linear arrays with zero redundancy; these are given in Table 4.1.

The limits in equations 4.15 and 4.17 suggest that as the number of array elements is increased, the minimum achievable value for R increases, therefore large arrays will always have some degree of redundancy.

For a regularly spaced linear array of $N > 1$ elements, the redundancy is

$$R_{regular} = \frac{N}{2} - 1 \qquad N > 1 \tag{4.19}$$

Linear arrays become increasingly redundant as their size is increased.

Given a difference basis $(a_1, a_2, \ldots, a_k)$ with respect to n, the number of difference pairs is $\frac{1}{2}k(k-1)$. Moffet [Moffet, 1968] gives the *redundancy*, R, of a minimum-redundancy linear aperture array as

$$R_{Moffet} = \frac{\frac{1}{2}k(k-1)}{n} \tag{4.20}$$

Under our definition the redundancy of a minimum-redundancy linear array is therefore

$$R_{MRA} = \frac{\frac{1}{2}k(k-1)}{n} - 1 \tag{4.21}$$

Figure 4.4 shows the relationship between number of apertures and redundancy for minimum-redundancy and regular arrays; these numbers are also provided in Table 4.1 for reference.

The performance of linear arrays in terms of spread and redundancy degrades linearly with an increase in the number of apertures. Unrestricted arrays provide slightly better

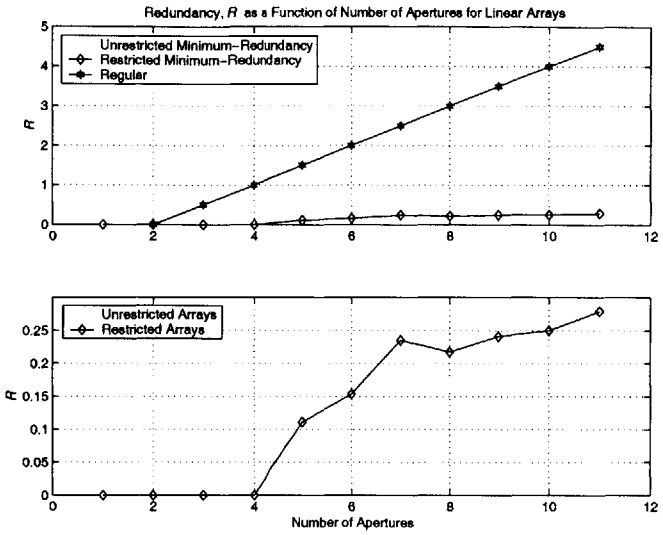

Figure 4.4 Array Redundancy

redundancy and spread performance than restricted arrays but generally have longer baselines.

Solutions for minimum redundancy arrays with a large ($k \gg 11$) number of apertures are discussed by [Ishiguro, 1980]. We do not consider such arrays because their cost makes them infeasible at the present time.

4.4.3 Application of MRAs to SPIR

Examination of the PSFs corresponding to MRAs reveals that they have periods that are significantly lower than the field-of-view. MRAs are constructed to sample spatial frequencies with minimum redundancy. Since all MRAs have $k_{min} = 1$, k_{max} is always relatively small. The period of the PSF is therefore small, and multiple points in the footprint are simultaneously sampled by array factor mainlobes. This causes ambiguity in the constructed signal.

4.5 A Design Approach to Reduce Ambiguity

Armed with equation 4.11, we can design an array that has a PSF with a given spatial period and hence reduce ambiguity in the constructed signal. The first step is to determine k_{max}, based on the system requirements. The number of apertures N is chosen based on high-level system requirements. Given k_{max} and N, the positions of the remaining $(N-2)$ apertures must be determined.

4.5.1 Determining the Maximum Spatial Frequency Number

The field-of-view of an aperture can be expressed as

$$FOV = 2\alpha \cdot \frac{\lambda}{D} \tag{4.22}$$

where D is the aperture diameter and α is a shape factor such that a rectangular aperture of width

$$D_{eff} = \frac{D}{\alpha} \tag{4.23}$$

has the same field-of-view as the actual aperture. For uniformly illuminated rectangular apertures $\alpha = 1$ and for uniformly illuminated circular apertures $\alpha = 1.22$.

The PSF matrix will be invertible if the period of the array factor is greater than the field-of-view of the aperture

$$\psi_p > FOV \tag{4.24}$$

Substituting equation 4.12 into this relation gives

$$k_{max} > FOV \cdot \frac{B}{\lambda} \tag{4.25}$$

The baseline is chosen to satisfy the cross-range resolution requirement Δ_{cross}, according to

$$B = \lambda \cdot \frac{r_{ground}}{\Delta_{cross}} \tag{4.26}$$

where r_{ground} is the range to the ground from the array.

The limiting case is taken by setting

$$k_{max} = \text{smallest ingeger} > FOV \cdot \frac{B}{\lambda} \tag{4.27}$$

Once k_{max} is determined, B and/or FOV must be adjusted to satisfy

$$FOV \cdot \frac{B}{\lambda} = k_{max} - 1 \tag{4.28}$$

By substituting (4.22) into the design relation (4.27), we obtain

$$k_{max} = \text{smallest ingeger} > \frac{B}{D_{eff}/2} \tag{4.29}$$

so that k_{max} is independent of the wavelength.

Given k_{max} and the number of apertures, how do we place the remaining $(N-2)$ apertures? In the next section we introduce a type of array that is guaranteed to result in a PSF with the desired period.

4.5.2 Optimally Irreducible Arrays

The period of the array factor is independent of the position of the array in space, therefore we place the first aperture at $k_1 = 0$, and the N^{th} aperture at $k_N = k_{max}$, with no loss of generality. From the earlier discussion (Section 4.3.1) on the reducibility of the k_{mn}, the only way to ensure irreducibility is to place the remaining $(N-2)$ apertures such that

$$k_{min} = 1 \tag{4.30}$$

The resulting array configuration, $\{k_1;k_2;\ldots;k_N\}$, will have an array factor with period ψ_p given by (4.11). We will refer to such an array as being (k_{max}, N) optimally irreducible. A (k_{max}, N) array that does not satisfy (4.30) will have a smaller period than the desired design value ψ_p.

The requirement that $k_{min} = 1$ means that the two most closely spaced apertures physically overlap. This limitation can be overcome by using phased-array antennas.

Arrays that have the same PSF are referred to as being *similar*; this is indicated by $array_1 \sim array_2$. Arrays that are mirror images of each other

$$\{0;k_2;\ldots;k_{N-1};k_{max}\} \sim \{0;k_{max}-k_{N-1};\ldots;k_{max}-k_2;k_{max}\} \tag{4.31}$$

are similar and are discarded when determining the configuration solution space.

With $N = 2$ apertures, there is only one difference pair $k_{12} = k_{max}$ and the irreducibility condition requires $k_{max} = k_{min} = 1$. Therefore there is one unique solution, $\{0;1\}$. The array factor has the same shape regardless of λ and the baseline. The period is given by

$$\psi_p\big|_{N=2} = \frac{\lambda}{B} \tag{4.32}$$

However, the design relation (4.29) between k_{max} and B implies that for $k_{max} = 1$ we must have $D_{eff} = 2B$, which is physically impossible. Therefore the two-element optimally irreducible array cannot be implemented in a real system.

With $N = 3$ apertures, one unique solution exists for each k_{max}, namely $\{0;1;k_{max}\}$. It is always possible to design an array with the required period, but the shape of the array factor is defined by k_{max}.

For $4 \le N \le k_{max}$ there is always at least one solution. The additional degrees of freedom provided by the fourth and higher apertures allows shaping of the array factor to the most desired form.

4.5.3 Aperture Placement

Minimum redundancy arrays provide good spatial frequency sampling, while optimally irreducible arrays guarantee invertible PSF matrices. All minimum redundancy arrays have $k_{min} = 1$, by appending an extra aperture to provide the desired k_{max}, an array with good spatial frequency sampling and PSF matrix characteristics should emerge.

$$\left[MRA \mid k_{max}\right] \tag{4.33}$$

4.6 Evaluation of Arrays

The design method suggested in the previous section guarantees that the resulting PSF has the desired period. Intuitively, placing an MRA inside the optimally irreducible array should result in good spatial frequency sampling. We begin by selecting a nominal set of system requirements to be used as a baseline when comparing different array designs, as shown in Table 4.3.

TABLE 4.3 System Parameters

Parameter	Value
Carrier wavelength λ (m)	0.03
Desired field-of-view ψ_p	0.06
Cross-range resolution Δ_{cross_range} (m)	100
Range to ground r_{ground} (km)	1000

Using the design relations given in Section 4.5.1, the baseline and k_{max} are

$$\begin{aligned} B &= 600 \text{ m} \\ k_{max} &= 1201 \end{aligned} \tag{34}$$

Note that the *shape* of the PSF array factor is independent of the baseline and wavelength and is entirely determined by the placement of apertures relative to one another, as discussed in Section 2.3.4.

The simplest array consists of two apertures, which requires $k_{max} = 1$ and is physically impossible, as shown in Section 4.5.2. We therefore begin with the smallest optimally irreducible array, consisting of three apertures.

4.6.1 Three Apertures

With three apertures, there is one unique array, namely $\{0;1;1201\}$. This corresponds to the two element minimum redundancy array with k_{max} appended at the end. The array factor is shown in Figure 4.5. The array factor has a clearly discernible envelope, with period equal to the design value ψ_p. The highest spatial frequency that is sampled is λ/B and gives rise to the high frequency variation inside the envelope.

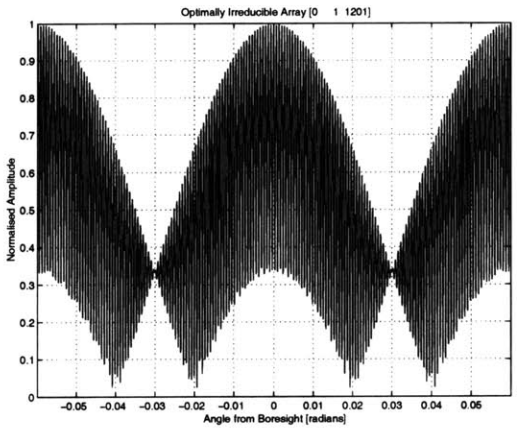

Figure 4.5 $\{0;1;1201\}$ Three Element Optimally Irreducible Array

4.6.2 Four Apertures

With four apertures there is one additional degree of freedom, which gives rise to a large solution space. To gain an appreciation for the variability in the array factor, we begin by arbitrarily placing the free aperture. Some examples are shown in Figure 4.6 through Figure 4.8.

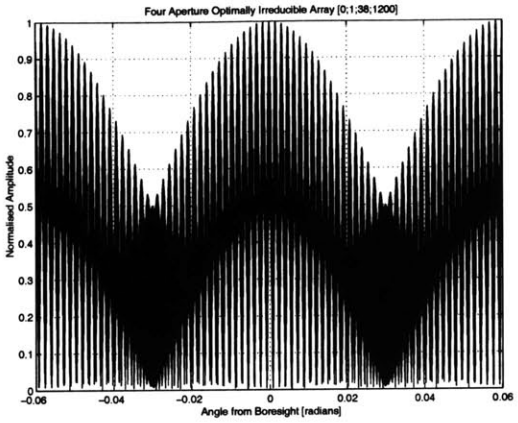

Figure 4.6 {0;1;38;1201} Four Element Optimally Irreducible Array

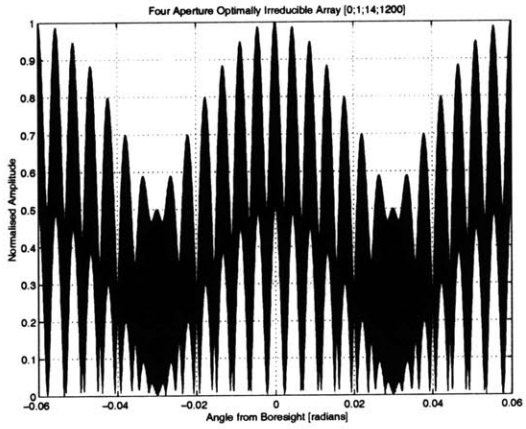

Figure 4.7 {0;1;14;1201} Four Element Optimally Irreducible Array

In each case there is a clearly discernible envelope, with period equal to the design value ψ_p. By changing the position of the third aperture in the array, the shape of the envelope is changed and additional lower frequency terms are introduced.

The {0;1;3;1201} array shown in Figure 4.8 has the narrowest envelope main lobe of the examples. Closer examination reveals that it contains a three-element MRA, with aperture positions {0;1;3}. What happens if we move the free aperture slightly to the left or the right? The resulting array factors are shown in Figure 4.9 and Figure 4.10.

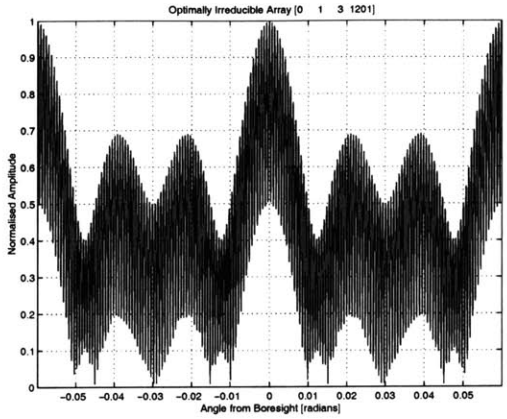

Figure 4.8 {0;1;3;1201} Four Element Optimally Irreducible Array

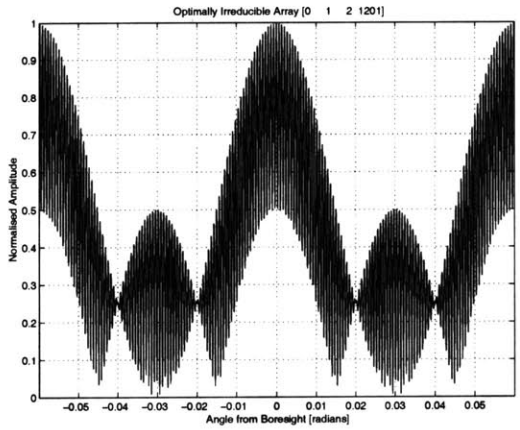

Figure 4.9 {0;1;2;1201} Four Element Optimally Irreducible Array

Moving the free aperture one elementary spacing to the left decreases the envelope side lobes, but increases the width of the main lobe. Conversely, moving the free aperture one elementary spacing to the right increases the envelope side lobes, but decreases the width of the main lobe. The array containing the three-element MRA is a compromise between these two configurations: the main lobe is fairly narrow while the side lobes are not excessively high.

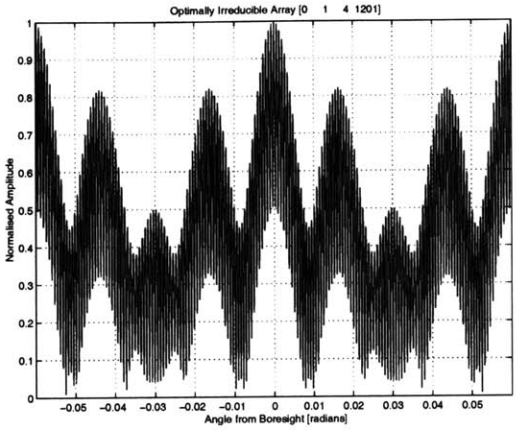

Figure 4.10 {0;1;4;1201} Four Element Optimally Irreducible Array

4.6.3 Five Apertures

With five apertures there are two additional degrees of freedom, which gives rise to an even larger solution space. Once again, we begin by arbitrarily placing the two free apertures. Two examples are shown in Figure 4.11 and Figure 4.12.

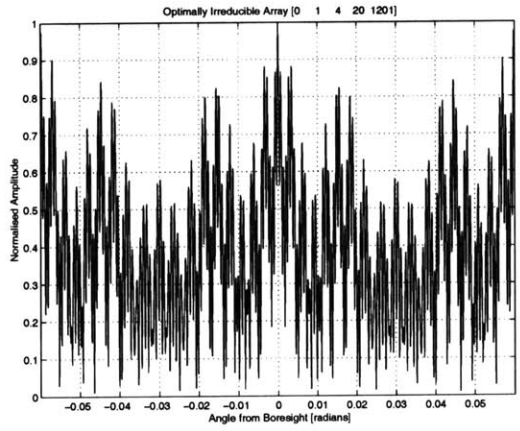

Figure 4.11 {0;1;4;20;1201} Five Element Optimally Irreducible Array

The main lobe in both cases is not well defined, while the side lobes are high. Can we obtain a "better" array factor using a four element MRA for the first four apertures? The

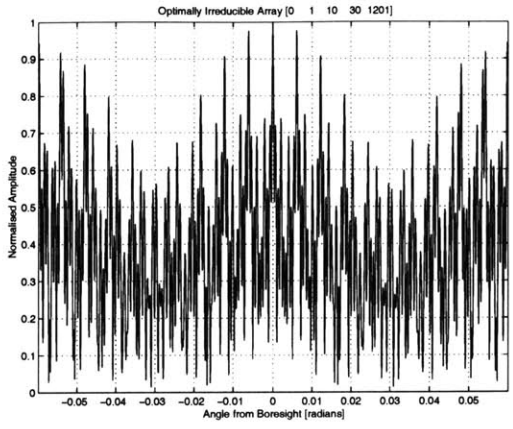

Figure 4.12 {0;1;10;30;1201} Five Element Optimally Irreducible Array

resulting array factor is shown in Figure 4.13. The main lobe is well defined and narrow and the side lobes are quite low. Once again, the configuration containing a minimally redundant subarray yields the best results.

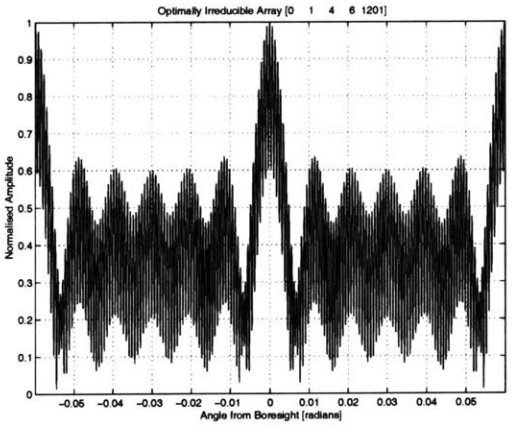

Figure 4.13 {0;1;4;6;1201} Five Element Optimally Irreducible Array

Appendix A provides results for optimally irreducible arrays of up to twelve elements.

4.6.4 Discussion

Intuition indicates that optimally irreducible arrays with minimum redundancy subarrays should yield PSF with the desired period and with good spatial frequency sampling characteristics. Minimum redundancy arrays of up to eleven elements exist and are tabulated in Table 4.1 and Table 4.2. Qualitative comparison of these arrays with optimally irreducible arrays having non-minimum redundancy subarrays indicates that minimum redundancy subarrays also result in PSFs with narrow main lobes and low side lobes. Although PSFs with narrower main lobes or lower side lobes can be found, the suggested arrays provide the best compromise between main lobe width and side lobe level. Increasing the number of apertures decreases the width of the main lobe and decreases the side lobe levels.

4.7 Summary

The physical configuration of the satellite cluster directly impacts SPIR performance. Beginning with an analysis of the point spread function (PSF), we obtained an understanding for the relation between PSF shape and cluster configuration. Minimum redundancy arrays are regarded as good candidates for linear interferometric system clusters; we showed that they can introduce ambiguity in the constructed signal.

An alternative design approach, based on the PSF characteristics, was proposed. The resulting arrays are guaranteed to have PSFs with the desired period. The design approach specifies the position of three apertures in the cluster. Any remaining apertures can be placed to yield the desired PSF shape. Placing these apertures so that the $N_{aperture}$ array has a minimum redundancy $(N_{aperture}-1)$ subarray results in PSFs with narrow main lobes and low side lobes.

Chapter 5

SPIR SYSTEM MODELING

As we have seen in Chapter 3, deconvolution reveals the true ground scene, at least in the ideal case. In this chapter we discuss the numerical issues related to performing deconvolution.

5.1 Solution of the SPIR System

In this section we re-examine the SPIR problem from a linear algebra perspective. We begin with an interpretation of deconvolution as a matrix operation. Next, we discuss solution of the system using conventional linear algebra techniques. In practice SPIR systems are so poorly conditioned that these techniques do not work. Section 5.2 presents a possible solution method.

5.1.1 SPIR as a Matrix Operation

The rows of the constructed signal matrix, $\mathbf{S}$, are created by superposing scaled and shifted versions of the complex interferometer response (the point spread function in complex form). A target or clutter ground cell located at cross-range angle θ_1 and frequency index k_1 with RCS $\sigma(\theta_1, k_1)$ creates a row in $\mathbf{S}$ at frequency index k_1, as defined by (3.15)

$$s(\theta_q,k_1) = \frac{1}{R^2}\sigma(\theta_1,k_1)\left(E(\theta_1)\sum_{i=1}^{N_{sats}} e^{-j2\pi u_i(\sin\theta_1 - \sin\theta_q)} \right) \qquad \theta_q \in FOV \qquad (5.1)$$

The term in parentheses is the complex PSF, shifted by $\sin\theta_1$. Additional targets at the same frequency index k_1 add additional shifted and scaled versions of the complex PSF to $s(\theta_q,k_1)$

$$s(\theta_q,k_1) = \frac{1}{R^2}\sum_{p=0}^{N_{cells}-1} \sigma(\theta_p,k_1)\left(E(\theta_p)\sum_{i=1}^{N_{sats}} e^{-j2\pi u_i(\sin\theta_p - \sin\theta_q)} \right) \qquad \theta_q \in FOV \qquad (5.2)$$

Each row of the complex PSF matrix is also a shifted version of the complex PSF. Therefore the target and clutter cross-range/doppler positions can ideally be recovered by multiplying $\mathbf{S}$ with the inverse of the complex PSF matrix.

A simple example serves to illustrate this concept. Define a 3×3 matrix $\mathbf{A}$ with cyclically permuted rows

$$\mathbf{A} = \begin{bmatrix} 8 & 1 & 6 \\ 1 & 6 & 8 \\ 6 & 8 & 1 \end{bmatrix} \qquad (5.3)$$

This represents the PSF matrix of the SPIR system.

Now define a 3×5 matrix $\mathbf{B}$ such that

$$\mathbf{B} = \begin{bmatrix} 8 & 1 & 8 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 6 & 0 & 6 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 6 & 0 & 0 \\ 6 & 0 & 8 & 0 & 0 \\ 8 & 0 & 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 9 & 0 & 14 & 0 & 0 \\ 7 & 0 & 9 & 0 & 0 \\ 14 & 0 & 7 & 0 & 0 \end{bmatrix} \qquad (5.4)$$

This represents the constructed signal for a particular ground scene.

The first column of $\mathbf{B}$ is the sum of the first and second rows of $\mathbf{A}$, while the third column is the sum of the first and third rows of $\mathbf{A}$. The position and contents of the rows of $\mathbf{B}$ can be recovered by solving $\mathbf{AP} = \mathbf{B}$. If the inverse of $\mathbf{A}$ is well-defined,

$$\mathbf{P} = \mathbf{A}^{-1}\mathbf{B} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \tag{5.5}$$

The columns of $\mathbf{P}$ correspond to columns of $\mathbf{B}$. The rows of $\mathbf{P}$ correspond to entries of $\mathbf{A}$.

The row position of each non-zero element indicates the row of $\mathbf{B}$ that contains a particular entry, while the column position indicates which entry of $\mathbf{A}$ it contains. In a SPIR system, the column position of each non-zero element represents the doppler shift, while the row position represents the cross-range position, of each target.

5.1.2 Deconvolution and System Conditioning

SPIR images are constructed in a similar way, using (3.18), restated here for convenience

$$\mathbf{S} = \frac{1}{R^2}\mathbf{G}\Sigma \tag{5.6}$$

Provided that $\mathbf{G}$ is well conditioned, and that $\mathbf{S}$ is accurately constructed, targets and clutter can be accurately recovered according to

$$\Sigma = R^2\mathbf{G}^{-1}\mathbf{S} \tag{5.7}$$

Noting that the entries of Σ are real and positive, and taking the norm on both sides of (5.7)

$$\Sigma = R^2\left|\mathbf{G}^{-1}\mathbf{S}\right| \tag{5.8}$$

If Σ only has one non-zero entry per column (doppler frequency), this is equal to

$$\Sigma = R^2|\mathbf{G}|^{-1}|\mathbf{S}| \tag{5.9}$$

$|\mathbf{G}|$ is the PSF matrix. The optimally irreducible arrays proposed in Chapter 4 provide the ability to design PSFs with a desired period. By setting the period to be just greater than the field-of-view of the apertures, the associated PSF matrix should be full rank. This is not always the case. This PSF matrix is susceptible to quantization effects. Although sampling the PSF at the Nyquist spatial frequency defined in Section 2.3.4 ensures that all the spatial frequency content is observed, it does not mitigate loss of accuracy due to quantization. It is necessary to experiment with values greater than the minimum number of cells required for sampling to find PSF matrices with full rank absolute values.

If the PSF matrix has full rank, the solution of (5.9) reveals the true ground scene when there is only one target per temporal frequency. If there is more than one object (target or clutter) at a particular doppler frequency, (5.9) is not valid, and (5.8) must be used instead.

The requirement to keep $\mathbf{G}$ and $\mathbf{S}$ in complex form for the deconvolution operation has severe consequences. The optimally irreducible arrays proposed in Chapter 4 yield full-rank $|\mathbf{G}|$ matrices, but unfortunately the complex form $\mathbf{G}$ has very low rank.

By looking at the singular values of a matrix, underlying properties may be revealed. Appendix B discusses the theory of singular value decomposition. Figure 5.1 shows the singular values of a typical $\mathbf{G}$ matrix, formed with three apertures, and with 119 cells. $\mathbf{G}$ has three singular values on the order 10^{1}, while the remaining 116 are smaller than 10^{-10}. The cutoff between the significant and non-significant singular values is clear and the resulting effective rank of $\mathbf{G}$ is 3. In contrast, the singular values for $|\mathbf{G}|$ are more evenly distributed. In this case it is not so easy to determine which singular values, if any, are not significant. One approach is to keep all the singular values except the last two or three, which are an order of magnitude smaller than each preceding singular value. This gives the rank as 116 or 117.

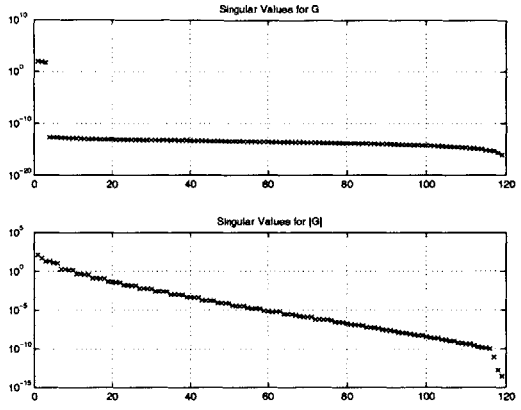

Figure 5.1 Singular Values for G and |G|

SVD is often used to reduce poorly conditioned systems and obtain approximate solutions, as discussed in Appendix B. Arrays with a low number of apertures have few significant singular values. The resulting reduced system has too few significant singular values to yield accurate results using conventional linear algebra techniques.

5.2 Deconvolution using CLEAN

The previous section has shown that the ill-conditioning of SPIR systems impairs algebraic deconvolution. Fortunately, this problem occurs in astronomical interferometry systems too, and alternative methods of solution have been found that yield good results. A particularly popular and effective technique is the CLEAN algorithm, which was developed to construct images using radio interferometers. We begin with a description of the algorithm. Next, we discuss its application to SPIR.

5.2.1 The CLEAN Algorithm

As we have seen, a sparse interferometer array with few apertures results in a poorly conditioned system. What does this mean in a physical sense? Sparse arrays sample selected spatial frequency components in the (u, v) plane. Incomplete filling of the (u, v) plane introduces artifacts and distortion in the image (ground scene). This is manifested as low

rank in the $\mathbf{G}$ matrices. We therefore seek a technique that compensates for poor (u, v) plane coverage.

One such technique is the CLEAN algorithm, which was first proposed by Högbom in 1974 [Thompson et al., 1994]. It has been extensively and successfully used in astronomical interferometry. There are many variations, and we present the simplest version here.

The algorithm conceptualises the image source (ground scene) as a collection of point sources against a dark background. Each point source has a complex point spread function associated with it, which appears in the constructed signal, shifted and scaled according to the position of the point source. Astronomical systems refer to the synthesized signal as the "dirty" signal. The goal is to determine the location and size of the point sources, or the "clean" signal. The main steps of the algorithm are shown in Figure 5.2.

The process begins by looking for the largest value, M, in the dirty signal. It is assumed to result from the convolution of a point source with the PSF. The amplitude of the point source is assumed to be $M\gamma$, where γ is a damping factor. The damping factor, also referred to as the loop gain, determines how deep the "CLEAN" goes, and is selected based on the type of signal. Experimentation with known signals of the expected type is usually necessary to determine the best value. A shifted and scaled PSF corresponding to the assumed point source is now subtracted from the dirty signal. Construction of the clean signal begins by placing a point source (delta function) of amplitude $M\gamma$ in the clean signal at the location of the maximum value. The process is repeated with the new dirty signal. The algorithm terminates when the maximum value of the dirty signal is less than a threshold value, ε. Finally, the residual contents of the dirty signal is added back to the clean signal.

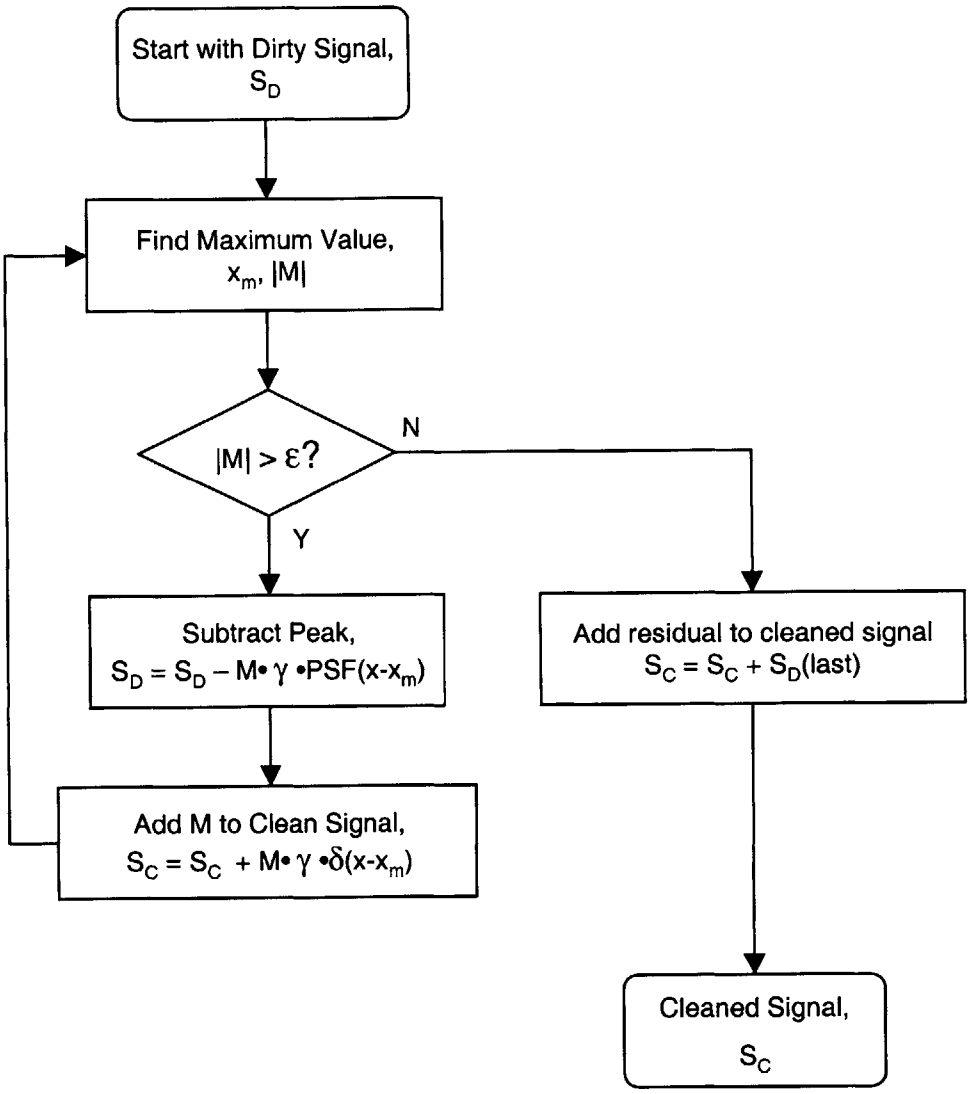

Figure 5.2 The CLEAN Algorithm

5.2.2 Application of CLEAN to SPIR

SPIR systems first determine range and temporal frequency using range gating and doppler filtering, respectively. Deconvolution is then used to determine the cross-range position of objects. Both range and temporal frequency may be ambiguous, the techniques discussed in Chapter 2 are used to resolve any ambiguities.

We begin by applying CLEAN to a cross-range strip at a single range and frequency. By examining the results under various source configurations, and with different values for γ and ε, we gain a quantitative appreciation of CLEAN performance. For display purposes the number of cells is limited by setting the cross-range resolution to be only 1000 m. The pertinent system parameters are shown in Table 5.1.

TABLE 5.1 System Parameters

Parameter	Value
Carrier wavelength λ (m)	0.03
Desired field-of-view ψ_p	0.06
Cross-range resolution Δ_{cross_range} (m)	1000
Range to ground r_{ground} (km)	1000
Baseline B (m)	30
k_{max}	61
Number of cross-range cells	119

Recall that the number of cross-range cells is chosen to correspond to the required sampling rate for the PSF (2.60). The actual resolution that can be obtained is determined by the array baseline, so that the number of cross-range resolution cells is half the number of points in the PSF (2.64). It is important not to confuse apparent resolution, which is an artifact of simulation, with the resolution that can be obtained with a real system. Accordingly, when testing SPIR performance using CLEAN, targets are placed with at least one empty cell between them.

Performance using CLEAN is unpredictable and can vary greatly between similar scenarios, as shown in the next sections.

Accurate Object Recovery

Ideally, objects are resolved at the resolution determined by the array baseline. An optimally irreducible array of three apertures in a $\{0;1;k_{max}\}$ configuration is used. There are two identical (equal amplitude) targets, spaced one cross-range cell apart, in cells 40 and 42. The damping factor γ is set to 0.1, and the threshold ε is 0.1. Figure 5.3 shows the PSF, constructed (dirty) signal, and the recovered (clean) signal. The positions of both targets are accurately determined, but the amplitudes of the recovered targets differ slightly. This happens because the algorithm detects one peak at a time. If the damping factor is increased the amplitudes of the two recovered targets will approach each other more closely. In this case the difference is about 4%, well within the noise level that would be experienced in a real system.

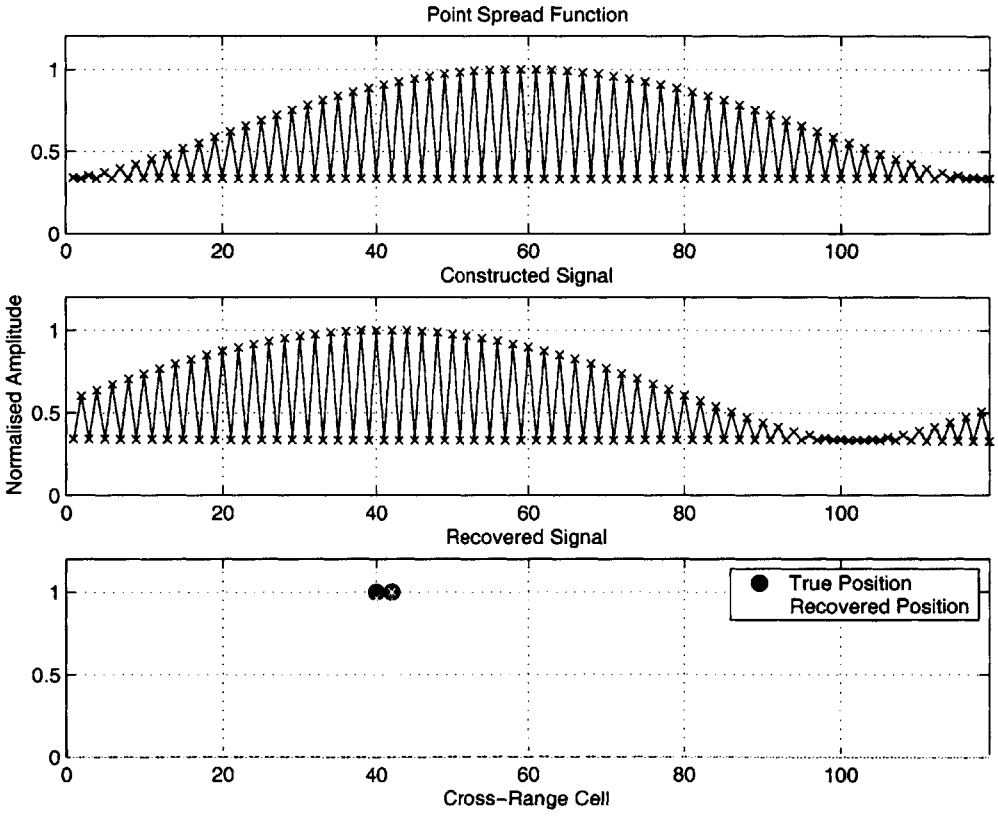

Figure 5.3 CLEAN Recovery of Two Closely Spaced Targets (Cells 40 and 42) (3 Apertures)

Inaccurate Object Position Recovery

If we change the positions of the two targets slightly, the performance degrades considerably. In Figure 5.4, the targets are now placed at cross-range cells 40 and 45. The recovered signal has two significant peaks, but these occur in cells 42 and 43, between the actual target positions.

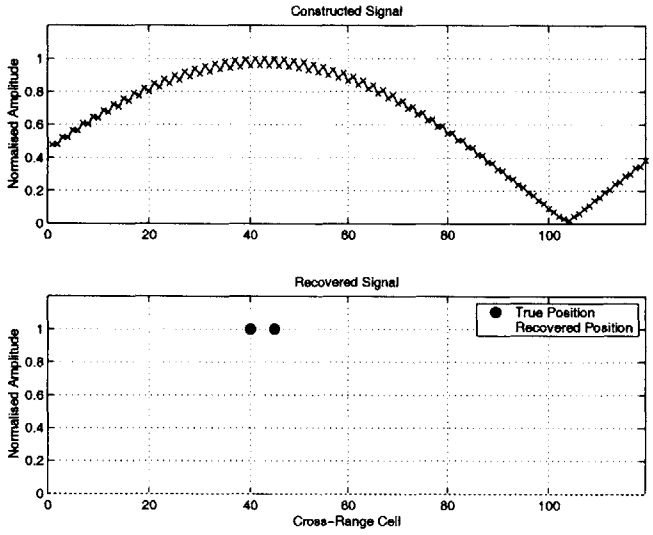

Figure 5.4 CLEAN Recovery of Two Closely Spaced Targets (Cells 40 and 45) (3 Apertures)

Unwanted Components

Figure 5.5 shows the constructed and recovered signals when there is one target at each edge of the footprint. The additional components of the recovered signal result from addition of the residual to the cleaned signal. By decreasing the threshold value ε these components can be attenuated. Decreasing ε increases the computation time since more passes through the CLEAN loop are required before termination occurs.

Figure 5.6 shows the result when $\varepsilon = 0.01$. The unwanted components have been reduced by a factor of about five. There is a computational price to pay however. Whereas

seven iterations were required to solve the system with $\varepsilon = 0.1$, this system required twenty-nine iterations.

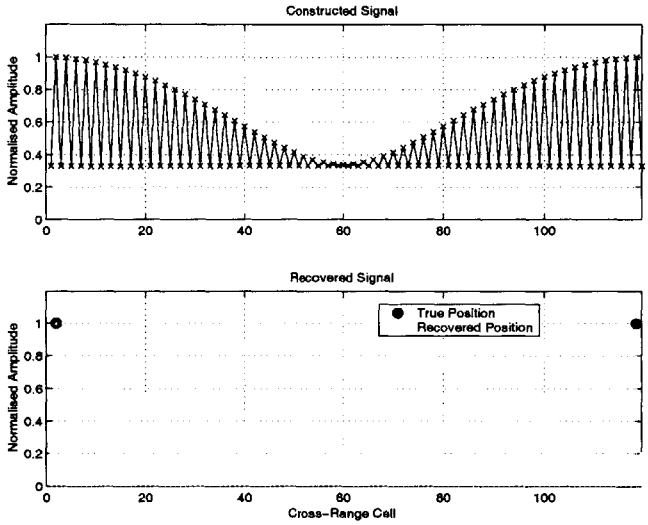

Figure 5.5 CLEAN Recovery of Two Targets at the Edges of the Footprint (3 Apertures)

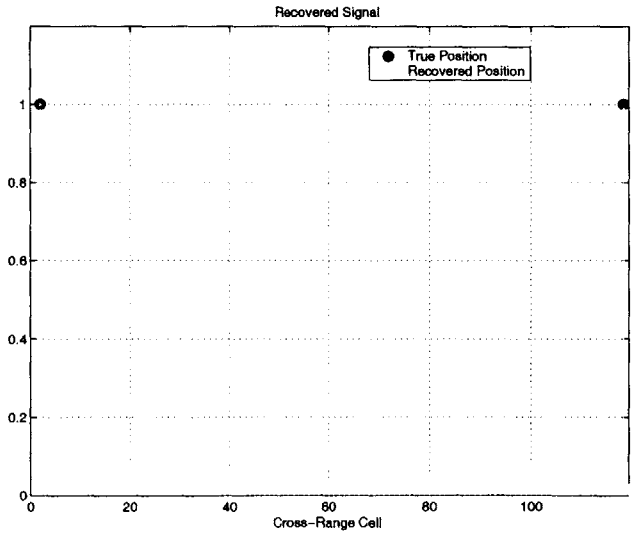

Figure 5.6 CLEAN Recovery of Two Targets at the Edges of the Footprint with Smaller Threshold (3 Apertures)

Failure to Recover Targets

In some cases CLEAN may not recover all the targets. Targets close to the centre of the cross-range strip result in poor recovery of other objects, as shown in Figure 5.7. In Figure 5.8 there are two targets at roughly equal intervals from the centre of the strip. Neither target is recovered, instead there is a false single target in between the true targets.

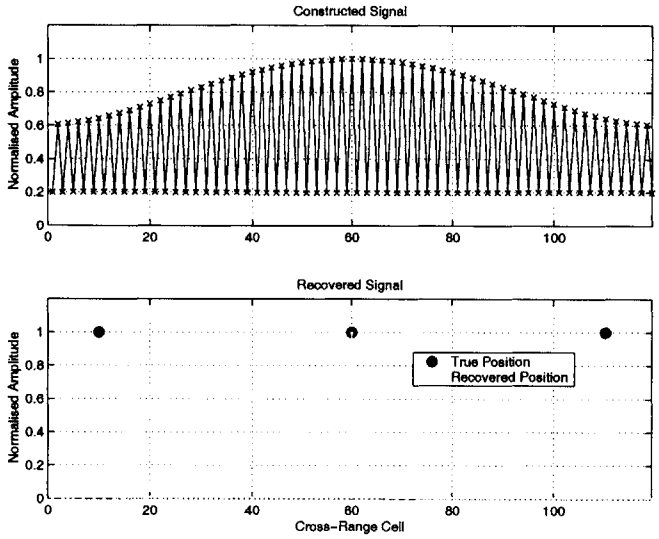

Figure 5.7 CLEAN Recovery with a Target at the Footprint Centre (3 Apertures)

CLEAN Performance in the Clutter Environment

Radar return from the ground will contain not only targets, but also clutter. Depending on the PRF, the clutter may appear several times at the same frequency and range, due to aliasing. It is therefore necessary to distinguish between several objects in a given cross-range strip. In our next example, we assume that clutter appears three times in the cross range strip, which also contains two targets. We begin by assigning equal sizes to both clutter and targets.The clutter resides in cells 30, 70 and 110, and the targets are in cells 20 and 22. The result is shown in Figure 5.9. The clean signal has little resemblance to the true target and clutter positions.

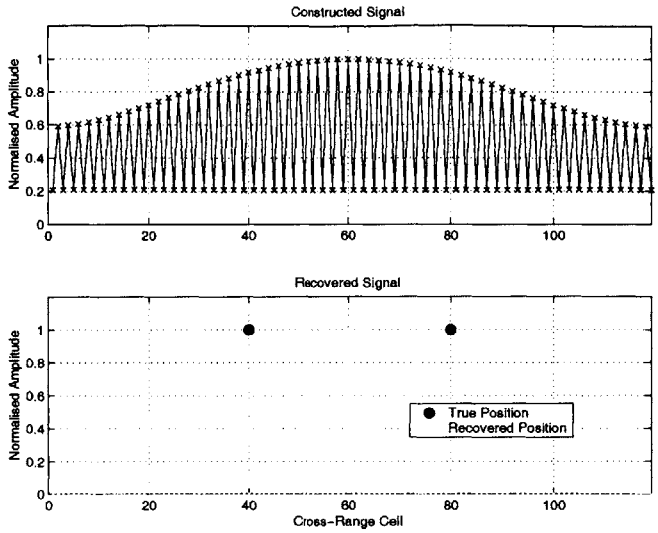

Figure 5.8 CLEAN Recovery with Centre Targets (3 Apertures)

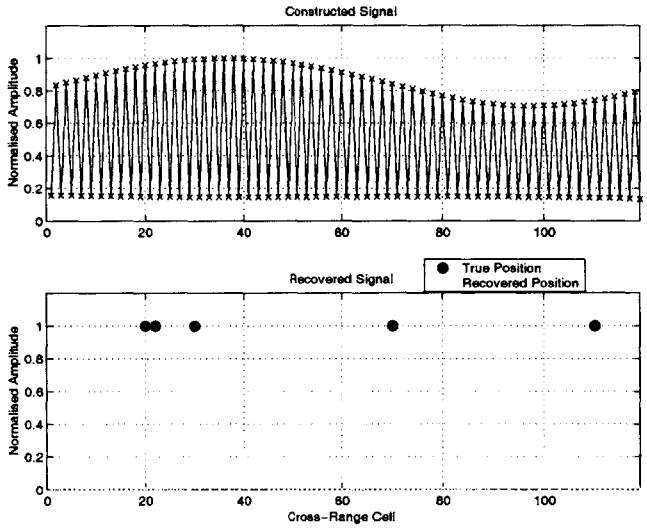

Figure 5.9 CLEAN Recovery with Targets and Clutter (3 Apertures)

By increasing the number of apertures, the result can be improved. Figure 5.10 shows the result with an optimally irreducible array of eight apertures. Two of the clutter spikes are accurately recovered. The remaining spike does not appear in the clean signal. One of the targets is recovered, but with a significantly smaller amplitude. The other target is not

recovered. The energy from the lost clutter spike and target appears in unwanted signal components.

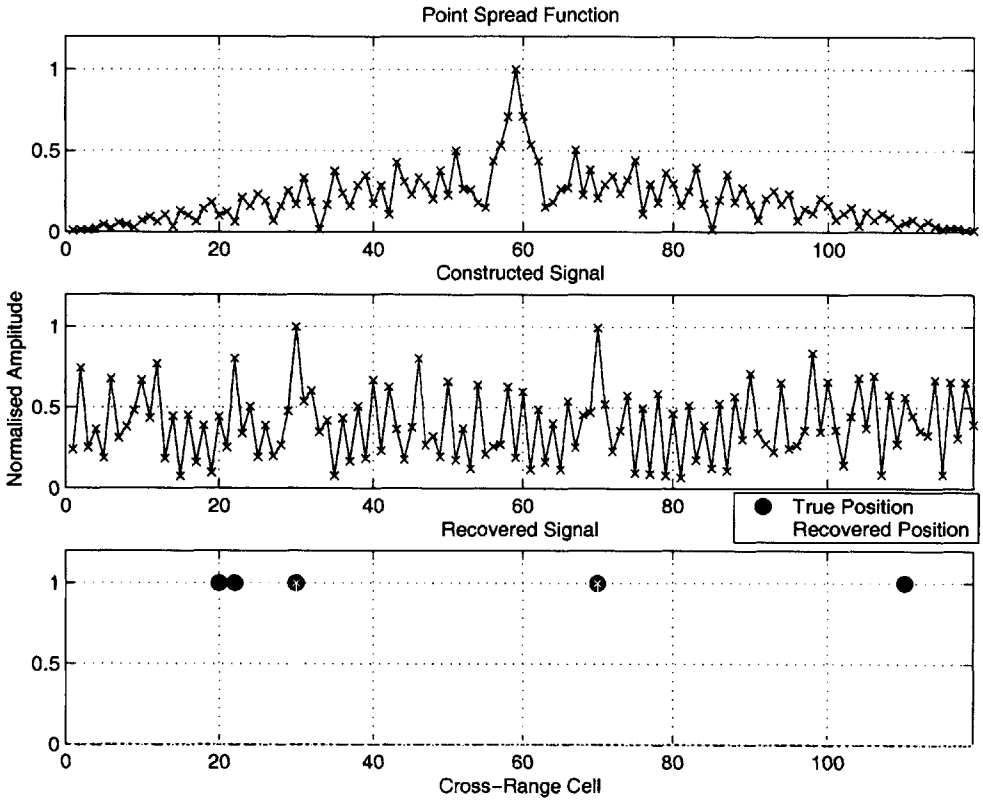

Figure 5.10 CLEAN Recovery with Targets and Clutter (8 Apertures)

Since the cross-range position of clutter at a given frequency is known, and clutter amplitude is effectively indeterminate, accurate recovery of the clutter amplitude is not critical. It is important that clutter energy appear in the correct cross-range cells, to avoid generating false targets.

Discussion

Target recovery performance using SPIR and CLEAN is highly variable. Certain positions and combinations of objects are accurately recovered. In other cases some targets may not be recovered at all, or appear in completely the wrong positions. The more complex the

ground scene is, the more likely the clean signal is to be confused. By increasing the number of apertures, the performance can generally be improved, but consistently satisfactory performance remains elusive. The best choice of γ and ε depends on the target and clutter positions. As the number of objects is increased, the clean signal is more likely to be confused.

We can gain a qualitative understanding of the behaviour by considering how the CLEAN algorithm works. As discussed in Section 5.1.2, poor spatial frequency sampling in the cross-range direction results in ambiguous PSFs. The CLEAN algorithm attempts to compensate for missing spatial frequencies. These components correspond to the nullspace of the system. The aperture array is blind to their contribution, and *any* values can be assigned at the missing spatial frequencies. CLEAN attempts to "guess at" the missing frequency content. When many spatial frequencies are sampled, this works well, but with few spatial frequencies there is too little information. The result is poor reconstruction of the image.

5.2.3 Improving SPIR Performance Using CLEAN

In practice, the results obtained using CLEAN are inconsistent and unpredictable. In this section we discuss some approaches to improving performance.

Evaluating Performance

The first step towards developing an effective system is to develop a means of evaluating the performance of candidate algorithms and parameters with respect to known inputs.

For a given aperture array, ground scene and set of tuning parameters, we could assign a numerical score based upon the following factors:

- The number of true detections
- The number of false detections
- Position error for each detection
- Amplitude error for each detection

By combining these measures into a single performance metric, performance for a specific scenario different system configurations can be evaluated. Using a consistent performance evaluation, an understanding of how well the system is likely to behave under a given scenario can be attained.

Clutter Reduction

As we have seen, the CLEAN algorithm is less effective when there are many sources in the dirty signal. The cross-range position of clutter is defined by its range and doppler frequency. Both clutter and targets create features in the synthesized (dirty) signal. Although the magnitude of the clutter is unpredictable, its position is known a priori. By adapting the CLEAN algorithm to first remove contributions from known locations, target recovery can be improved.

The adapted algorithm is shown in Figure 5.11. A clutter "mask" is used to mask out cells in the dirty signal that do not contain clutter. CLEAN is then used to locate the clutter sources in the dirty signal, and construct a new dirty signal that has been largely cleared of clutter, referred to as the *reduced clutter signal.* The new dirty signal has fewer sources, so that applying CLEAN to it is more effective.

This approach is applied to the example in Figure 5.10, and the two targets are revealed. Figure 5.12 shows the cleaned target signal as well as the cleaned clutter signal. The cleaned target signal has the targets in the correct positions. The recovered targets' amplitudes are smaller than the input targets, but they are still clearly discernible. The cleaned clutter signal has clutter in the correct positions, although the amplitudes are not correct. Clutter amplitude is random in any case, so this is not a great concern. The amplitudes of the recovered clutter cells are significantly larger than those of the recovered targets. This is because the clutter cells are cleaned first.

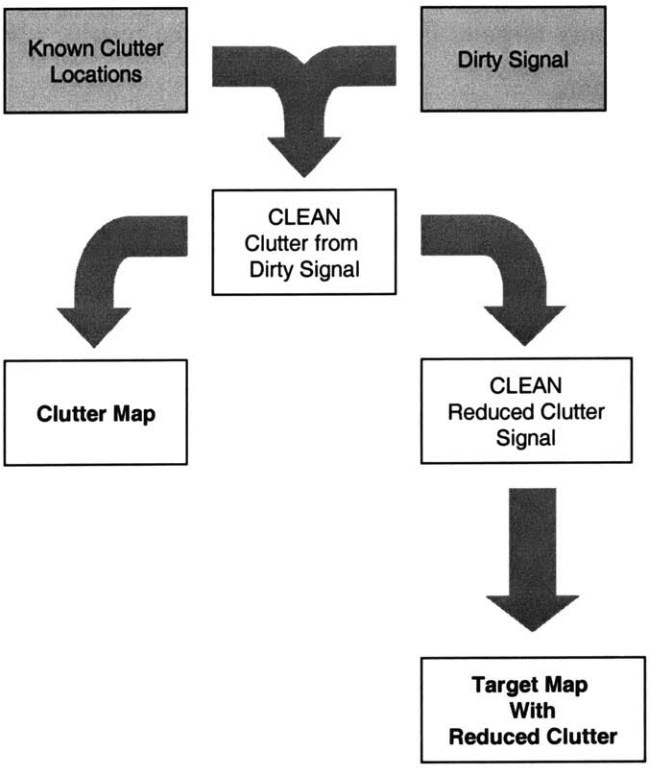

Figure 5.11 Clutter Reduction using CLEAN

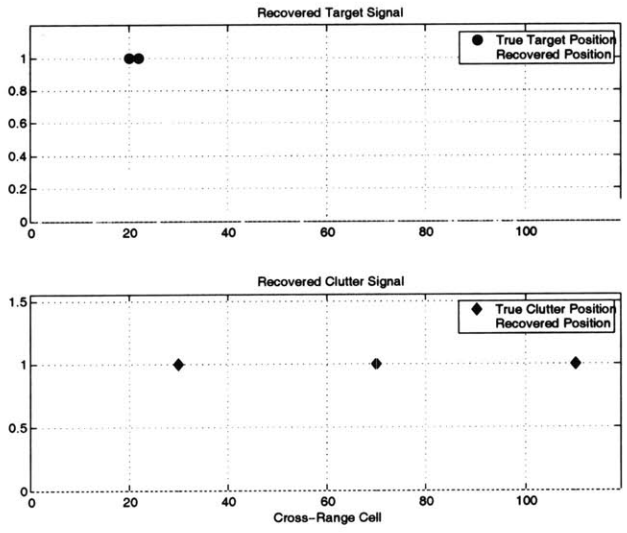

Figure 5.12 Improved CLEAN Recovery with Clutter (12 Apertures)

When the ground resolution cells are large, the clutter cross-section may be significantly larger than the RCS of any targets. Figure 5.13 shows the results when the clutter is ten times larger than the targets.

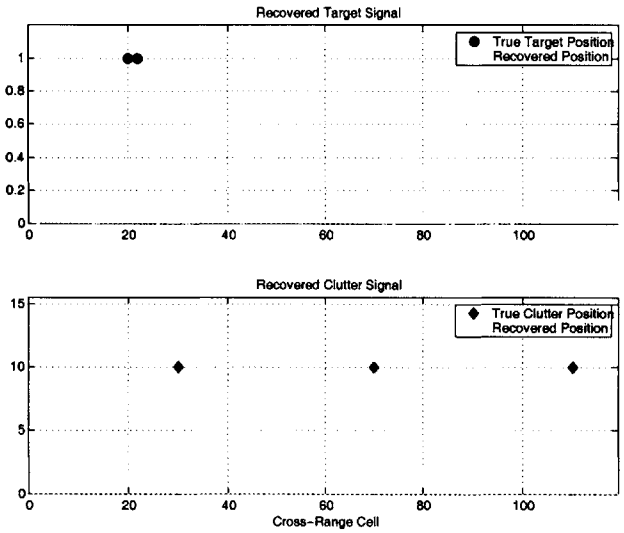

Figure 5.13 Improved CLEAN Recovery with Large Clutter (12 Apertures)

In addition to being larger than the targets' amplitudes, the clutter amplitude is expected to vary randomly from one clutter cell to the next. The improved CLEAN algorithm does not require the clutter amplitude to be constant either, as shown in Figure 5.14.

We have assumed that the clutter from a particular ground cell has a single doppler frequency component. This is an idealisation, since the continuous extent of the ground produces a continuous band of doppler shift for each cell. In a cross-range strip, the clutter therefore appears as bands. The number of bands depends on the PRF. The clutter components that make up a band in cross-range are aliased from different parts of the clutter spectrum and therefore bear no relation to each other. Figure 5.15 shows the recovered target and clutter signals when the clutter is assumed to be spread across three cells at each aliased doppler shift. The cleaned signal has components corresponding to the targets, and

components in the clutter cells. There are some false responses of similar or greater magnitude, which will be mistaken for targets.

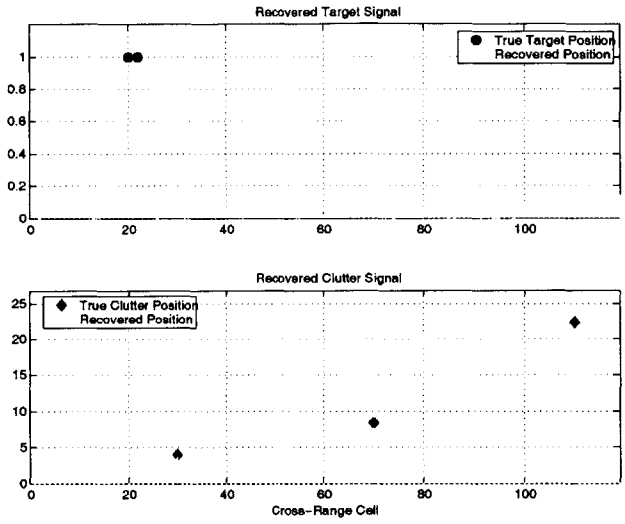

Figure 5.14 Improved CLEAN Recovery with Targets and Randomly Varying Large Clutter (12 Apertures)

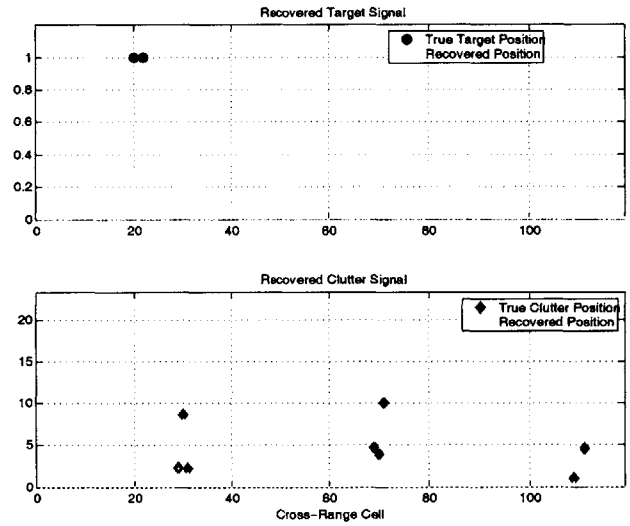

Figure 5.15 Improved CLEAN Recovery with Targets and Randomly Varying Large Clutter Spread Across Multiple Cells (12 Apertures)

Target Recovery Improvement

Target recovery can be improved by ensuring that CLEAN does not look for targets in cells where targets cannot occur, and by forcing CLEAN to look for targets in cells where targets were previously found.

First we determine whether there are any cross-range "dead" cells, where targets cannot be found. The maximum speed of the ground moving targets defines the doppler shift bandwidth within which targets will appear. This bandwidth is centered about the centre clutter doppler shift for a given ground resolution cell. For a side-looking configuration, the bandwidth of frequencies that a target can appear in is given by

$$\frac{2\sin\theta}{\lambda}[v_{cluster} \pm v_{max}] \tag{5.10}$$

where v_{max} is the maximum possible target velocity and θ is the angle from the interferometer boresight, or cross-range angle. At a particular clutter centre frequency f_d, targets will appear in an angle range $[\theta_{min}, \theta_{max}]$ given by

$$\begin{aligned} \sin\theta_{min} &= f_d \cdot \frac{\lambda}{2(v_{cluster} + v_{max})} \\ \sin\theta_{max} &= f_d \cdot \frac{\lambda}{2(v_{cluster} - v_{max})} \end{aligned} \tag{5.11}$$

Due to aliasing, there are multiple target bands, centered about each aliased clutter frequency component, as shown in Figure 5.16. Note that the bands may have different widths, since they correspond to different f_d that appear in the same frequency bins due to aliasing.

If the PRF is chosen sufficiently high that these bands do not overlap, there will be sections of each cross-range strip that contain only noise. By forcing CLEAN to ignore these sections, target recovery performance may be improved.

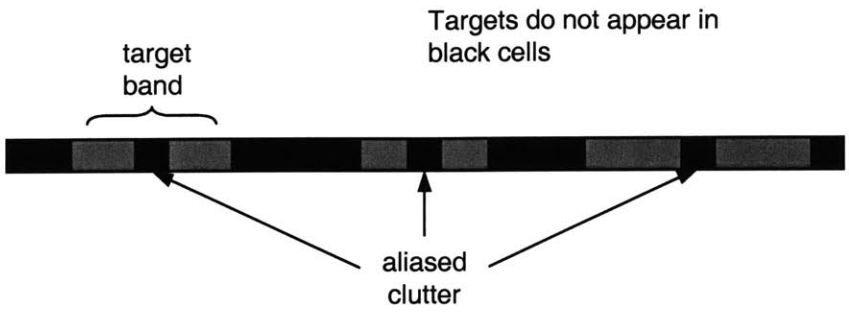

Figure 5.16 Target Appear in Bands Centered about Clutter

Further improvement could be attained using prior target position information. Using a similar technique to the clutter reduction discussed earlier, the CLEAN algorithm first searches for maxima in the vicinity of known previous target positions. If the number of unknown signal sources is reduced, CLEAN is less likely confuse the remaining sources.

5.3 Summary

This chapter has discussed the numerical difficulties of solving SPIR systems. The numerical conditioning of SPIR systems is poor, so that conventional linear algebra techniques cannot be used.

The CLEAN algorithm, developed in the field of radio interferometry, was suggested as an alternative. In tests, the performance of CLEAN is highly varied. Using the geometric relationship between clutter doppler frequency and position, a clutter reduction technique was developed. This technique significantly improves CLEAN performance. The cross-range cells in which a target can appear are also determined by geometry. By forcing CLEAN to ignore cross-range cells that cannot contain targets, further improvement is possible. Prior target information can be used to improve target recovery, in a similar manner to clutter reduction.

Chapter 6

SUMMARY AND FUTURE WORK

This thesis has developed and analysed a new spaceborne radar interferometry system to be used for Ground Moving Target Indication (GMTI).

6.1 Summary

Spaceborne radar systems offer significant advantages over airborne systems. Airborne systems require large support infrastructures and significant deployment time and risk the safety of pilots. In contrast, spaceborne radar can offer nearly continuous global coverage without appreciable operator risk.

Monolithic spaceborne systems are expensive, and subject to single-point failure. Interferometric radar systems use several spacecraft to deliver reliable and cost-effective performance. Whereas single aperture systems have to compromise between resolution and field-of-view, multiple aperture interferometers can provide both high resolution and wide coverage. The resolution of the interferometer is determined by the maximum distance (baseline) between apertures, while the field-of-view is determined by the individual aperture response.

This thesis has discussed a novel processing algorithm for use in interferometric radar, referred to as Scanned Pattern Interferometric Radar (SPIR). SPIR uses the concept of deconvolution, borrowed from radio interferometry, to reconstruct the actual ground

scene. Deconvolution is the process of compensating for the effect of the interferometer response, or point spread function (PSF), on the received signals. Radar return from the ground can be characterised in terms of its range, cross-range and doppler shift. SPIR processing uses radar techniques to determine the range and doppler shift of the return signal. Deconvolution is used only to determine the cross-range target positions at each range and doppler frequency.

The theory behind SPIR is derived from the fields of radar and interferometry. Chapter 2 provided the required background with a summary of basic radar principles, Fourier techniques and the fundamentals of interferometry. SPIR is not the only interferometric technique. Two other techniques, Classical Displaced Phase Centre Antenna (DPCA) and Space Time Adaptive Processing (STAP) were reviewed.

In Chapter 3 the SPIR concept was introduced. Following an intuitive discussion, a mathematical derivation was presented. Next, the implementation of a SPIR system was discussed. Techniques for dealing with range and frequency ambiguities were proposed. High resolution ground scenes require heavy computation. Efficient distribution of the processing and communication requirements is essential to ensure timely image generation, and some possible schemes were discussed.

The physical configuration of the satellite cluster determines the PSF and therefore directly impacts SPIR performance. Chapter 4 discussed the relation between cluster configuration and PSF shape. Minimum redundancy arrays are generally regarded as good candidates for linear interferometric system clusters, but were shown to result in undesirable PSFs for SPIR systems. An alternative cluster design approach, based on an understanding of the construction of the PSF, was proposed.

Chapter 5 discussed the numerical solution of SPIR systems. Attempts to generate true ground scenes using example SPIR systems revealed that the associated systems of linear equations are ill-conditioned. Conventional linear algebra solution techniques are therefore not appropriate. An alternative solution technique, the CLEAN algorithm, which was

developed for use in astronomical interferometry, is used instead. In practice, the performance of CLEAN is highly varied. Some techniques to improve performance were proposed. These include using the geometrically determined clutter position, as well as prior target information, to improve the recovery of both existing and new targets.

6.2 Future Work

The results obtained using simplified SPIR simulations have been promising, but further work is required to more accurately analyse and improve SPIR performance. To this end, an implementation of SPIR on a real-time testbed is currently in process.

The effect of noise has not been explicitly considered. Accurate estimation of the expected noise level under different scenarios is required so that an effective noise mitigation strategy can be developed.

Aperture response side lobe levels were assumed to be negligible. This is in general not the case, and the effect of side lobe energy on the reconstructed ground scene must be considered.

The conditioning of the complex PSF matrix is poor, but the reason for this is not well understood. In-depth analysis of the interferometer cluster and associated PSF is required to determine why the conditioning is poor. Better understanding may enable the design of better conditioned systems.

This thesis has highlighted some of the issues that cause poor performance. In future work, strategies to improve performance should be investigated. These can be divided into two main areas, system level strategies and computational strategies.

System level strategies adapt the system configuration to deliver the best performance. One possibility is to use a multiple look strategy to increase the number of apparent apertures in the cluster. Preliminary analysis indicates that increasing the number of measurements during a look will improve the system conditioning. Another possibility is to

change the cluster shape in real time, to more completely sample the (u, v) plane. Improved (u, v) plane coverage should result in improved ground scene reconstruction.

Computational strategies are aimed at improving the algorithms used to reconstruct the ground scene. This thesis has focused on the deconvolution step, but other areas, such as thresholding, should be considered in conjunction with deconvolution. Some strategies to improve CLEAN performance were proposed. These need to be developed further to assess their impact.

REFERENCES

[Bracewell, 1962] Bracewell, R. N., "Radio Astronomy Techniques", *Handbuch Der Physik*, Vol. 54, Springer, Berlin, pp. 42-129, 1962.

[Buderi, 1996] Buderi, R., The Invention that Changed the World: The Story of Radar from War to Peace, Abacus, London, 1996.

[Hacker, 2000] Hacker, T. L., *Performance Analysis of a Space-Based GMTI Radar System Using Separated Spacecraft Interferometry*, MIT S.M. Thesis in Aeronautical and Astronautical Engineering, May 2000.

[Jilla et al., 1997] Jilla, C.D., Miller, D.W., *A Reliability Model for the Design and Optimization of Separated Spacecraft Interferometer Arrays*, Proceedings of the 11th AIAA/USU Conference on Small Satellites, SSC97-XI-2, Logan, UT, September 1997.

[Ishiguro, 1980] Ishiguro, M., *Minimum Redundancy Linear Arrays for a Large Number of Antennas*, Radio Science, Volume 15, pp. 1163-1170, 1980.

[Klemm, 1999] Klemm, R., *Introduction to Space-Time Adaptive Processing*, Electronics and Communication Engineering Journal, pp. 5-12, February 1999.

[Kong, 1998] Kong, E. M., *Optimal Trajectories and Orbit Design for Separated Spacecraft Interferometry*, MIT S.M. Thesis in Aeronautical and Astronautical Engineering, November 1998.

[Leech, 1956] Leech, J., *On the Representation of 1,2,..., n by Differences*, Journal of the London Mathematical Society, Volume 31, pp. 160-169, 1956.

[Moffet, 1968] Moffet, A. T., *Minimum-Redundancy Linear Arrays*, IEEE Transactions on Antennas and Propagation, AP-16(2), pp. 172-175, 1968.

[Oppenheim et al., 1999] Oppenheim, A. V., Schafer, R. W., Buck, J. R., *Discrete-Time Signal Processing*, Second Edition, Prentice Hall, New Jersey, 1999.

[Ramo et al., 1984] Ramo, S., Whinnery, J. R., Van Duzer, T., *Fields and Waves in Communication Electronics*, Second Edition, John Wiley & Sons, New York, 1984.

[Rosen et al., 2000] Rosen, P. A., Hensley S., Joughin, I. R., Li F.K., Madsen S.N., Rodríguez E., Goldstein R.M., *Synthetic Aperture Radar Interferometry*, Proceedings of the IEEE, Volume 88, Number 3, pp. 333-381, 2000.

[Shaw, 1998] Shaw, G. B., *The Generalized Information Network Analysis for Distributed*

Satellite Systems, MIT Ph.D. Thesis in Aeronautical and Astronautical Engineering, October 1998.

[Skolnik, 1980] Skolnik, I. M., *Introduction to Radar Systems, Second Edition*, McGraw Hill, Boston, 1980.

[Staelin et al., 1994] Staelin, D. H., Morgenthaler, A. W., Kong, J. A., *Electromagnetic Waves*, Prentice Hall, 1994.

[Stimson, 1998] Stimson, G. W., Moran, J. M, Swenson, G. W. Jr., *Introduction to Airborne Radar, Second Edition*, SciTech, Mendham, 1998.

[Thompson et al., 1994] Thompson, A. R., Moran, J. M., Swenson, G. W. Jr., *Interferometry and Synthesis in Radio Astronomy*, Krieger Publishing Company, Malabar, Florida, 1994.

[Trefethen, 1997] Trefethen, L. N., Bau, D., *Numerical Linear Algebra*, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, 1997.

[Ward, 1994] Ward, J., Space-Time Adaptive Processing for Airborne Radar, Lincoln Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts, ESD-TR-94-109, 1994.

[Wertz and Larson, 1999] Wertz, J. R., Larson, W. J., Space Mission Analysis and Design, Third Edition, Space Technology Library, Microcosm Press, Torrance, California, 1999.

Appendix A

EXAMPLE LINEAR CLUSTERS

This appendix illustrates the Point Spread Functions (PSFs) for arrays of six to twelve apertures. Arrays containing both restricted and unrestricted minimum redundancy subarrays are shown.

System parameters are as given in Table 4.3.

Note that the shape of the array factor does not change significantly between different minimum redundancy subarrays. As the number of apertures is increased the width of the main lobe decreases and the side lobes become lower.

A.1 Six Apertures

A.1.1 Unrestricted

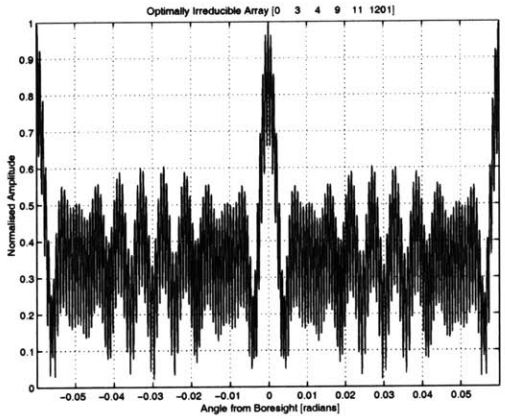

Figure A.1 {0;3;4;9;11;1201} Six Element Optimally Irreducible Array (Unrestricted)

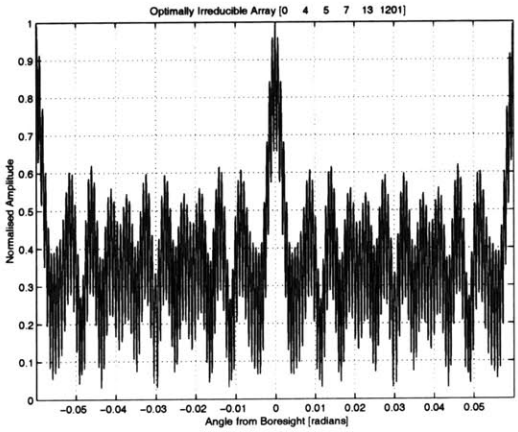

Figure A.2 {0;4;5;7;13;1201} Six Element Optimally Irreducible Array (Unrestricted)

A.1.2 Restricted

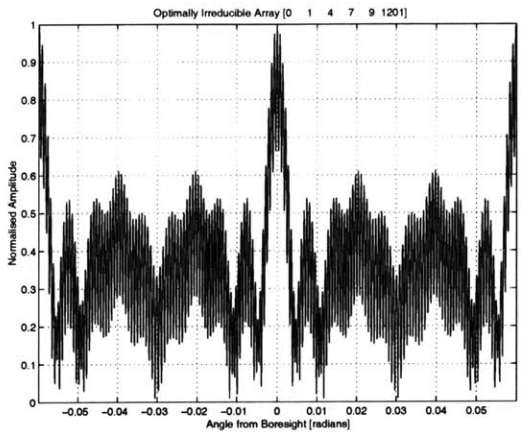

Figure A.3 {0;1;4;7;9;1201} Six Element Optimally Irreducible Array (Restricted)

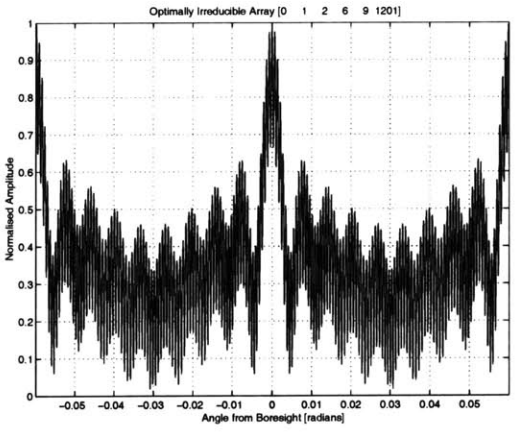

Figure A.4 {0;1;2;6;9;1201} Six Element Optimally Irreducible Array (Restricted)

A.2 Seven Apertures

A.2.1 Unrestricted

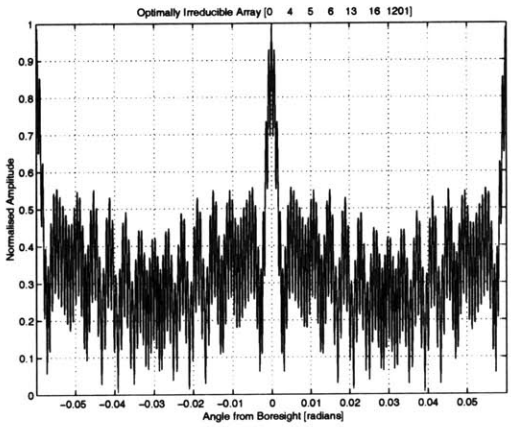

Figure A.5 {0;4;5;6;13;16;1201} Seven Element Optimally Irreducible Array (Unrestricted)

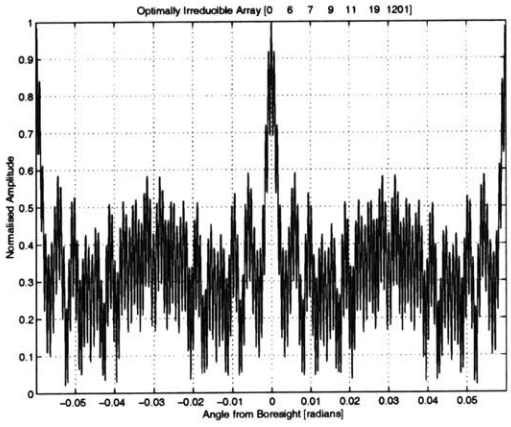

Figure A.6 {0;6;7;9;11;19;1201} Seven Element Optimally Irreducible Array (Unrestricted)

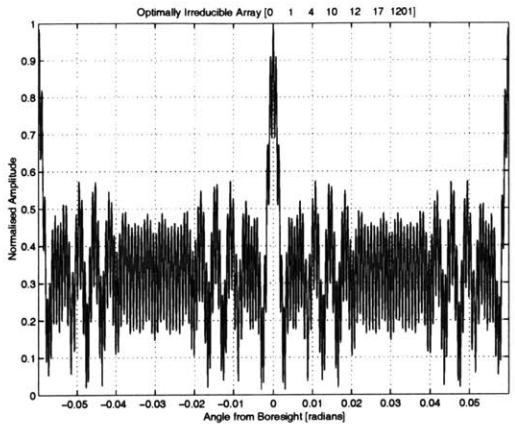

Figure A.7 {0;1;4;10;12;17;1201} Seven Element Optimally Irreducible Array (Unrestricted)

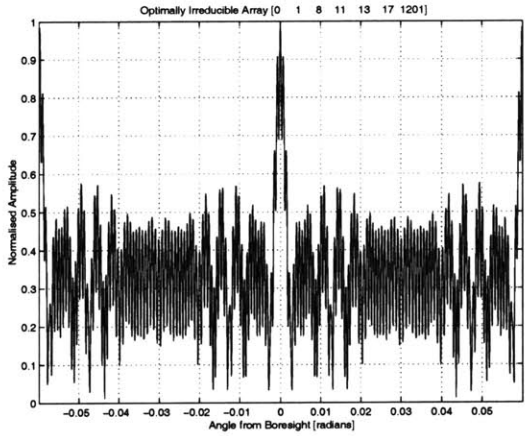

Figure A.8 {0;1;8;11;13;17;1201} Seven Element Optimally Irreducible Array (Unrestricted)

A.2.2 Restricted

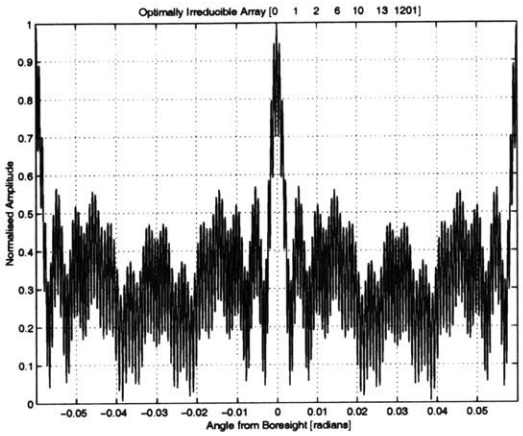

Figure A.9 {0;1;2;6;10;13;1201} Seven Element Optimally Irreducible Array (Restricted)

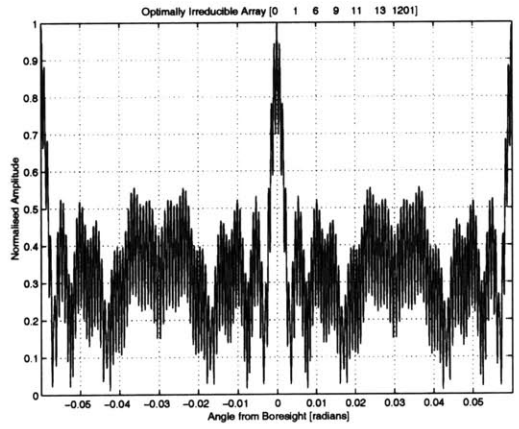

Figure A.10 {0;1;6;9;11;13;1201} Seven Element Optimally Irreducible Array (Restricted)

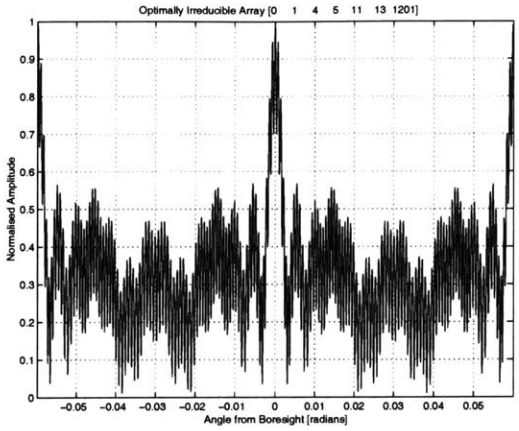

Figure A.11 {0;1;4;5;11;13;1201} Seven Element Optimally Irreducible Array (Restricted)

A.3 Eight Apertures

A.3.1 Unrestricted

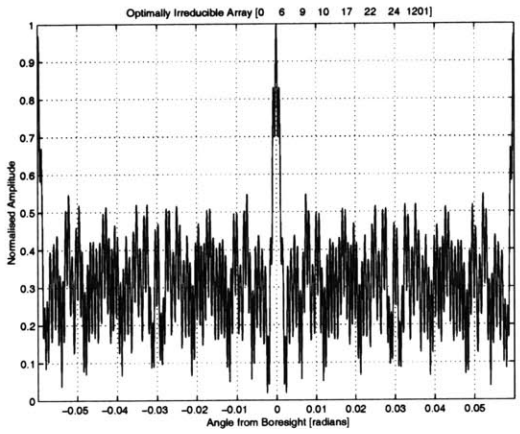

Figure A.12 {0;6;9;10;17;22;24;1201} Eight Element Optimally Irreducible Array (Unrestricted)

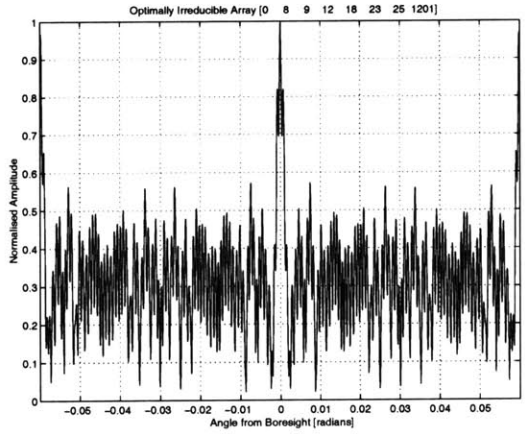

Figure A.13 {0;8;9;12;18;23;25;1201} Eight Element Optimally Irreducible Array (Unrestricted)

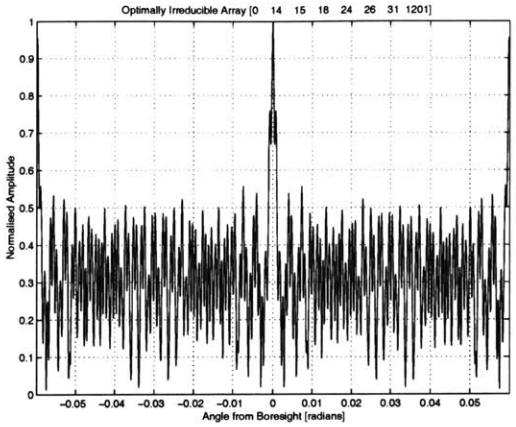

Figure A.14 {0;14;15;18;24;26;31;1201} Eight Element Optimally Irreducible Array (Unrestricted)

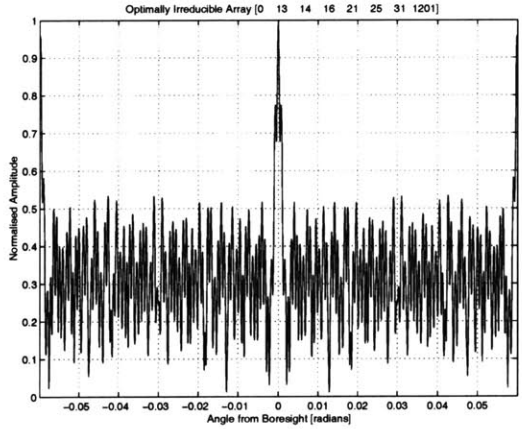

Figure A.15 {0;13;14;16;21;25;31;1201} Eight Element Optimally Irreducible Array (Unrestricted)

A.3.2 Restricted

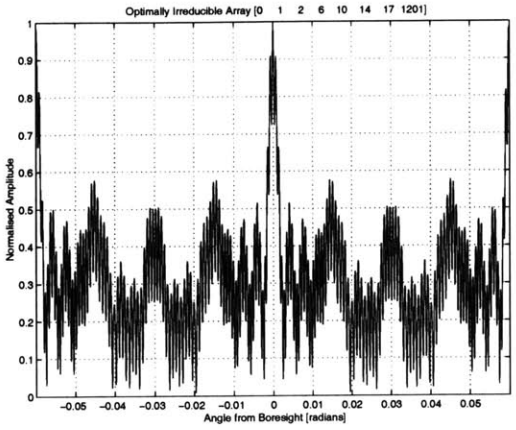

Figure A.16 {0;1;2;6;10;14;17;1201} Eight Element Optimally Irreducible Array (Restricted)

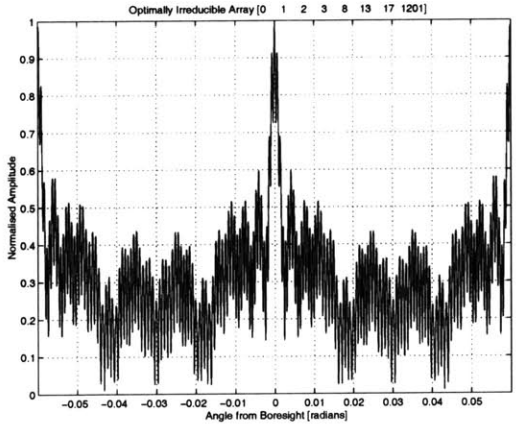

Figure A.17 {0;1;2;3;8;13;17;1201} Eight Element Optimally Irreducible Array (Restricted)

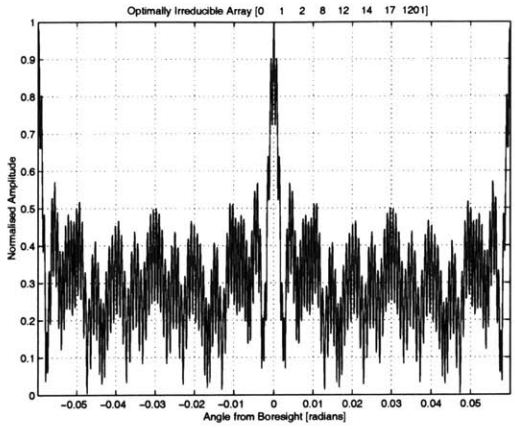

Figure A.18 {0;1;2;8;12;14;17;1201} Eight Element Optimally Irreducible Array (Restricted)

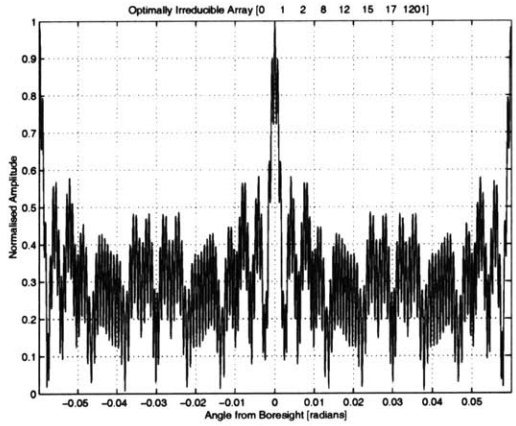

Figure A.19 {0;1;2;8;12;15;17;1201} Eight Element Optimally Irreducible Array (Restricted)

A.4 Nine Apertures

A.4.1 Unrestricted

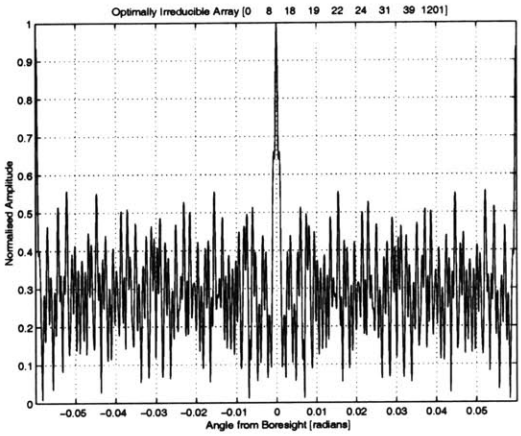

Figure A.20 {0;8;18;19;22;24;31;39;1201} Nine Element Optimally Irreducible Array (Unrestricted)

A.4.2 Restricted

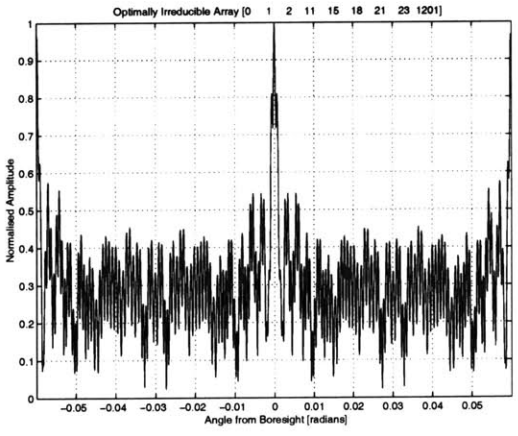

Figure A.21 {0;1;2;11;15;18;21;23;1201} Nine Element Optimally Irreducible Array (Restricted)

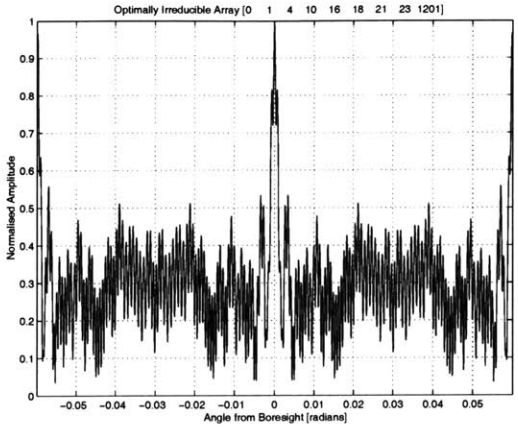

Figure A.22 {0;1;4;10;16;18;21;23;1201} Nine Element Optimally Irreducible Array (Restricted)

A.5 Ten Apertures

A.5.1 Unrestricted

There are no nine element unrestricted minimum redundancy arrays.

A.5.2 Restricted

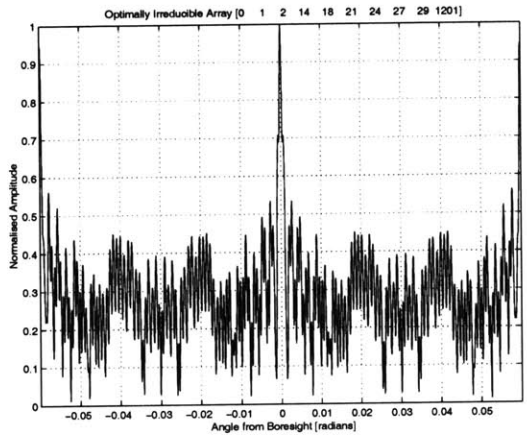

Figure A.23 {0;1;2;14;18;21;24;27;29;1201} Ten Element Optimally Irreducible Array (Restricted)

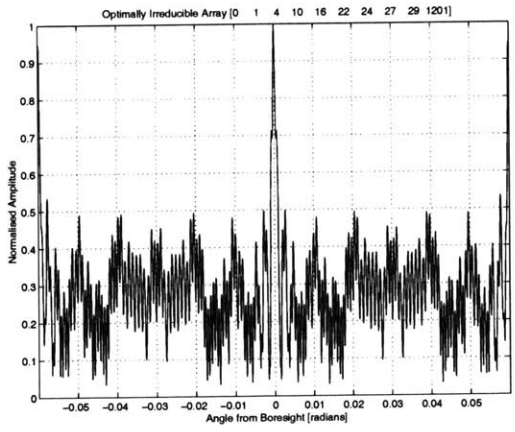

Figure A.24 {0;1;4;10;16;22;24;27;29;1201} Ten Element Optimally Irreducible Array (Restricted)

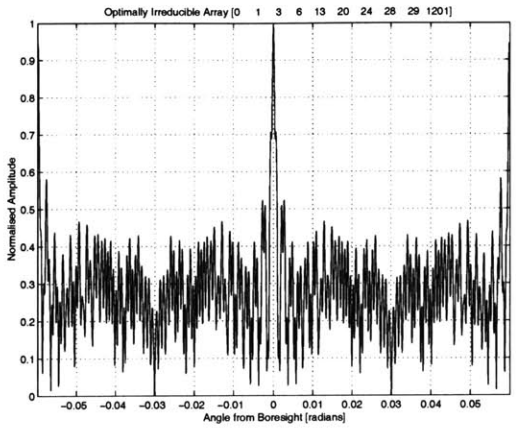

Figure A.25 {0;1;3;6;13;20;24;28;29;1201} Ten Element Optimally Irreducible Array (Restricted)

A.6 Eleven Apertures

A.6.1 Unrestricted

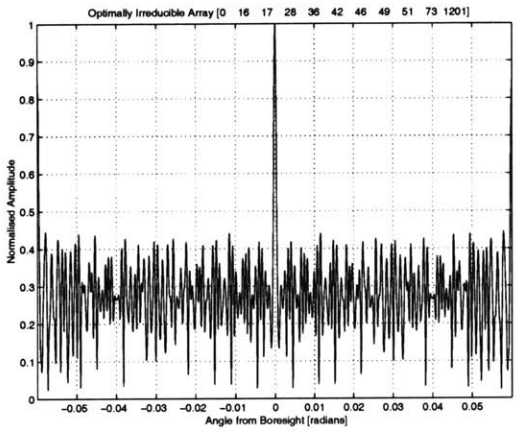

Figure A.26 {0;16;17;28;36;42;46;49;51;73;1201} Eleven Element Optimally Irreducible Array (Unrestricted)

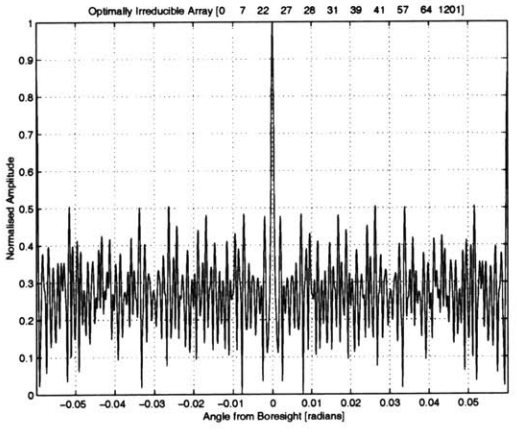

Figure A.27 {0;7;22;27;28;31;39;41;57;64;1201} Eleven Element Optimally Irreducible Array (Unrestricted)

A.6.2 Restricted

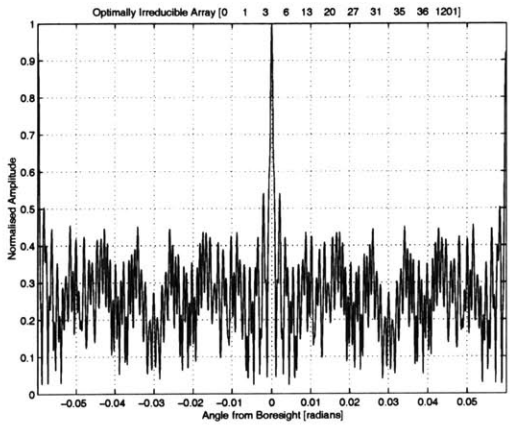

Figure A.28 {0;1;3;6;13;20;27;31;35;36;1201} Eleven Element Optimally Irreducible Array (Restricted)

A.7 Twelve Apertures

A.7.1 Unrestricted

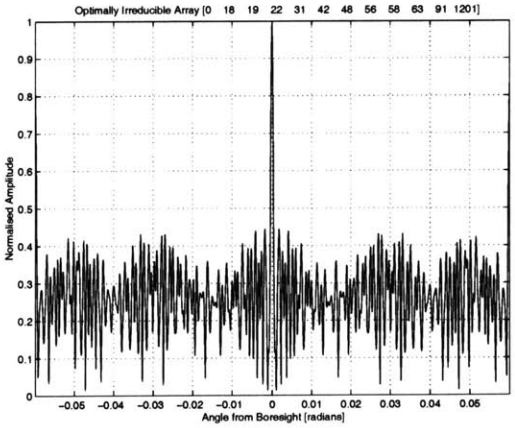

Figure A.29 {0;18;19;22;31;42;48;56;58;63;91;1201} Twelve Element Optimally Irreducible Array (Unrestricted)

A.7.2 Restricted

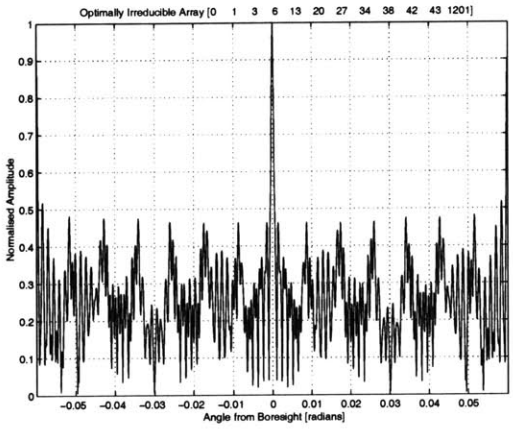

Figure A.30 {0;1;3;6;13;20;27;34;38;42;43;1201} Twelve Element Optimally Irreducible Array (Restricted)

Appendix B

SINGULAR VALUE DECOMPOSITION

The singular value decomposition (SVD) is a matrix factorisation that reveals underlying matrix properties [Trefethen, 1997]. It is particularly useful when dealing with poorly conditioned systems. In this appendix we introduce the theory behind SVD and its application to systems of linear equations.

B.1 The Singular Value Decomposition

SVD is motivated by the geometrical observation that the image of the unit sphere under any $m \times n$ matrix is a hyperellipse. It is applicable to both real and complex matrices, but for the sake of visualisation, we assume that the matrix is real. A hyperellipse is the generalised form of an ellipse, in m dimensions ($\Re^m$). An ellipse in two dimensions is defined by its major and minor axes. Similarly, a hyperellipse in $\Re^m$ can be defined by a set of *principal semiaxes* $s_1, \ldots, s_m$, which determine the extent of the hyperellipse in m orthogonal directions $u_1, \ldots, u_m \in \Re^m$, where $\|u_i\| = 1$. The number of non-zero principal semiaxes is equal to the rank of the matrix.

It is not obvious that the unit sphere maps onto a hyperellipse under an $m \times n$ matrix. Figure B.1 illustrates the concept for a 2×2 matrix $\mathbf{A}$, which is assumed to be full rank.

Let $\mathbf{S}$ be the unit sphere in $\Re^n$, and take any $\mathbf{A} \in \Re^{m \times n}$ with $m \geq n$. The image $\mathbf{AS}$ is a hyperellipse in $\Re^m$. Some properties of $\mathbf{A}$ can be defined in terms of the shape of $\mathbf{AS}$.

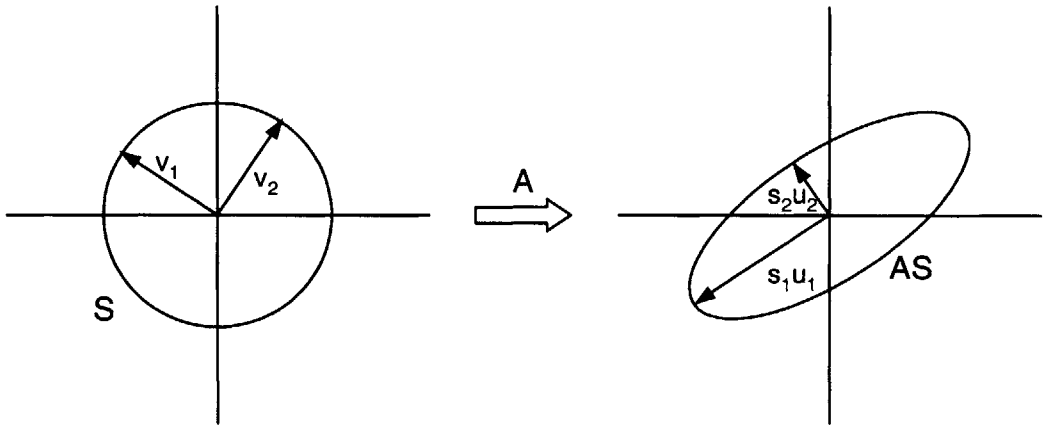

Figure B.1 SVD of a 2x2 Matrix (adapted from [Trefethen, 1997])

The n *singular values* of $\mathbf{A}$ are the lengths of the principal semiaxes of $\mathbf{AS}$, $s_1, \ldots, s_n$. The convention is to write the singular values in descending order, $s_1 \geq s_2 \geq \ldots \geq s_n$. The mapping $\mathbf{AS}$ extends into as many dimensions of $\Re^m$ as there are non-zero singular values. For the 2×2 $\mathbf{A}$ shown in Figure B.1, if $s_2 = 0$, the ellipse becomes a line along the direction u_1. Each zero singular value corresponds to a dimension in $\Re^m$ that cannot be reached. The relative size of the singular values in comparison with each other indicates the relative gain that different components in $\Re^n$ are subject to under the mapping to $\Re^m$.

The unit vectors $\{u_1, u_2, \ldots, u_n\}$ are the n *left singular vectors* of $\mathbf{A}$. These are the unit vectors along the direction of each principal semiaxis of $\mathbf{AS}$, numbered to correspond to the singular values. The vector $s_i u_i$ is the i^{th} largest principal semiaxis.

The unit vectors $\{v_1, v_2, \ldots, v_n\} \in \mathbf{S}$ are the n *right singular vectors* of $\mathbf{A}$. These are the unit vectors that are the preimages of the principal semiaxes of $\mathbf{AS}$, numbered so that

$$\mathbf{A}v_j = s_j u_j \qquad 1 \leq j \leq n \tag{B.1}$$

This collection of vector equations can be expressed as

$$\mathbf{AV} = \hat{\mathbf{U}}\hat{\mathbf{S}} \tag{B.2}$$

If we assume that $\mathbf{A}$ is real and has full rank, $\hat{\mathbf{S}}$ is an $n \times n$ diagonal matrix with positive real entries, $\hat{\mathbf{U}}$ is an $m \times n$ matrix with orthonormal columns, and $\mathbf{V}$ is an $n \times n$ matrix with orthonormal columns. Therefore $\mathbf{V}$ is unitary, with $V^* = V^{-1}$ and we obtain

$$\mathbf{A} = \hat{\mathbf{U}}\mathbf{S}\hat{\mathbf{V}}^* \tag{B.3}$$

This is the *reduced* singular value decomposition of $\mathbf{A}$. The *full* decomposition is constructed by appending an additional $m-n$ orthonormal columns onto $\hat{\mathbf{U}}$, so that the appended matrix $\mathbf{U}$ is unitary and its columns form a basis for $\Re^m$. Since $\hat{\mathbf{U}}$ has changed, $\hat{\mathbf{S}}$ must change too. To keep the product unaltered, appended columns in $\mathbf{U}$ must be multiplied by zero. $\mathbf{S}$ is constructed by appending $m-n$ additional rows of zeros to the bottom of $\hat{\mathbf{S}}$. The new full decomposition is

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^* \tag{B.4}$$

Under the full decomposition we can discard the assumption that $\mathbf{A}$ is full rank. If $\mathbf{A}$ has rank $r < n$ there are r non-zero singular values, and only r of the left singular vectors u_i are determined by the geometry of the hyperellipse. The unitary matrix $\mathbf{U}$ is now constructed by adding $m-r$ arbitrary orthonormal columns to the r orthonormal columns of $\hat{\mathbf{U}}$. Similarly, $n-r$ arbitrary orthonormal columns replace the last $n-r$ columns of $\mathbf{V}$.

The first r columns of $\mathbf{U}$ span the column space of $\mathbf{A}$ and the remaining $m-r$ columns span the nullspace of $\mathbf{A}^T$. The first r columns of $\mathbf{V}$ span the row space of $\mathbf{A}$ and the remaining $m-r$ columns span the nullspace of $\mathbf{A}$.

If $\mathbf{A}$ is square and full rank, the inverse is

$$\mathbf{A}^{-1} = \mathbf{V}\mathbf{S}^{-1}\mathbf{U}^* \tag{B.5}$$

For non-square $\mathbf{A}$ with all non-zero singular values, replacing $\mathbf{S}^{-1}$ with $\mathbf{S}^{\chi} = \mathrm{diag}(s_i^{-1})$ yields the pseudoinverse $\mathbf{A}^{\chi}$.

Singular Values and Eigenvalues. In general, the singular values are not the eigenvalues of $\mathbf{A}$.The squares of the singular values, s_i^2, are the eigenvalues of $\mathbf{A}^T\mathbf{A}$ (and $\mathbf{A}\mathbf{A}^T$). These are symmetric matrices, and their orthogonal eigenvectors are the left or right singular vectors of $\mathbf{A}$. If $\mathbf{A}$ is Hermitian the eigenvectors of $\mathbf{A}$ are orthogonal and the singular values are the same as the eigenvalues.

Singular Values and Matrix Rank. We have seen that the rank of a matrix is equal to the number of non-zero singular values. Some singular values may be very small relative to the largest singular values. It is important to realise that these singular values do not contribute significantly to the rank of the matrix. Examination of the full set of singular values is necessary to determine which singular values contribute to the rank. In cases where the singular values are smoothly distributed over a wide range, the rank may not be well defined. Calculation of the singular values for large poorly conditioned matrices may result in a large range of singular values. Singular values which should be zero are assigned very small values by the numerical calculation. It may be impossible to determine which singular values are zero and which are very small.

B.2 Solution of Non-Full Rank Systems using SVD

Approximate solutions to non-full rank systems of the type

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{B.6}$$

can be found by eliminating the insignificant singular values. Determining which singular values are significant depends on the specific problem and the distribution of singular values. The idea is to discard singular values which are orders of magnitude smaller than the largest singular values. Where there is a clear jump in the distribution of singular values selection of the significant singular values is obvious. However, where the range of singular values is large but the distribution is smooth, selection is not so obvious.

We discuss the case where $\mathbf{A}$ is $n \times n$. The process is illustrated in Figure B.2. Begin by selecting the r significant singular values in $\mathbf{S}$ (the grey part of the diagonal). The r columns in $\mathbf{U}$ and $\mathbf{V}$ which correspond to these singular values are also selected.

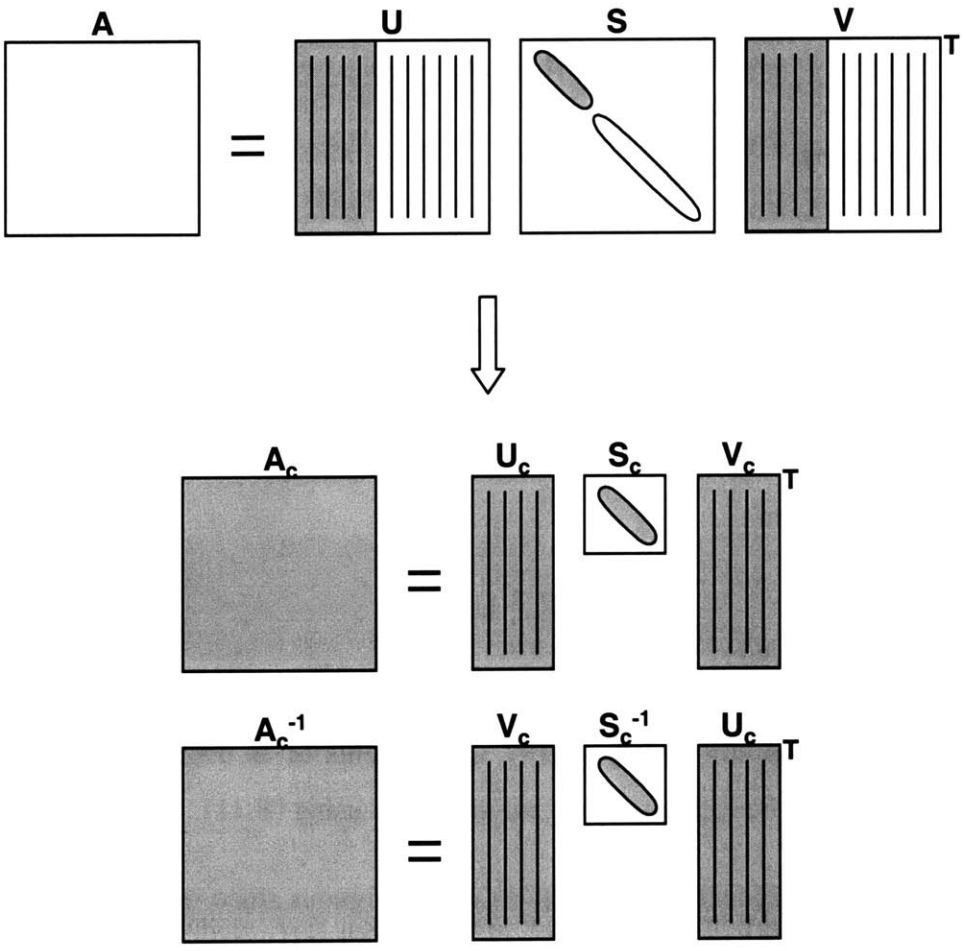

Figure B.2 Singular Value Removal

Components of input vectors $\mathbf{x}$ which fall in the space spanned by $\{v_{r+1}, \ldots, v_n\}$ experience low relative gain under the mapping $\mathbf{A}$. Similarly, the output vector $\mathbf{b}$ does not have significant components in the space spanned by $\{u_{r+1}, \ldots, u_n\}$.

Recall from the earlier discussion that only the first r columns are determined by the geometry of the hyperellipse. The remaining $n - r$ columns of $\mathbf{U}$ span the left nullspace

of $\mathbf{A}$. Vectors $\mathbf{x}_h$ that fall in the nullspace can be added to (B.1) without affecting the result, since $\mathbf{A}\mathbf{x}_h = \mathbf{0}$

$$\mathbf{A}(\mathbf{x} + \mathbf{x}_h) = \mathbf{A}\mathbf{x} = \mathbf{b} \tag{B.7}$$

Retaining only the first r singular values, the "cropped" matrix $\mathbf{A}$ is now given by

$$\mathbf{A}_c = \mathbf{U}_c\mathbf{S}_c\mathbf{V}_c^T \tag{B.8}$$

Its inverse is given by

$$\mathbf{A}_c^{-1} = \mathbf{V}_c\mathbf{S}_c^{-1}\mathbf{U}_c^T \tag{B.9}$$

The original system $\mathbf{A}\mathbf{x} = \mathbf{b}$ can now be written as

$$\mathbf{U}_c\mathbf{S}_c\mathbf{V}_c^T\mathbf{x} = \mathbf{b} \tag{B.10}$$

and $\mathbf{x}$ can be found from

$$\mathbf{x} = \mathbf{V}_c\mathbf{S}_c^{-1}\mathbf{U}_c^T\mathbf{b} \tag{B.11}$$

The solution $\mathbf{x}$ is not unique. Additional vectors $\mathbf{x}_h$ that fall in the span of $\{v_{r+1}, \ldots, v_n\}$ can be added to $\mathbf{x}$ and still satisfy (B.10). Components of $\mathbf{b}$ that fall in the span of $\{u_{r+1}, \ldots, u_n\}$ cannot be captured when solving for $\mathbf{x}$ using (B.11).

Systems with few significant singular values are ambiguous since information is lost in the forward mapping $\mathbf{A}\mathbf{x}$ as well as the reverse mapping $\mathbf{A}^{-1}\mathbf{b}$.

Appendix C

MATLAB CODE

C.1 Master File

```
%eval_one_freq.m
%Karen Marais
%August 2001
%Evaluate SPIR performance at a single frequency

close all

new_psf = input('new psf? ');
if new_psf == 1

clear all

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%
%psf matrix

ii = sqrt(-1);

%setup
lambda = 0.03;
f_c = 10E9; %carrier frequency, Hz
r_ground = 1000E3;
d_cross = 1000;
D = 1;
B = lambda*r_ground/d_cross       %this set based on resolu-
tion (min value)
FOV = 2*asin(lambda/D);

k_max = ceil(FOV*B/lambda)
B = (k_max-1)*lambda/FOV

%sample at the sampling rate
```

```
n_cells = floor(2*B/lambda*FOV);
%n_cells = floor(B/lambda);
n_cells = n_cells + 1 - mod(n_cells, 2)     %force number of
cells to be odd, this ensures psi = 0 is included

%-------------------------------------------------------------
---------------------------
%unrestricted
%3 apertures
%x_raw = [0 1 k_max];                                    %58
%4 apertures
%x_raw = [0 1 4 k_max];                                  %123
%5 apertures
%x_raw = [0 1  4  6 k_max];                              %123
%6 apertures
%x_raw = [0 3  4  9  11 k_max];                          %123
%x_raw = [0 4  5  7  13 k_max];
%7 apertures
%x_raw = [0 4  5  6  13 16 k_max];                       %123!!!
%x_raw = [0 6  7  9  11 19 k_max];
%x_raw = [0 1  4  10 12 17 k_max];
%x_raw = [0 1  8  11 13 17 k_max];
%8 apertures
%x_raw = [0 6  9  10 17 22 24 k_max];
%x_raw = [0 8  9  12 18 23 25 k_max];
%x_raw = [0 14 15 18 24 26 31 k_max];
%x_raw = [0 13 14 16 21 25 31 k_max];
%9 apertures
%x_raw = [0 8  18 19 22 24 31 39 k_max];
%11 apertures
%x_raw = [0 16 17 28 36 42 46 49 51 73 k_max];
%x_raw = [0 7  22 27 28 31 39 41 57 64 k_max];
%12 apertures
x_raw = [0 18 19 22 31 42 48 56 58 63 91 k_max];

%restricted
%6 apertures
%x_raw = [0 1 4 7  9 k_max];
%x_raw = [0 1 2 6  9 k_max];
%7 apertures
%x_raw = [0 1 2 6  10 13 k_max];
%x_raw = [0 1 6 9  11 13 k_max];
%x_raw = [0 1 4 5  11 13 k_max];
%8 apertures
%x_raw = [0 1 2 6  10 14 17 k_max];
%x_raw = [0 1 2 3  8  13 17 k_max];
%x_raw = [0 1 2 8  12 14 17 k_max];
%x_raw = [0 1 2 8  12 15 17 k_max];
%9 apertures
%x_raw = [0 1 2 11 15 18 21 23 k_max];
%x_raw = [0 1 4 10 16 18 21 23 k_max];
```

```
%10 apertures
%x_raw = [0 1 2 14 18 21 24 27 29 k_max];
%x_raw = [0 1 4 10 16 22 24 27 29 k_max];
%x_raw = [0 1 3 6  13 20 24 28 29 k_max];
%11 apertures
%x_raw = [0 1 3 6 13 20 27 31 35 36 k_max];
%12 apertures
%x_raw = [0 1 3 6 13 20 27 34 38 42 43 k_max];
x_raw

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%
%PSF Generation
N_sats = length(x_raw);                                    %number
of apertures
x = (x_raw*B./(max(x_raw) - min(x_raw)))';

psi = linspace(-0.5*FOV, 0.5*FOV, n_cells + 2);%once FOV
psi = psi(2:end - 1);
psi_mat = repmat(sin(psi), [n_cells 1]);
psf_mat = zeros(n_cells);

for i_sat=1:length(x) %add the contributions from each aper-
ture
                psf_mat    =    psf_mat    +    exp(+ii*2*pi/
lambda*x(i_sat)*(psi_mat - psi_mat'));
end

psf_mat = sinc(D/lambda*psi_mat).*psf_mat;
psf_mat2 = psf_mat';

end %new psf

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%
%Signal Generation
%-------------------------------------------------------------
-----------------------------
%construct the rcs vector sigma
randn('state',sum(100*clock));
clutter_cells = [29 30 31 69 70 71 109 110 111];
clutter_size = 10*abs(randn(length(clutter_cells),1));

tgt_cells = [20 22];
tgt_size = [1 2];

sigma = zeros(n_cells, 1);
sigma(tgt_cells) = tgt_size;
sigma(clutter_cells) = clutter_size + sigma(clutter_cells);
```

```
%-------------------------------------------------------------
----------------------------
%Construct the composite time signal
%remember m is a vector!

rr  = find(sigma > 0);
spq = zeros(n_cells, 1);
for i = 1:length(rr)
    spq = spq + sigma(rr(i))*psf_mat(:, rr(i));
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%
%Target Recovery
gamma = 0.1;
e = 0.01;

%for clutter cancellation
mask = zeros(n_cells, 1);
mask(clutter_cells) = 1;

[cleaned    ,    cluttered]    =    clean_clutter(real(spq),
real(psf_mat), gamma, e, mask);
cleaned = abs(cleaned);
cluttered = abs(cluttered);

%clean_no_comp = clean(spq, psf_mat, gamma, e);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%
%plotting
subplot(311)
psf_vec = abs(psf_mat(floor(n_cells/2), :));
psf_vec = psf_vec/max(psf_vec);
plot(psf_vec, '-x')
title('Point Spread Function')
%xlabel('Cross-Range Cell')
%ylabel('Normalised Amplitude')
grid on
axis([0 n_cells 0 1.2])

subplot(312)
plot(abs(spq)/max(abs(spq)), 'b-x')
%xlabel('Cross-Range Cell')
ylabel('Normalised Amplitude')
title('Constructed Signal')
grid on
axis([0 n_cells 0 1.2])

subplot(313)
sigma2 = sigma;
```

```
sigma2(find(sigma == 0)) = NaN;
plot(sigma2, 'o', 'MarkerEdgeColor','b','MarkerFaceColor',
'b', 'MarkerSize', 8)
hold on
%plot(cleaned, 'r-x')
stem(cleaned, 'rx')
legend('True Position', 'Recovered Position')
%ylabel('Amplitude')
xlabel('Cross-Range Cell')
grid on
axis([1 n_cells 0 1.2])
title('Recovered Signal')
axis([0 n_cells 0 1.2])

%-------------------------------------------------------------
---------------------------
figure
sigma2 = sigma;
sigma2(find(sigma == 0)) = NaN;
plot(tgt_cells, sigma2(tgt_cells), 'o', 'MarkerEdge-
Color','b','MarkerFaceColor', 'b', 'MarkerSize', 8)
hold on
plot(clutter_cells, sigma2(clutter_cells), 'd', 'MarkerEdge-
Color','g','MarkerFaceColor', 'g', 'MarkerSize', 8)
%plot(cleaned, 'r-x')
stem(cleaned, 'rx')
legend('True Target Position','True Clutter Position'
,'Recovered Position')
%ylabel('Amplitude')
xlabel('Cross-Range Cell')
grid on
axis([1 n_cells 0 1.2])
title('Recovered Signal')
axis([0 n_cells 0 1.2*max(sigma)])

%-------------------------------------------------------------
---------------------------
figure
subplot(211)
plot(abs(spq)/max(abs(spq)), 'b-x')
%xlabel('Cross-Range Cell')
ylabel('Normalised Amplitude')
title('Constructed Signal')
grid on
axis([0 n_cells 0 1.2])

subplot(212)
sigma2 = sigma;
sigma2(find(sigma == 0)) = NaN;
plot(sigma2, 'o', 'MarkerEdgeColor','b','MarkerFaceColor',
'b', 'MarkerSize', 8)
```

```
hold on
%plot(cleaned, 'r-x')
stem(cleaned, 'rx')
legend('True Position', 'Recovered Position')
ylabel('Normalised Amplitude')
xlabel('Cross-Range Cell')
grid on
axis([1 n_cells 0 1.2])
title('Recovered Signal')
axis([0 n_cells 0 1.2])

%-------------------------------------------------------------
----------------------------
figure
subplot(211)
plot(tgt_cells,    sigma2(tgt_cells),    'o',    'MarkerEdge-
Color','b','MarkerFaceColor', 'b', 'MarkerSize', 8)
hold on
stem(cleaned, 'rx')
legend('True Target Position','Recovered Position')
%ylabel('Amplitude')
%xlabel('Cross-Range Cell')
grid on
axis([1 n_cells 0 1.2])
title('Recovered Target Signal')
max_val = max(max(sigma2(tgt_cells)),max(cleaned))
axis([0 n_cells 0 1.2*max_val])

subplot(212)
plot(clutter_cells, sigma2(clutter_cells), 'd', 'MarkerEdge-
Color','g','MarkerFaceColor', 'g', 'MarkerSize', 8)
hold on
stem(cluttered, 'rx')
legend('True Clutter Position' ,'Recovered Position')
%ylabel('Amplitude')
xlabel('Cross-Range Cell')
grid on
%axis([1 n_cells 0 1.2*max(sigma)])
title('Recovered Clutter Signal')
max_val = max(max(sigma2(clutter_cells)),max(cluttered))
axis([0 n_cells 0 1.2*max_val])
```

C.2 Implementation of CLEAN

```
function Y=clean(spq,psf_mat,gamma,e)
%Karen Marais
%August 2001
%Implementation of CLEAN algorithm

[cross, doppler] = size(spq);
I_clean=zeros(cross,doppler);

for i_dopp = 1:doppler
    %each frequency is CLEANED independently
    vec_to_clean = spq(:, i_dopp);
    vec_in_progress = vec_to_clean;
    max_value = max(abs(vec_to_clean));
          max_value_location  =  find(abs(vec_to_clean)  ==
max_value);
     max_value_location = max_value_location(1);       %when
there are two points with same max
    max_value = vec_to_clean(max_value_location);
    flag = 0;
    loop_cnt = 0;

    while ((abs(max_value) > e) & flag == 0)
        loop_cnt = loop_cnt + 1;
        old_max_value = max_value;
        psf_vec = psf_mat(:, max_value_location);
        psf_vec = psf_vec/max(abs(psf_vec));
                    I_clean(max_value_location,  i_dopp)  =
I_clean(max_value_location, i_dopp) + max_value*gamma;
                    vec_in_progress  =  vec_in_progress  -
max_value*gamma*psf_vec;
        max_value = max(abs(vec_in_progress));
          max_value_location = find(abs(vec_in_progress) ==
max_value);
       max_value_location = max_value_location(1);     %when
there are two points with same max
        max_value = vec_in_progress(max_value_location);

        if (abs(max_value) > abs(old_max_value))
            flag = 1;
            disp('unstable')
        end
    end
          I_clean(:,  i_dopp)  =  I_clean(:,  i_dopp)  +
vec_in_progress;
end

%close
loop_cnt
Y = I_clean;
```

C.3 Implementation of CLEAN with Clutter Reduction

```
function [Y,Z]=clean_clutter(spq,psf_mat,gamma,e, mask)
%Karen Marais
%August 2001
%Implementation of CLEAN algorithm

[cross, doppler] = size(spq);
I_clean=zeros(cross,doppler);
I_clutter=zeros(cross,doppler);

for i_dopp = 1:doppler
    %each frequency is CLEANED independently
    vec_to_clean = spq(:, i_dopp);
    vec_in_progress = vec_to_clean;
    masked_vec = mask.*vec_in_progress;

    max_value = 0;
    old_max_value = Inf;
    [max_value max_value_location] = max(masked_vec);
    max_value = masked_vec(max_value_location);

    loop_cnt1 = 0;
    flag = 0;
    disp('clutter')
    %locate clutter
    while ((abs(max_value) > e) & flag == 0)
        loop_cnt1 = loop_cnt1 + 1;

        psf_vec = psf_mat(:, max_value_location);
        psf_vec = psf_vec/max(abs(psf_vec));
                  I_clutter(max_value_location,  i_dopp)  =
I_clutter(max_value_location, i_dopp) + max_value*gamma;
                    vec_in_progress  =  vec_in_progress  -
max_value*gamma*psf_vec;

        masked_vec = mask.*vec_in_progress;
        old_max_value = max_value;
        [max_value max_value_location] = max(masked_vec);
        max_value = masked_vec(max_value_location);

%         plot(abs(masked_vec))
%         drawnow
%         max_value
%         %masked_vec(max_value_location)
%         pause

        if (abs(max_value) > abs(old_max_value))
            flag = 1;
            disp('unstable clutter')
```

```
        end
    end
    I_clutter(:, i_dopp) = I_clutter(:, i_dopp) + masked_vec;

    %renormalise
    temp_norm = max(abs(vec_in_progress));
    vec_in_progress = vec_in_progress/temp_norm;

    max_value = 0;
    old_max_value = Inf;
    [max_value max_value_location] = max(vec_in_progress);
    max_value = vec_in_progress(max_value_location);

    %locate targets
    loop_cnt2 = 0;
    flag = 0;
    disp('targets')
    while ((abs(max_value) > e) & flag == 0)

        loop_cnt2 = loop_cnt2 + 1;
        psf_vec = psf_mat(:, max_value_location);
        psf_vec = psf_vec/max(abs(psf_vec));
                    I_clean(max_value_location,   i_dopp)   =
I_clean(max_value_location, i_dopp) + max_value*gamma;
                    vec_in_progress   =   vec_in_progress   -
max_value*gamma*psf_vec;
        old_max_value = max_value;
                        [max_value   max_value_location]   =
max(vec_in_progress);
        max_value = vec_in_progress(max_value_location);

        if (abs(max_value) > 1.1*abs(old_max_value))
            flag = 1;
            disp('unstable targets')
        end
    end

          I_clean(:,   i_dopp)   =   I_clean(:,   i_dopp)   +
vec_in_progress;
end

%close
loop_cnt1
loop_cnt2

%normalise to retain relative amplitudes
norm_factor   =   (sum(I_clutter.^2)   +   sum(I_clean.^2))/
sum(abs(spq).^2);

Z = I_clutter*norm_factor;
Y = I_clean*norm_factor*temp_norm;
```

3231-36