BIOLOGICALLY-INSPIRED RADAR SENSING

by

GALEN MAXWELL REICH

A thesis submitted to the University of Birmingham for the degree of
DOCTOR OF PHILOSOPHY

Department of Electrical Engineering
School of Engineering
College of Engineering and Physical Sciences
University of Birmingham
September 2019

Abstract

The natural world has an unquantifiable complexity and natural life exhibits remarkable techniques for responding to and interacting with the natural world. This thesis aims to find new approaches to radar systems by exploring the paradigm of biologically-inspired design to find effective ways of using the flexibility of modern radar systems. In particular, this thesis takes inspiration from the astonishing feats of human echolocators and the complex cognitive processes that underpin the human experience. Interdisciplinary research into human echolocator tongue clicks is presented before two biologically-inspired radar techniques are proposed, developed, and analyzed using simulations and experiments. The first radar technique uses the frequency-diversity of a radar system to localize targets in angle, and the second technique uses the degrees-of-freedom accessible to a mobile robotic platform to implement a cognitive radar architecture for obstacle avoidance and navigation.

Dedication

In memoriam.

'Gran'

Sheila Reich

1927 - 2018

Steve Baker

1963 - 2014

Acknowledgements

I would like to thank all of the people who have helped and supported me during the course of my research.

Thank you to my supervisors, Dr. Mike Antoniou and Prof. Chris Baker, for encouraging my ideas, guiding my development, and keeping me and my research on track.

Thank you to my parents, for the years of love and care which have shaped my development, and enabled me to become the inquisitive and diligent person that I am today.

Thank you to Jo, for taking this journey with me, and providing the unceasing support, care, and encouragement which has kept me going through the most difficult moments.

Thank you to Alan Yates and Alison Morey, for the invaluable support and advice which helped my time in the department go safely and smoothly.

Thank you to all my fellow students in the Microwave Integrated Systems Laboratory (MISL) for the enthusiasm, assistance, and friendship which helped make my time in the research group rewarding and enjoyable.

Thank you to the Engineering and Physical Sciences Research Council (EPSRC) for the doctoral stipend which enabled this research.

Table of Contents

List of Figures

List of Tables

List of Acronyms

AWGN additive white Gaussian noise.

BTD binaural timbre difference.

EM electro-magnetic.

FFT fast Fourier transform.

FPGA field programmable gate array.

GPS global positioning system.

HRTF head-related transfer function.

IFFT inverse fast Fourier transform.

IID inter-aural intensity difference.

ILD inter-aural level difference.

IPD inter-aural phase difference.

ITD inter-aural time difference.

LFM linear frequency modulated.

MIMO multiple-input multiple-output.

PRF pulse repetition frequency.

PSD power spectral density.

PSF point spread function.

RCS radar cross-section.

RMS root-mean-squared.

SCAT spectrogram correlation and transformation.

SNR signal-to-noise ratio.

SPL sound pressure level.

STFT short time Fourier transform.

TDOA time difference of arrival.

VNA vector network analyzer.

WAF wideband ambiguity function.

Chapter 1

Introduction

> There are as many worlds as there are kinds of days, and as an opal changes its colors and its fire to match the nature of a day, so do I.
>
> — John Steinbeck
>
> *Travels with Charley: In Search of America, 1962*

1.1 Overview

A man stands on the side of a busy street, head cocked, listening to the bustle around him. He describes the wooden facade of the buildings behind him as a tour guide would. Pausing a moment, he waits for a gap in the traffic before crossing the road, avoiding a car parked in his path. On the other side he points to the low bushes surrounding a play park and, after opening the gate, explains the layout of the metal play equipment inside it. Throughout the tour, this man has been clicking purposefully; a series of tongue clicks illuminating his environment, allowing this blind expert echolocator to perceive and interact with his environment.

Fundamentally, radar is echolocation, and there is much that radar engineers can learn from natural echolocators. The hope for this thesis is that, by understanding the different ways that nature approaches the task of perception using echolocation, novel radar techniques can be developed. This thesis sets out natural approaches to echolocation and localization, and it then presents research that developed the fundamental understanding of human echolocation. Building on these ideas, the thesis will go on to present and develop two biologically-inspired approaches to radar: the first taking inspiration from the frequency diversity of echolocator calls, and the second taking inspiration from the cognitive processes seen in humans and other echolocating mammals.

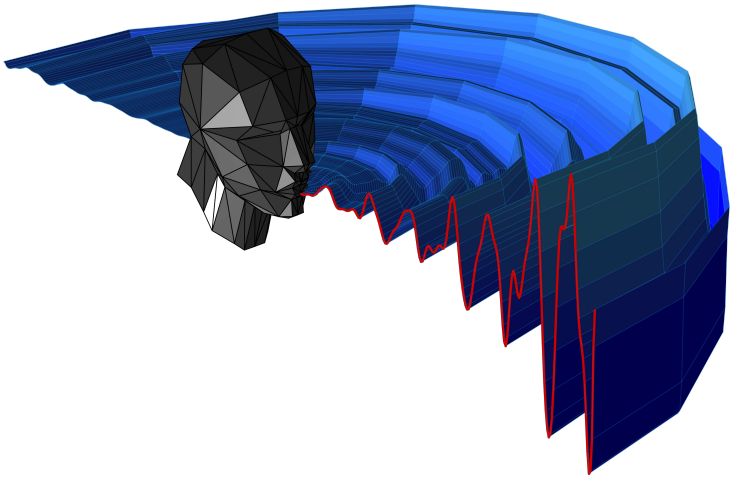


Figure 1.1: Visualization of a human echolocator tongue click.

1.2 Motivation and Aims

Traditionally, designers of radar systems have sought to increase the operational range of their systems, pushing the limits of hardware and signal processing to achieve systems with low false alarm rates and high signal-to-noise ratios (SNRs). In pushing the frontiers of radar, the transmit power of radar systems has generally increased and antennas have been increasingly optimized to maximize gain and balance power budgets, ultimately achieving optimal performance characteristics for the task at hand. However, as radar technology develops into smaller form factor systems that operate at higher frequencies, there is a dramatically increasing number of potential applications. Ranging from autonomous vehicles to gesture recognition, mobile robotics to structural analysis, the field of radar is further from its origins than ever before. This wide variety of short-range sensing applications brings with it new goals and constraints that are no longer satisfied by historic optimizations. As radar diversifies, so too should its design paradigms: increasing detection range and decreasing false alarm rate should no longer be the motivation for these new short-range technologies for which a new paradigm may be better suited.

That is not to say that this thesis eschews historic radar principles; traditional principles and techniques provide a solid foundation on top of which new approaches can be built. Because of the wide variety of adverse environments involved with traditional radar approaches, high performance hardware was developed that provides modern short-range systems with more flexibility than was previously possible. When combined with the advances of digital signal processing that have made trialling, tweaking, and implementing new sensing approaches easier and more dynamic than ever before, this increased flexibility provides scope for exploration of new techniques to solve the challenges of a new generation of radar systems.

But, even with the increased flexibility of modern systems, how should researchers address the varied and complex challenges that modern radar systems and applications require?

The natural world has an unquantifiable complexity and natural life exhibits remark-

able techniques for responding to and interacting with the natural world. These techniques have often evolved over millennia, and the evolutionary process has overcome a plethora of natural optimization challenges. From ants that optimize routing from food sources to their colonies, to the streamlined beaks of kingfishers that optimize quiet entry into water, we are surrounded by natural systems that are able to excel in complex environments. By taking inspiration from these incredible biological systems, researchers can gain a fresh perspective and insight into modern engineering challenges.

The research presented in this thesis aims to find and develop new approaches to radar systems by exploring the paradigm of biologically-inspired design to find effective ways of using the potential of modern radar systems. Others have taken this approach in the field of radar and many interesting and powerful techniques have been developed as a result, demonstrating the efficacy of the approach. Part of the challenge of this thesis is to develop techniques that are novel and take inspiration from relatively unexplored areas of the natural world (from the perspective of a radar engineer).

This author takes inspiration from the astonishing feats of human echolocators and the complex cognitive processes that underpin the human experience. If it is possible to gain insight into the varied mechanisms of human perception that result in a rich understanding of our environments, then perhaps it is possible to develop radar systems that achieve a richer perception of their environments.

1.3 Thesis Layout

This thesis finds biological inspiration in the passive techniques used for sound localization by humans (and other mammals) as well as the active techniques for localization used by mammalian echolocators. Accordingly, a review of research into different approaches to localization is presented in Chapter 2. Both passive and active localization approaches are considered, as well as existing approaches to biologically-inspired radar to outline the research landscape in which this work is placed.

Throughout this thesis, various theoretical ideas and approaches are used, some are standard in radar research and some are alien to it. The details of these ideas and approaches is not part of the novel contribution of this research, but explanations of them are provided for completeness in Chapter 3.

Human echolocation is an area of relatively new research in acoustics, psychology, and engineering. Chapter 4 provides the first detailed description of the properties of human echolocator tongue clicks. This description is then used to synthesize artificial tongue clicks that share the key properties of the echolocator tongue clicks and can be easily used by others.

Variants of a novel wideband target localization technique are introduced in Chapter 5, which also presents simulated and experimental results of using these techniques. These approaches are inspired by the properties of human echolocator tongue clicks and the human echolocation process, utilizing radar analogues of human sound localization cues.

A cognitive radar system is introduced in Chapter 6. This system takes biological inspiration from the cognitive structures of human brains, and implements a closed-loop autonomous control system for a mobile robotic platform. Simulated and experimental results into the use of memory in cognitive radar systems are also detailed, and allows initial observations about the use of memory in cognitive radar systems to be made.

Finally, Chapter 7 outlines conclusions from the preceding chapters, and draws conclusions based on the aggregation of research in this thesis. It also provides a series of suggestions for future avenues of research work which naturally extend from the presented material.

1.4 Novel Contributions

Novel parametrization of human echolocator tongue clicks was conducted which allowed an analytical click model to be derived, and open access synthetic human echolocator tongue clicks to be constructed and released. This parametrization was based on original analysis of a large database of human echolocator tongue clicks which was conducted in unprecedented detail [1, 2]. This analysis allowed several observations to be made. Firstly, that the frequency-components of the clicks have a constant frequency as a function of time and span the 2-13 kHz band, with a primary spectral peak in the 2-4 kHz band. Secondly, that the time domain profile of the clicks has a rapid onset and a typical duration of 3 ms. Thirdly, that the clicks have a broad beamwidth and illuminate a wide area before the echolocator. Finally, that the clicks have a good range resolution of approximately 2 cm, but are ambiguous in Doppler. This research is detailed in Chapter 4.

The theoretical basis for a family of closely-related radar techniques was developed and is presented in Chapter 5 [3, 4, 5]. These techniques use a combination of waveform frequency diversity and spatial encoding by the receiver antennas to localize a target. The techniques were verified through simulation and experiment, confirming the fundamental possibility that they can operate successfully using single measurements. Furthermore, analysis of each technique's performance is provided.

An autonomous cognitive radar system architecture for navigation was developed and tested using simulated and experimental results to assess its performance [6, 7]. To the author's best knowledge this work, presented in Chapter 6, is the first investigation into the effects of using memory in cognitive radar systems. This work finds that some memory is useful for a cognitive radar, but that too much memory can hinder such a system.

1.5 Publications Arising

The novel contributions presented in the previous section led to the following publications.

[1] L. Thaler, G. Reich, X. Zhang, D. Wang, G. Smith, M. Cherniakov, C. Baker, D. Kish, and M. Antoniou, "Mouth-clicks used by blind expert human echolocators - signal description and model based signal synthesis," *PLoS Computational Biology*, vol. 13, Aug. 2017

[2] X. Zhang, G. M. Reich, M. Antoniou, M. Cherniakov, C. J. Baker, L. Thaler, D. Kish, and G. E. Smith, "Human echolocation: waveform analysis of tongue clicks," *Electronics Letters*, vol. 53, no. 9, pp. 580–582, 2017

[3] G. M. Reich, M. Antoniou, and C. J. Baker, "Biologically-inspired wideband target localisation," *Sonar Navigation IET Radar*, vol. 12, no. 12, pp. 1410–1418, 2018

[4] G. M. Reich, M. Antoniou, and C. J. Baker, "Bio-inspired techniques for target localization," in *2018 IEEE Radar Conference (RadarConf18)*, pp. 1239–1244, Apr. 2018

[5] G. M. Reich, M. Antoniou, and C. J. Baker, "Frequency-dependent target localization," in *International Conference on Radar Systems (Radar 2017)*, pp. 1–6, Oct. 2017

[6] G. M. Reich, M. Antoniou, and C. J. Baker, "Memory-Enhanced Cognitive Radar for Autonomous Navigation," 2019. Manuscript submitted for publication

[7] G. M. Reich, M. Antoniou, and C. J. Baker, "MIMO Radar for Cognitive Robot Platform Control and Navigation," in *2019 International Radar Conference (RADAR)*, pp. 1–6, 2019. In Press

Chapter 2

Approaches to Localization

> The clearest way into the Universe is through a forest wilderness.
>
> — John Muir
>
> *Diary Entry, July 1890*

2.1 Introduction

Because this research aims to develop novel biologically-inspired radar techniques, this chapter provides overviews of the mechanisms for sound localization seen in nature; the current state of research into echolocation in bats and humans; and the existing biologically-inspired approaches developed by the radar community.

By understanding the acoustic approaches taken by creatures in the natural world and the overlap with existing radar techniques, we can attempt to incorporate these techniques and designs into the engineering of electromagnetic systems. Initially, this chapter will consider the 'passive' situation that the majority of humans experience on a daily basis: locating the position of a sound source. Secondly, the 'active' case of echolocation and the additional information that it provides will be explored. Finally, existing biologically-inspired radar and sonar techniques will be examined.

This chapter does not provide a general overview of acoustics as this is beyond the scope of this thesis. However many good acoustic reference books are available for the interested reader [8, 9, 10, 11]. Although this chapter is primarily a literature review, it does not aim to provide a comprehensive review of all bat and human echolocation research, but does examine key aspects that are relevant to the research presented in this thesis.

2.2 Psychoacoustics: Passive Localization

In our daily lives as humans we encounter many situations where we are required to locate the origin of a sound. For example, a lost mobile phone may be found by ringing it and listening to the sounds it produces. We are not the only animals to localize sounds; many birds of prey are reliant on sound localization for hunting, and must be able to determine the precise location of prey from the faint sounds that they produce. This section investigates the mechanisms for sound localization, particularly focusing on humans, but drawing on evidence from the wider natural world where appropriate.

2.2.1 Binaural Cues for Localization

The psychology of perception of sound by humans has been a subject of much interest and investigation by scientists and philosophers who have been eager to better understand how we, as humans, perceive our environments. Some of the earliest formal scientific investigations of human perception of sound sources were carried out by John William Stutt, 3rd Baron Rayleigh, more commonly referred to as Lord Rayleigh. Rayleigh conducted a series of open-air experiments in the 1870s which asked that an observer with their eyes closed, point towards the source of a sound. Many different sources of sound were used, including normal speech; individual consonants and vowels; and tuning forks with resonant boxes to amplify the purest of tones possible. Rayleigh's acoustical observations were extensive and go far beyond the scope of this chapter, but he did identify several cues which are examined in more detail in the following sections, and they are:

- Inter-aural level difference (ILD)
- Inter-aural time difference (ITD)
- Inter-aural phase difference (IPD)
- Binaural timbre difference (BTD)

Rayleigh came to several deeply insightful conclusions about binaural localization cues (those reliant on a difference between the ears) which persisted through the development of psychoacoustics and are still relevant today [12, 13, 14, 15, 16].

2.2.2 Inter-aural Level Difference

Rayleigh concluded that "[there] could be no doubt but that relative intensities at the two ears play an important part in the localization of sound" [16, p.218]. This conclusion was reached through the observation that the location of a tuning fork could be reliably determined if the observer was asked to discriminate between fork positions to their right and to their left, but that the same confidence in fork position was absent when the

positions were before and behind the observer. Rayleigh deduced that the observer had no way to tell the difference between signals arriving from forks placed before and behind because the intensity difference between the sound arriving at the ears would be identical. It is worth noting that the observer initially plugged their ears to try and remove any possibility of the sound onset delay at the ears influencing their decision. Once the sound was started, the observer unplugged their ears. The relative difference in intensity of a sound source is widely recognized today as one of the important cues for the localization of sound sources and is often referred to as the ILD or inter-aural intensity difference (IID). Further, Rayleigh observed that, due to the size of the human head, the ILD should only be an effective cue for localizing the source of high-pitched tones, which have a wavelength much smaller than the dimensions of the head and are therefore significantly attenuated by it. This is shown by Equation (2.1), which is given by Rayleigh to estimate the maximum intensity difference between the two ears for a sound of wavelength, λ, and a person of head circumference, $2\pi c$ [16, p.216]. For a tone of 128 Hz, Equation (2.1) indicates that the ILD should be less than 1% of the signal intensity, which Rayleigh believed would be imperceptable to the human ear.

$$\Delta I = \frac{3}{4}\left(\frac{2\pi c}{\lambda}\right)^4 \tag{2.1}$$

Subsequent research on the sound field at the ears by Sivian and White [17] showed that the ILD varied as a function of frequency and azimuth which indicated that the ILD would be an effective cue for frequencies above 300Hz [18]. Work by Feddersen et al. [19] describes the difficulty associated with making accurate measurements of the ILD due to difficulties in reliably positioning a microphone within the ear. Their work controlled for this however, and they were able to publish some of the first reliable ILD measurements for different frequencies and angles. When a sound is presented directly to one side of an observer the ILD varies according to the data given in Table 2.1 which is consistent with the work by Sivian and White, and supports Rayleigh's conclusion that low-frequency localization using an ILD cue is unlikely [19].

Table 2.1: ILD measurements for a sound source positioned to one side of a subject [19, Fig 2.].

Frequency (kHz)	ILD (dB)
0.2	0
0.5	3
1	4
1.8	9
2.5	9
3	9
4	8
5	14
6	14

2.2.3 Inter-aural Time Difference

Rayleigh knew that the commencement of a sound source could also be used as a cue for localization, and it was a cue that he sought to remove from his experiments into ILD by temporarily plugging his participants' ears for the moment of sound onset. The ITD is the second important cue used in sound localization and is particularly useful for localizing sound sources in the azimuth plane, discriminating between the left and right directions reliably [18, pp.520-522]. For reference, Figure 2.1 shows the orientation of the planes that are referred to in this thesis.

A first-order approximation of the ITD is a situation analogous to optical parallax. Figure 2.2 shows the geometry for a sound source located sufficiently far away from the listener for the wavefront to be considered parallel, which gives Equation (2.2)[18, p.521] for the distance difference for a sound to arrive at both ears.

$$D_s = r\left(\theta + \sin\theta\right) \tag{2.2}$$

Research by Woodworth found that in the horizontal plane the ITD could be used to discriminate between sound sources separated by 3^o which corresponded to an ITD of $30\mu s$, indicating that very small time differences can be used by humans for sound localization [18, p.522].

Work by Kuhn [20] investigated the frequency-dependence of the ITD by using a

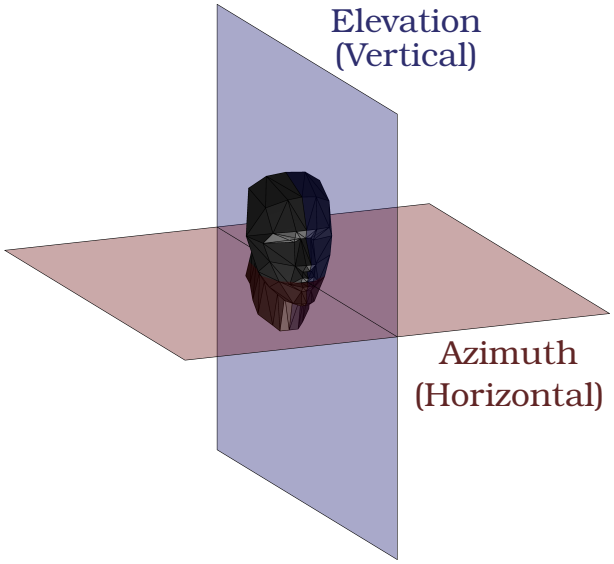


Figure 2.1: The azimuth and elevation planes defined relative to head orientation.

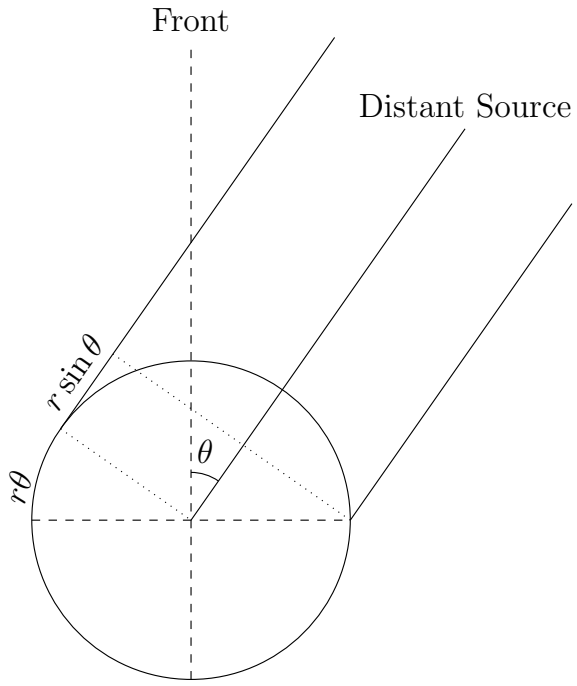


Figure 2.2: ITD geometry for a sound source sufficiently far from the observer for the wavefront to be considered parallel. Adapted from [18, Fig. 143. p.521].

mannequin, resulting in data which shows that at low frequencies (below 500 Hz) the ITD is proportional to $3\sin\theta$ due to diffraction effects of low-frequency waves around the head. At high frequencies (above 3 kHz) the ITD becomes proportional to $2\sin\theta$ due to the creeping wave effects around the head, which corresponds to Woodworth's geometrical model. Over a mid-range of frequencies the proportionality of the ITD varies between 3

and 2 times the sine of the incident angle [20].

For low frequencies the ITD is largest and varies straightforwardly with angle. In the mid-region of frequencies the ITD is frequency-dependent and localization performance using the ITD alone decreases. At high frequencies the ITD is shorter than at low frequencies, but returns to a frequency-independent behaviour. Kuhn's observations show that Woodworth's model, Equation (2.2), is a frequency-independent model that is valid for high frequency ITDs and those for transient clicks where detection is made on the leading edge, both of which rely on the creeping wave phenomena across the surface of the head [20].

2.2.4 Inter-aural Phase Difference

In contrast to the ITD, the IPD refers specifically to the difference of phases at the two ears, not the time difference of sound onset (though many publications draw an equivalence between them as they are closely related and often equivalent). Rayleigh reluctantly concluded that when "a pure tone of low pitch is recognized as being on the right or the left, the only alternative to the intensity theory is to suppose that the judgement is founded upon the difference of phases at the two ears" [16, p.218]. The path difference between the ears introduces a phase difference between them, in addition to a time difference of arrival, which can manifest as a beat between the sounds received at the two ears. By using two tuning forks of slightly different frequencies producing sound in each ear of a participant, Rayleigh was able to control out the ILD and ITD, he observed that the perceived location of the sound source shifted from side to side, corresponding to the phase of the beat frequency. In this way Rayleigh's experiments showed that the perception of left and right *could* be derived solely by the beat resulting from the IPD.

The extent to which humans directly use the IPD cue on a regular basis is unknown, but there is evidence to suggest that the IPD can be used indirectly by causing an ILD, which can then be used for localization [21].

2.2.5 Binaural Timbre Difference

The final cue used by humans for sound source localization is the BTD, which was alluded to by Rayleigh in his work when he observed "that the possibility of detecting whether a source of sound was in front or behind depended [...] upon the compound character of the sound" [12, p.76]. The "compound character" of a sound refers to a sound that is a combination of multiple tones, or alternatively one with a wide frequency bandwidth. It would appear that the complex spectrum of a wide bandwidth sound improves localization ability, but the exact mechanism of this was not known to Rayleigh. One of Rayleigh's contemporaries, a Dr. Stone remarked that Rayleigh's work did not investigate "the pinna, which was a very powerful guiding apparatus [...] to distinguish where a sound came from."[12, p.82]. Dr. Stone's words were extremely insightful as subsequent investigation into how humans localize sound sources revealed that the frequency-dependent effects of the head and the pinna (which is the outer ear) were of vital importance for localizing sound sources in elevation [22, 23].

Experiments into the effects of the pinna showed that, in the absence of ILD and ITD cues for a sound source on the elevation plane (where there is no path length difference between the ears), that participants with their pinna occluded are unable to localize the sound source with any certainty; participants with different degrees of pinna occlusion were able to localize sound sources with performance increasing as occlusion decreased, indicating that the spectral filtering effects of the pinna are vital for this type of sound localization [23, 24].

Research into the precise effects of the pinna, mostly conducted within the field of acoustics, revealed a function termed the head-related transfer function (HRTF) which describes how the bulk of the head and shape of the pinna alter the magnitude of sound signals arriving at the ears as a function of frequency [25, 26, 27]. Because of the 3D structure of the head and pinna, the HRTF is not strictly independent of the range to a sound source. However, the effects of range are negligible compared to the angular effects, and vanish completely if the sound source satisfies the far field approximation that the

incident wavefronts are parallel.

Figure 2.3 shows an example HRTF for a human's left ear, and captures how the gain of the ear varies with frequency and azimuth angle of the incident plane wave. For example, a tone of 10 kHz located at -50^o has a normalized gain of 0 dB, the peak of the HRTF, whereas a tone of 10 kHz located at 90^o (placing it on the far side of the head from the ear) undergoes around 25 dB of attenuation. The spectral features in the HRTF modulate a sound signal with information about its angle of origin. By comparing the tone of an incoming sound received at the ears, information about the sound location can be extracted.

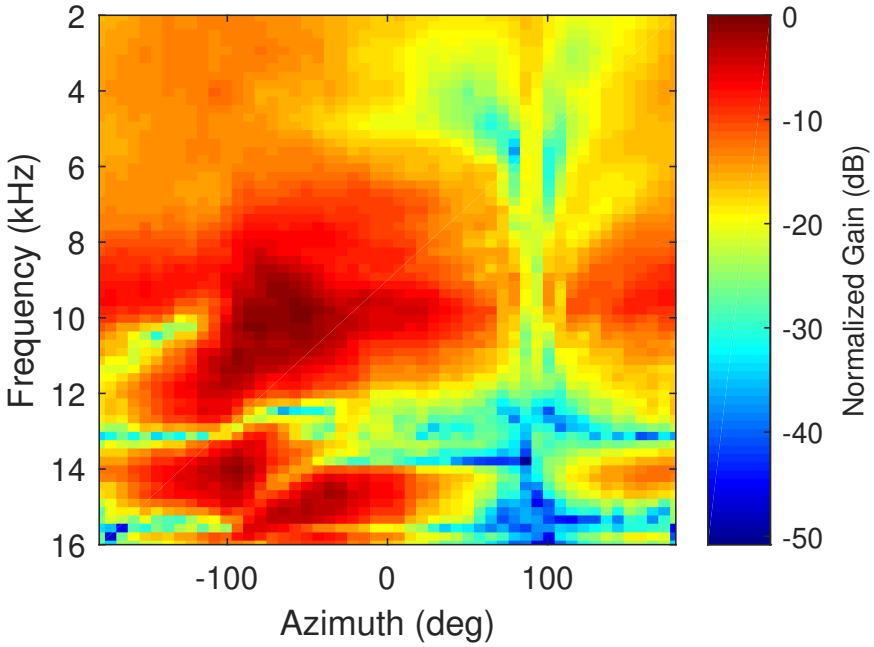


Figure 2.3: The HRTF of the left ear of a human, taken from the CIPIC HRTF Database (Subject 3) [27]

The spectral filtering introduced by the HRTF is very interesting, and is the basis of a monaural sound localization cue discussed in the following section. However, researchers have found that the brain also compares the spectral cues at each ear for determining the elevation of a sound source, and that the brain weights the comparison based on the perceived azimuth angle to the sound source [28]. This work highlights one particular approach to integrating the perception of timbre at the two ears to localize a sound source.

The term 'binaural timbre difference' is used by Woodworth [29, p. 525] and is the term preferred by the author to distinguish this cue from the ILD which is treated as a narrow bandwidth cue only. Many authors prefer to use a wide bandwidth version of the ILD which is equivalent to the BTD. For clarity in this work the ILD refers only to a narrow bandwidth cue, whereas the BTD refers only to a wide bandwidth cue.

These different binaural cues are processed simultaneously in the brain, and there is evidence to suggest that these different cues may be combined in higher-level processing to form an improved perception of a sound source [30, 31, 32]. However, the mechanisms and details of binaural cue integration is an area of active research in psychoacoustics and is beyond the scope of this work.

2.2.6 Other Cues for Localization

The previous section focused on binaural cues for sound localization which will be used heavily in the rest of this thesis. However, it is also useful to note the presence of other cues and cue combinations for sound localization that are less central to the thesis.

Early work into monaural localization by Angell and Fite found that pure tones were unlocalized using monaural hearing, but that complex sounds could be localized with accuracy [33, 34]. They speculated that the pinna was an important factor in the localization process, indicating that it altered the quality of the incoming sound in a way that enabled localization of wide bandwidth sound sources. Further research confirmed the role of the pinna in monaural localization of wide bandwidth sounds. Artificially removing a person's pinna inhibits their monaural localization ability, while the addition of an artificial pinna improves it [22, 23].

Research into the bandwidth necessary for monaural sound localization using spectral cues found that a spectral 'notch' is introduced by the shape of the pinna, and that the frequency at which this notch occurs corresponds to the elevation of the sound source. Because of this correspondence, the notch could be used by an observer to localize a sound source in elevation, provided that the sound source consisted of a wide bandwidth signal

(such as the white noise used by the researchers) [35].

The idea of creatures using a notch in the perceived spectrum of a sound source for localization is an important one, and researchers found that the ears of domestic cats introduce a notch into the received spectrum in the same way to the human ear. The cat's ear however, does not encode the notch at a frequency that is purely dependent on elevation. Rather, the frequency of the notch introduced by the ear is a function of both azimuth and elevation. This coupling appears counter-intuitive but considering the integration of both ears reveals a remarkable effect shown in Figure 2.4: the frequency at which the notches in the left and right occur in the spectra act as a coordinate system that uniquely corresponds to the azimuth and elevation of a sound source [36].

Another important cue is introduced by the use of head movements by listeners to localize different types of sound source. By moving the head a listener alters the ITD, ILD, and BTD of a sound, gaining information about the scene which usually leads to an increase in localization performance [37, 38, 39, 40, 41].

2.2.7 Neural Correlates

With the development of modern neuroscience, research into sound localization has increasingly turned to searching for neural correlates to the previously discussed cues. Researchers have found evidence for ILD and ITD cues in the owl's auditory pathway [42, 43]; the integration of binaural cues in owls [44, 45]; direction sensitive neurons in the cat [46]; binaural cue tuned neurons in the cat [47]; ITD sensitive neurons in rabbits [48]; and spatially sensitive neurons in rhesus monkeys [49]. For a review of sound localization in the brain see [50].

2.2.8 Localization by Blind Humans

In the context of monaural localization, research indicates that blind people have better monaural localization abilities than sighted individuals [51], which indicates the importance of spectral-based monaural cues for localization in blind individuals. In general

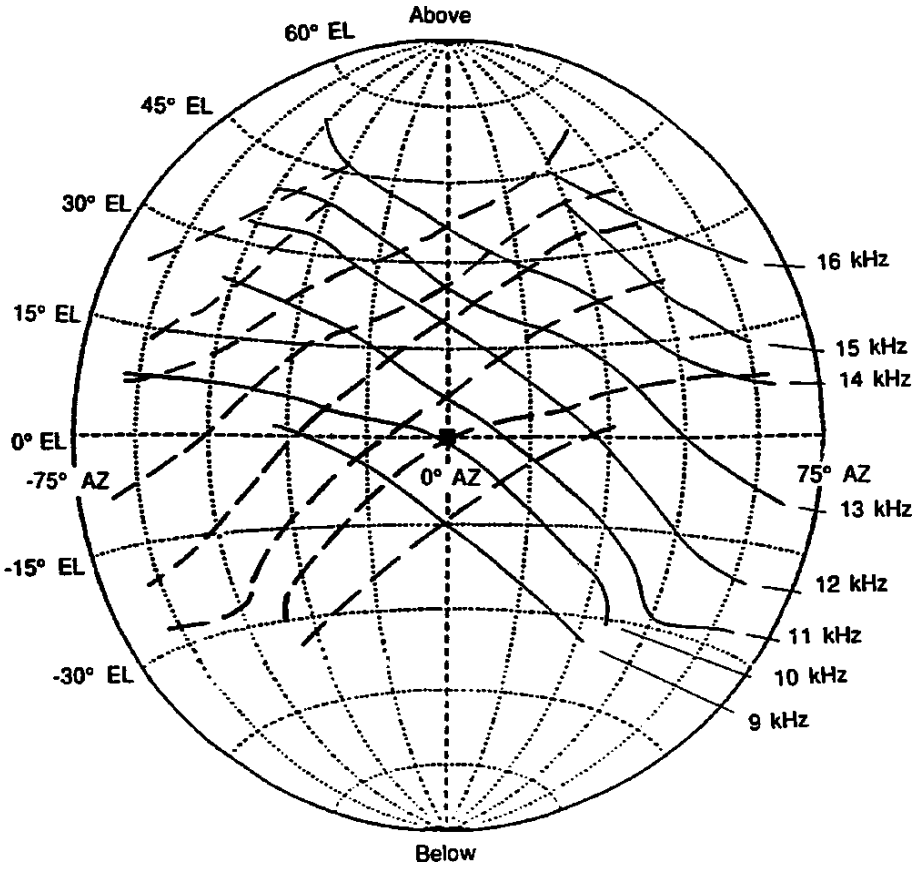


Figure 2.4: Iso-frequency contours for the notch introduced to the spectrum by the shape of the ears of a domestic cat. Solid lines correspond to the contours for the right ear, and dashed lines correspond to the contours of the left ear. The frequency at which the notches occur in the left and right spectra act as a coordinate system that uniquely corresponds to the azimuth and elevation of a sound source. Reprinted from Hearing Research, 58/2, Rice et al., Pinna-based spectral cues for sound localization in cat, 132-152, Copyright 1992, with permission from Elsevier. [36]

however, when all cues are available there is no performance difference in laboratory conditions between blind and sighted individuals [52]. Indeed, in some circumstances, performance of blind individuals in sound localization tasks is poorer on average than the average performance of sighted individuals [53].

2.3 Natural Echolocation: Active Localization

Echolocation is an ability developed by many creatures in the natural world which gives them a unique perception of their environment that is not reliant on the visual senses.

The passive reception of sound signals and localization of their origins using the auditory senses has already been discussed, but echolocation is an *active* technique that, as with radar and sonar, requires the active transmission of a sound signal into the environment. The transmitted signal propagates into the environment and reflects from objects therein, returning to the ears of the echolocating creature as a series of echoes that may be blurred, distorted, and overlapped. The echolocator must interpret these echoes in order to make sense of their environment, and this process occurs in the complex neural circuitry of the brain.

Understanding the techniques used by natural echolocators and the types of perception that they can form is important for taking a biologically-inspired approach to designing radar systems. The following sections briefly explore some of the key ideas of natural echolocation to complement and support the analysis presented in Chapter 4, and aims to identify to the reader the key ideas from which inspiration will be drawn in Chapter 5 and Chapter 6.

2.3.1 Bats

One of the most prolific echolocating animals is the bat; famed for their high-frequency echolocation calls, bats are able to navigate around caves, forests, and scrub. There are many different types of bats with different echolocation calls, but many bat species use short duration calls that use a wide band of frequencies with frequency modulation over the duration of the call, and often with multiple harmonics within the call. Many bats also possess the ability to control the parameters of their calls, such as duration, pitch, chirp rate, and pulse interval, depending on the situation in which they are placed. For example, the big brown bat increases its pulse repetition frequency, pitch, and chirp rate as it approaches its prey [54].

Another dynamic behaviour of bats is their ability to alter pulse intervals in order to mitigate the effects of overlapping pulse reflections [55]. By varying the pulse interval, real targets will still be perceived as having a constant position, but spurious targets

(resulting from the overlapping reflections) will instantaneously vary in position, allowing the bat to distinguish between real and spurious auditory images. The use of cognitive parameter control is an active area of interest within the radar community [56, 57].

Research into the neural mechanisms of bat echolocation have shown that bats possess delay-tuned neurons which selectively respond to echoes at particular time delays [58, 59, 60]. These neurons also selectively respond to different elements of the bat's echolocation calls, indicating that different features of the calls could be used for different purposes. One of the descriptions of the processing of bat echolocation calls that mimics the delay-tuned neurons is the spectrogram correlation and transformation (SCAT) model [61], which approaches the computational modelling of the echolocation process using functional blocks which relate to the physiology and neural structures of the bat.

Some bats also appear to have a remarkable tolerance to overlapping echoes from target and clutter sources, and are still able to successfully intercept prey in close proximity to vegetation [62]. One of the explanations for this exceptional ability to separate target and clutter is the wide bandwidth calls that the bats transmit. These calls have a beamwidth that narrows as the frequency of transmission increases. This change in beamwidth means that objects directly in front of the bat are strongly ensonified by all of the transmission spectrum whereas objects away from the transmission axis are strongly ensonified by the lower frequencies only. The higher frequency components are still present in the transmission off-axis, but the losses introduced by signal propagation cause these frequency components to fall below the bat's hearing threshold. It is possible that the perceived absence of high-frequency components causes the bat's perception of an echo to defocus, indicating echoes away from the boresight direction [63].

In addition to being able to avoid collisions with clutter in their environments, bats also posses the ability to identify sources of food, often in the form of flowers for nectar-feeding bats. It has been suggested that this ability is a result of the spectral composition of the echoes that those bats receive as they approach the flower [64, 65, 66]. Research into the use of wideband spectral content in echoes for classification of targets found that

performance of a classifier of both flowers and classical radar targets improved significantly with multiple look angles [67, 68, 69].

The use of wide bandwidth signals by bats is also interesting for radar because of the reliance on broadband signals for the pinna-related cues discussed previously. Experiments have been conducted to assess the impact of impairing bats' pinnae. These experiments found that the localization performance of the bats suffered when their pinnae were occluded, and that they altered their strategy to compensate for the pinna impairment [70]. The use of wideband signals and pinna-type cues is of significant importance for the work presented in Chapter 5.

Some bats also appear to adopt strategies for target interception that are consistent with an acoustic flow model of behaviour [71]. This model has been used for research into echoic flow for radar, sonar, and mobile robotics [72, 73, 74, 75, 76], and provides the inspiration for the work presented in Chapter 6.

2.3.2 Humans

In addition to commonly known and researched echolocating species, some humans have developed the ability to echolocate. In general, there are three categories of echolocating humans: early-blind individuals who have been expert echolocators for much of their lives; late-blind individuals who have learnt echolocation after losing their sight; and sighted individuals who have been taught some elements of echolocation, often while blindfolded.

Of the blind echolocators, many regularly use tongue clicks and other self-generated noises in order to perceive and navigate their environments. These sounds for human echolocation vary, with the most common types being the tongue click and a noise-like hissing sound. A detailed examination of novel research into the properties of human echolocation tongue clicks is presented in Chapter 4, but an overview of human echolocation is given here. Detailed reviews into human echolocation are available and cover a variety of themes including: early research and acoustics [77]; contemporary research and echolocation cues [78]; and contemporary neuroscientific research [79].

Early investigation into the ranging abilities of blind human echolocators showed that echolocators were able to discriminate between closely-spaced targets with a reliability much better than sighted subjects, who performed only slightly better than chance [80]. More recent experiments have shown that human echolocators are able to separate objects spaced as close as $2-3.5^o$ [81]. Experimental results also indicate that human echolocators perceive size consistently, regardless of the absolute range to an object [82].

There is evidence that echolocation can be taught to sighted individuals [83], indicating that the ability to echolocate is not predicated on blindness. However, some early-blind echolocators have been shown to exhibit signals in the brain that correspond to the shape of the object they are perceiving through echolocation. These neural signals were not observed for late-blind echolocators, suggesting that the brains of early-blind echolocators are performing additional processing, not present in other echolocators[84].

Researchers into the localization abilities of blind and sighted humans using echolocation carefully emphasize the importance of not over-generalizing the conclusions of experiments conducted under laboratory or other restricted conditions [85, p.63].

Human echolocators appear to use head motion in conjunction with tongue clicks in order to increase the quality of their perception of their environments, and in some cases to enable object identification [86, 87, 88]. It is possible that this is comparable to the multiple look angles used by bats to correctly identify flowers.

While there has been a significant volume of research into the capabilities of human echolocators, the body of work focused on the properties of human echolocator tongue clicks is very restricted. Chapter 4 presents novel research focused on understanding human echolocator tongue-click properties with a view to increase the overall understanding of human echolocation.

2.4 Biologically-Inspired Radar and Sonar Approaches

Throughout the following chapters of this thesis, novel biologically-inspired approaches to radar systems are presented. This section sets out some of the pre-existing approaches to biologically-inspired radar and sonar that helped to guide the development of this thesis.

2.4.1 Spectrum-Based Localization with Sonar

A binaural spectrum-based approach to localization relies on an understanding of the BTD introduced in Section 2.2.1. Researchers working with acoustic transducers in the sonar domain showed that they could achieve angular localization of a target by considering the BTD between two angularly offset receiving transducers [89].

The spectrogram of a wideband acoustic signal at the output of a receiver is modulated by the beampattern of the receiver. As the beampattern changes significantly as a function of frequency, the received signal is modulated in a manner unique to that angle. If a monaural spectrum is considered, the received signal is strongly dependent on the range to the target and the reflective properties of the target itself. However, researchers found that considering a binaural cue caused the target-dependent effects to cancel out, leaving a spectral cue that was theoretically only a function of the receiver beampatterns.

By comparing the ratio of received signal spectrograms to the system's prior knowledge about the beampatterns of the receivers, researchers were able to localize a single target in angle over an angular range of operation of $\pm 30^o$ from the transducer pointing direction and at a range of 1m. The approach presented in Chapter 5 is structurally similar to this acoustic approach. However, Chapter 5 presents research which focuses on a radar system built to operate with electromagnetic waves, and contends with the fundamental physical challenges and opportunities that this presents.

Other researchers have proposed a related family of techniques which rely on the information encoded in the audio signal by the receiving transducer. This approach estimates the angle to the target independently for each receiver for a set of frequency extents that

comprise the total bandwidth, and extracts the angle to the target based on the most frequently estimated position common to both receivers and across all frequency extents [90]. This method assumes that there is no significant spectral modulation introduced by the target.

2.4.2 Spectrogram Transformation and Correlation

SCAT is a computational model of the functional processes that govern perception of echolocation calls in certain bat species [61]. It treats the steps used for echolocation in three functional blocks: cochlear processing, spectrogram correlation, and spectrogram transformation.

The cochlear processing step consists of a bank of band pass filters tuned to centre frequencies corresponding to the frequency range of interest. The filters feed into half-wave rectifiers, and then into a bank of low pass filters with a cut-off frequency lower than the lowest frequency of interest. This cochlear bank emulates the tonotopic organization of hair cells and neurons in the auditory cortex, which respond selectively to particular frequency bands [91, p.213].

The spectrogram correlation block operates in the time-domain and performs a cross correlation between the transmitted and reflected signals (after the processing of the cochlear block). This correlation process mimics the function of delay-tuned neurons in bat brains [58, 59, 60], which can be modelled as tapped delay lines in the brain [92]. The cross correlation process results in sharp peaks corresponding to the echo delay (and hence the range to the target) for well-separated echoes. However, when there is overlap between echoes, the response is less sharp and can have spurious peaks in the region of the targets.

The spectrogram transformation block transforms regions of the signal of interest into the time-domain. The region of interest could be an area where there is a response from the spectrogram correlation block that is not clear (as is the case for closely-spaced targets). This process is equivalent to performing an inverse fast Fourier transform (IFFT) over

the time delays of interest. By transforming into the time domain, closely-spaced targets may be resolved, with a performance equivalent to the output of a matched filter.

More recent work on SCAT [93] has developed a baseband version that can be used with the very high frequencies involved in radar systems (in contrast to the lower frequencies in sonar systems). Development of the radar-based technique found that it is possible to improve the resolution capabilities of the spectrogram transformation block in the case of two closely-spaced targets [94]. The baseband SCAT model has led to the development of an equivalent receiver based on the matched filter [95].

2.4.3 Adaptive Waveforms

As introduced in Section 2.3, some bats alter the waveforms that they transmit as they approach a target [54]. In a similar way, radar systems can be designed that adapt their waveforms as a target is approached [96]. Balleri et al. designed a radar system that adapted its waveforms on a pulse-by-pulse basis to optimize the system's ability to track and intercept a target, and minimize the energy cost to the system.

The approach combined a technique for intercepting a target in the presence of noise [97] and a technique for optimising the transmitted waveform to maximise tracking performance [98]. By using these strategies on a pulse-by-pulse basis in simulation, the interceptor was able to intercept a moving target in the required time frame. Over the course of the flight the interceptor decreases its pulse duration, which sacrifices Doppler resolution for accuracy in the time delay measurement which helps to reduce the covariance of target in range and Doppler [96]. This strategy has some correspondence with the terminal buzz of certain bat species [54], which are also seen to reduce their call pulse duration.

2.4.4 Echoic Flow

Echoic flow is a flow theory that is adapted for use in Radar and Sonar systems. Flow theory was first introduced for optical flow [99, 100], and was later developed into a

general model of flow theories [101]. The central parameter of flow theories is tau, τ, which represents the time to close a gap in an arbitrary coordinate, X, as given in Equation (2.3). Further, the taus for different parameters may be coupled to force gaps in each parameter to close simultaneously. An example of coupled taus is given in Equation (2.4), where m is a coupling constant that relates the kinematics of the two parameters.

$$\tau_X = \frac{X}{\dot{X}} \tag{2.3}$$

$$\tau_Y = m\tau_X \tag{2.4}$$

While optical flow is of interest in the field of computer vision, the focus for radar research was into acoustic flow, which bats have been shown to use when navigating their environments [71, 102]. The bats in the experiments were found to close the angular gap and the range gap simultaneously when approaching a perch location, consistent with acoustic flow models. Echoic flow was introduced to radar in the context of braking strategies when approaching a target [74], and has since been used to perform target interception [73] and robotic guidance [76].

In the context of mobile robotics, echoic flow was implemented using a robot equipped with ultrasonic transducers. The robot always attempted to equalize the tau measured between two lateral sensors. This simple control rule led to sophisticated behaviour which allowed the robot to navigate successfully through a corridor and to avoid obstacles placed before it [76].

Echoic flow is powerful in its simplicity and allows the emergence of relatively complex behaviours from apparently simple control rules, a hallmark of cognitive systems. Echoic flow is limited when it comes to objects that remain a fixed distance away from the platform, such as a target matching the robot's speed, or walls running parallel to the robot's motion vector. Chapter 6 explores some of the ideas of echoic flow, and presents a related set of control rules that enables a cognitive radar platform to avoid obstacles

and reach a goal position.

2.4.5 Cognitive Target Tracking

One of the more recent developments in cognitive radar is the Cognitive Radar Experimental Workspace (CREW) based at Ohio State University. The CREW is a powerful system with 4 transmitters and 4 receivers that can use an arbitrary waveform on a pulse-to-pulse basis and operate coherently across the 92-96 GHz frequency band [57].

Researchers working on the CREW used a Bayesian approach to implementing a cognitive radar architecture. By describing the current state of observations and the costs of using various parts of the system, the processor decides which parameters and combination of sensors to use for the task at hand. In a simple target tracking case, the system developed its own strategy for maintaining a track of the target [56]. In other experiments, the radar was able to alter its pulse repetition frequency (PRF) and number of integrated pulses to reduce the range and Doppler errors while preventing the target from aliasing [103]. The use of cognitive processes to optimize performance metrics has been shown to outperform more conventional optimization techniques [104].

This approach favours a dynamic response of the system to its environment over a static set of parameters chosen at the radar design stage. This approach has the advantage of the radar being able to respond to a variety of situations dynamically and without the need for human operation, with the radar able to strategically allocate its resources to achieve a goal. A disadvantage of this approach is that it requires the system hardware to be highly adaptable compared to traditional systems. This disadvantage is lessening with the advance of digital approaches and technologies such as field programmable gate arrays (FPGAs), that make highly reconfigurable radar systems more widely available.

2.5 Conclusion

This chapter has explored some of the approaches to localization in nature and in electrical engineering. Active and passive mechanisms for sound localization in nature have been examined and the use of wideband binaural cues, such as the BTD, have been identified as important and potentially useful for a radar system. An overview of research into echolocation in nature has been presented which indicated significant connections to existing radar approaches, highlighted the use of wide bandwidth waveforms by bats, and identified that a detailed description of human echolocator tongue click properties is absent from the literature.

A review of biologically-inspired radar and sonar techniques has been presented, and this included multiple techniques in the acoustic domain which consider the spectral content of echo signals. The review also identified that equivalent techniques in the radar domain are absent. Existing research into cognitive radar was briefly presented, which shows the potential for cognitive systems to out-perform conventional systems, but that systems that explicitly implement a memory function are absent.

These related observations provide the foundations of, and biological-inspiration for, the research presented in the later chapters of this thesis, with Chapter 4 exploring human echolocator tongue click properties, Chapter 5 investigating a radar version of some acoustic techniques related to the BTD, and Chapter 6 developing a cognitive radar system that explicitly makes use of memory to improve performance.

Chapter 3

Foundational Ideas and Approaches

L'étude approfondie de la nature est la source la plus féconde de découvertes mathématiques.

Profound study of nature is the most fertile source of mathematical discoveries.

— Joseph Fourier

Théorie De La Chaleur, 1822

Table 3.1: Reference for mathematical symbols.

Symbol	Definition	Units
f	frequency	Hertz
P_{Rx}	signal power received at the receiving antenna	Watts
P_{Tx}	signal power transmitted to the transmitting antenna	Watts
G_{Tx}	gain of the transmitting antenna	-
G_{Rx}	gain of the receiving antenna	-
θ'	angle to target	rads
c	speed of light in a vacuum	m s^{-1}
σ	radar cross section (RCS) of the target	m^2
L	losses	-

3.1 Introduction

This chapter aims to present the theoretical foundations of this thesis, which are a variety of distinct techniques and mathematical tools. All of these ideas are relied on in subsequent chapters and an understanding of them is useful for a complete understanding of this thesis. This chapter does not attempt to cover the foundational principles of radar in general, for which many excellent reference books exist [105, 106, 107, 108].

3.2 The Radar Equation

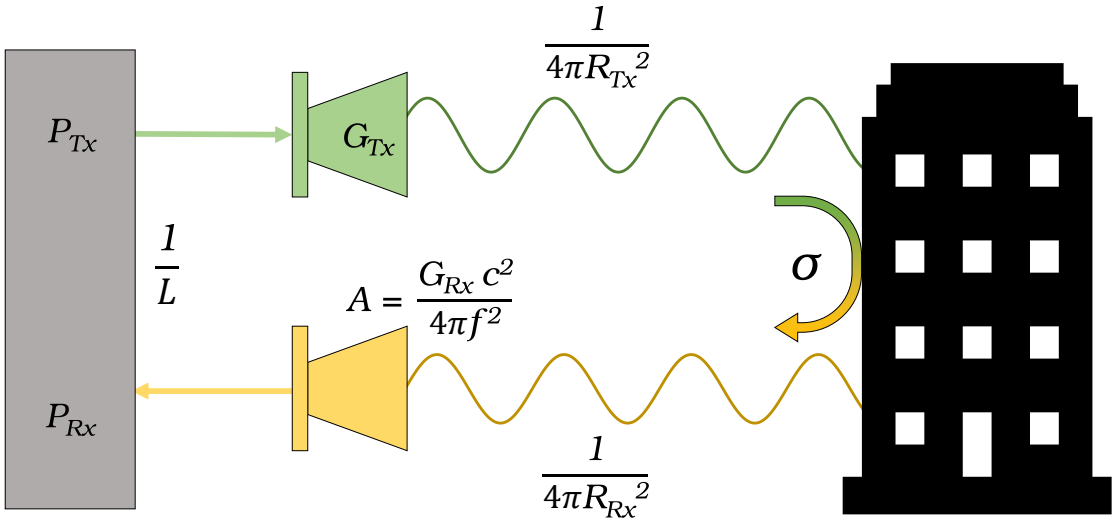


Figure 3.1: Schematic to show the physical significance of terms in the radar equation.

The radar equation is one of the cornerstone mathematical descriptions used when approaching the design of a radar system, and is introduced in many radar textbooks [109]. The research presented in Chapter 5 relies on a modified form of the radar equation, and so the conventional free space form of the equation is presented here and developed into the modified version for use later in this thesis.

The radar equation describes the power received by a radar system as a function of system and target parameters and is usually written as some variant of Equation (3.1) for a bistatic radar configuration [108, p.631]. Figure 3.1 shows the different factors present in the radar equation and the point in the signal journey that they occur.

$$P_{Rx} = \frac{P_{Tx} G_{Tx}(\theta) G_{Rx}(\theta) c^2 \sigma(\theta)}{(4\pi)^3 R_{Tx}^2 R_{Rx}^2 f^2 L} \tag{3.1}$$

The signal with transmit power, P_{Tx}, is passed to an antenna. The antenna amplifies the signal by a factor corresponding to the antenna gain, G_{Tx}, in the direction of the target. The signal propagates through the air from the antenna and incurs attenuation due to the inverse-square law over the distance to the target, R_{Tx}. The signal is reflected by the target, which has an effective area described by the radar cross-section (RCS), σ. The signal travels to the receive antenna and incurs inverse-square law attenuation over the distance to the receiver, R_{Rx}. The effective area of the receiving antenna, A, determines how much of the signal power is received and is a function of the receive antenna gain, G_{Rx}, and the frequency of the signal, f. With the addition of a loss term, L, which arises due to propagation losses (such as scattering and absorption) and imperfections in hardware, the signal's journey is complete and is received with a power, P_{Rx}.

Normally, radar systems operate with a narrow bandwidth relative to their centre frequency, and in this form of the radar equation, the antenna gains, G_{Tx} and G_{Rx}, and RCS, σ, are considered to be functions of angle only. However, in Chapter 5 a wideband signal will be used that spans a 2-16 GHz band of frequencies. In this instance, the frequency-dependence of terms in the radar equation become vital for understanding how to process the return signals. This section introduces a modified version of the radar

equation, which is suitable for working with arbitrarily wide bandwidth signals such as the ones used in Chapter 5.

3.2.1 Frequency-dependent Radar Cross-Section

The RCS of a target, denoted by σ, describes the effective cross-sectional area of the target for reflection of an incident electromagnetic signal. The RCS is related to, but is not purely a function of, the physical size of an object. Other factors that interact to affect a target's RCS include its material and structure, and all of these factors can introduce complex frequency dependence to the RCS.

For example, a target that is an ideal metal sphere has an RCS that varies with frequency as shown in Figure 3.2. Here, the low-frequency region exhibits Rayleigh scattering, the high-frequency region exhibits optical scattering, and the intermediate region exhibits Mie scattering [110]. In this situation the RCS of the sphere is a function of frequency.

Due to the relative orientation of the radar and the target, the angle of incidence also influences the RCS. To capture frequency-dependence is straightforward; it is possible to rewrite the RCS, $\sigma(\theta)$ from Equation (3.1), as a function of frequency, $\sigma(\theta, f)$.

3.2.2 Frequency-dependent Antenna Gain

As with the the reflectivity of a target, it is often convenient to approximate the behaviour of an antenna as being frequency-independent. Equation (3.2) gives the antenna gain for an ideal antenna, where A is the effective antenna area, which shows how the gain of the antenna is frequency-dependent [111, p.104]. When using a narrowband signal however, the change in antenna pattern across the frequency band is often negligible compared to other experimental factors, and so the assumption of frequency-independence is both useful and valid. However, because of physical factors of antenna design, antenna patterns

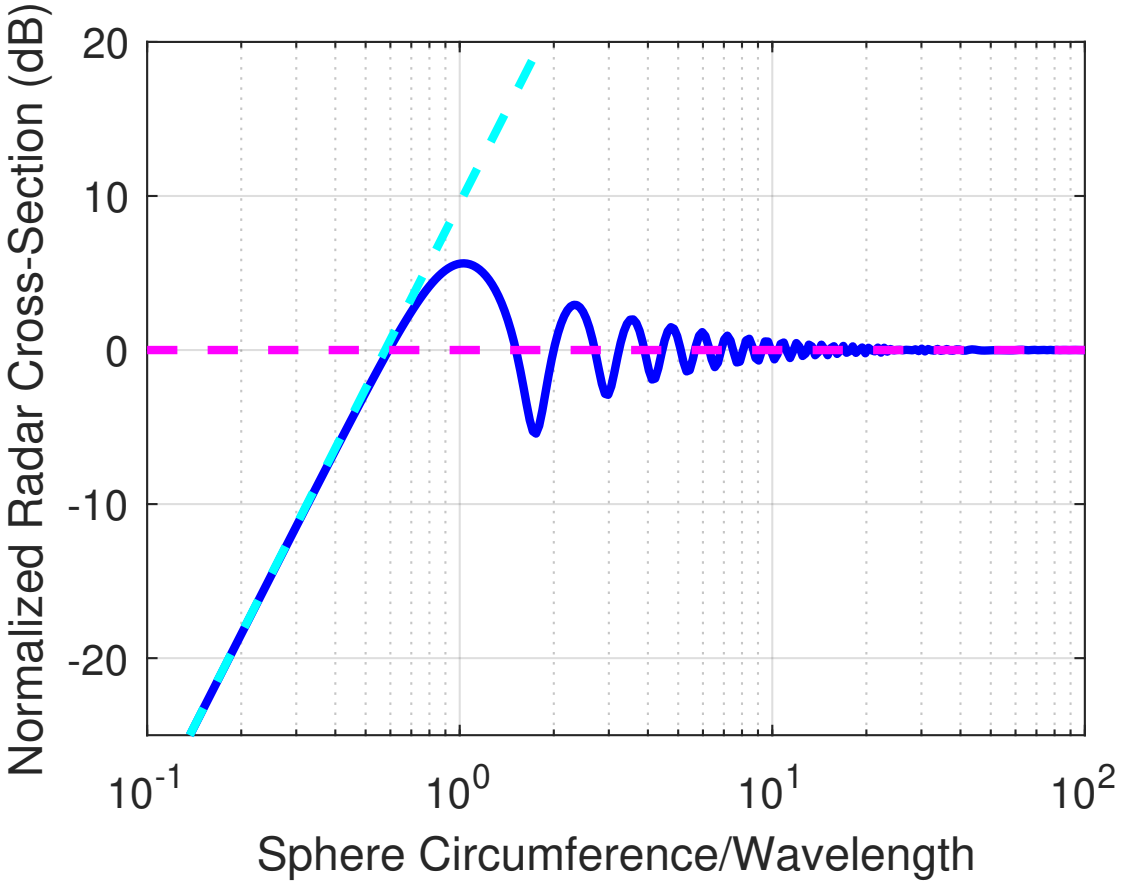


Figure 3.2: The normalized RCS (blue) of an ideal metallic sphere as a function of the ratio of circumference and wavelength; as the frequency increases, the ratio increases. The dashed lines represent the low-frequency Rayleigh scattering limit (cyan) and the high-frequency optical limit (magenta) [110].

must be treated as frequency-dependent when using a large bandwidth.

$$G(f) = \frac{4\pi A f^2}{c^2} \tag{3.2}$$

Antennas have a fixed geometry and the different structures within the antenna result in different propagation characteristics for different frequency electro-magnetic (EM) waves. As a result, different frequency EM waves are focused by different amounts and result in different beam patterns. In principle, the beam pattern of an antenna can be determined analytically by solving Maxwell's equations for the antenna design. However, for real antennas this analysis is complex and time-consuming, and it is often easier to measure an antenna's beam pattern. Figure 3.3 shows antenna patterns for the off-the-

shelf wideband horn antenna that is used in Chapter 5 [112], which shows how the beam pattern varies as a function of frequency.

To denote this frequency variation, the antenna gain terms in the radar equation must be written as a function of frequency as well as angle, $G(\theta, f)$.

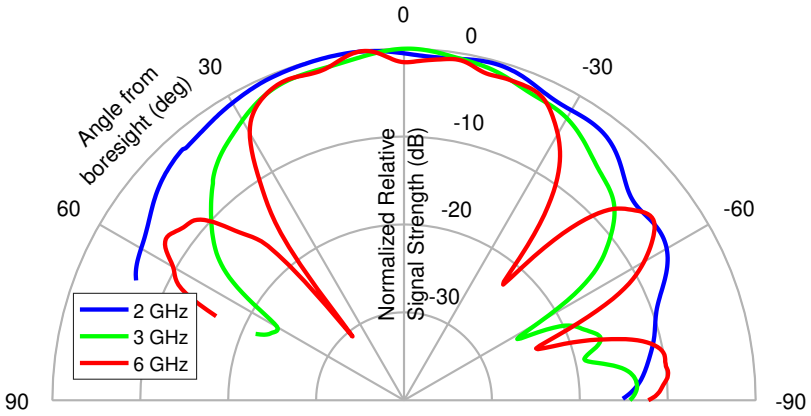


Figure 3.3: Antenna patterns of the Q-par Angus (WBH1-18)[112] wideband horn antennas, showing the different antenna patterns at different frequencies over a wide band.

3.2.3 Frequency-dependent Radar Equation

Combining the frequency-dependent terms from the previous section results in the frequency-dependent radar equation given in Equation (3.3).

$$P_{Rx}(\theta, f) = \frac{P_{Tx} G_{Tx}(\theta, f) G_{Rx}(\theta, f) c^2 \sigma(\theta, f)}{(4\pi)^3 R_{Tx}^2 R_{Rx}^2 f^2 L(f)} \tag{3.3}$$

3.2.4 The Wideband Phase-Only Radar Equation

The radar equation discussed in the previous section is primarily focused on understanding the power of a received signal. However, Chapter 5 will also consider the phases of a

received signal independently, and so it is useful to formulate a version of the radar equation to use when considering the phase of a signal in lieu of the power.

There is much discussion in psychoacoustics about the extent to which humans and animals use phase differences to locate sound sources [16, 43, 113, 114, 21]. However, for radar systems phase is typically present in the signal processing and may therefore be used as the basis for a radar-only cue. A phase-only form of the radar equation is shown in Equation (3.4), where ϕ_{Rx} and ϕ_{Tx} are the phases of the received signal and transmitted signal respectively, $\angle G_{Tx}$ and $\angle G_{Rx}$ are the phase shifts introduced by the transmitting and receiving antennas respectively, and $\angle\sigma$ is the phase shift introduced by the target. The final term describes the phase shift introduced by the bistatic path length from transmitter to receiver. This equation represents the ideal scattering scenario and ignores the effect of noise introduced by imperfect scattering and atmospheric conditions.

$$\phi_{Rx} = \phi_{Tx} + \angle G_{Tx}(\theta, f) + \angle G_{Rx}(\theta \pm \theta_0, f) + \angle\sigma(\theta, f) + \frac{2\pi\left(R_{Tx} + R_{Rx}\right) f}{c} \tag{3.4}$$

Conventionally, antenna gain is written as a real quantity, however over the wide frequency band used by this technique the phase centre of each antenna can vary. This phase centre variation occurs due to the EM reflections within the very near-field of the antenna structure which modulate the received signal with a frequency- and angle-dependent phase-shift. Phase centre variations are often treated in global positioning system (GPS) research where they provide a deviation from the ideal behaviour which must be corrected [115, 116, 117, 118, 119].

However, for the application in Chapter 5, we are interested in the angle-dependent coding of space by the antennas and phase centre variations perform this function. The variations can be incorporated into the antenna gain terms by considering them as complex quantities (modifying both the amplitude and phase of the transmitted or received signal). The angles of the complex gains represent the angle-dependent phase-shift that

the antennas introduce.

3.3 Conventional Monopulse

Monopulse is a technique for localizing a target in both range and angle using a single pulse and is the subject of many books [120, 121]. Accordingly, only a brief overview of the relevant material is presented here. It is particularly useful for tracking radars due to the high localization accuracy that it can achieve. Monopulse radars require multiple antennas that can be used coherently for transmission, but are capable of receiving separately. This requirement can be fulfilled by an array radar that is split into two halves, or by a pair of feedhorn antennas connected to a monopulse duplexer.

The two antennas transmit the same pulse coherently, forming the sum channel which corresponds to the blue antenna pattern in Figure 3.4a and Figure 3.4b. Provided that this pulse illuminates the target successfully, the reflected energy is received independently by the two antennas. The cyan and magenta beam patterns in Figure 3.4a represent the receiver patterns of these two antennas.

Comparing the energy received at the two antennas results in the response shown in Figure 3.4c, which is the difference channel of the monopulse radar. The difference channel is uniquely valued when the target is within $\pm 20^o$ of the boresight direction (in this example), and can be used to drive the motion of the antennas in a closed-loop process which has a stable equilibrium when the angle to the target from the antenna boresight is 0^o.

The example given here was for localization in a single angular dimension, such as in azimuth. However, monopulse can localize in both azimuth and elevation by adding additional antennas and comparing the received signals in a similar manner.

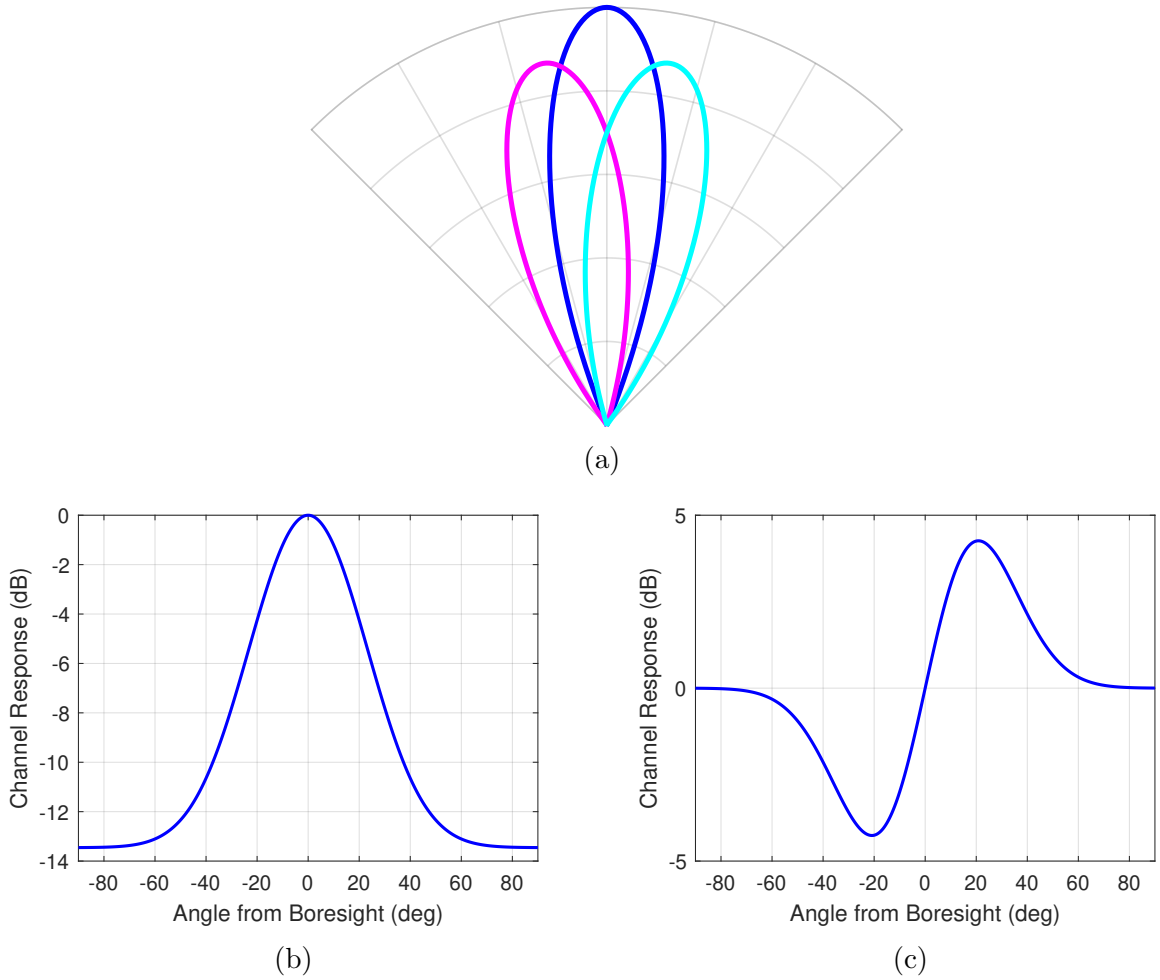


Figure 3.4: Monopulse radar operation. (a) Polar representation of two beams (magenta and cyan) and sum beam (blue) used for monopulse in one angular dimension. (b) Sum channel response. (c) Difference channel response.

3.4 Vector Network Analyzer Operation

Vector network analyzers (VNAs) are designed to measure the characteristics of electrical networks, and are also capable of generating very wide bandwidth signals. A VNA will be used to generate a very wide bandwidth signal in Chapter 5, and so a brief description of VNA operation is given here.

VNAs often measure the scattering parameters (or S-parameters) between multiple terminals of an electrical network. The scattering parameters describe the ratio of outgoing signals and incoming signals, such that, $S_{ij} = \frac{b_i}{a_j}$. Where b_i and a_j represent the outgoing complex signal from VNA port i, and incoming signal to VNA port j. In general,

the scattering parameters can be written as in Equation (3.5) [122, pp.3.27-3.28].

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \tag{3.5}$$

In the context of the bistatic radar system shown in Figure 3.5 a VNA applies an voltage of a particular magnitude and frequency to an output port that is connected to an antenna, the changing voltage induces currents in the antenna. These currents cause a self-sustaining EM wave that can propagate into the environment. The reflected wave returns to the antennas and induces a signal voltage at the VNA input ports, forming a complete electrical network. To determine the scattering parameters, the received signals are mixed with the transmitted signal, in a way that is equivalent to performing a matched filter in a conventional radar system. For example, S_{21}, is the scattering parameter that describes the ratio of received signal at the output of antenna 2, b_2, to the transmitted signal at the input of antenna 1, a_1.

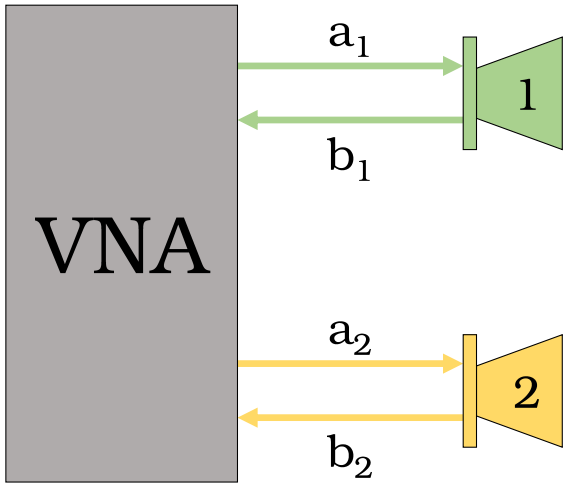


Figure 3.5: Schematic to show the configuration of a VNA for radar measurements.

By stepping through a wide range of frequencies and making separate measurements of the scattering parameters at each frequency, the data from the VNA is equivalent to a frequency-domain measurement of a radar chirp made at baseband (i.e. without

downconverting and sampling at the intermediate frequency). Using a flexible frequency oscillator means that the VNA is able to sample over very large bandwidths provided that the oscillator is configurable to that frequency and that there is enough time to make the sequence of measurements. Figure 3.6 presents a block diagram of the VNA signal processing.

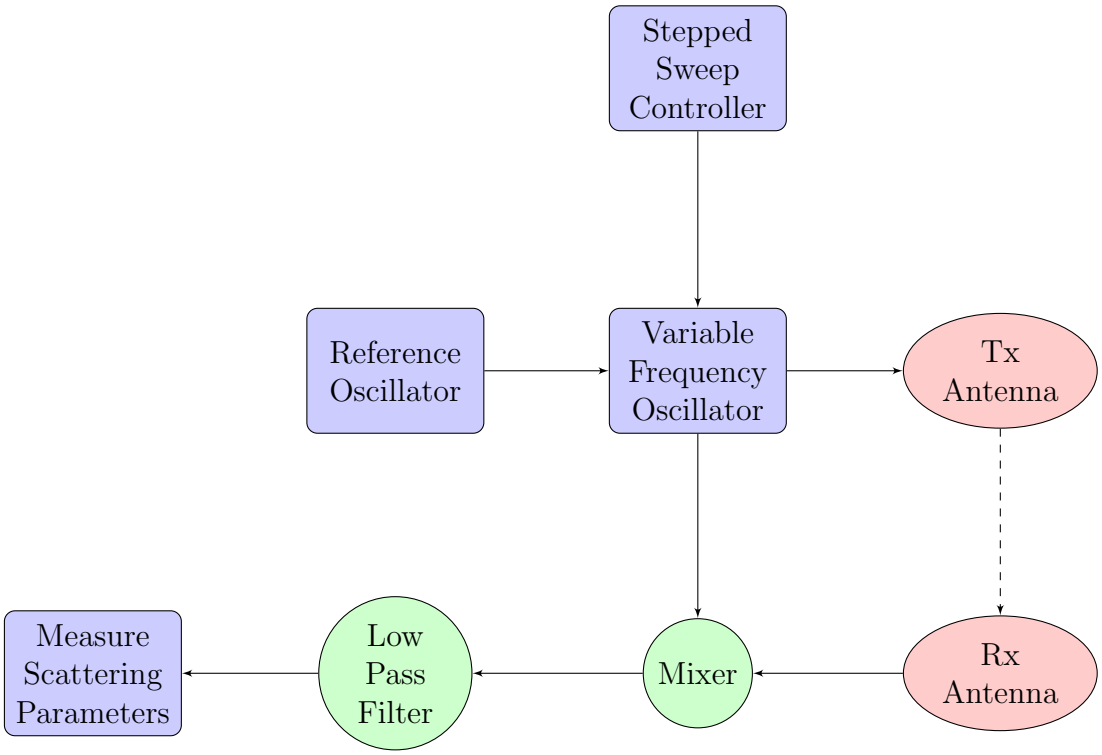


Figure 3.6: Flowchart to show the processing chain of the VNA.

3.5 Radar and Array Processing

This section presents processing that will be relied upon in Chapter 6, it presents the signal processing required to form a range profile from a linear frequency modulated (LFM) signal, and then extends this signal processing for the array antenna case, allowing a range-angle map to be formed. A range-angle map is a map of radar echo intensity as a function of range and angle.

3.5.1 The Range Profile

This section details the signal processing of a LFM signal using stretch processing to form a range profile.

The transmit signal given in Equation (3.6) is an LFM upchirp with centre frequency, f_c, chirp rate, k, and duration, τ. A target at a (two-way) range, r_d, which corresponds to a time delay of, t_d, introduces a delay and phase shift into the received signal as written in Equation (3.7). In these equations, $\Pi(t)$ is the rectangular function defined in Equation (3.8).

$$x\left(t\right) = e^{2\pi i\left(f_c t + \frac{1}{2}kt^2 + \phi\right)}\Pi\left(\frac{t}{\tau}\right) \tag{3.6}$$

$$y\left(t\right) = Ae^{2\pi i\left(f_c(t-t_d) + \frac{1}{2}k(t-t_d)^2 + \phi\right)}\Pi\left(\frac{t-t_d}{\tau}\right) \tag{3.7}$$

$$\Pi(t) = \begin{cases} 0, & \text{if } |t| > 0.5 \\ 0.5, & \text{if } |t| = 0.5 \\ 1, & \text{if } |t| < 0.5 \end{cases} \tag{3.8}$$

Stretch processing mixes the received signal with a reference chirp, so that the frequency difference between the two chirped signals corresponds to the target range. The reference signal is given in Equation (3.9) and the baseband signal at the output of the stretch processing is given in Equation (3.10). The rectangular function is omitted from the reference signal for simplicity, but it should be understood that there is a practical restriction on the bandwith of the reference.

$$h\left(t\right) = e^{2\pi i\left(f_c t + \frac{1}{2}kt^2 + \phi\right)} \tag{3.9}$$

$$v\left(t\right) = h\left(t\right) \times y^{\star}\left(t\right) = Ae^{2\pi i\left(f_c t_d - \frac{1}{2}kt_d^2 + kt_d t\right)}\Pi\left(\frac{t-t_d}{\tau}\right) \tag{3.10}$$

As the baseband signal has a fixed frequency that is proportional to the target range, applying a Fourier transform can be used to extract the target range. The output of the Fourier transform is given in Equation (3.11) and is a function in the frequency domain. By taking the magnitude of the frequency-domain function as shown in Equation (3.12), a sinc function (defined as $\operatorname{sinc} x \equiv \sin x/x$) centred at kt_d is extracted, where β is the bandwidth of the chirp, $\beta = k\tau$. This function is the range profile and can be converted to the range domain using the definitions, $f = kr/c$ and$t_d = r_d/c$.

$$\begin{aligned} w(f) = \mathcal{F}[v(t)] &= Ae^{2\pi i\left(f_c t_d - \frac{1}{2}kt_d^2\right)} \int_{-\infty}^{\infty} e^{-2\pi i\left(f - \frac{1}{2}kt_d\right)t} \Pi\left(\frac{t - t_d}{\tau}\right) \mathrm{d}t \\ &= A\tau e^{2\pi i\left(f_c t_d + kt_d^2 - t_d f\right)} \operatorname{sinc}(\pi\tau(f - kt_d)) \end{aligned} \tag{3.11}$$

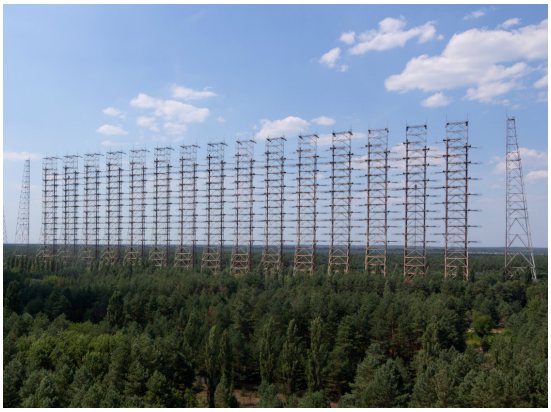
$$|w(f)| = A\tau \operatorname{sinc}(\pi\tau(f - kt_d)) = A\tau \operatorname{sinc}\left(\frac{\pi\beta}{c}(r - r_d)\right) \tag{3.12}$$

3.5.2 Array Processing

An array radar is a radar system that is comprised of multiple antenna elements, often arranged linearly, and in which the signals received at the elements are processed coherently. Figure 3.7 shows two example antenna arrays for reference. The research presented in Chapter 6 uses an array radar (albeit a *much* smaller one) and so the theory underpinning its operation is presented here.

A linear array configuration enables a variety of techniques that are not possible when using a single antenna. For example, by comparing the signals received at the antenna elements it is possible to derive a 2-dimensional range-angle map of the area in front of the radar. The formation of a range-angle map for a generic array is demonstrated by deriving the point spread function (PSF). This process builds on the derivation of the range-profile given in the previous section, and uses multiple range profiles (one for each element of the array) to derive the PSF.

(a)



(b)

Figure 3.7: Array Antennas. (a) Australia Telescope Compact Array, Image by CSIRO. (b) Duga Array Radar, Image by Ingmar Runge.

For a 1-dimensional array, the signal arriving at each position along the array baseline is slightly different due to the small difference in range to the target. If the incident waves are treated as plane waves from a distant target as shown in Figure 3.8, the time to the target can be written as a function of array element position, t_a, as, $t_d + t_a \sin\theta_d$, where θ_d is the angle to the target. Using this description of time to the target and Equation (3.11) gives Equation (3.13) which describes how the complex range profiles vary as a function of array element position, where L is the length of the array. Applying the far-field condition, $t_a \sin\theta_d << t_d$ to the argument of the sinc function causes the t_a dependence to vanish, and neglecting the $\frac{1}{2}k\left(t_d + t_a \sin\theta_d\right)^2$ term as vanishingly small, results in the approximation given in Equation (3.14).

$$\begin{aligned} w\left(f, t_a\right) = & A\tau e^{2\pi i\left(\phi_2 - \phi_1 + (f_c - f)(t_d + t_a \sin\theta_d) + \frac{1}{2}k(t_d + t_a \sin\theta_d)^2\right)} \\ & \operatorname{sinc}\left(\pi\tau\left(f - k\left(t_d + t_a \sin\theta_d\right)\right)\right) \Pi\left(\frac{t_a}{L}\right) \end{aligned} \tag{3.13}$$

$$w\left(f, t_a\right) = A\tau e^{2\pi i\left(f_c t_d + \frac{1}{2}kt_d^2 - f\right)} \operatorname{sinc}\left(\pi\tau\left(f - kt_d\right)\right) e^{2\pi i f_c t_a \sin\theta_d} \Pi\left(\frac{t_a}{L}\right) \tag{3.14}$$

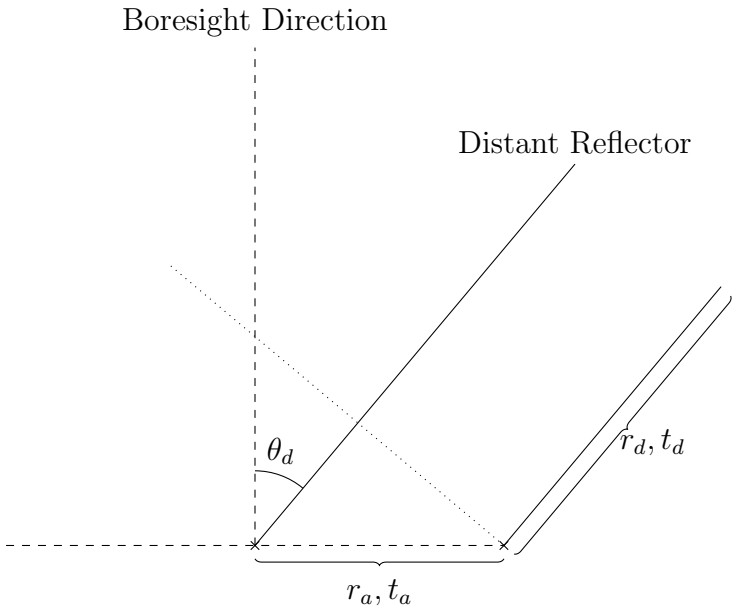


Figure 3.8: Geometry for array processing.

Equation (3.14) represents a signal as a function of the array element position with a fixed spatial frequency proportional to $\sin\theta_d$. As with deriving the range profile, a Fourier transform is able to extract the frequency of interest. In this case, the Fourier transform acts to form beams across all angles. The result of applying a Fourier transform across the array dimension is given in Equation (3.15), where f_a is the spatial frequency across the array dimension. Taking the magnitude of the beamformed complex signal results in the map given in Equation (3.16), which can be transformed to the range-angle map, Equation (3.17) using the previously used transformation and, $\sin\theta = f_a/f_c$. In the range-angle map, the sinc characteristics of the PSF are present independently in both dimensions as expected.

$$\begin{aligned} m(f, f_a) =& \mathcal{F}[w(f, t_a)] \\ =& A\tau e^{2\pi i\left(f_c t_d + \frac{1}{2}kt_d^2 - f\right)} \operatorname{sinc}(\pi\tau(f - kt_d)) \int_{-\infty}^{\infty} e^{2\pi i f_c t_a \sin\theta_d} \Pi\left(\frac{t_a}{L}\right) \mathrm{d}t_a \\ =& A\tau L e^{-2\pi i f} \operatorname{sinc}(\pi\tau(f - kt_d)) \operatorname{sinc}(\pi L(f_a - f_c \sin\theta_d)) \end{aligned} \tag{3.15}$$

$$|m(f, f_a)| = A\tau L \operatorname{sinc}(\pi\tau(f - kt_d)) \operatorname{sinc}(\pi L(f_a - f_c \sin\theta_d)) \tag{3.16}$$

$$|m(r, \theta)| \propto \operatorname{sinc}\left(\frac{\pi\beta}{c}(r - r_d)\right) \operatorname{sinc}(\pi L f_c(\sin\theta - \sin\theta_d)) \tag{3.17}$$

3.5.3 Co-located Multiple-Input Multiple-Output

Co-located Multiple-input multiple-output (MIMO) radar is a particular variation of array radar in which the array is virtual and is formed by pairs of real transmit and receive elements that are spaced along an array aperture in a specific configuration. The fundamental physical principle of MIMO radar is that every pair of transmit and receive antennas is equivalent to a single virtual antenna that is located halfway between them.

To form a virtual array, the transmit signals from each of the transmitter elements must be orthogonal so that the signals at the receiving antennas may be assigned to the appropriate virtual element of the array. One method for achieving orthogonality is to separate the transmissions from each element in time, allowing the received signals to be unambiguously associated with the correct transmitter element.

As MIMO will be used as a technique rather than the subject of research, this thesis does not attempt to compare or contrast different MIMO approaches (such as different techniques for ensuring orthogonality of the transmit signals). Instead, this section presents the theoretical basis for monostatic MIMO, in which the antenna elements are co-located and the RCS of a target is the same across all elements.

Conceptualized in terms of path length, this means that the distance travelled by the radar signal must be the same for the transmit-receive pair of antennas as for the virtual antenna. Figure 3.9 indicates the geometry in this situation, where r_{Tx}, r_{Rx}, and r_V are the ranges to a target from the transmit antenna, the receive antenna, and the virtual antenna respectively. Using this schematic, Equation (3.18) can be written to relate the path lengths to each other, from which Equation (3.19) can be written to show that the

path lengths are the same.

$$\begin{aligned} r_V &= r_{Tx} - \frac{r_a}{2}\sin(\theta) \\ &= r_{Rx} + \frac{r_a}{2}\sin(\theta) \end{aligned} \tag{3.18}$$

$$r_{Tx} + r_{Rx} = 2r_V \tag{3.19}$$

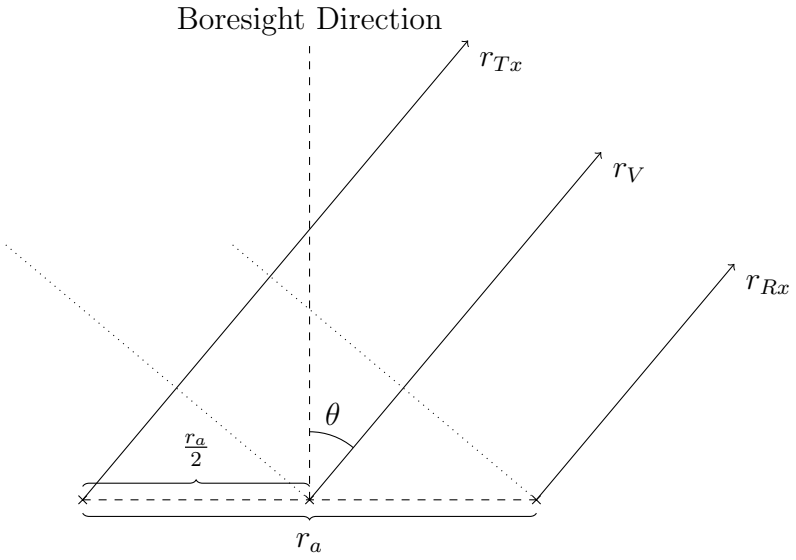


Figure 3.9: Geometry for MIMO array forming.

By increasing the number of transmit and receive elements, the number of virtual elements scales favourably. Each pair of transmit and receive elements generates a virtual element, and so the number of virtual elements is the product of the number of transmit elements and the number of receive elements. In a 4-by-8 MIMO configuration, meaning 4 transmit elements and 8 receive elements, there are 32 virtual elements but only 12 real elements as shown in Figure 3.10. Once the virtual array is formed by arranging the data from each transmitter-receiver pair appropriately, the signal processing to form a range-angle map is exactly the same as for a real array, given in Section 3.5.2.

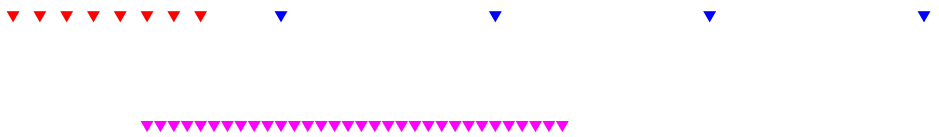


Figure 3.10: An example 4-by-8 MIMO array, with 4 transmit elements (blue) and 8 receive elements (red). The corresponding virtual array with 32 elements is also shown (magenta).

3.6 Wideband Ambiguity Function

The wideband ambiguity function (WAF) is a function that describes the result of cross-correlation of a wideband waveform with a time-delayed and Doppler-compressed version of itself. It is a useful function for understanding the range and Doppler resolution of a wideband waveform, as well as the interplay between these factors and is used for the analysis of human echolocator clicks in Chapter 4.

The WAF is the general function which allows the narrowband ambiguity function to be derived. The narrowband ambiguity function is specifically for narrowband signals, and is given in Equation (3.20), where $s(t)$ is the transmit signal, tau is the time delay of the received signal, and δf is the Doppler frequency shift. For a narrowband signal, the Doppler frequency shift is approximately equal across the whole signal, and so a single term, δf, can be used to account for this [123].

However, in the wideband case, the Doppler compression results in very different Doppler frequency shifts for the different frequency components of the signal. For wideband signals, the WAF can be formed, as in Equation (3.21), by using the Doppler compression factor defined in Equation (3.22). Here, c_0 is the speed of propagation in the medium, and v is the radial velocity that causes the Doppler shift. This factor compresses the whole signal, resulting in an accurate description of a signal after Doppler shift.

$$\chi(\tau, \delta f) = \int_{-\infty}^{\infty} s(t) s^{\star}(t-\tau) e^{2\pi i \delta f t} \mathrm{d}t \tag{3.20}$$

$$\chi(\tau, \alpha) = \sqrt{|\alpha|} \int_{-\infty}^{\infty} s(t) s^{\star} (\alpha t - \tau) \, \mathrm{d}t \tag{3.21}$$

$$\alpha = \frac{c_0 - v}{c_0 + v} \tag{3.22}$$

The width of the WAF along the range dimension corresponds to the range-resolution of the signal and the width along the Doppler dimension corresponds to the Doppler resolution. Because of these properties, it makes the ambiguity function useful in assessing the resolution properties of a waveform.

3.7 Sound Pressure Level

The sound pressure level (SPL) is a measure of the pressure of a sound wave as it causes compressions and rarefactions in the air and is used in Chapter 4 for analysing the signal-to-noise ratio (SNR) of audio recordings. The SPL is the root-mean-squared (RMS) measurement of the sound pressure averaged over a period of time as given in Equation (3.23), where p_i is the i^{th} sample of sound pressure and N is the number of samples in the time period. It is often expressed logarithmically (due to the logarithmic nature of perception), and is referenced to 20 μPa as given in Equation (3.24), where $p_0 = 20$ μPa. 20 μPa is the chosen reference as this is approximately the threshold of human hearing [124].

Example SPLs for common situations are given in Table 3.2 for reference and to provide context to the values quoted in this thesis.

$$p = \frac{1}{\sqrt{N}} \sqrt{\sum_{i=1}^{N} p_i^2} \tag{3.23}$$

$$L_p = 20 \log_{10} \frac{p}{p_0} \tag{3.24}$$

Table 3.2: Example SPL values for common situations [125, p.17].

Source / observing situation	**Typical sound pressure level (db SPL)**
Hearing threshold	0
Leaves fluttering	20
Whisper in an ear	30
Normal speech conversation for a participant	60
Cars/vehicles for a close observer	60-100
Airplane taking-off for a close observer	120
Pain threshold	120-140

Chapter 4

Characterizing Blind Expert Echolocator Tongue Clicks

There are things known and there are things unknown,
and in between are the doors of perception.

— Aldous Huxley

The Doors of Perception, 1954

4.1 Introduction

Echolocation is an ability shared by many species in nature and involves the transmission of a signal which is then reflected by objects in the environment. The echoes return to the echolocator, where they are received at the ears and used to build up a perception of the echolocator's environment. It has long been established that certain species of bats and marine mammals use echolocation to navigate and locate prey [71, 62, 126, 127], but research has also demonstrated that humans are capable of echolocation (for reviews see [77, 78, 79]). There are some blind people who have learned to use mouth-clicks to achieve extraordinary levels of echolocation performance [51, 82], in some cases rivalling performance of bats [81].

As with any natural echolocation, human echolocation shares many parallels with man-made sonar systems, and relies similarly on signal transmission (of tongue-clicks) and signal reception (at the ears). This chapter examines a large number of human echolocator tongue clicks to build a model representation that captures their essential characteristics. The tongue-click is one of the essential elements of the human echolocation system and there have been several broad descriptions of these tongue-clicks which estimate some of their important parameters (such as click duration and peak frequency in the forward direction) [83, 85], but the descriptions of other parameters (including click spectral content and time-domain profiles) are restricted to small datasets [128, 129, 83]. A more complete discussion of echolocation and previous work on human echolocation is presented in Chapter 2.

This chapter aims to present the synthesis of artificial human echolocator tongue-clicks. The motivation for creating artificial clicks is that conducting experiments with human echolocators is difficult for researchers due to the logistical and financial challenges of finding echolocators that are able to participate. For some experiments, it may be helpful for researchers to pilot a trial using artificially synthesized clicks that mimic the clicks of human echolocators in a simulation before they take steps to arrange for human echolocators to participate in person.

To synthesize the clicks, a detailed description of human echolocator tongue clicks is required. This chapter provides a more detailed description of human echolocator tongue clicks than has previously been reported and is used to suggest ideas of how human echolocation could inform the design and signal processing of a biologically-inspired radar system.

This chapter describes the experiments and processing necessary for building a database of thousands of human echolocator tongue clicks, and presents detailed descriptions of the acoustic properties of the human echolocator tongue-clicks in the database, including their spectral-, spatial-, and time-domain representations. These different representations of the clicks are discussed in more detail later in the chapter.

This chapter begins by providing details of the experiments required to capture thousands of human echolocator tongue-clicks in Section 4.2, and then presents the techniques used to assemble the database from the experiments in Section 4.3. Section 4.4 details the approach and results of analyzing the database of clicks using a variety of techniques, and Section 4.5 presents the synthesized tongue clicks and compares them to individual tongue clicks. A summary of results is included in Section 4.6.

This chapter is based on published work by the author [1].

4.2 Experimentation

The goal of the experiments referred to in this chapter was to construct a significantly large database of human echolocator tongue click recordings that could be processed and analyzed to extract the key parameters of the expert echolocator tongue clicks and hence inform the design of a realistic model representation of them. The approach was to record stationary expert human echolocators producing tongue-clicks in a controlled laboratory environment and to process the audio recordings into a database. There are numerous challenges in doing this, and care must be taken to ensure that the audio recordings are of sufficient quality to enable the subsequent processing and analysis. To

ensure useable quality, the signal-to-noise ratio (SNR) must be sufficiently large and the instrumentation must be designed to reduce the number of unusable clicks in the recording. The following sections address the participants of the experiments as well as the equipment and methodology used to produce a useful click database.

4.2.1 Participants

Three blind people with expertise in echolocation participated in the trials conducted at Durham University. The participation of three expert echolocators in the trials enabled comparisons between the echolocator clicks to be made to help understand the key features of their clicks. Throughout this chapter, the expert echolocators are referred to as EE1, EE2, and EE3. EE1 is male, was 49 years at time of testing, was enucleated in infancy because of retinoblastoma, and is reported to have used echolocation for as long as he can remember. EE2 is male, was 33 years at time of testing, lost sight aged 14 years due to atrophy of his optic nerve, and is reported to have used echolocation on a daily basis since he was 15 years old. EE3 is male, was 31 years at time of testing, he lost sight gradually from birth due to glaucoma. From approximately 3 years old EE3 was only able to detect bright light and is reported to have used echolocation daily since he was 12 years old.

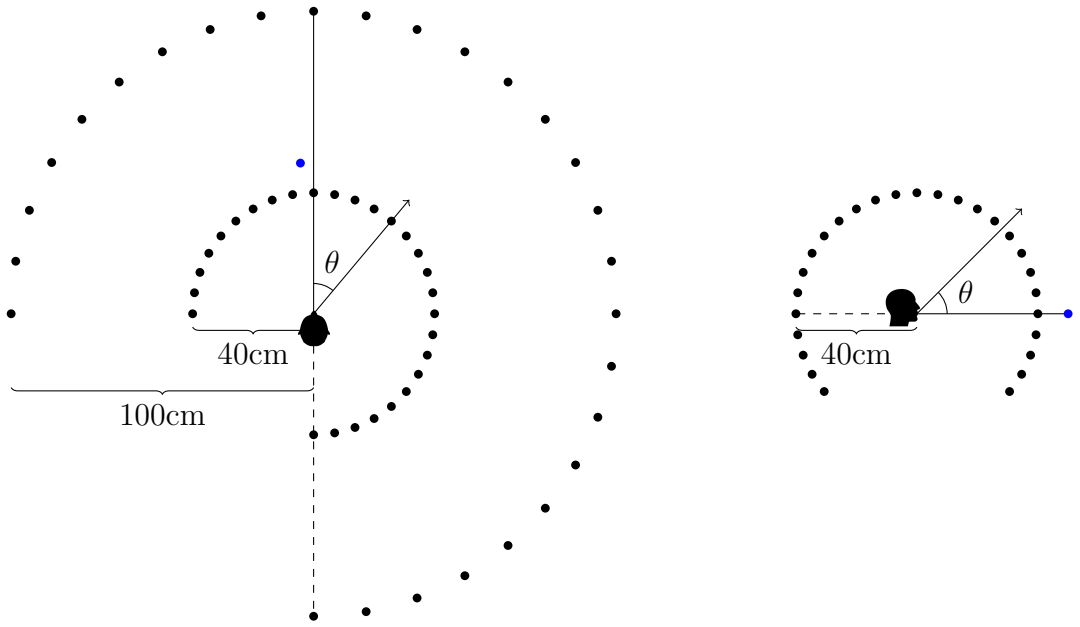
One of the limitations of this research is that it is based on recordings of three echolocators, as the incidence of echolocators in the population is extremely low. As such, care must be taken when attempting to make general conclusions about all human echolocators. Each echolocator's clicks will be treated separately in the analysis and only features common to all three will be generalized into the synthetic model.

4.2.2 Apparatus

The trials were conducted in a sound-insulated and echo-acoustic dampened room to reduce the effects of clutter noise from the busy university department outside the room. The room was rectangular, with dimensions of approximately 2.9m x 4.2m x 4.9m and was lined with fabric acoustic dampeners that effectively absorbed frequencies above 315 Hz.

Participants were positioned in the centre of the room. The elevation of a participant's mouth with respect to the floor was: EE1 - 154 cm, EE3 - 170 cm, EE2 - 143 cm. The floor was covered with foam baffles to reduce multipath reflections from the floor.

In order to record the expert echolocators, a reference microphone was placed 50 cm in front of the participant, while a second microphone was placed at either 40 cm or 100 cm distance and moved around the participant to capture variation in clicks as a function of range, azimuth, and elevation. The participants remained stationary during the trial. The smallest distance to the microphone was selected to be 40 cm to reduce the likelihood of a click saturating the microphone input which would result in a 'clipped click' that would be unusable in the analysis. Figure 4.1 indicates the microphone positions which were measured; in the azimuthal plane a span of 270^o in 10^o steps starting to the left of the participant was measured and in the elevation plane a span of 260^o in 10^o steps starting 40^o below the horizontal plane to the front was measured. The microphones used were DPA SC4060 [130] (with protective grid removed) and TASCAM DR100-MKII [131] operating at 96 ksps with 24 bit recording.



(a) Azimuth plane viewed from above. (b) Elevation plane viewed from the side.

Figure 4.1: The microphone positions for the click recording trials. The position of the reference microphone is shown in blue, and the possible positions of the signal microphone are shown in black. The signal microphone was moved sequentially between these positions to record a series of clicks at each position.

4.3 Building the Click Database

The experimental campaign resulted in a series of audio files of the human echolocators. These files contained background noises from the building's air conditioning, coughs and sniffs from the echolocators, and low frequency and low magnitude signals likely derived from the recording equipment power supply. These background noises needed to be filtered out for the subsequent analysis of the clicks. Additionally, each recording contained many sequential clicks which needed to be processed into a database of individual click recordings. Figure 4.2 shows a schematic overview of the processes used to build the click database.

The audio recordings were filtered using a finite impulse response equiripple filter with the parameters given in Table 4.1 to remove background noises including the low frequency hum of air conditioners in the building and increase the SNR. The magnitude response of the filter is shown in Figure 4.3, the linear phase response is shown in Figure 4.4 and

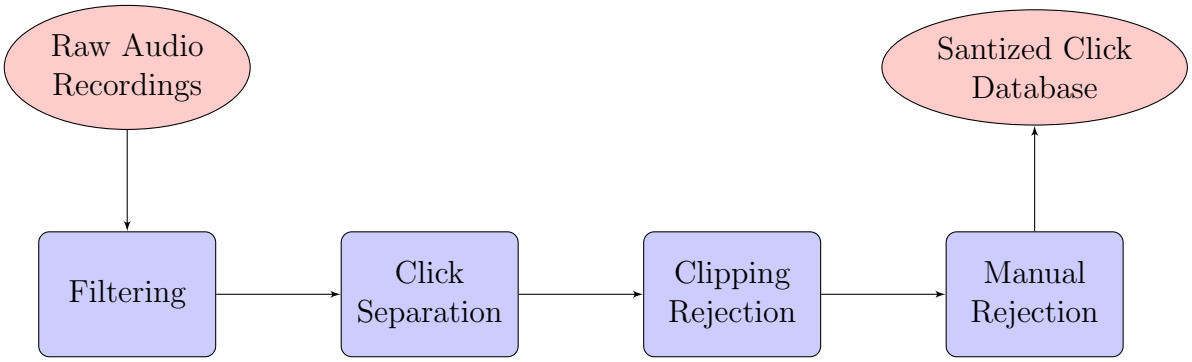


Figure 4.2: Flowchart to show the steps taken to build the sanitized click database.

examples of the audio file before and after filtering are shown in Figure 4.5 which shows the removal of the low-frequency oscillations from the signal, with minimal impact on the expert echolocator tongue clicks.

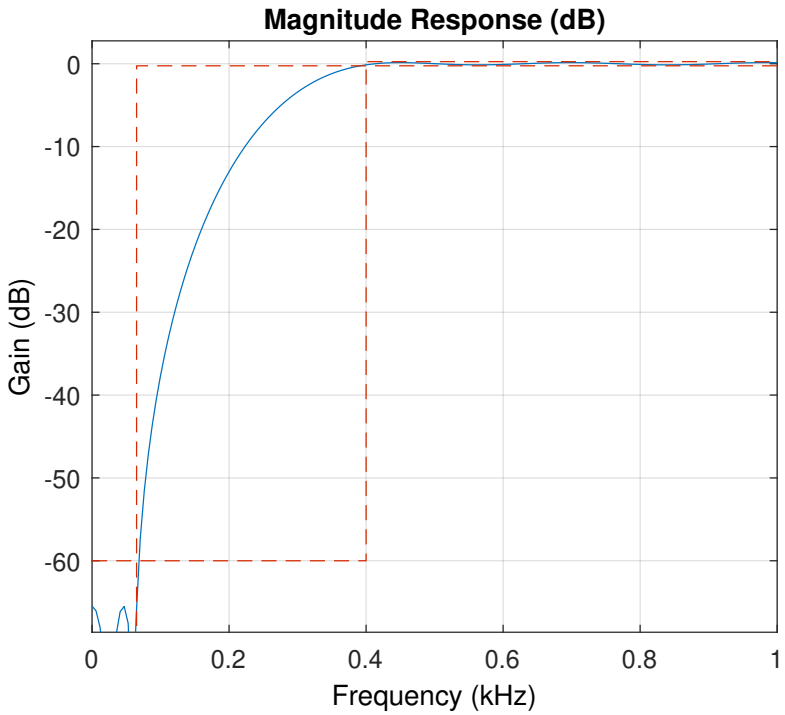


Figure 4.3: The magnitude response of the equiripple filter.

To separate the long audio recordings into short recordings of the clicks, the location of the clicks needed to be known. Clicks were identified based on the sound level exceeding a threshold of 9.5% of the maximum amplitude observed in the file, and were windowed by a fixed period in time (10ms before the onset and 30ms after) to capture the full spread

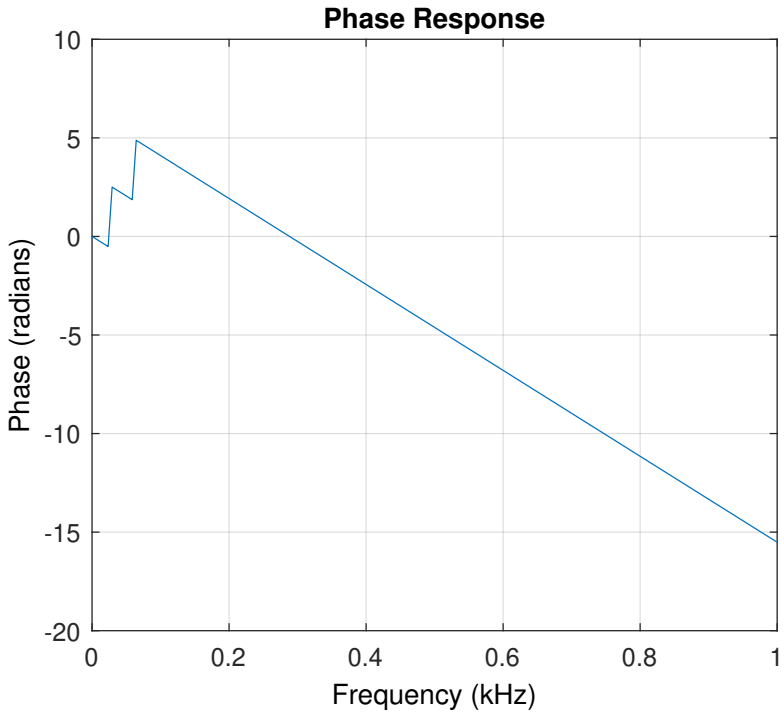


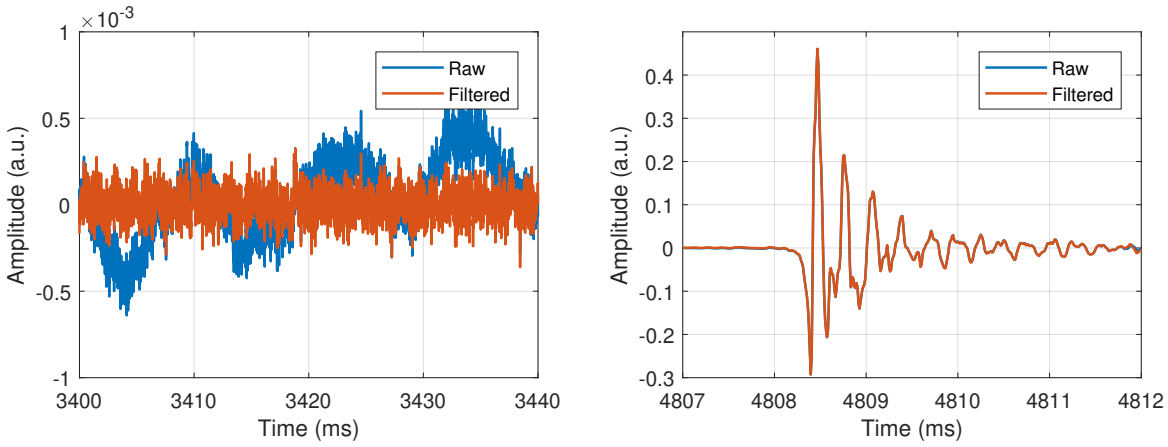
Figure 4.4: The phase response of the equiripple filter.

Table 4.1: The parameters used to design a finite impulse response equiripple filter using MATLAB to implement the Parks-McClellan algorithm [132].

Parameter	**Value**	**Units**
Stopband Frequency	65	Hz
Passband Frequency	400	Hz
Stopband Attenuation	60	dB
Passband Ripple	0.5	dB

of the click. These clicks were then manually reviewed to reject other sounds including coughs and sniffs from the database.

Clicks that saturated the microphone and were therefore clipped in the recording were also excluded as the full waveform was not available for analysis in these cases. The number of sanitized clicks (meaning clicks after filtering, clipping rejection, and manual rejection) for each echolocator and each recording position are given in Appendix A, which shows that 9579 clicks were available for analysis, varying between 19 and 84 clicks at each angular position.



(a) (b)

Figure 4.5: Time series plots showing the recorded audio signal before and after filtering. (a) Quiet section of the recording which shows how filtering removes the low-frequency oscillations present in the raw signal. (b) Section of the recording containing a tongue click which shows how filtering has a minimal effect on the overall click.

4.3.1 Signal-to-Noise Ratio

Figure 4.6 shows a selection of individual clicks taken from the database, which shows that the click waveforms of each echolocator are visually different, but with some commonalities between clicks. These sample waveforms are indicative of the waveforms stored within the click database.

The sound pressure level (SPL) is one of the standards used for measuring magnitude of sound sources, as reviewed in Chapter 3. SPL measurements for the background noise of the room used for the trials place the noise power before filtering at 38 dB (SPL), which is decreased to 30 dB (SPL) by the filtering process. The recorded clicks have peak SPLs around 100 dB (SPL) and this is unchanged by the filtering process. The SNR for the dataset after filtering is approximately 70 dB (SPL) which is sufficient for the subsequent analysis.

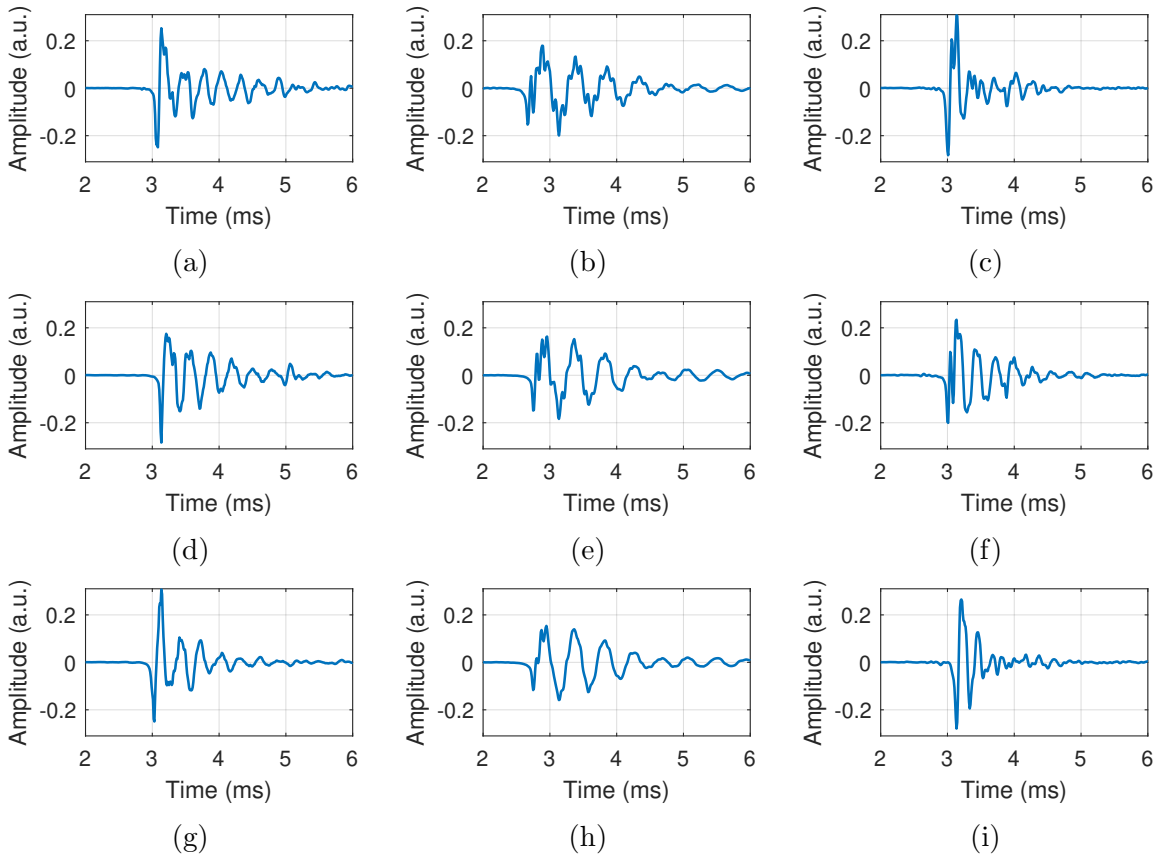


Figure 4.6: Sample tongue clicks. The left column is EE1, the middle column is EE2, and the right column is EE3.

4.4 Results

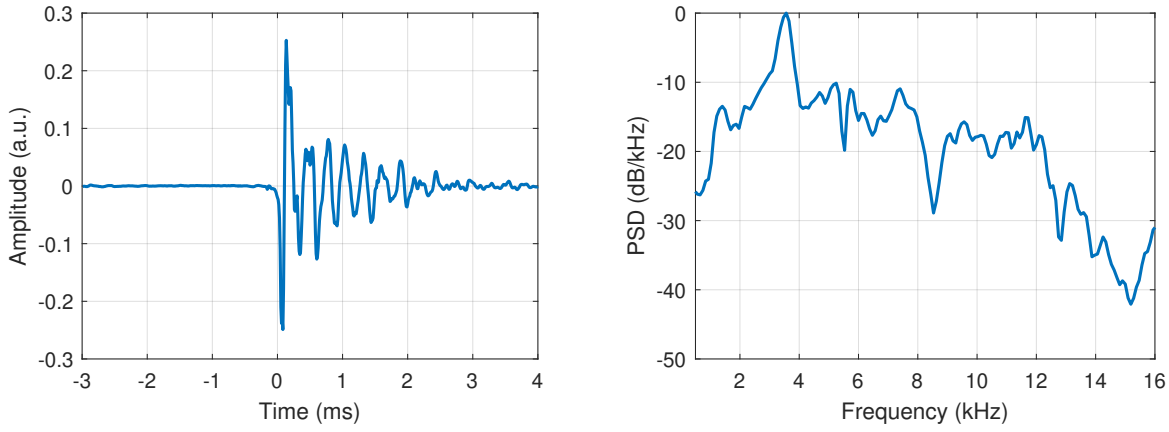
The clicks in the database were found to have little to no correlation in the time domain, despite having some visually similar features. This low correlation was likely due to different clicks containing different frequency components (as different frequency sinusoids are orthogonal, this would decrease the time-domain correlation). Instead of conducting detailed time-domain analysis of the clicks, the different elements of the click were analyzed independently so that the analysis could be performed in the most suitable domain for each component. The following sections present the analysis of the spectral content, the spatial distribution, and time-domain envelope of the clicks, each of which was performed in a different domain (frequency-, spatial-, and time-domain respectively). This section also presents results considering the wideband ambiguity function (WAF) of the clicks. The WAF is a useful tool often used in radar for determining the range-Doppler ambigu-

ities that would arise due the waveform of a signal, and provides information about the range and Doppler resolution of a signal.

4.4.1 Spectral Content

The spectral content of the clicks, meaning the spread of frequency information encoded within them, was of interest because frequency modulation is of great importance to radar and sonar systems, as it determines the system's ability to resolve the range and Doppler of targets, and is seen in many bat echolocation calls. Further, understanding the frequency content of the clicks is essential for creating a model representation of them.

Two techniques were used to access the spectral content of the tongue clicks in the database. The first technique was to use the periodogram power spectral density (PSD) estimate [133], which uses the fast Fourier transform (FFT) to identify the overall spectral components of the tongue click. The second technique was to take the short time Fourier transform (STFT) of the clicks to look for time-varying frequency-modulation within the clicks.



(a) (b)

Figure 4.7: (a) Example Click. (b) Spectrum of Example Click.

The PSD for clicks belonging to a single echolocator were found to contain features that were present in the majority of the echolocator's tongue clicks. These spectral peaks varied significantly between the echolocators, but were reasonably consistent for a single

echolocator, and can be thought of as a 'fingerprint' for each echolocator's tongue clicks. An example click and its corresponding PSD estimate are shown in Figure 4.7 for reference, which shows an initial peak just below 4kHz, with a spectrum of frequencies decreasing away from the peak. The PSD estimates were made with an FFT of length 1024 (which is the next power of 2 greater than the data length of 700) with a rectangular window on the signal sampled at 96 ksps.

The top row of Figure 4.8 shows averaged PSD estimates calculated for each echolocator. These average PSD plots show the average spectrum in blue, with the grey area representing the standard deviation of the logarithm of the spectrum (which was found to be log-normally distributed at each frequency). From these plots, it can be seen that the differences in clicks between the echolocators are much greater than the variation between clicks for a single echolocator. It is likely that these differences originate due to different physical characteristics (such as mouth shape) between the echolocators as well as technique differences between them.

The middle row of Figure 4.8 shows waterfall plots with a set of individual click PSD estimates which are combined to form the averaged plots discussed above. These plots show the similarities between an echolocator's clicks graphically.

Average STFTs of the clicks are shown on the bottom row of Figure 4.8 and show how the spectral components of the clicks vary as a function of time. These plots show that the key components within the clicks consist of single monotones that are amplitude modulated, manifesting as a horizontal peak in the spectrograms. An example of a bat spectrogram which exhibits structural frequency-modulation is given in Figure 4.9 for comparison. The STFT representation of the clicks was calculated using an FFT length of 1024, with a moving Kaiser window that was 220 samples (2.3 ms) long with a beta value of 3.

As no structural frequency-modulation was identified in the human echolocator tongue clicks, attention was instead focused on the identification and modelling of the key static spectral components of the tongue clicks. Figure 4.8 also shows that, for EE1 and EE2,

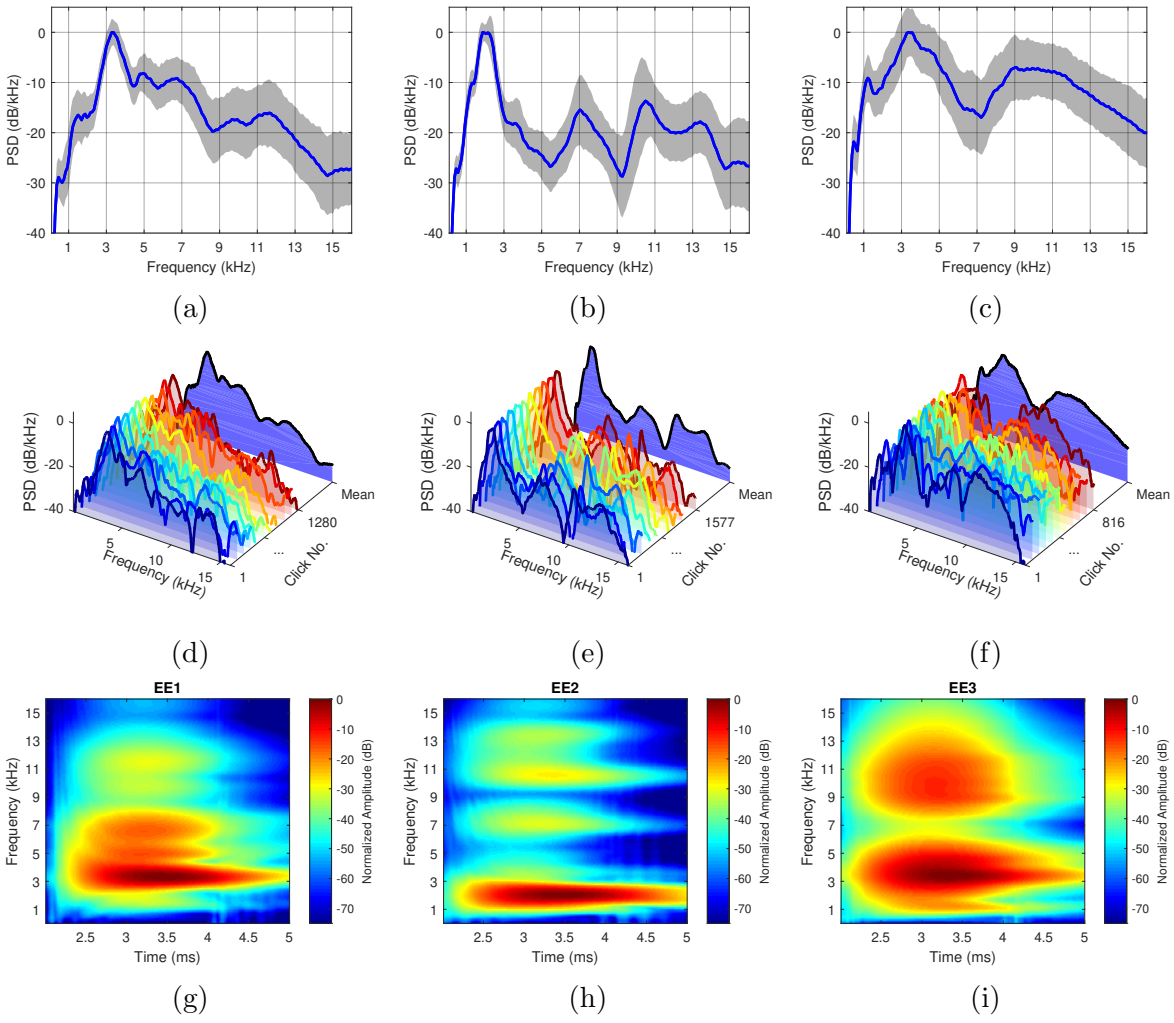


Figure 4.8: Plots indicating the frequency content of the expert echolocator clicks at a distance of 40 cm. (a)-(c) Average periodogram PSD estimates for each expert echolocator in blue with ± 1 standard deviation of the logarithm of the spectrum in the shaded grey region. (d)-(f) Waterfall plots for a sample selection of clicks indicating the consistency of frequency components between a single echolocator's tongue clicks. (g)-(i) Average spectrograms for each expert echolocator, showing that there is no time-dependent frequency-modulation over the duration of the tongue clicks.

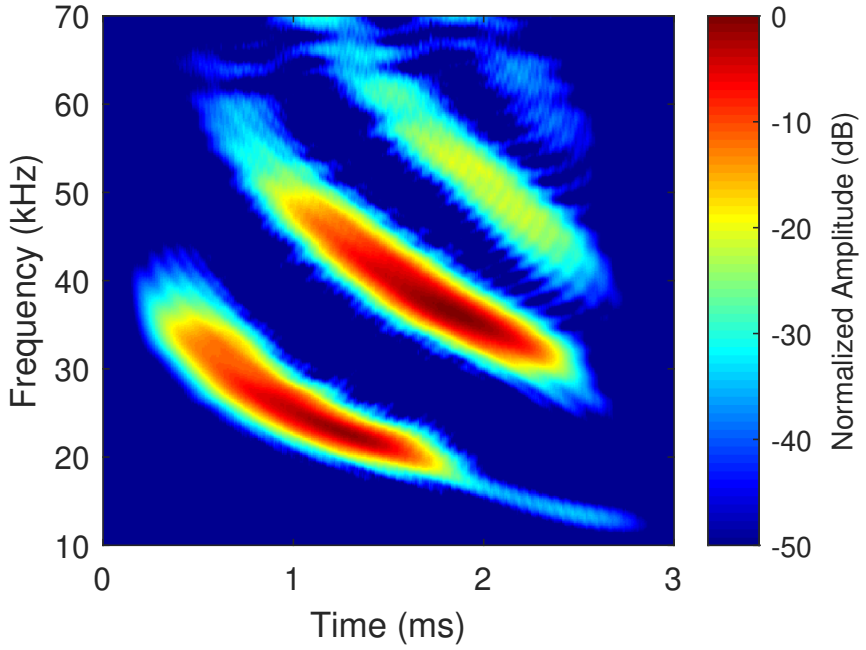


Figure 4.9: A Large Brown Bat (*Eptesicus Fuscus*) echolocation call exhibiting structural frequency-modulation. The author wishes to thank Curtis Condon, Ken White, and Al Feng of the Beckman Center at the University of Illinois for the bat data and for permission to use it in this paper [134].

the main frequency peak varies less than the other frequency components. For EE3, the peaks are more broadly spread, indicating that the monotone frequency locations within the frequency domain vary more than for the other echolocators.

4.4.1.1 Parametrization

To parametrize the spectral content of the tongue clicks, only the most significant (largest amplitude) frequency components were considered. As these components did not vary in frequency as a function of time, they were treated as monotone sinusoids that could be summed independently. The sum of monotones captures a set of key frequency components and their relative amplitudes and phases, as described in Equation (4.1). In Equation (4.1), A_m describes the amplitude of the m^{th} frequency component, f_m is the

frequency of the component, and ϕ_m is the relative phase of the component.

$$F(t) = \sum_{m=1}^{M} A_m \cos(2\pi f_m t + \phi_m) \tag{4.1}$$

This model is capable of capturing the most significant m monotones within the clicks, as well as their relative amplitudes and phases which are required to capture the timbre of the clicks. This parametrization will be used later in Section 4.5 to help synthesize an artificial click by contributing a frequency-domain component to the synthetic click.

4.4.2 Spatial Distribution

The spatial distribution of the clicks refers to how the energy of the click is distributed in space and is analogous to the beam pattern of an antenna in radar. For an antenna, the beam pattern is determined by its physical construction. Similarly for human echolocators, the main contributing factor to the click distribution is expected to be the bulk of the head, which is known to effect how humans perceive sound sources [135, 136]. Additionally, it is likely that the shape of the mouth is a contribution factor in the tongue click spatial distribution.

The echolocators produced several tongue clicks with the microphone at each measurement position (shown in Figure 4.1) and the set of recorded clicks were used to extract the click spatial distribution. This measurement technique assumes that the clicks had broadly similar spatial distributions over the course of the trial, and the echolocators were accordingly asked to produce uniform clicks, though it is not possible to validate this assumption.

The signal energy for each click, E, was found using Equation (4.2), where S is an audio recording of T samples. The energy of each click at the signal microphone was then normalized to the energy of the click as recorded at a reference microphone (denoted by subscripts s and r respectively), and averaged over all the clicks recorded at the same angular position to determine the average gain at that angle as indicated in Equation (4.3).

This processing allowed for natural variations in click volume to be compensated, and was repeated for every angular position (in azimuth or elevation), θ, to build up an overall directivity pattern as described in Equation (4.4).

$$E = \sum_{t=1}^{T} S\left(t\right)^2 \tag{4.2}$$

$$G = \frac{1}{N} \sum_{n=1}^{N} \frac{E_s\left(n\right)}{E_r\left(n\right)} \tag{4.3}$$

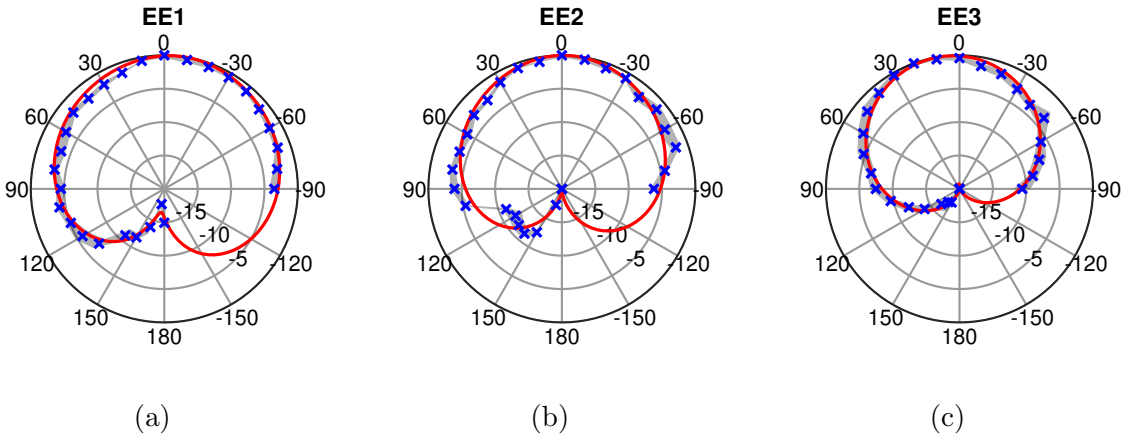
$$D\left(\theta\right) = \frac{1}{N\left(\theta\right)} \sum_{n=1}^{N(\theta)} \frac{E_s\left(n,\theta\right)}{E_r\left(n,\theta\right)} \tag{4.4}$$

The azimuthal directivity patterns in Figure 4.10 show similar spatial distributions between the expert echolocators. The reason for the near-cardioid behaviour is likely due to the bulk of the head increasingly attenuating the tongue click in the reverse direction, in a similar manner to the head-shadow effect and the head-related transfer function (HRTF) [27, 137]. As the head is approximately symmetrical in the azimuthal plane, it is expected that these patterns would also be symmetrical. The beamwidth of the directivity patterns is also of note and could be influenced by the method of click production; EE1's clicks are the broadest with a 3 dB beamwidth of approximately 80^o, EE2's clicks are the most symmetrical and narrower at roughly 60^o beamwidth, and EE3's clicks are the most directive and least symmetrical, with a beamwidth of approximately 50^o.

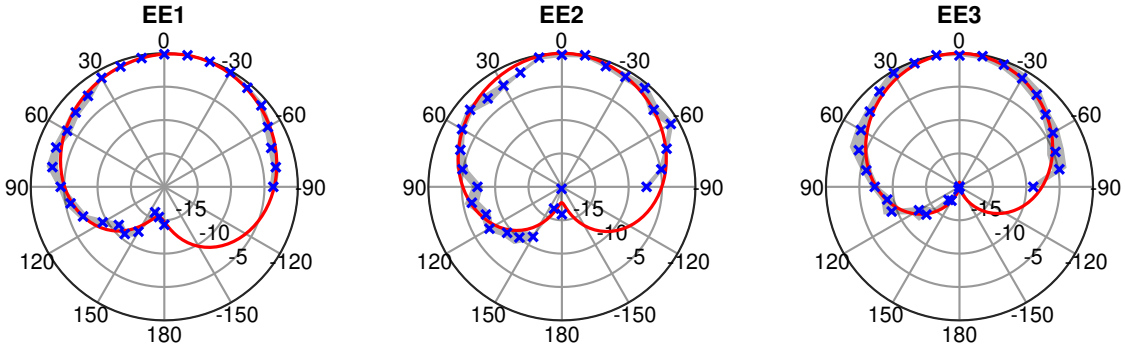
The elevation directivity patterns in Figure 4.11 show similar spatial distributions between the echolocators, with asymmetry between the angles above and below the azimuth plane, which is expected due to the asymmetry of the head in the vertical plane.

4.4.2.1 Parametrization

The directivity pattern in azimuth was parametrized by using a cardioid modified with an ellipse as given in Equation (4.5). Initially a cardioid was used as it is a function which commonly occurs in acoustics when a diaphragm is positioned next to a bulk object, such

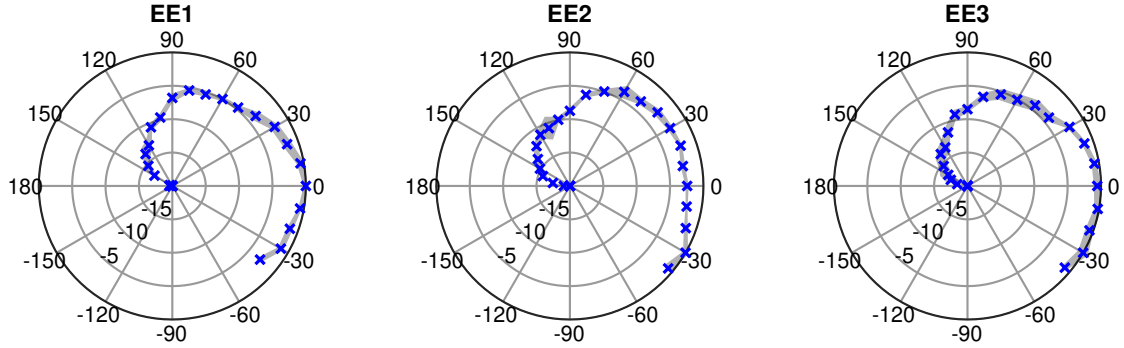


(a) (b) (c)



(d) (e) (f)

Figure 4.10: Tongue click directivity patterns measured in the azimuthal plane for three expert echolocators at a distance of 40 cm (top row) and 100 cm (bottom row). The blue crosses mark the mean power ratio (in dB) at that angle, the grey region shows ±1 standard deviation about the mean, and the red line shows the fit of a modified cardioid described in Section 4.4.2.1



(a) (b) (c)

Figure 4.11: Tongue click directivity patterns measured in the elevation plane for three expert echolocators at a distance of 40 cm. The blue crosses mark the mean power ratio (in dB) at that angle and the grey region shows ±1 standard deviation about the mean.

as with many loudspeakers and microphones. With these objects, sound destructively interferes in the 'reverse' direction due to different paths from/to the diaphragm, and this seemed to be a reasonable hypothesis for the human head with the mouth acting as the sound source.

The shape of the click directivity patterns appeared to be cardioid, but this description was found to lack the degrees of freedom need to accurately model the spatial distribution of the clicks. The addition of elliptical modulation allows for the broadness of the click to be captured in the model and for the overall model to be a good fit to the data. This parametrization of the azimuthal directivity pattern will be used later in Section 4.5 to help synthesize an artificial click by contributing a spatial-domain component to the synthetic click.

$$D(\theta) = \frac{-(1+\cos\theta)}{\sqrt{\alpha^2\cos^2\theta + \beta^2\sin^2\theta}} \tag{4.5}$$

For elevation, no physically meaningful parametrization was attempted due to the complex shape of the head in the vertical plane.

4.4.2.2 Frequency-Dependent Analysis

It was also of interest to investigate if the frequency components of the tongue clicks (extracted in the previous section) shared a spatial distribution or if, as with electromagnetic antennas, the directivity of the component varied with frequency.

The frequency-dependent spatial distribution of the tongue clicks was estimated in a similar way to the overall distribution. To estimate the gain for a particular set of frequency components, the periodogram estimate of power spectral density was used. The periodogram of a real signal is equivalent to the squared magnitude of the one-sided Fourier transform, $\mathcal{F}$, of the same signal. Equation (4.6) describes how the periodogram was used to estimate the energy contained within a frequency band, f_{lo} to f_{hi}, normalized to the bandwidth. The bands used are given in Table 4.2. The directivity pattern for each band was evaluated as described previously, using Equation (4.3) and Equation (4.4), and

Table 4.2: The frequency bands used to estimate the frequency-dependent click directivity.

Band	f_{lo} **(kHz)**	f_{hi}**(kHz)**
1	0.1	2
2	2	4
2	4	8
4	8	16

resulted in the directivity patterns given in Figure 4.12 and Figure 4.13.

$$E = \sum_{f=f_{lo}}^{f_{hi}} \frac{|\mathcal{F}\{S(t)\}(f)|^2}{f_{hi} - f_{lo}} \tag{4.6}$$

Both the azimuth and elevation frequency-dependent directivity patterns given in Figure 4.12 and Figure 4.13 show that the higher frequencies of the tongue click experience greater attenuation in the reverse direction than the lower frequencies which is a more pronounced effect compared with prior literature focused on speech [138, 139]. In general, it appears that the click frequency-components distribute similarly in space, and that the only significant difference between different frequencies is the increased attenuation by the mass of the head in the reverse direction.

4.4.3 Time-domain Envelope

The time-domain form of the audio signal is of interest because it is the factor, in combination with the spectral content, that determines the range resolution capabilities of the tongue clicks. In order to understand the time-domain form of the clicks, it is useful to determine the envelope of the signal. An envelope is a function which modulates the amplitude of underlying frequency components of a signal.

Peaks in the signal magnitude were identified and interpolated using cubic interpolation, to estimate the click envelope. Figure 4.14 shows an example click and the result of envelope extraction by peak detection, which shows that the envelope extracted by peak detection successfully captures both the click onset and the signal decay.

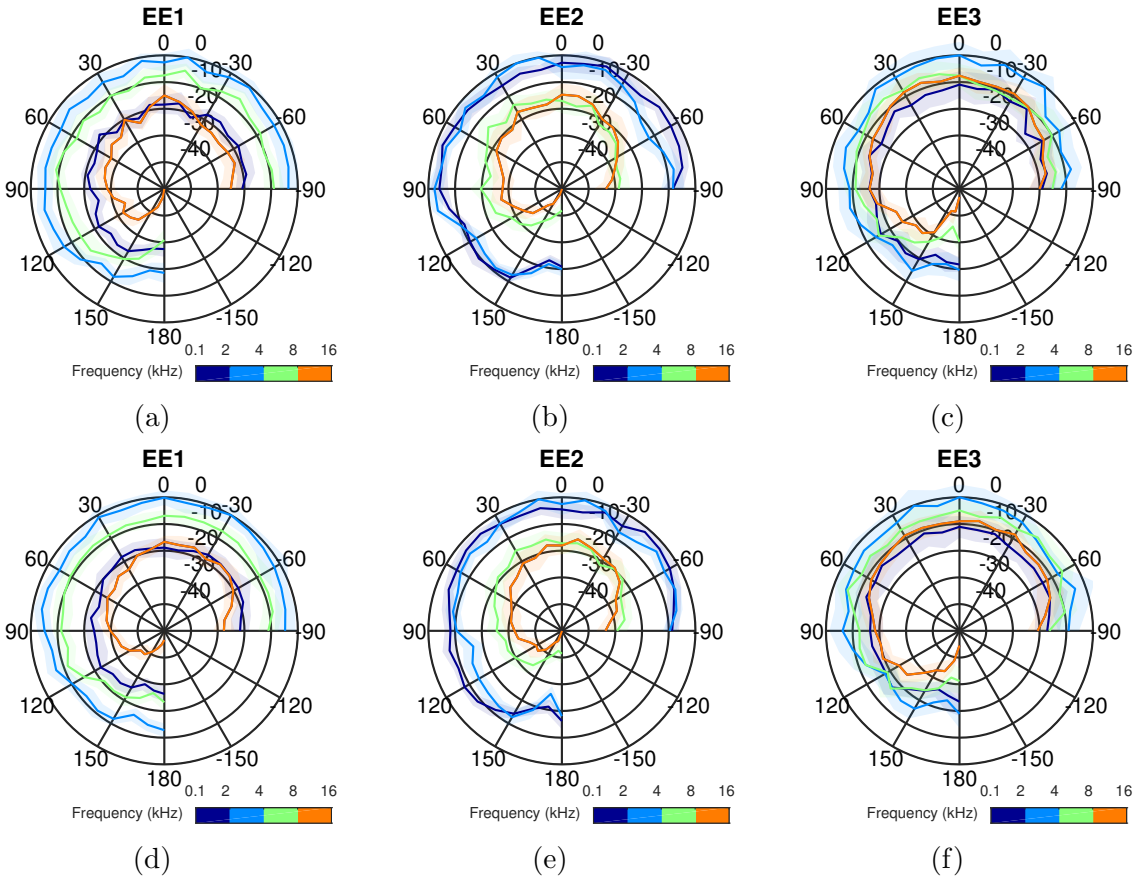


Figure 4.12: Tongue click directivity patterns for four frequency bands measured in the azimuthal plane for three expert echolocators at a distance of 40 cm (top row) and 100 cm (bottom row). The lines indicate the mean power ratio (in dB) and the shaded regions show ±1 standard deviation about the corresponding mean.

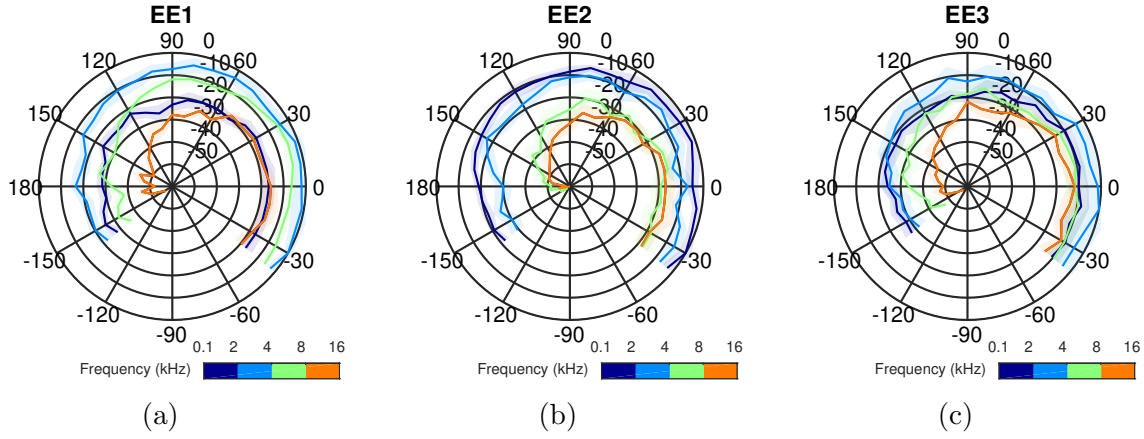


Figure 4.13: Tongue click directivity patterns for four frequency bands measured in the elevation plane for three expert echolocators at a distance of 40 cm. The lines indicate the mean power ratio (in dB) and the shaded regions show ±1 standard deviation about the corresponding mean.

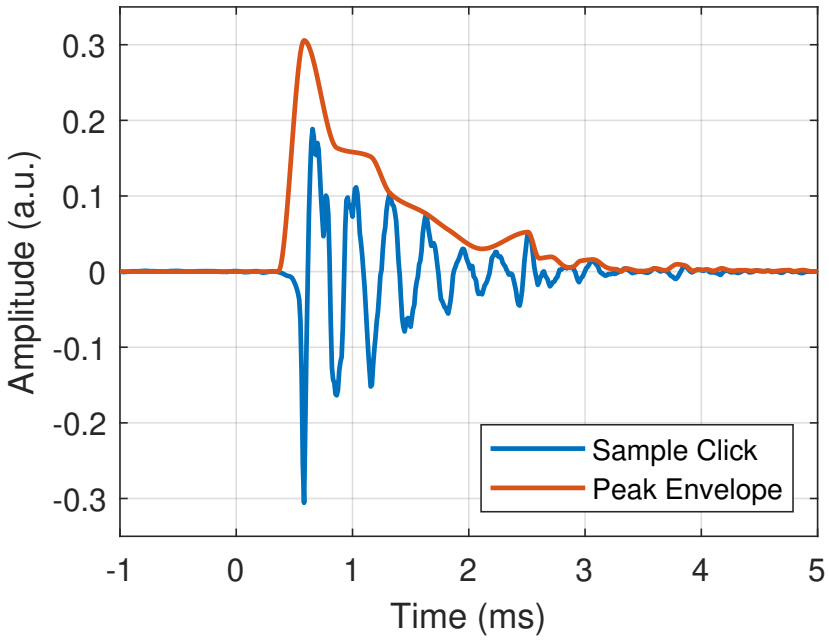


Figure 4.14: An example click (blue) and the envelope extracted using peak interpolation (red).

4.4.3.1 Parametrization

In order to parametrize the click envelope a function that captured both the initial rapid onset and the longer decay was required. The model that most simply satisfied these conditions was an exponential decay function mediated by a Heaviside step function and is given in Equation (4.7). In this model the step function term captures the initial sharp onset, and the exponential term captures the decay, resulting in a model that is able to well describe the click envelope.

$$E(t) = a \exp(-bt - c) H(t - c) \tag{4.7}$$

This parametrization will be used later in Section 4.5 to help synthesize an artificial click by contributing a time-domain component to the synthetic click.

4.4.4 Wideband Ambiguity Function

The WAF is a function which describes the ambiguity of a waveform at the output of a correlation process in terms of time delay caused by the range to a target and Doppler caused by a relative instantaneous velocity. The equation for the WAF for an analytic signal in the time domain, $s(t)$, is given in Equation (4.8) in terms of the time delay, τ, and the Doppler compression factor α. The analytic signal is achieved by taking the Hilbert tranform of the real signal. The Doppler compression factor is defined in Equation (4.9) where c_0 is the speed of propagation in the medium and v is the radial velocity of the target. More information about the WAF is available in Chapter 3.

$$\chi(\tau, \alpha) = \sqrt{|\alpha|} \int_{-\infty}^{\infty} s(t)\, s(\alpha t - \tau)\, \mathrm{d}t \tag{4.8}$$

$$\alpha = \frac{c_0 - v}{c_0 + v} \tag{4.9}$$

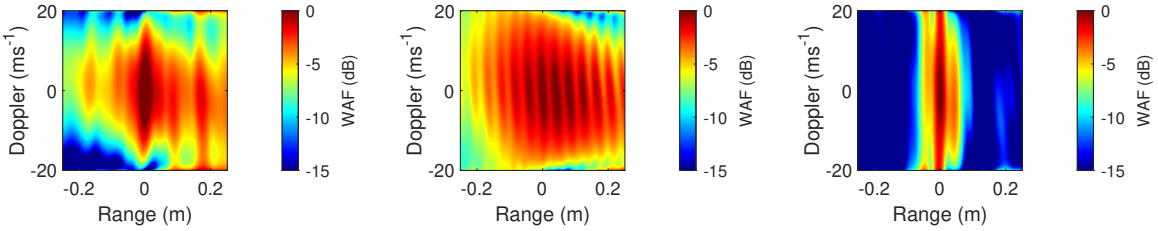
A set of sample WAFs for the different expert echolocators are given in Figure 4.15. These sample plots have a narrow peak in the surface along the range dimension at zero Doppler, which shows that the echolocator clicks have the ability to resolve targets that are closely spaced in range. Because this peak is a few tens of centimetres in width it indicates that the resolving capabilities of echolocator clicks are to within a few tens of centimetres. Previous work has reported that, under certain conditions, human echolocators can resolve range changes to within 10-20 cm [80, 87] and the resolutions predicted by the WAFs are consistent with this, but indicate lower average range resolution capability.

The sample WAFs also show that the Doppler resolution of the clicks is poor, indicated by a very wide peak in the Doppler dimension at zero range. At the time of writing, there is no other research into the Doppler resolution capabilities of human echolocators, and it is speculated that this is because a human echolocator is more likely to resolve motion through changes in range measurement between sequential clicks, than by sensing the Doppler shift of a single echo.

Table 4.3: The mean 3 dB peak widths of the WAF and their standard deviations. These numbers represent the resolving ability of the echolocator clicks in range and Doppler.

	Range (cm)		Velocity (msˆ-1)	
	Mean	Std Dev	Mean	Std Dev
EE1	30.0	16.2	60.8	16.2
EE2	58.0	16.7	73.9	15.6
EE3	37.5	27.3	62.3	25.3

The mean range and Doppler resolutions for each echolocator are given in Table 4.3. These resolutions are calculated by considering the 3 dB width of the peak in the zero-Doppler and zero-range cuts of the WAF. The WAF is not useful for developing synthetic clicks, but it is important for characterizing the resolutions of human echolocator tongue clicks.



(a) EE1 (b) EE2 (c) EE3

Figure 4.15: Sample WAFs.

4.5 Artificial Click Synthesis

The main goal of this chapter is to produce a model for an artificial human echolocator tongue click that could be used by other researchers for their own modelling of human echolocator trials, and this goal assisted in determining which elements of the tongue clicks to investigate. The artificial click should capture the important features of the real clicks, and this information is contained within the parameters detailed in the previous sections. The frequency- and time-domain parameters were evaluated for each click individually, and were then combined to produce the parameters for a typical tongue click. The aggregated spatial-domain information was then added as an overall angle-dependent

amplitude modulation to the synthetic click.

The equation for the synthetic click was developed by combining Equation (4.5), Equation (4.1), and Equation (4.7) as given in Equation (4.10).

$$\begin{aligned} C(\theta,t) &= E(t)\,F(t)\,D(\theta) \\ &= a\exp(-bt-c)H(t-c)\sum_{m=1}^{M} A_m\cos(2\pi f_m t+\phi_m)\,10^{\wedge}\frac{-(1+\cos\theta)}{10\sqrt{\alpha^2\cos^2\theta+\beta^2\sin^2\theta}} \end{aligned} \tag{4.10}$$

Table 4.4: The parameters for the synthetic click of each echolocator.

	f (kHz)	A (a.u.)	ϕ (radians)	a (-)	b (t^{-1})	c (-)	α (dB)	β (dB)
EE1	3.52	6.50	1.59	6.70	1.56×10^3	2.03	0.130	0.282
	5.25	2.66	1.60					
	6.88	2.49	1.65					
	9.89	0.93	1.72					
	11.82	1.06	1.39					
EE2	2.20	8.40	1.46	2.23	1.05×10^3	1.97	0.101	0.185
	7.20	1.22	1.57					
	10.78	1.68	1.57					
	13.26	0.98	1.53					
EE3	3.67	5.21	1.59	6.57	1.56×10^3	2.03	0.963	0.104
	10.01	2.70	1.56					

To capture the click frequency content, the average frequency, amplitude, and phase of the peaks in the frequency-domain form of the tongue clicks were extracted from the database and are given in Table 4.6. Similarly, to capture the click directivity, the parameters of the directivity patterns were evaluated by fitting the model to the magnitude data extracted from the database. The time domain form of the clicks was determined by fitting the exponential decay model, Equation (4.7), to each click and extracting a distribution of parameters. The outliers of these distributions were excluded, and the median envelope parameters were evaluated. All the extracted parameters are given in Table 4.4 for each echolocator, and were used to generate the typical clicks for each echolocator. A summary of the parameters is given in Table 4.5.

The synthetic tongue clicks are shown in Figure 4.16, which shows that the tongue

Table 4.5: The parameters used in the synthetic click model and their meanings.

Parameter	**Meaning**	**Units**
f	Frequency of each monotone component.	kHz
A	Amplitude of each monotone component.	Arbitrary units $\propto$ volts
ϕ	Phase angle of each montone component.	radians
a	Scale factor of the exponential decay.	Unitless
b	Time constant of the decay	t^{-1}
c	Offset for the decay	Unitless
α	Semi-major width of elliptical modulation	dB
β	Semi-minor width of elliptical modulation	dB

clicks capture the key spectral and time domain features of the expert echolocator tongue clicks. A tool for producing these tongue clicks was published by the author as supplementary material to a journal paper on the topic [1]. For comparison, the time-domain form, PSD, and spectrogram of a sample click from EE1 are provided in Figure 4.17, which shows the similarity between the synthetic and recorded clicks.

Table 4.6: The mean peak frequencies and amplitudes for the monotones model.

	Dist (cm)	Frequency (kHz)		Amplitude (a.u.)		Phase (radians)
		Mean	Std Dev	Mean	Std Dev	Mean
EE1	40	3.52	0.16	6.50	1.08	1.59
		5.25	0.43	2.66	0.68	1.60
		6.88	0.48	2.49	0.74	1.65
		9.89	0.60	0.93	0.51	1.72
		11.82	0.47	1.06	0.55	1.39
	100	3.51	0.16	6.58	1.02	1.62
		5.20	0.45	2.69	0.67	1.56
		6.86	0.51	2.44	0.70	1.62
		9.84	0.60	0.88	0.42	1.78
		11.81	0.46	0.97	0.45	1.38
EE2	40	2.20	0.22	8.40	0.89	1.46
		7.20	0.37	1.22	0.67	1.57
		10.78	0.37	1.68	0.93	1.57
		13.26	0.53	0.98	0.55	1.53
	100	2.10	0.19	8.53	0.76	1.32
		7.18	0.39	1.18	0.72	1.78
		10.82	0.39	1.26	0.83	1.56
		13.30	0.55	0.79	0.43	1.61
EE3	40	3.67	0.44	5.21	1.40	1.59
		10.00	0.91	2.70	0.90	1.56
	100	3.82	0.51	4.75	1.21	1.59
		10.05	1.00	2.53	0.82	1.54

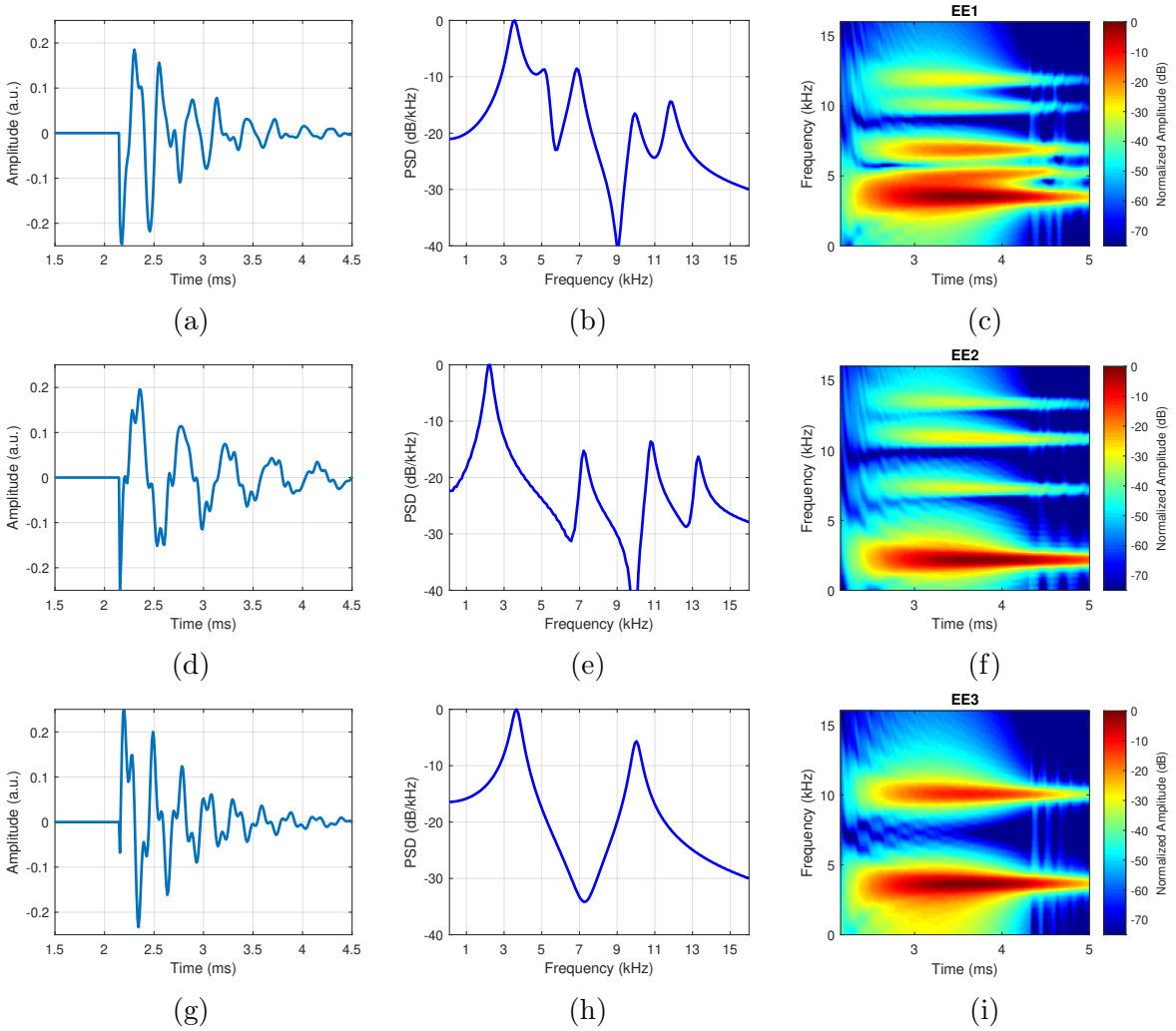


Figure 4.16: Synthetic tongue clicks representations in the time domain (left column), frequency domain (middle column), and time-frequency domain (right column). (a)-(c) Based on EE1. (d)-(f) Based on EE2. (g)-(i) Based on EE3.

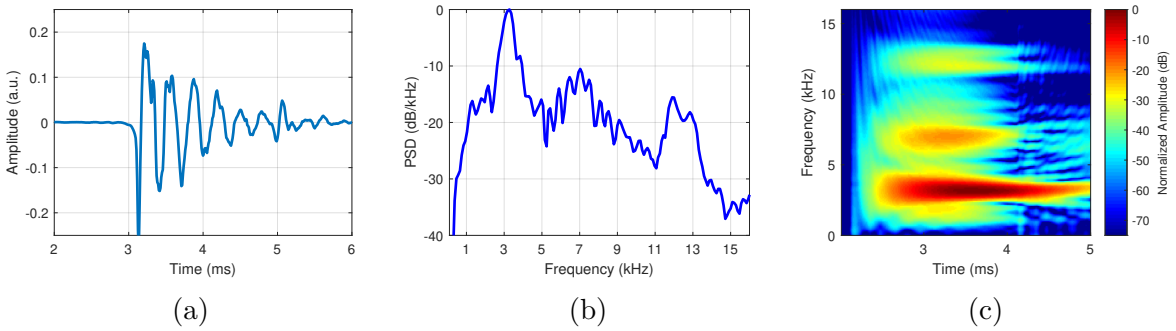


Figure 4.17: Sample click from EE1 for comparison with the synthetic tongue clicks.

4.6 Summary

The previous sections have provided details of how to extract the key parameters from human echolocator tongue clicks from the time-, frequency-, and spatial- domains. All of these differing characteristics have been incorporated into the synthetic tongue click model, allowing realistic rendering of human echolocator tongue clicks for broader subsequent use.

The frequency-components of human-echolocator tongue clicks are exclusively of constant frequency as a function of time. This property is contrasted with many bats that use both constant frequency waveforms and frequency modulated sweeps in their calls. The primary peak frequency for all echolocators was different, but fell into the 2-4 kHz range. Similarly to bats, human echolocator tongue clicks are comprised of several frequency components that span a broad frequency band, ranging from 2-13 kHz. For localization of sound sources, discussed in depth in Chapter 2, the bandwidth of the sound affects a human's ability to localize it, with broader band signals being easier to localize. The broad band of frequencies in an echolocation call is likely to be of importance in localizing the sources of echoes in a similar way. The use of broadband signals is of particular importance when considering the frequency-dependent characteristics of the HRTF, which are developed and utilized in the next chapter.

The time-domain profile of human echolocator tongue clicks is very sharp, and it is thought that the sharp initial onset is a key feature for successful echolocation and allows for resolution of objects in range. The duration of tongue-clicks was consistent between echolocators and is approximately 3 ms, which means that objects separated by 0.9 m in range would result in fully-separated echoes at the ears. Because of the sharp onset and exponential decay of the tongue clicks, it is unlikely that human echolocators require echoes to be fully separated in range in order to be able to make a determination about their origins. Consequently it is reasonable to suggest that the range resolution of human echolocators is better than 0.9 m, an assertion supported by the literature [80, 87]. The click durations used by the echolocators are not assumed to be optimal, but are likely

the shortest duration sounds with a sharp onset that can be easily made with the mouth. Because human echolocators operate with acoustic signals with a much lower speed of propagation than light, the observations about click duration and envelope are of less utility for developing a biologically-inspired radar system.

The spatial-behaviour of the tongue clicks was found to be slightly more directional than speech, and the main source of attenuation was the shadowing caused by the bulk of the head. Higher-frequency components were found to be slightly more directive than lower frequency components, and this is speculated to be due to the shape of the mouth when the click is produced. Despite these observations, the clicks possess a very broad beam, and illuminate a wide area in front of the echolocator. This broad illumination is suggested to be useful for achieving a wide perception of an echolocator's environment. Other sources [88, 86, 87] have reported motion of the head by human echolocators when performing experimental trials. In combination with the observation of a broad click beamwidth, it seems that this motion may be less about directing the boresight of the tongue-click and more about changing the pointing direction of the ears, which have a much more directional receiving pattern [27, 136, 137].

The tongue clicks were found to be very good for resolving closely separated targets in range using a correlation process, but were also found to be ambiguous in the Doppler domain. The average range resolution of the echolocator clicks, estimated using the WAF, was between 30 cm and 60 cm, which is in agreement with the range resolution reported in previously conducted experiments [80, 87]. It seems unlikely that humans process Doppler information to determine velocity directly, and instead are likely to compare range measurements on a pulse-to-pulse basis for determining the velocity of objects. It is also possible that expert human echolocators have developed neural structures similar to the delay-tuned neurons found in other echolocating mammals [58, 59] as this mechanism would lead to similar resolution performance as the correlation process.

4.7 Conclusion

This chapter has explored the properties of human echolocator tongue clicks, has presented novel research into those properties, and has used these properties to develop a set of synthetic tongue clicks. The broad angular distribution of the clicks mean that the area in front of the echolocator is approximately uniformly illuminated, and echoes can therefore originate from any angle in the forward direction. The wide bandwidth of the echolocation clicks means that the echolocator receives wideband spectral echoes from the environment, which introduces the possibility of echolocators using spectral cues to localize echoes in angle.

The next chapter will move away from the acoustic domain and into the electromagnetic domain to develop a radar system. This radar system will take inspiration from human echolocation and will embed some of the tongue click properties into the system design. Using this biologically-inspired radar system, the next chapter will explore the challenges of localizing echoes in angle using only the spectral information that they contain.

Chapter 5

Frequency Diverse Target Localization

Jeder hält das Ende seines Gesichtskreises für das der Welt.

Every man takes the limits of his own field of vision for the limits of the world.

— Arthur Schopenhauer

Parerga und Paralipomena, 1851

5.1 Introduction

The techniques used to localize the positions of objects have been of great interest to the seemingly disparate fields of psychoacoustics and radar for many years; areas of psychoacoustics seek to explain how animals localize sounds sources and carry out echolocation, and many areas of radar development seek to design systems and techniques to localize targets (amongst other functions) to the best of their ability. The problem of localizing targets in range, azimuth, elevation, and Doppler has been very well treated in radar; with powerful techniques for array processing and beamforming, it is possible to believe that there is nothing new to add to conventional radar localization. However, looking to sound localization in nature (see Chapter 2) we see that both the inter-aural level difference (ILD) and inter-aural time difference (ITD) have direct analogues in radar with amplitude monopulse techniques and time difference of arrival (TDOA) respectively, but the binaural timbre difference (BTD) has not previously had a radar dual. This chapter seeks to investigate the BTD and use it to develop novel frequency-diverse techniques for target localization using radar. As with natural biological systems which may use many cues simultaneously, these techniques are not intended to be treated as standalone approaches and should be considered as ways of accessing additional information which may complement existing techniques.

Many echolocating species including bats and dolphins have been shown to use wide bandwidth echolocation calls. Chapter 4 explored the spectral properties of human echolocator tongue clicks and found that human echolocators also use wide bandwidth echolocation clicks. This wide bandwidth introduces frequency-diversity into the clicks, and can be used by human echolocators to improve their perception of their environment. The radar system used in this chapter will mimic natural echolocation approaches by implementing a wide bandwidth signal.

As introduced in Chapter 2, the binaural configuration of the head is important for sound localization in the natural world. Having two ears enables humans and other animals to access the ITD, ILD, and BTD. As these cues rely on comparing the signal

received at the two ears, any radar system hoping to investigate equivalent cues in the electro-magnetic (EM) domain must have two receiving antennas and be able to compare the signals between them.

This chapter will develop amplitude- and phase-based radar techniques analogous to sound localization cues found in nature and will present the underlying theory behind these techniques; the simulations and experiments used to investigate them; the results of these investigations; and an analysis based on discussion of these results.

Sections of this chapter are taken verbatim from novel works written entirely by the author [5, 4, 3].

5.1.1 The Head-Related Transfer Function and the BTD

The Head-related transfer function (HRTF) is a function often used in the field of psychoacoustics [27, 140, 26, 35] [141, pp.165-168] and describes the acoustic filtering process that aids humans (and other creatures) in localizing the sources of sounds. An explanation and example of the HRTF is included in Chapter 2.

Without a reference for the original spectral content of a sound, the spatial encoding of sound signals by the HRTF is not sufficient for localization of sound sources as it is impossible to know exactly which elements have been attenuated and which have not. Monaural localization experiments show a localization performance better than chance [34, 33, 142, 143] (but still worse than binaural localization performance) and suggest that there is a learning process which allows humans to learn the spectral patterns of common sounds. Having a reference for common sounds allows us to determine their sources of origin when we encounter them. The addition of learning behaviours to this system would be extremely interesting, but is outside the scope of this chapter.

However, the BTD cue is still relevant: by considering the *difference* in the frequency-content of the sounds received between two ears, rather than the absolute information received at either, the requirement for information about the original signal is removed which enables localization of the sound source.

An equivalent cue can be accessed using a radar system by considering the difference in signals received at two antennas. In the instance of a radar system, it is not the original frequency content of the transmission which is unknown, but instead the modulation applied to it by the frequency-dependent radar cross-section (RCS) of the target.

5.2 Theory

Before developing a mathematical description of the biologically-inspired radar techniques, it is useful to define the elements of the localization problem, which will also be used as the geometry for the experimental work presented later in this chapter.

5.2.1 Geometric Model

Figure 5.1 shows schematically the relative locations of the target; transmitter, Tx, with phase centre located at the origin; and a pair of receivers, R_1 and R_2, located such that all three antennas are collinear. The receiving antennas are separated by a baseline, d. For a target located at a point, U, the time taken for a signal to reach each antenna can be expressed as in Equation (5.1), where c is the speed of light in a vacuum and the subscript i can take a value of either 1 or 2 to denote the relevant receiving antenna.

$$t_i = \frac{r_i + r_{\text{tx}}}{c} \tag{5.1}$$

5.2.2 Signal Description

The signal used for localization given in Equation (5.2) is a linearly up-chirped signal of duration T, chosen because it is a standard radar waveform, distributes energy equally across all frequency components of the band, and can be experimentally implemented

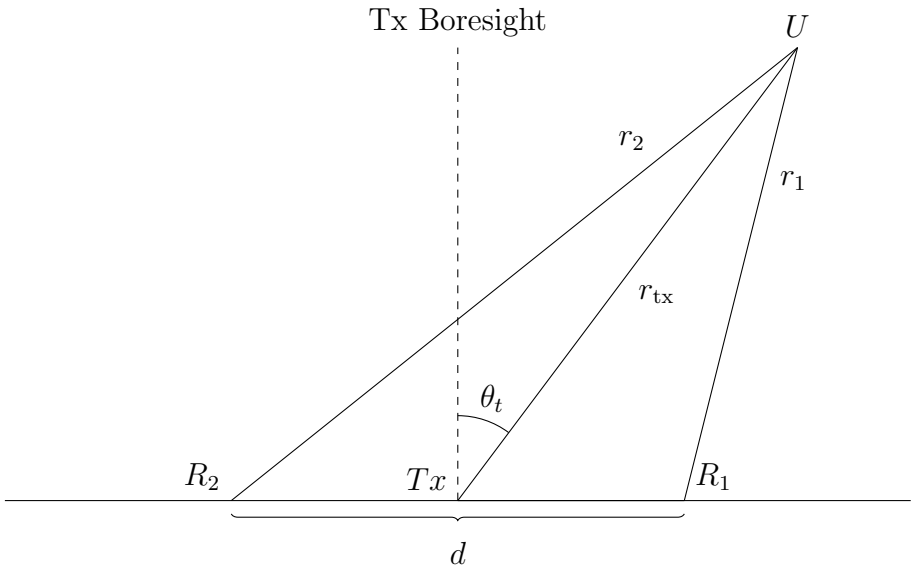


Figure 5.1: System geometry for two receivers and a single transmitter in a binaural configuration.

with a large bandwidth using a vector network analyzer (VNA).

$$s_{tx} \propto \exp\left(j\pi K t^2\right) \exp\left(j2\pi f_c t\right) \Pi\left(t, T\right) \tag{5.2}$$

Where f_c is the up-chirp starting frequency, K is the chirp rate (Hz/s), and $\Pi\left(t, T\right)$ is a boxcar function defined in Equation (5.3).

$$\Pi(t, T) = \begin{cases} 1 & \text{if } 0 \leq t \leq T \\ 0 & \text{otherwise.} \end{cases} \tag{5.3}$$

Recalling Equation (5.1) for the round-trip time taken by the signal and assuming a delta function target response allows us write Equation (5.4) for the received signal, which incorporates a time delay and a phase shift into the signal due to the range of the target.

$$s_i^2 = P_{Ri} \exp\left(j\pi K (t - t_i)^2\right) \exp\left(j2\pi f_c (t - t_i)\right) \Pi\left(t - t_i, T\right) \tag{5.4}$$

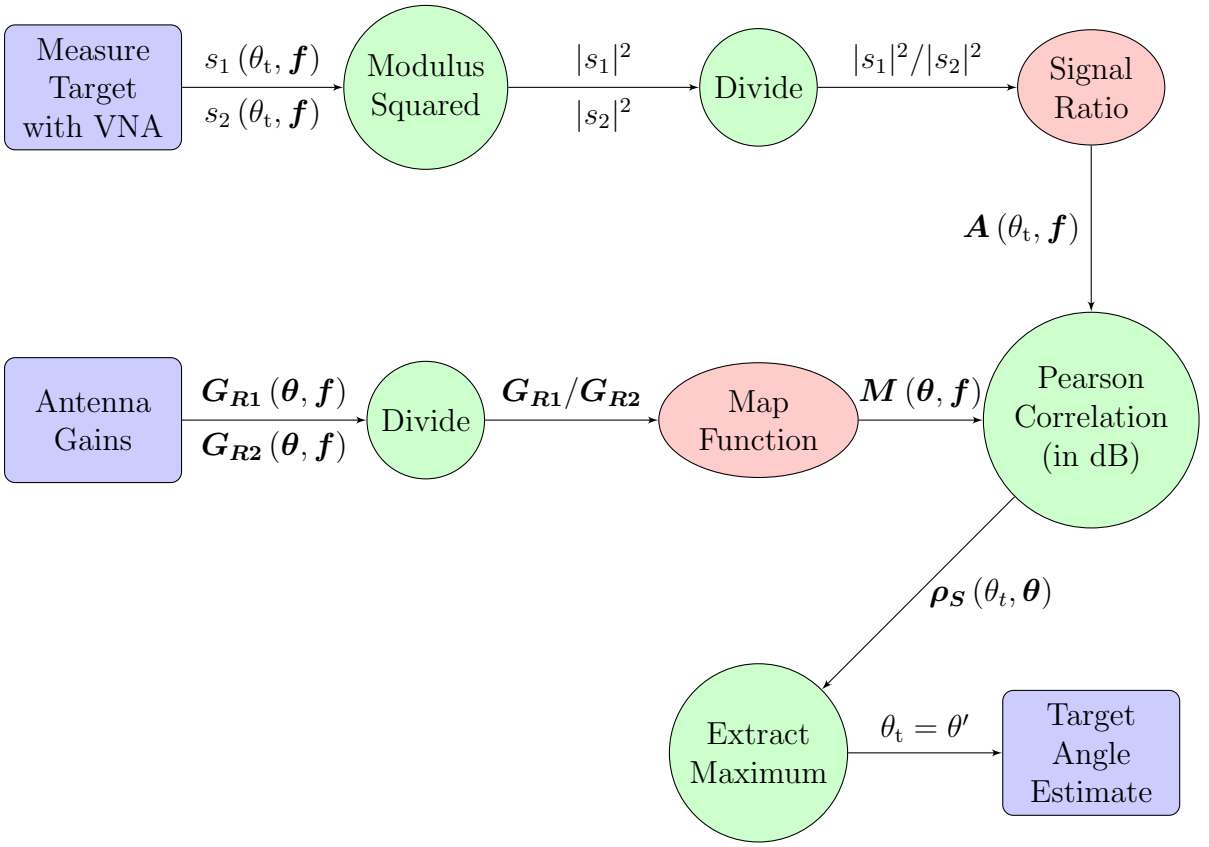


Figure 5.2: Flowchart to show the processing chain of the power-based approach.

5.2.3 Signal Power Ratio Approach

By taking inspiration from the ILD and the BTD (detailed in Chapter 2), we can formulate a power-based angular localization technique [5]. Beginning with the signal description from the previous section, the radar equation for the power at the output of the receiving antenna is given in Equation (5.5), with the relevant parameters described in Table 5.1. More detail of this formulation of the radar equation can be found in Chapter 3. The following signal processing is summarized in Figure 5.2.

$$P_{Ri} = \frac{P_{Tx} G_{Tx}(\theta_t, \boldsymbol{f}) G_{Ri}(\theta_t \pm \theta_0, \boldsymbol{f}) c^2 \sigma(\theta_t, \boldsymbol{f})}{(4\pi)^3 (r_{tx} + r_i)^4 \boldsymbol{f}^2 L(\theta_t)} \tag{5.5}$$

For a single measurement, there are several parameters that can vary with frequency (including the target RCS and the attenuation over the propagation space), but provided that the antenna baseline, d, is sufficiently small ($d << r_{tx}$), the difference in these

Table 5.1: Reference for mathematical symbols.

Symbol	Definition	Units
θ_{t}	target angle from transmitter boresight	radians
θ_0	receiver angle from transmitter boresight	radians
f	frequency	Hertz
P_{Ri}	signal power received at the i^{th} receiving antenna	Watts
P_{Tx}	power transmitted to the transmitting antenna	Watts
G_{Tx}	gain of the transmitting antenna	-
G_{Ri}	gain of the receiving antenna	-
c	speed of light in a vacuum	m s^{-1}
σ	radar cross section (RCS) of the target	m^2
L	losses	-

The subscript i is used throughout and can take a value of either 1 or 2 to denote the relevant receiving antenna.

terms between the two receiving antennas is also sufficiently small and can be taken as negligible. Considering the ratio of received signal powers, $|s_1|^2$ and $|s_2|^2$, between R_1 and R_2 respectively yields Equation (5.6), which gives the result that the signal ratio is independent of target range or reflectivity. This is useful because it means that the ratio of signals is independent of target scattering, and is purely a function of measurable hardware parameters.

The theory presented here is given in the absence of noise, which introduces dependence on target range and reflectivity, however the variation is on the order of the noise power and is treated as additive white Gaussian noise (AWGN) for the purpose of simplifying the analysis. This assumption will later be seen to be sufficiently valid to achieve useful results.

$$\begin{aligned}\frac{|s_1|^2}{|s_2|^2} =& \frac{\boldsymbol{P}_{R1}\left(\theta_{\mathrm{t}}+\theta_0, \boldsymbol{f}\right)}{\boldsymbol{P}_{R2}\left(\theta_{\mathrm{t}}-\theta_0, \boldsymbol{f}\right)} = \frac{\boldsymbol{G_{R1}}\left(\theta_{\mathrm{t}}+\theta_0, \boldsymbol{f}\right)}{\boldsymbol{G_{R2}}\left(\theta_{\mathrm{t}}-\theta_0, \boldsymbol{f}\right)} \\ =& \boldsymbol{A}\left(\theta_{\mathrm{t}}, \boldsymbol{f}\right)\end{aligned} \tag{5.6}$$

Therefore, by having prior information about the ratio of receiver gains across all angles of interest and all frequencies in the band, it is possible to build up a map function

which describes the expected result of a measurement in the presence of a target. This map function depends only on the angle to the target and the antenna patterns which are a known system characteristic, and is given in Equation (5.7), where $\boldsymbol{\theta}$ represents a set of all possible angles to a target.

$$\frac{\boldsymbol{G_{R1}}\left(\boldsymbol{\theta}+\theta_0,\boldsymbol{f}\right)}{\boldsymbol{G_{R2}}\left(\boldsymbol{\theta}-\theta_0,\boldsymbol{f}\right)}=\boldsymbol{M}\left(\boldsymbol{\theta},\boldsymbol{f}\right) \tag{5.7}$$

In this approach the signal ratio is the cue (and, as with its acoustic counterpart, requires some rudimentary processing) and the map function represents the prior information held by the system (analogous to information held in memory). What is required after a measurement is made is some way of relating the measured signal ratio to the prior information held by the system. To do this, the Pearson correlation coefficient is calculated between the signal ratio and the frequency profile across each angle in the map function as formulated in Equation (5.8). The Pearson correlation coefficient was chosen because it is a mathematically straightforward correlation measure which is able to discriminate between positive and negatively correlated data. Where $\sigma_{\boldsymbol{A}}$ and $\sigma_{\boldsymbol{M}}$ are the standard deviations of the signal ratio, $\boldsymbol{A}\left(\theta_{\mathrm{t}},\boldsymbol{f}\right)$, and the map function, $\boldsymbol{M}\left(\boldsymbol{\theta},\boldsymbol{f}\right)$, respectively. θ' represents a single angle from the set $\boldsymbol{\theta}$.

$$\boldsymbol{\rho}_A\left(\theta_t,\theta'\right)=\frac{\mathrm{cov}\left(\boldsymbol{A}\left(\theta_{\mathrm{t}},\boldsymbol{f}\right),\boldsymbol{M}\left(\theta',\boldsymbol{f}\right)\right)}{\sigma_{\boldsymbol{A}}\sigma_{\boldsymbol{M}}} \tag{5.8}$$

The Pearson correlation coefficients represents the degree of similarity between the measured signal ratio and the expected profile at each candidate angle. By extracting the peak from this likelihood profile, the angle to the target can be estimated.

5.2.4 Phase Difference Approach

This technique uses the same algorithmic structure as the power-based approach presented in Section 5.2.3, but instead relies on the phase of the received signals. Similaraly, a summary of the signal processing is given in Figure 5.3. A phase-only form of the radar

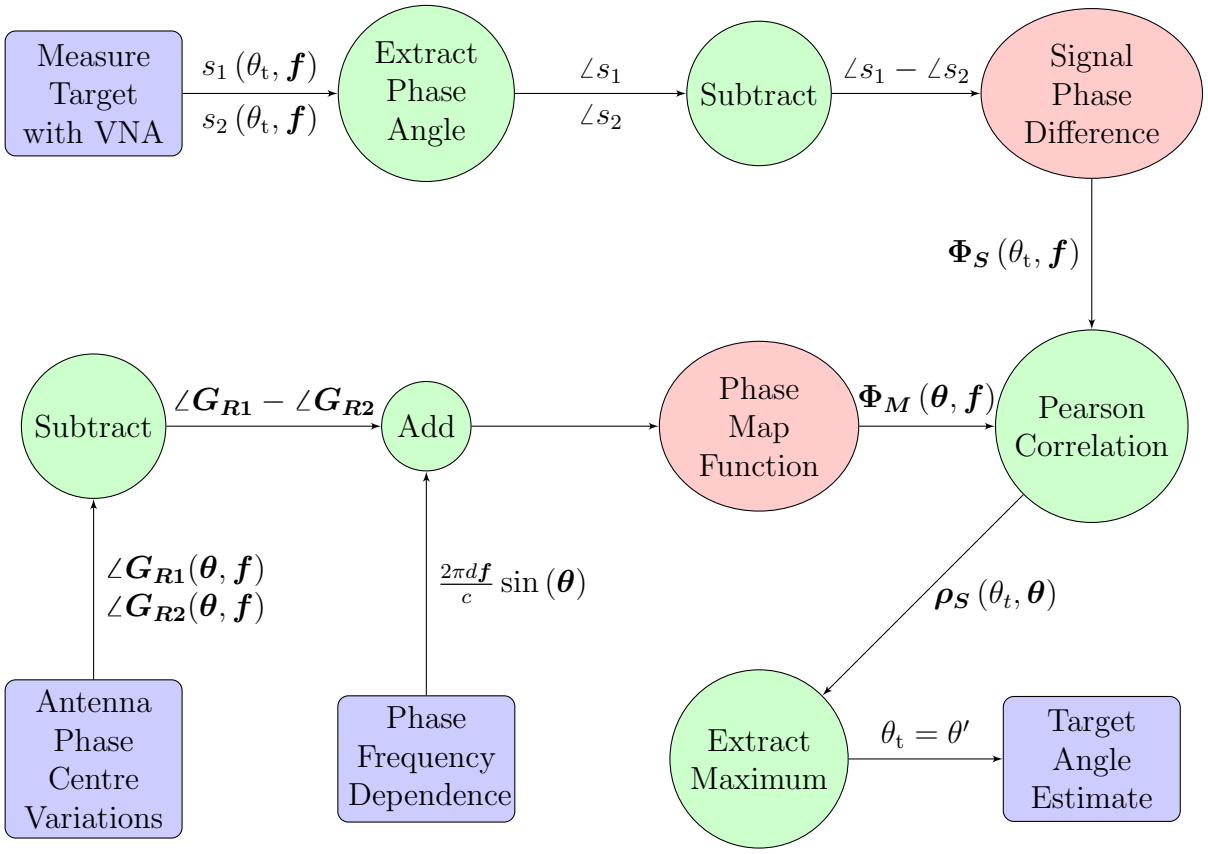


Figure 5.3: Flowchart to show the processing chain of the phase-based approach.

equation is shown in Equation (5.9), where $\boldsymbol{\phi}_{Ri}$ is the phase of the signal received at antenna i, and $\boldsymbol{\phi}_{Tx}$ is the phase of the transmitted signal. More detail of this phase-only formulation of the radar equation is presented in Chapter 3.

$$\begin{aligned} \boldsymbol{\phi}_{Ri} = \boldsymbol{\phi}_{Tx} + \angle \boldsymbol{G_{Tx}}(\theta_t, \boldsymbol{f}) + \angle \boldsymbol{G_{Ri}}(\theta_t \pm \theta_0, \boldsymbol{f}) \\ + \angle \sigma(\theta_t, \boldsymbol{f}) + \frac{4\pi (r_{tx} + r_i) \boldsymbol{f}}{c} \end{aligned} \tag{5.9}$$

The measured signal phase difference between the two receiving antennas is evaluated in Equation (5.10), which shows that target-dependent phase effects cancel-out in a similar

way to the target-dependent amplitude effects in Equation (5.6).

$$\begin{aligned}\boldsymbol{\Phi_S}\left(\theta_{\mathrm{t}}, \boldsymbol{f}\right) =& \boldsymbol{\phi}_{R1} - \boldsymbol{\phi}_{R2} \\ =& \angle \boldsymbol{G_{R1}}(\theta_{\mathrm{t}} + \theta_0, \boldsymbol{f}) - \angle \boldsymbol{G_{R2}}(\theta_{\mathrm{t}} - \theta_0, \boldsymbol{f}) \\ &+ \frac{4\pi r_1 \boldsymbol{f}}{c} - \frac{4\pi r_2 \boldsymbol{f}}{c}\end{aligned} \tag{5.10}$$

This formulation means that the measured signal is a function of prior information, the angle to the target, and the bistatic ranges to the target. However, the range dependence can be approximated as $d\sin(\theta)$ provided that the condition $d << r_{tx}$ is satisfied. This condition is straightforward to satisfy as the receiving antennas may be moved arbitrarily close together, in the extreme case causing the bistatic range terms to be equal and vanish. The approximation for the bistatic range terms replaces the range dependence and hence a phase map function (purely a function of angle and frequency) can be formulated as in Equation (5.11). The phase map function gives the expected phase difference between the antennas for a target at a given angle and across all transmitted frequencies.

$$\begin{aligned}\boldsymbol{\Phi_M}\left(\boldsymbol{\theta}, \boldsymbol{f}\right) =& \angle \boldsymbol{G_{R1}}(\theta_{\mathrm{t}} + \theta_0, \boldsymbol{f}) - \angle \boldsymbol{G_{R2}}(\theta_{\mathrm{t}} - \theta_0, \boldsymbol{f}) \\ &+ \frac{2\pi d \boldsymbol{f}}{c} \sin\left(\theta_{\mathrm{t}}\right)\end{aligned} \tag{5.11}$$

$$\boldsymbol{\rho_S}\left(\theta_t, \theta'\right) = \frac{\mathrm{cov}\left(\boldsymbol{\Phi_S}\left(\theta_{\mathrm{t}}, \boldsymbol{f}\right), \boldsymbol{\Phi_M}\left(\theta', \boldsymbol{f}\right)\right)}{\sigma_{\boldsymbol{\Phi_S}} \sigma_{\boldsymbol{\Phi_M}}} \tag{5.12}$$

In this approach, the signal phase difference is the utilized cue and the phase map function contains the prior information about the system characteristics. It is then possible to compare the measured signal phase difference with each frequency profile in the phase map function again using the Pearson correlation coefficient. The Pearson correlation coefficient for the phase comparison cue is given in Equation (5.12), where $\sigma_{\boldsymbol{\Phi_S}}$ and $\sigma_{\boldsymbol{\Phi_M}}$ refer to the standard deviations of the measured signal phase difference, $\boldsymbol{\Phi_S}$, and the

phase map function, $\boldsymbol{\Phi_M}$, respectively. The highest (closest to 1) value of the coefficient represents the angle at which the target is most likely to be located.

5.2.5 On-Boresight Detection

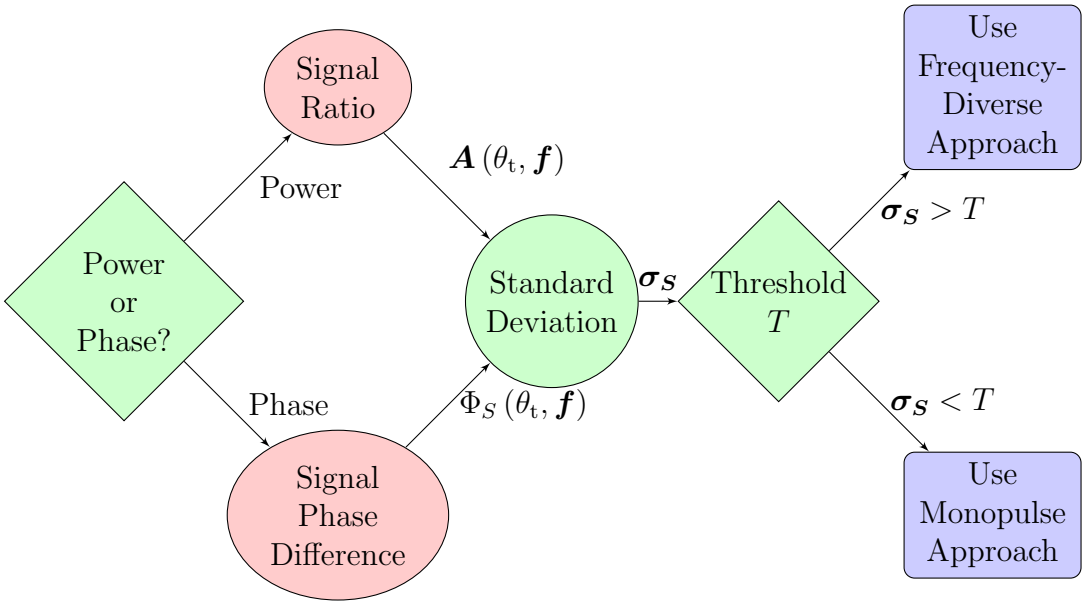


Figure 5.4: Flowchart to show the processing chain of the on-boresight detection technique.

The techniques presented in the previous sections rely on the Pearson correlation coefficient to estimate the angle to the target. This method can be confidently predicted to be deficient in the boresight direction because the Pearson correlation coefficient is independent of the powers of the functions being correlated. In the boresight direction, the difference between the measured signals (in both power and phase approaches) is zero across all frequencies, and so the effect of measurement noise is amplified in the correlation process, resulting in systematic error in the localization process. To compensate for this deficiency, a technique is introduced whereby the system identifies that it has a target in the region of poor performance, and switches to a traditional phase-monopulse approach which performs well in the boresight direction. This technique is summarized in Figure 5.4.

To identify if the system is in the region of poorest performance, a measure of angular variation is used because the signal difference is approximately zero in the boresight

direction. The angular variation was captured by using the standard deviation of the measured frequency profile, given in Equation (5.13) where $\boldsymbol{S}(\theta_{\mathrm{t}}, \boldsymbol{f})$ is either the power-based signal ratio, $\boldsymbol{A}(\theta_{\mathrm{t}}, \boldsymbol{f})$, or the phase difference, $\Phi_S(\theta_{\mathrm{t}}, \boldsymbol{f})$, depending on the technique being used. $\bar{S}$ is the mean value of $\boldsymbol{S}$ across all measured frequencies, N_f is the number of frequency components in $\boldsymbol{S}$, and f_i is the i^{th} component of the frequency vector $\boldsymbol{f}$.

$$\boldsymbol{\sigma_S} = \sqrt{\frac{\sum_{i=1}^{N_f} \left(S(\theta_{\mathrm{t}}, f_i) - \bar{S}\right)^2}{N_f}} \tag{5.13}$$

If this metric falls below a pre-defined threshold, a traditional phase-comparison monopulse technique is used. The phase-comparison monopulse technique used [144] is shown in Equation (5.14), where f_m is the single frequency chosen for monopulse.

$$\theta_{\mathrm{t}} = \arcsin\left(\frac{c\Phi_S(f_m)}{2\pi d\boldsymbol{f}}\right) \tag{5.14}$$

5.2.6 Summary of Theory

The theory presented in the above sections has developed two frequency-diverse localization techniques, one reliant on amplitude and the other on phase, that are inspired by the direction-dependent filtering of the HRTF in the biological world. In both instances the antennas introduce variations (in amplitude or phase) to the signal at the output of the antenna that are dependent on the incidence angle of the target. By considering a radar equivalent of a binaural cue, the target-dependent effects are negated, and the remaining signal is purely a function of the difference between the antenna patterns. It is worth noting that, while a large bandwidth is implied, it is not necessary for the functioning of this technique, provided that the variation introduced by the antennas is sufficiently complex across the range of frequencies in the band. By augmenting the bio-inspired techniques with traditional phase monopulse techniques, the performance of the overall system is expected to be greater than the performance of either technique in isolation.

Table 5.2: Experiment parameters and their values.

Parameter	Value	Units
θ_0	24	degrees
P_{Tx}	5	dBm
f_{step}	100	MHz
f_{start}	2	GHz
f_{stop}	6	GHz
θ_{step}	0.5	degrees
θ_{start}	-90	degrees
θ_{stop}	6	degrees

5.3 Experiment

Having developed the theory for two biologically-inspired frequency-diverse localization techniques, it is now useful to test these theories and to quantify the performance of each technique. An overview of the parameters used in the experiment is given in Table 5.2.

5.3.1 Measurement Method

In order to mimic the binaural hearing configuration of echolocators two spatially-separated identical receiving antennas were used, as presented in Section 5.2.1. Because of the use of a binaural receiver configuration, the HRTF based-cue discussed in Section 5.1.1 may be utilized by the radar system. To complete the biological mimicry a third identical antenna was used exclusively for the transmit signal, mimicking the central placement of the mouth, the origin of echolocator clicks. The transmit antenna pattern was not a factor in the signal processing due to its absence in Equation (5.6) and Equation (5.10), but it did affect the signal-to-noise ratio (SNR) of the received signals.

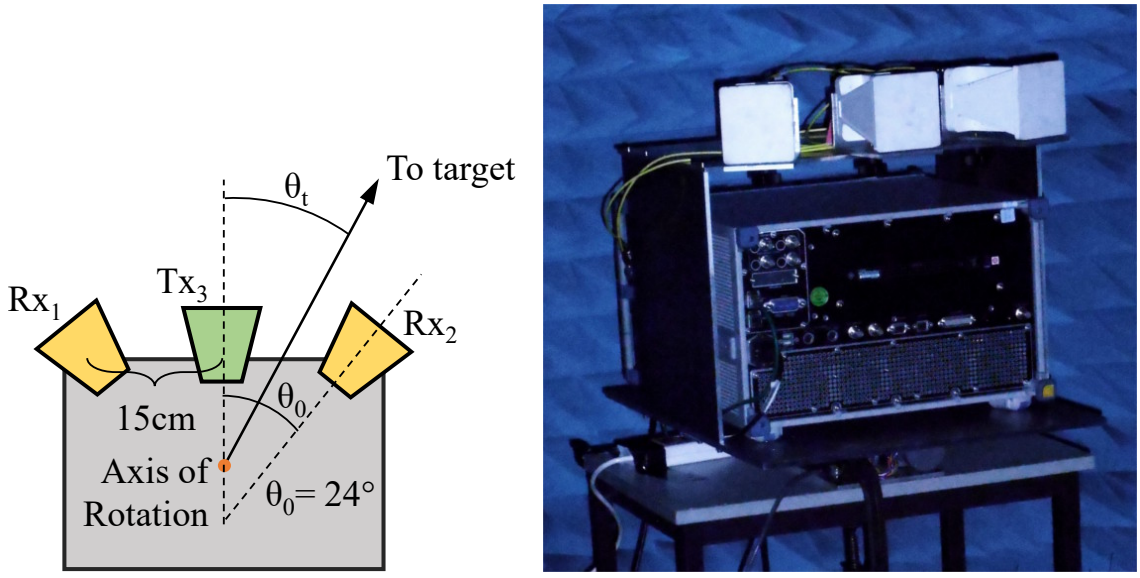
The antennas were mounted in a configuration broadly mimicking nature and consistent with the geometry introduced in Figure 5.1, as shown in Figure 5.5a. The assemblage was placed together with a target in an anechoic chamber so that the far-field condition was satisfied which is discussed in more detail in Section 5.3.2. Due to the limited space

available in the chamber, the target could not be moved through the required range of azimuth angles. Instead the antennas were mounted on a rotation table such that, relative to the transmit antenna boresight direction, measurements could be made over the desired range of angles to the target. The target, a single mirrored sphere of 36 cm diameter, was placed on a plinth to raise it into the same plane as the antennas, at a distance of approximately 3 m, placing it in the far field region for the antennas (see Section 5.3.2 for more detail).

The physical size of the antennas places a lower limit on their separation. When this restriction is combined with the maximum target range in the anechoic chamber (approximately 3 m), the assumption that $d << r_{tx}$ is no longer valid and introduces an error term to Equation (5.11). To mitigate this, the antenna calibration was performed for this range, in order to cancel out the error. For this technique to be viable at short ranges the antenna separation should be minimal, as the target range increases the geometric errors are expected to reduce.

A VNA was used to generate the require band of frequencies. The VNA was placed on the rotation table below the antennas as shown in Figure 5.5b to maintain phase coherence (by avoiding flexing of cables). The VNA used was a 4-port Rohde & Schwarz ZVA-67 which has a very low noise floor (at approximately -120 dBm). The rotation table used was a Parker 200RT which has a positioning accuracy of approximately 0.03^o and was therefore suitable to make measurements at 0.5^o intervals across a range of -90^o to $+90^o$ to the target.

The measurements made consisted of s_1 and s_2 measured across a frequency band of 2-6 GHz with a frequency step of 10 MHz. This band of frequencies was selected based on the characteristics of the antenna presented in the following section. In order to minimise the clutter response of the chamber, an initial background measurement of the environment (across -90^o to $+90^o$ and across the 2-6 GHz band) was made in the absence of the target and was subtracted from all subsequent measurements made in the presence of a target. Figure 5.6 shows the measurements of the chamber with and without the



(a) (b)

Figure 5.5: (a) Schematic of the biologically-inspired radar configuration including key angles. (b) Photograph of the vector network analyzer (ZVA-67) and three wideband horn antennas in the anechoic chamber.

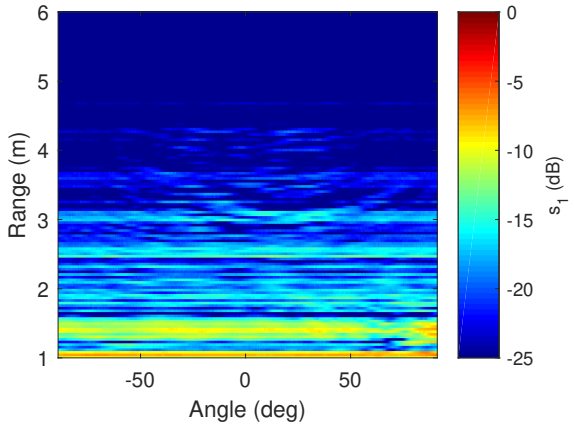
target, as well as the background-subtracted measurement, in which the target is clearly visible.

5.3.2 Antenna Characterization

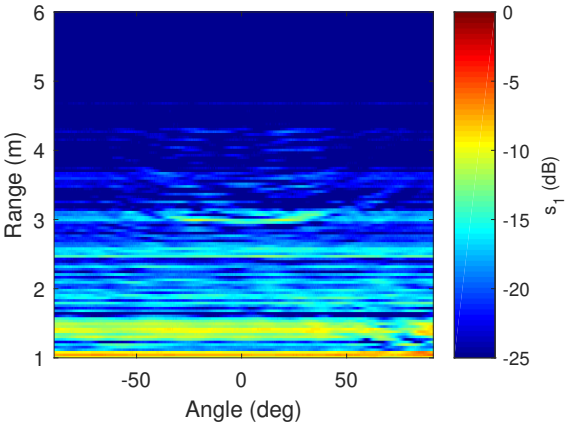
Horn antennas from Q-par Angus (WBH1-18)[112] which satisfied the requirements of a wide beamwidth and a wide operational bandwidth were used; Figure 5.7a shows how the antenna beamwidth varies across the 2-6 GHz frequency band, presenting a broad range of beamwidths from approximately 120^o at 2 GHz to 60^o at 6 GHz.

It was important to ensure that the measurements were made in the far-field region of the antenna, to prevent reactive effects at shorter distances. Equation (5.15) gives the formula used for the far-field distance as a function of wavelength, λ, and antenna aperture size, D [105, p.229]. As the antenna aperture was 96 mm [112], the far-field began at 0.1 m for 2 GHz and at 0.6 m for a transmit frequency of 10 GHz.

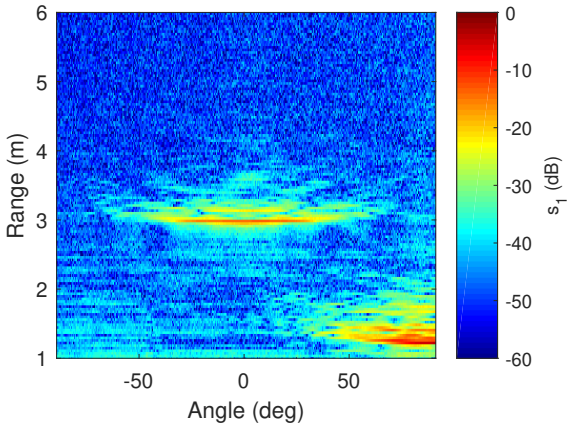
$$R_f = \frac{2D^2}{\lambda} \tag{5.15}$$



(a) Background measurement



(b) Target measurement



(c) Target measurement with Background removed

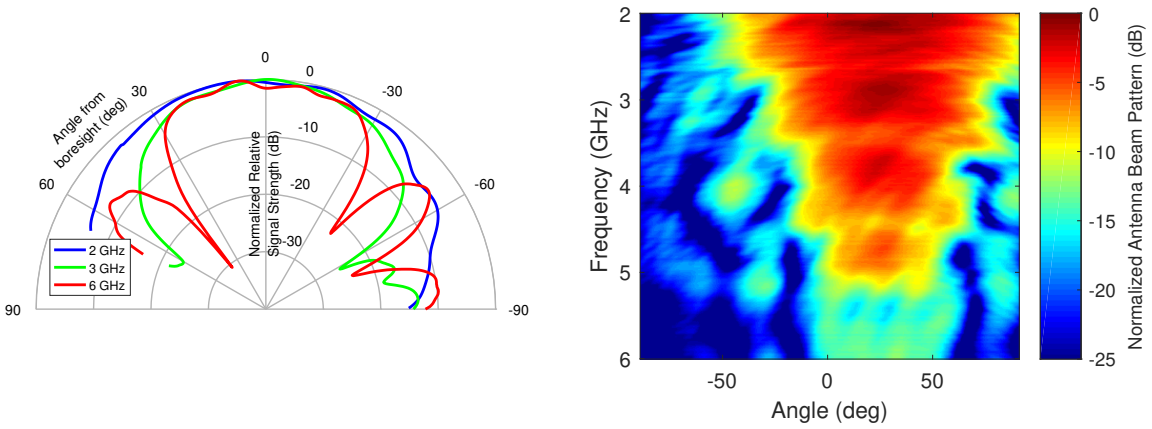
Figure 5.6: Range-azimuth maps of the anechoic chamber. (a) The chamber without the target, this is the background measurement which is subtracted from other measurements. (b) The chamber in the presence of a target, very slight differences can be seen compared to the background only measurement. (c) The chamber in the presence of a target with the background removed. Here, clutter is visible at 1-2 m between 50^o and 90^o which is attributed to the position of the anechoic chamber door being slightly different between measurements.

The wideband horn antenna was chosen, not because of any optimum beam pattern or beam pattern variation over the band of frequencies of interest, but because it was a readily-available off-the-shelf antenna with acceptable gain across the wide frequency band. There are likely to be many ways to optimize the design of an antenna to better encode space with spectral features for use with this technique. However, antenna optimization is beyond the scope of this chapter which seeks to present the theory of this technique alongside proof-of-concept results to demonstrate its viability.

In order to use the techniques described in the previous section it was essential to have a detailed characterization of the antenna beam patterns at a range of angles and over a range of frequencies. To measure the antenna patterns, the configuration in the previous subsection was modified within the anechoic chamber; a single antenna was removed from its mounting point and was placed on a plinth approximately 3 m away facing towards the VNA and the remaining antennas. As with the measurement technique described above, the VNA was rotated on the turntable through 180^o in steps of 0.5^o and the direct signal from the antenna on the plinth was measured. As the plinth antenna did not rotate, this technique was able to describe the variation in the antenna gain as a function of angle (in the azimuthal plane) and frequency as shown in Figure 5.7b. This characterization of the antenna gain was repeated over several experiment rounds and was found to be consistent between calibration measurements.

5.3.3 Signal Processing

The received signals, s_1 and s_2, were measured in the frequency domain, and were transformed into the time domain using the fast Fourier transform (FFT). In the time domain, direct-path signals (signals received without reflection directly from the transmitter) and reflections from the experimental apparatus corrupt the measured signal. A range-azimuth map in the presence of the target (an with the background removed) is shown in Figure 5.6c which shows the effect of clutter in the environment. Some clutter is not removed by subtracting the background because of small variations in the scene, such as



(a) (b)

Figure 5.7: Antenna patterns of the Q-par Angus (WBH1-18)[112] wideband horn antennas. (a) Three sample beam patterns, measured across a 2-6 GHz band, showing 3 dB beamwidths of approximately 120^o at 2 GHz and 60^o at 6 GHz. (b) Wideband characterization surface measured across a 2-6 GHz band, showing the continuous variation of the antenna pattern as a function of frequency and azimuth angle. This peak of the surface narrows as the frequency increases, consistent with a reduction in beam width expected at higher frequencies.

the door of the chamber being positioned slightly differently. In this range-azimuth map, the target is clearly visible (above -20 dB) over an angular region of approximately 60^o and is slightly visible (above -40 dB) over an angular range of approximately 120^o which is in agreement with the antenna pattern presented in Figure 5.7a which has a beamwidth of approximately 60^o at 6 GHz and 120^o at 2 GHz. A range-azimuth map of the chamber without a target is shown in Figure 5.6a, which indicates that clutter and noise are present in the signals. The target was then windowed within 25 range bins (0.935 m) in order to reduce the effects of the clutter and noise. The windowed region was extracted and zero-padded before transforming back into the frequency domain through the use of an inverse fast Fourier transform (IFFT). The resulting frequency-domain signal had 292 sampling points.

The measurements of the antenna patterns were made by placing an antenna on the plinth at the far end of the chamber, pointed directly at the receiving antennas. As a result of this configuration, the directly received signal dominates the return and clutter is not significant. Before correlating with the measured signal profile, the map function

was down-sampled to match the dimensions of the up-sampled measured signal.

5.4 Results and Discussion

5.4.1 Signal Power Ratio

By taking the ratio of the powers of the direct signals as described in Equation (5.7), the power-based map function was estimated and is shown in Figure 5.8a.

5.4.1.1 Ideal Behaviour

Because the power-based map function in Figure 5.8a has a high SNR, it represents the ideal result of measurement of a target and by using this property it is possible to estimate the ideal correlation surface for the localization technique. The correlation surface is a visual representation of the Pearson correlation coefficients between two sets of measurements. If the two sets of measurements used are the same, the self-correlation surface will represent the mutual information contained between each pair of measurement angles. This self-correlation surface was generated and is shown in Figure 5.8b, which shows a correlation of 1 along the leading diagonal (which corresponds to an autocorrelation at each angle). The width of the peak along the leading diagonal represents an estimate of the robustness of the technique at that particular angle. For example, near the boresight direction to around 20^o the peak is broad and so worse performance is expected than from a measurement at 45^o (where the peak is narrower). The poor performance in the boresight direction is expected due to the symmetry of the antennas and the flat frequency-profile in this direction.

5.4.1.2 Simulated Behaviour

By degrading the power-based map function with AWGN it is possible to simulate the signal ratio which would result from such a measurement. The example given in Figure 5.8c is for a received SNR of 10 dB in all directions. From the simulated signal ratio

and the map function, the correlation surface can be estimated for a particular value of SNR and is given in Figure 5.8d for the 10 dB case. From the correlation surface, the resulting accuracy of the technique under particular SNR conditions may be estimated. A more detailed exploration of the effects of SNR on this localization technique is given in Section 5.4.5.

5.4.1.3 Real Target Behaviour

The measurement made for a single target in the anechoic chamber resulted in a correlation surface as shown in Figure 5.8f. This correlation surface is in good agreement with Figure 5.8b but is corrupted by measurement noise that increases away from the boresight direction. By extracting the peak correlation value for each measurement angle, the performance of the technique can be quantified, and is compared to the results of simulation in Figure 5.9. There is good agreement between the measured and simulated results, with a region of systematic error in the $0-20^o$ region. This error is expected as the correlation surface is broadest in this region. In the $20-70^o$ region, the standard deviation of the error is 1.83^o, and is 1.14^o in the $20-40^o$ region. Above 70^o the localization completely fails due to the low SNR of the signal at these angles, and so no attempt is made to quantify the error in this region.

Figure 5.9 demonstrates the fundamental viability of using the frequency-diversity of a signal and the natural spatial encoding that occurs due to the physical characteristics of antennas to localize a target in angle over a wide angular range of operation.

5.4.2 Phase Difference

As with the power-based approach of the previous section, a map function can be determined for the phase-based approach by taking the phase angle difference between the two receiving antennas in the direct signal case, as described in Equation (5.11). This method yields the phase-based map function shown in Figure 5.10a. To determine the ideal and simulated behaviours the same approach is used as the power-based approach and so the

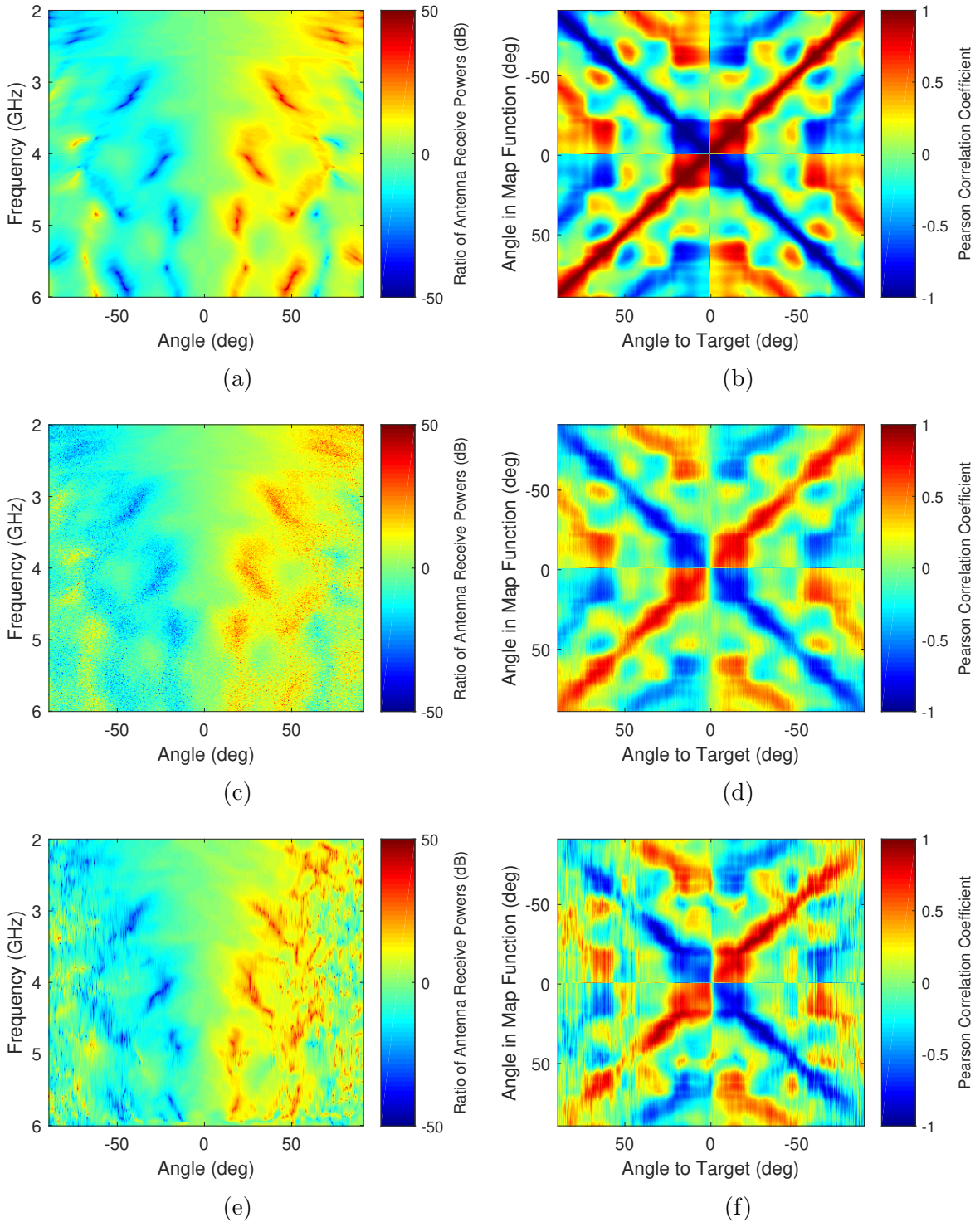


Figure 5.8: Power-based map functions and signal ratios (left) and correlation surfaces (right). (a)(b) Measurements of the direct signal. (c)(d) The direct signal measurements degraded by 10 dB of AWGN. (e)(f) Measurements with a target present.

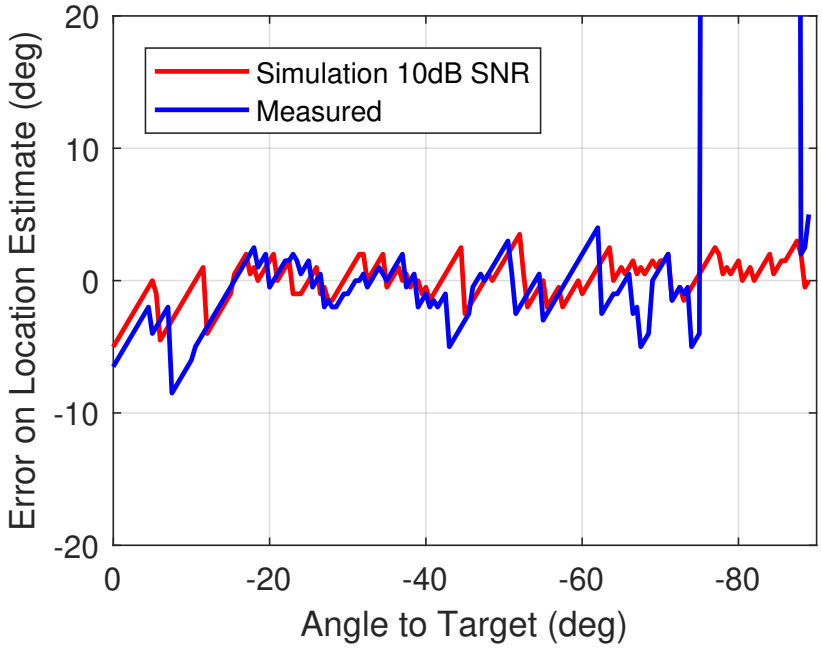


Figure 5.9: Simulated and measured errors of the power-based localization technique as a function of angle. In the $20 - 70^o$ region, the standard deviation of the measured error is 1.83^o, and is 1.14^o in the $20 - 40^o$ region.

following sections are kept concise to avoid repetition.

5.4.2.1 Ideal Behaviour

The ideal behaviour is estimated by self-correlating the phase-based map function to generate the correlation surface shown in Figure 5.10b, which shows a tight peak along the lead diagonal. Because the width of the peak represents the robustness of the technique at that angle, the correlation surface indicates that the technique is also expected to perform poorly at angles close to the system boresight (where the phase difference between the antennas is close to zero), with performance expected to improve at angles greater than 5^o where the peak is narrow, and with performance expected to decrease slowly away from the boresight direction.

5.4.2.2 Simulated Behaviour

The phase-based map function is degraded with AWGN to simulate a measurement (as with the power-based approach) of the signal ratio as shown in Figure 5.10c. The SNR on the signal ratio decreases away from the boresight direction due to the decrease in the antenna gains at these angles. The signal ratio surface is correlated with the original map function to determine the correlation surface given in Figure 5.10d. The simulated correlation surface shows a drop in correlation at angles furthest from the boresight, which is expected due to the decrease in SNR to 10 dB.

5.4.2.3 Real Target Behaviour

The measurement made for a single target in the anechoic chamber resulted in the correlation surface shown in Figure 5.10f. This correlation surface is in good agreement with Figure 5.10b but, as with the power-based approach, is corrupted by measurement noise that increases away from the boresight direction. The performance of the technique is quantified and compared to the results of simulation in Figure 5.11. There is good agreement between the measured and simulated results, with a region of systematic error in the $0 - 5^o$ region. This error is expected as the correlation surface is broadest in this region. In the $20 - 70^o$ region, the standard deviation of the error is 0.82^o, and is 0.36^o in the $20 - 40^o$ region. As with the power-based approach, the localization performance suffers significantly above 70^o and so no attempt is made to quantify the error in this region.

It is worth noting that further than 40^o from the boresight a small systematic error is present, which manifests as a small offset from 0^o error on the localization estimate. This effect is not visibly present in the results for the power-based approach and is possibly due to slight variations in environmental conditions between calibration and experiment. The presence of this error indicates the high sensitivity of the phase-based approach.

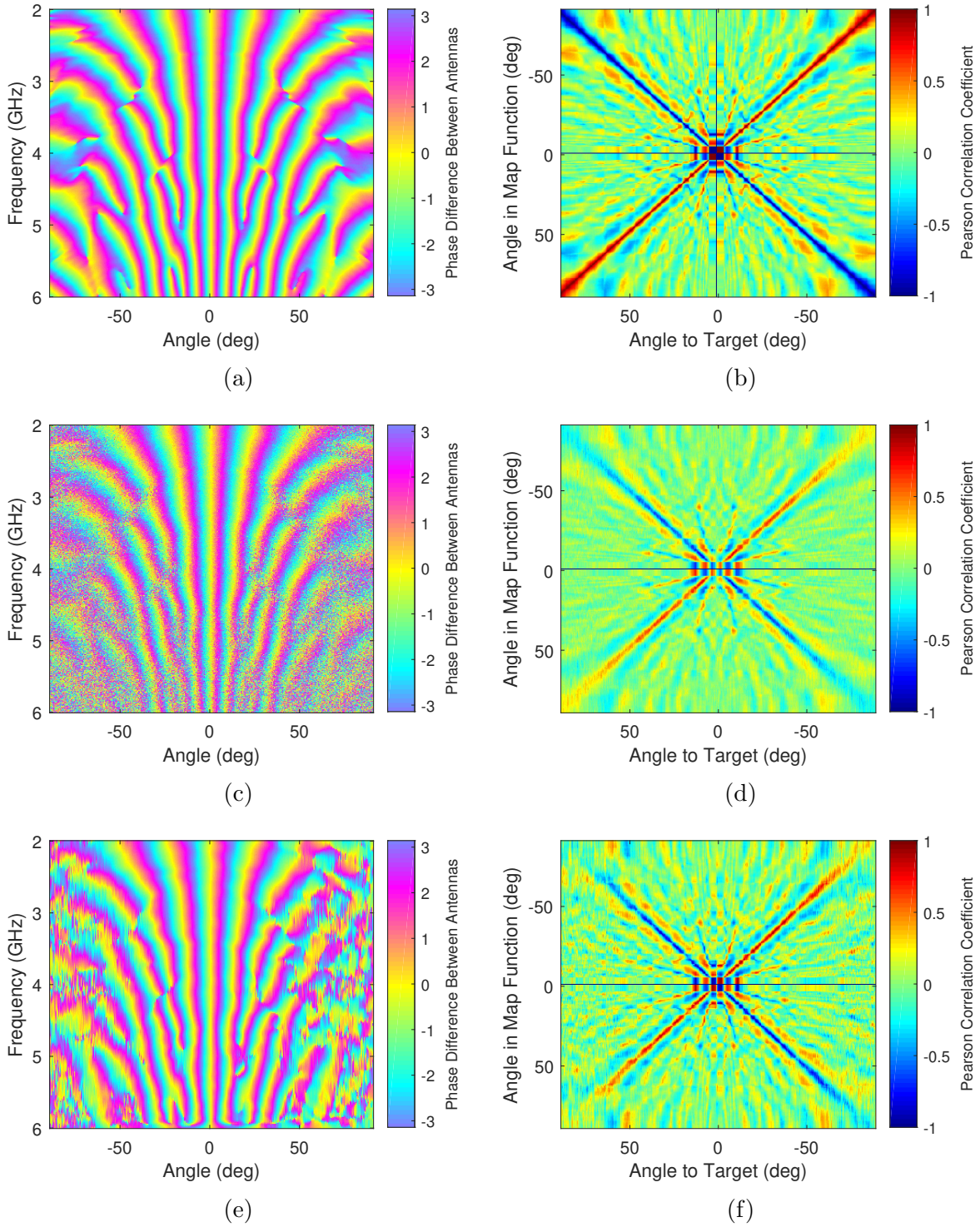


Figure 5.10: Phase-based map functions and signal ratios (left) and correlation surfaces (right). (a)(b) Measurements of the direct signal. (c)(d) The direct signal measurements degraded by 10 dB of synthesized AWGN. (e)(f) Measurements with a target present.

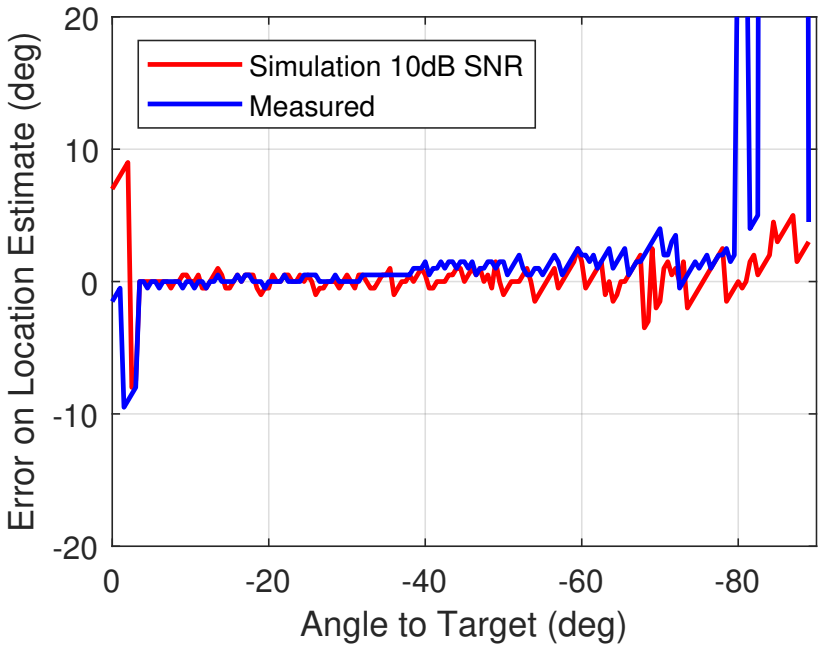


Figure 5.11: Simulated and measured errors of the phase-based localization technique as a function of angle. In the 20-70^o region, the standard deviation of the measured error is 0.82^o, and is 0.36^o in the $20-40^o$ region.

5.4.2.4 Comparison of Power- and Phase-based Approaches

The previous sections have presented the results of using the power- and phase-based techniques, both of which yield good results. The phase-based approach gives much better localization performance, but it is also reliant on phase coherency between the two receiver channels, and is much more sensitive to noise due to the finer encoding of space. Therefore it is likely that in the real-world, where phase-coherence over such a large bandwidth may not be practical, the power-based approach is the more easily realized of the two techniques.

The reason that the phase-based approach performs better in laboratory conditions than the power-based approach is largely due to the decorrelation between closely spaced angular positions. This decorrelation is visible in Figure 5.10f as a blue region (negative correlation) on either side of the leading diagonal (positive correlation), which narrows the width of the main peak across all angles and results in a localization result that is more stable to noise. The correlation surface for the phase-based approach also has fewer

other peaks than the power-based approach, meaning that different angles are less likely to be confused. The correlation surface has fewer features because the phase-based map function is a function of both antenna gains and spatial separation (see Figure 5.10a). The spatial separation term of this equation causes every angle to be characterized by a very different frequency profile which reduces the likelihood of correlation between returns from different angles.

5.4.3 On-Boresight Detection

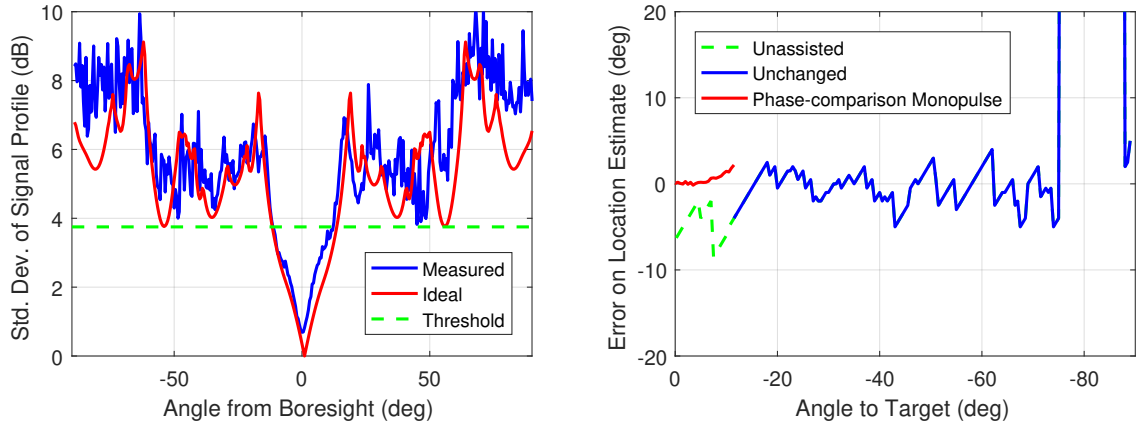
Using the technique described in Section 5.2.5 for detecting a target in the boresight direction it was possible to effectively switch techniques in this zone of degraded performance. The results of the on-boresight detection for the power- and phase-based approaches are given in Figure 5.12 and Figure 5.13 respectively. These figures show that, by setting an appropriate threshold for the number of standard deviations in the signal profile or in the phase difference profile, the boresight condition can be effectively determined, and the technique switched to improve performance.

This switching is indicative of how this technique could be integrated with existing systems that use alternative techniques, allowing the switching between techniques when one is known to be performing poorly.

5.4.4 Performance Effects of Different Frequency Bands

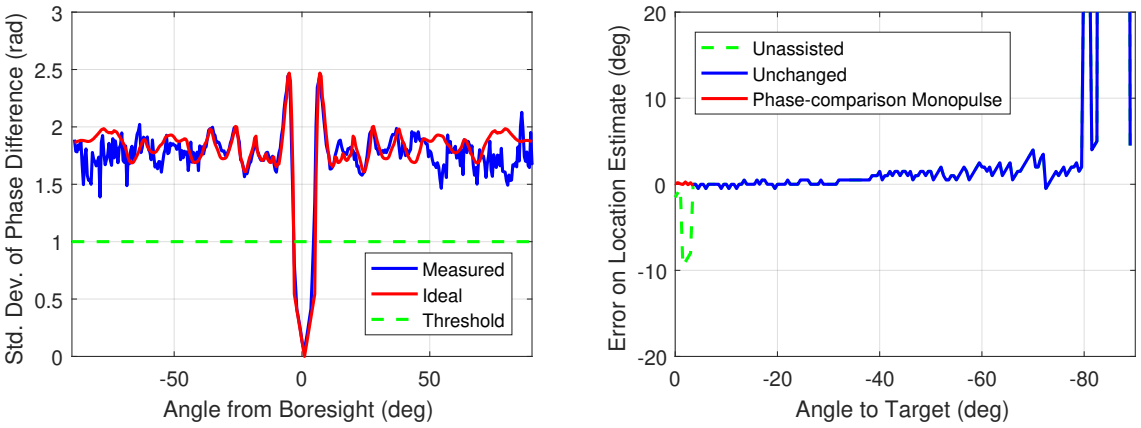
This section considers the effects of different frequency bands for both power- and phase-based approaches. In general, the band of frequencies used must elicit substantial variation in the beam patterns of the antennas. As introduced in Section 5.2.6, there is nothing unique about the absolute band or bandwidth used, but rather it is the variation in the encoding of space by the antenna across of the frequencies of the band that is being used.

The results shown in this section are determined by processing the data recorded by the VNA over the desired frequency bands, which are 2-4 GHz, 2-5 GHz, 2-6 GHz, 2-10 GHz, 4-8 GHz, and 6-10 GHz. These particular bands were selected because they



(a) (b)

Figure 5.12: Power-based results for the on-boresight identification technique. (a) Shows the standard deviation of the signal profile across the full bandwidth. (b) Shows the result of using the on-boresight detection to improve performance at low angles. The green-dashed line shows localization performance of the power-based technique near the boresight (where performance is poor), by switching to a phase-comparison monopulse approach, the results in red are attained. The results of the power-based approach further from the boresight are shown in blue and remain unchanged.



(a) (b)

Figure 5.13: Phase-based results for the on-boresight identification technique. (a) Shows the standard deviation of the signal profile across the full bandwidth. (b) Shows the result of using the on-boresight detection to improve performance at low angles. The green-dashed line shows localization performance of the phase-based technique near the boresight (where performance is poor), by switching to a phase-comparison monopulse approach, the results in red are attained. The results of the phase-based approach further from the boresight are shown in blue and remain unchanged.

allow comparison of the effects of varying the start frequency of the band as well as the bandwidth as four results have the same starting frequency and three results have the same bandwidth.

5.4.4.1 Power-based Approach

Figure 5.14 shows that when the bandwidth is low, as in the case of 2-4 GHz, there is a large ambiguous region in the boresight direction up to around 30^o, which manifests as a 'chequerboard' square region in the centre of the correlation surface. This region occurs because the frequencies in this smaller band have broad beamwidths, meaning that there are fewer spectral features to discriminate between angular positions near the boresight, and consequently ambiguity occurs in that region. The 2-5 GHz results demonstrate that by increasing the bandwidth to include frequencies with narrower beamwidths, the extent of the angular ambiguity is decreased (to around 20^o). This pattern of reduced ambiguity towards the boresight direction continues for the 2-6 GHz and 2-10 GHz results.

If it is higher frequencies that help discriminate between different angular positions, it might be expected that using the same bandwidth but with a higher central frequency would yield better performance. In Figure 5.14 the results for 2-6 GHz show a narrow peak on the correlation surface from $20 - 70^o$ from the boresight direction, which translates to relatively good localization performance in this region. Increasing the band of frequencies to 4-8 GHz does reduce the ambiguity in the boresight direction for the reasons discussed previously. However, the technique now only works well up to about 40^o, this occurs because the sidelobes of the higher frequencies introduce ambiguity into the signal power ratio. Above 40^o, the ratio of signal powers is no longer sufficiently different at every angular position to enable good localization performance. This degradation of performance continues if the centre frequency is further increased to 6-10 GHz. At this point, notable angular ambiguities occur in the $20 - 30^o$ range, which further reduce the performance.

The final important effect is related to the SNR of the signal. As the lower frequen-

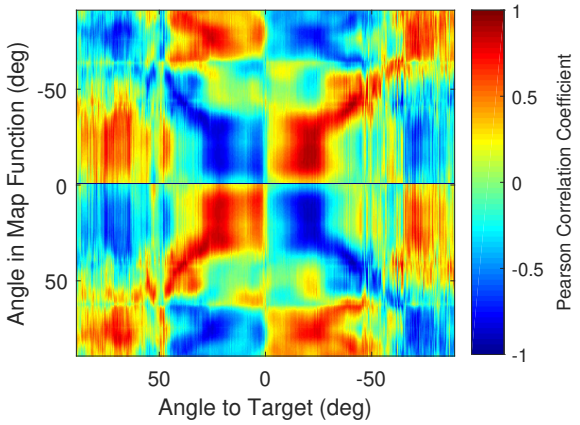
cies are increasingly omitted from the band, the performance at angles furthest from the boresight decreases. Part of this reduction in performance is due to the structural ambiguity of the beam patterns discussed previously, and part is due to a decrease in SNR. The lower frequencies correspond to wider beams, and so the incident power on a target far from the boresight direction is greater when lower frequencies are used. The SNR effects are responsible for the majority of dissimilarities between simulated and measured results because the simulation was performed with a constant SNR and did not take into account the reduction of SNR away from the boresight direction. It is therefore possible to separate the effects due to structural ambiguity and those due to noise. The results for 6-10 GHz show structural ambiguity for angles greater than 40^o , whereas the 4-8 GHz results show small regions of structural ambiguity at 40^o, 70^o, and 80^o, which manifest as spikes in the simulated result. More detail of noise-dependent effects is discussed in Section 5.4.5

5.4.4.2 Phase-based Approach

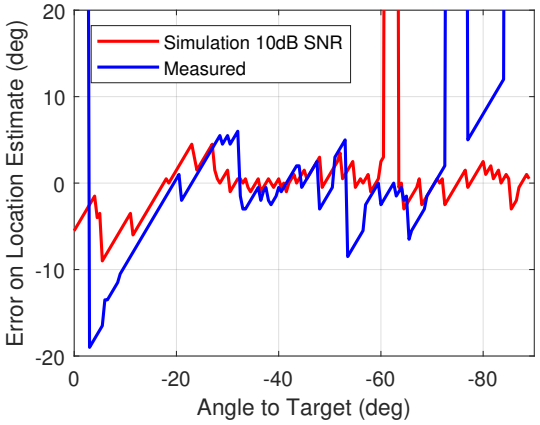
The phase-based approach exhibits similar behaviour to the power-based approach, with results for different bands shown in Figure 5.15. For 2-4 GHz, the 'chequerboard' ambiguity effect near the boresight is present in the correlation surface once more. The regions of ambiguity are smaller than for the power-based approach, and this is due to the lack of spectral features in the phase map function in the boresight direction (see Figure 5.10e). As with the power-based approach, the chequerboard region decreases in extent as the bandwidth increases to 2-5 GHz, and further still to 2-6 GHz. The extent of the ambiguity is reduced with the inclusion of higher frequencies because the phase of the component changes more rapidly at higher frequencies, encoding space more finely, and allowing for more accurate localization.

The finer encoding of space by higher frequencies is not solely beneficial, and introduces new ambiguities to the system. The new ambiguity manifests in the correlation surface as high correlation regions radiating from the origin, not along the leading diagonal. This

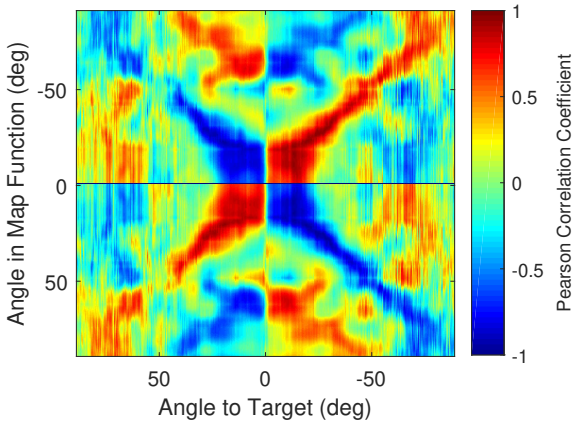
effect can be seen in the results for 4-8 GHz, and results in low reliability of localization as the SNR drops beyond around 40^o from the boresight. Because space is encoded more finely by the higher frequencies, localization performance is far more sensitive to noise than at the lower frequencies. The results for 6-10 GHz show an extreme case where localization is almost impossible due to the ambiguous banding present in the correlation surface.


Angle in Map Function (deg)
-50
0
50
50
0
-50
Angle to Target (deg)
1
0.5
0
-0.5
-1
Pearson Correlation Coefficient

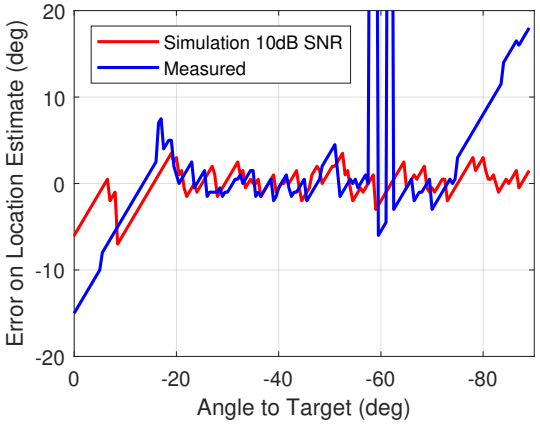
(a)


Simulation 10dB SNR
Measured
Error on Location Estimate (deg)
20
10
0
-10
-20
0
-20
-40
-60
-80
Angle to Target (deg)

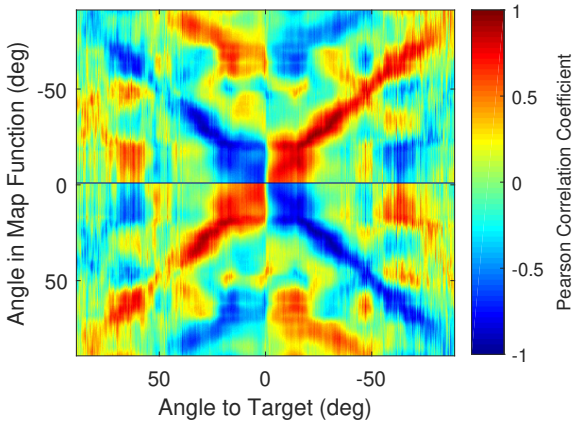
(b)


Angle in Map Function (deg)
-50
0
50
50
0
-50
Angle to Target (deg)
1
0.5
0
-0.5
-1
Pearson Correlation Coefficient

(c)


Simulation 10dB SNR
Measured
Error on Location Estimate (deg)
20
10
0
-10
-20
0
-20
-40
-60
-80
Angle to Target (deg)

(d)


Angle in Map Function (deg)
-50
0
50
50
0
-50
Angle to Target (deg)
1
0.5
0
-0.5
-1
Pearson Correlation Coefficient

(e)


Simulation 10dB SNR
Measured
Error on Location Estimate (deg)
20
10
0
-10
-20
0
-20
-40
-60
-80
Angle to Target (deg)

(f)

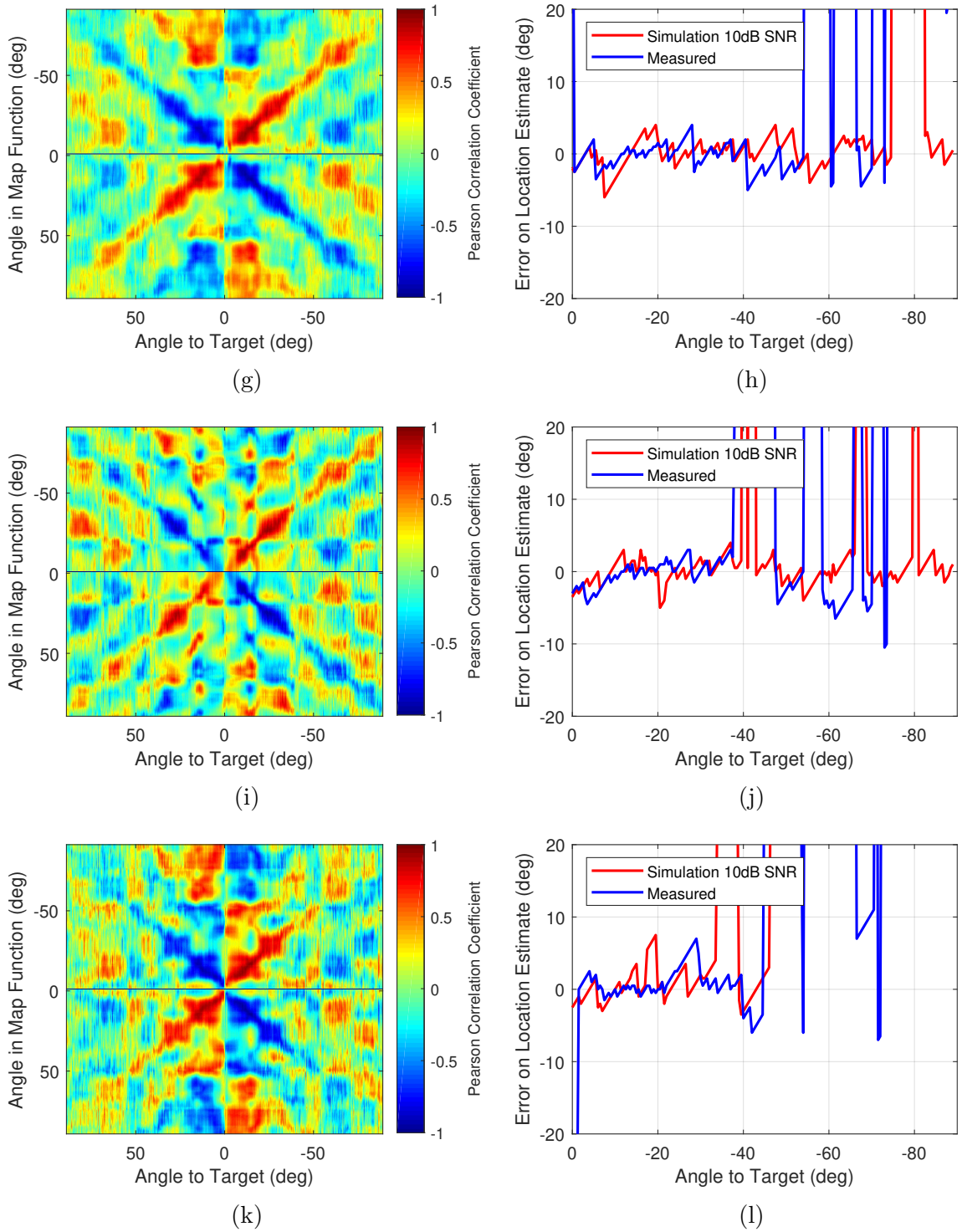
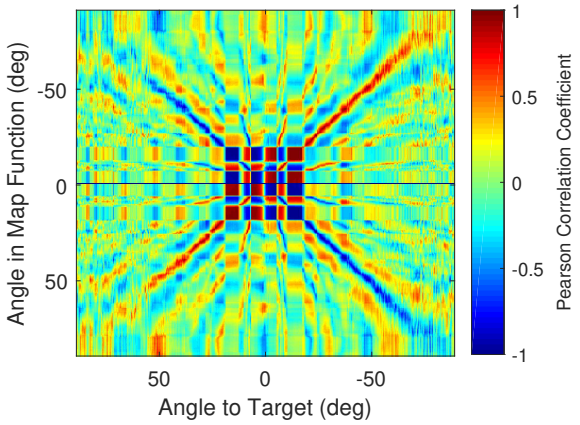
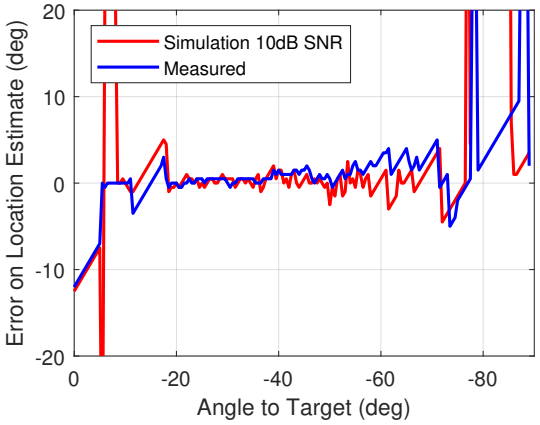


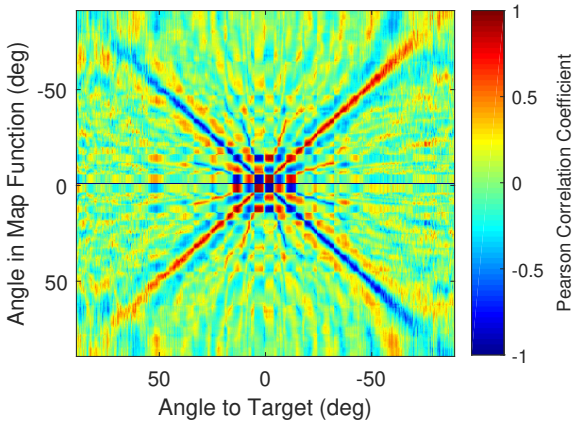
Figure 5.14: Power-based correlation surfaces and results for different frequency bands. (a)(b) 2-4 GHz, (c)(d) 2-5 GHz, (e)(f) 2-6 GHz, (g)(h) 2-10 GHz, (i)(j) 4-8 GHz, (k)(l) 6-10 GHz.


Angle in Map Function (deg)
-50
0
50
50
0
-50
Angle to Target (deg)
1
0.5
0
-0.5
-1
Pearson Correlation Coefficient

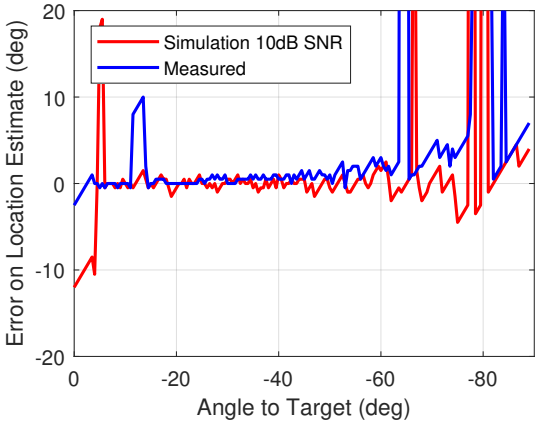
(a)


Error on Location Estimate (deg)
20
10
0
-10
-20
0
-20
-40
-60
-80
Angle to Target (deg)
Simulation 10dB SNR
Measured

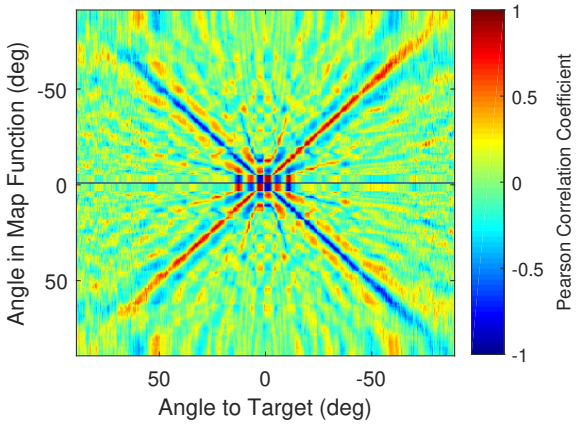
(b)


Angle in Map Function (deg)
-50
0
50
50
0
-50
Angle to Target (deg)
1
0.5
0
-0.5
-1
Pearson Correlation Coefficient

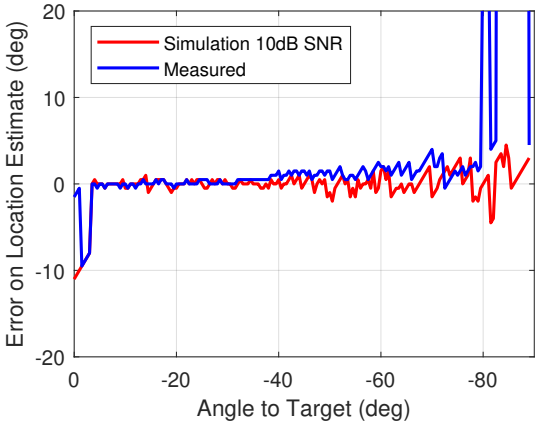
(c)


Error on Location Estimate (deg)
20
10
0
-10
-20
0
-20
-40
-60
-80
Angle to Target (deg)
Simulation 10dB SNR
Measured

(d)


Angle in Map Function (deg)
-50
0
50
50
0
-50
Angle to Target (deg)
1
0.5
0
-0.5
-1
Pearson Correlation Coefficient

(e)


Error on Location Estimate (deg)
20
10
0
-10
-20
0
-20
-40
-60
-80
Angle to Target (deg)
Simulation 10dB SNR
Measured

(f)

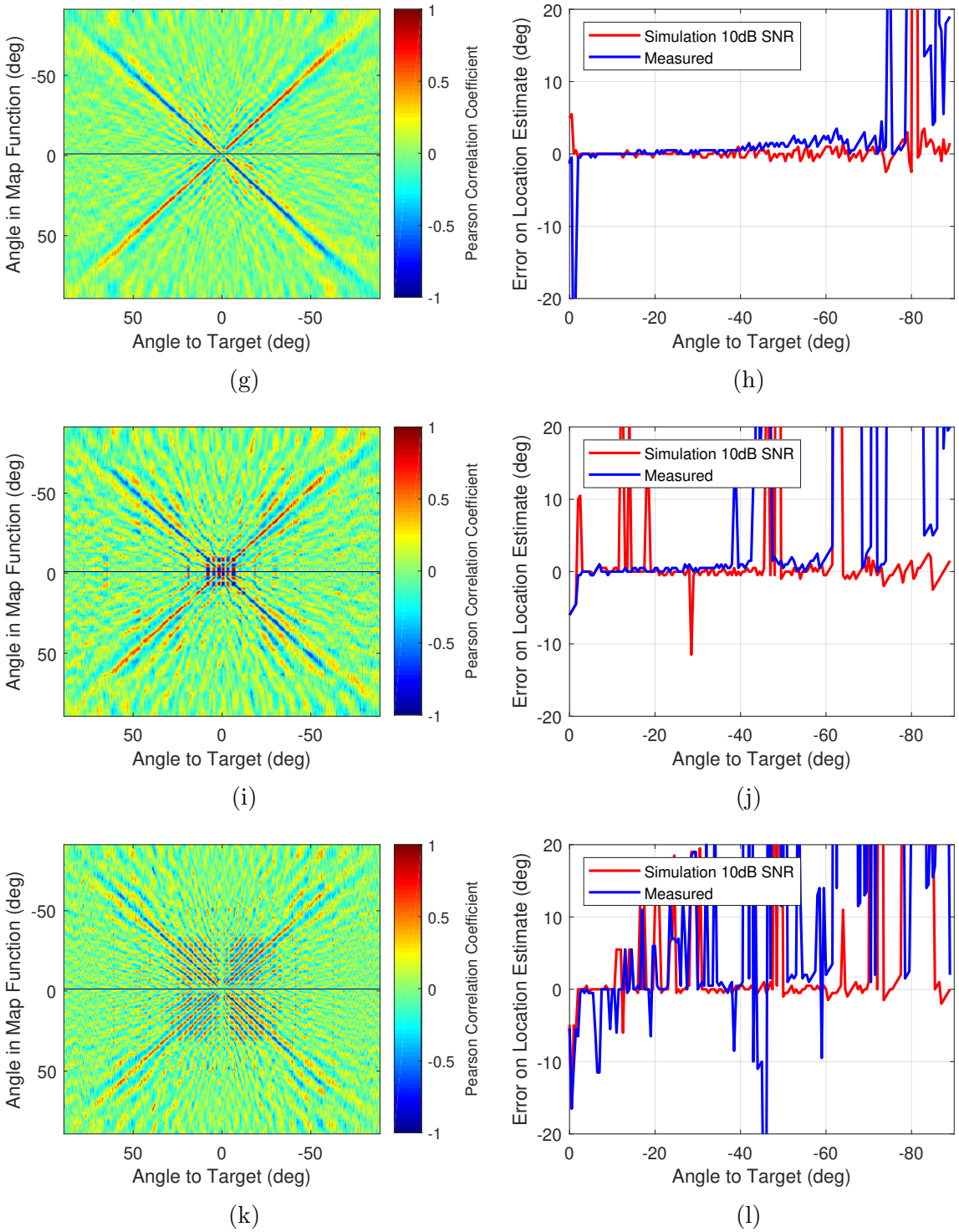


Figure 5.15: Phase-based correlation surfaces and results for different frequency bands. (a)(b) 2-4 GHz, (c)(d) 2-5 GHz, (e)(f) 2-6 GHz, (g)(h) 2-10 GHz, (i)(j) 4-8 GHz, (k)(l) 6-10 GHz.

5.4.4.3 Summary

One of the features that suggests itself when considering the results presented in Figure 5.14 and Figure 5.15 is that the lowest-frequency components (which correspond to a very wide beam and to low-frequency phase variations) are required to disambiguate the returns from different azimuth positions. Conversely, the higher frequencies are required to improve the azimuth resolution, particularly at the angles near the system boresight direction. The use of different frequency components to discriminate between targets at different angular locations is similar to an approach possibly used by bats to discriminate targets and clutter [63] introduced in Chapter 2, and the findings presented here could possibly support that research. These two factors also have to be balanced against the increased admittance of noise and clutter into the system that comes with a wider frequency band, as well as the greater sensitivity to noise that comes with the higher frequencies. An increase of noise admitted into the system degrades the performance of the correlation process, and localization performance suffers as a consequence.

5.4.5 Signal-to-Noise Ratio

In order to understand the performance effects of SNR on the techniques it is important to quantify the SNR at various angles from the system, as the SNR is expected to drop at angles away from the boresight direction. The SNR was estimated by comparing the return signals in the presence of a target to signals recorded when no target or clutter was present. The measurements of SNR on the signals on the receiver channels, s_1 and s_2, can be used to estimate the SNR on the power-based joint pattern using Equation (5.16) [5] and the result of this is shown in Figure 5.16. The drop of SNR away from the boresight direction of each antenna is expected from the beam patterns; as the frequency increases the antenna beamwidth narrows causing the signal magnitude to fall closer to the noise floor at angles away from the boresight and outside of the antenna main lobe. From this estimate it can be seen that the SNR drops to below 10 dB at approximately -60^o, which is consistent with the measured performance for both techniques suffering for angles

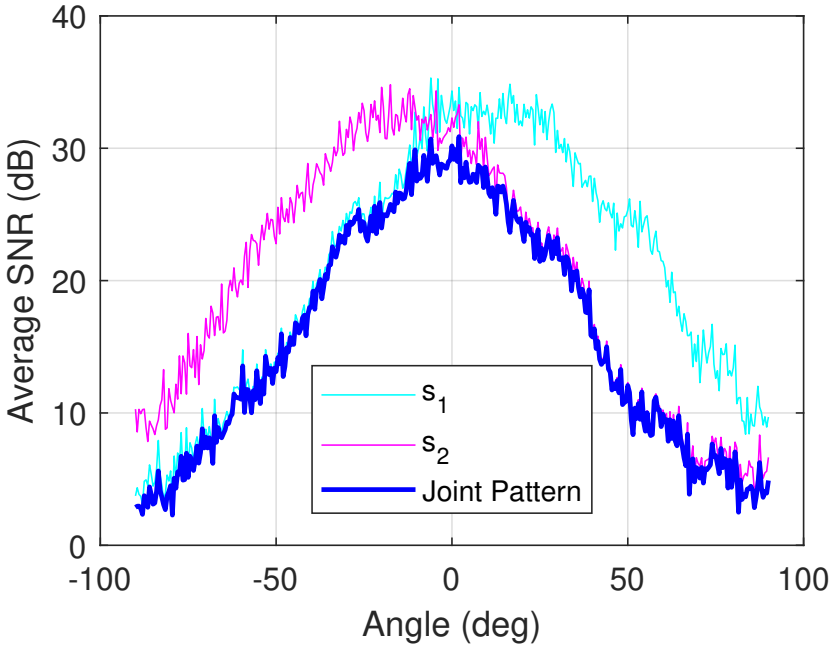


Figure 5.16: The SNR on the signal used for the power-based approach.

further from the boresight.

$$\mathrm{SNR}_{\mathrm{joint}} = \left(\mathrm{SNR}_1{}^{-2} + \mathrm{SNR}_2{}^{-2}\right)^{-0.5} \tag{5.16}$$

5.4.5.1 SNR Filtering

As discussed above, the SNR decreases with angle and varies across the wide frequency band, causing a significant decrease in average SNR at angles away from the boresight. If the average SNR was increased, it is likely that performance of the previously presented technique could be improved. By setting a signal magnitude threshold for both channels, only the frequency-components with an SNR greater than the threshold are included in the correlation process. Figure 5.17 shows the improvement in average SNR on the joint pattern as a function of angle for an SNR threshold of 10 dB, at angles greater than approximately 40^o from the boresight an SNR improvement can be seen. At the angles furthest from the boresight an SNR improvement of about 8 dB is achieved.

The signal ratio and correlation surface for the SNR filtering are shown in Figure 5.18

which show that the threshold excludes an increasing number of frequency components as the angle from the boresight direction increases and that the result of this is a sharper correlation surface. Figure 5.19 shows the result of localization using this technique. The performance achieved here is very high and is better than would be expected in a real-world environment. This high performance occurs because the clutter and noise floor of the anechoic chamber was static and well-known. However, the ability to improve performance by using a threshold to remove low SNR measurements is a useful one, and these results demonstrate the fundamental viability of processing any portion of the frequency band in this way.

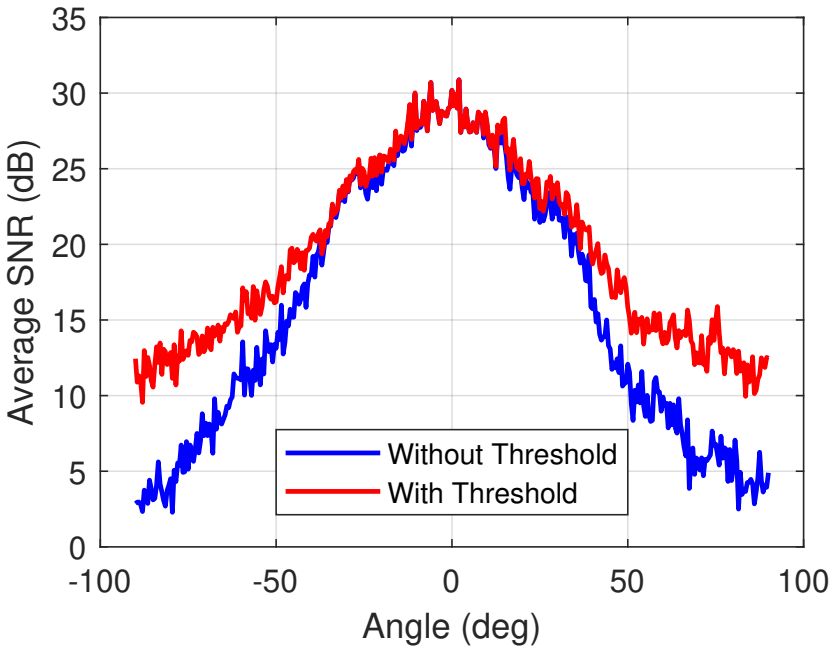
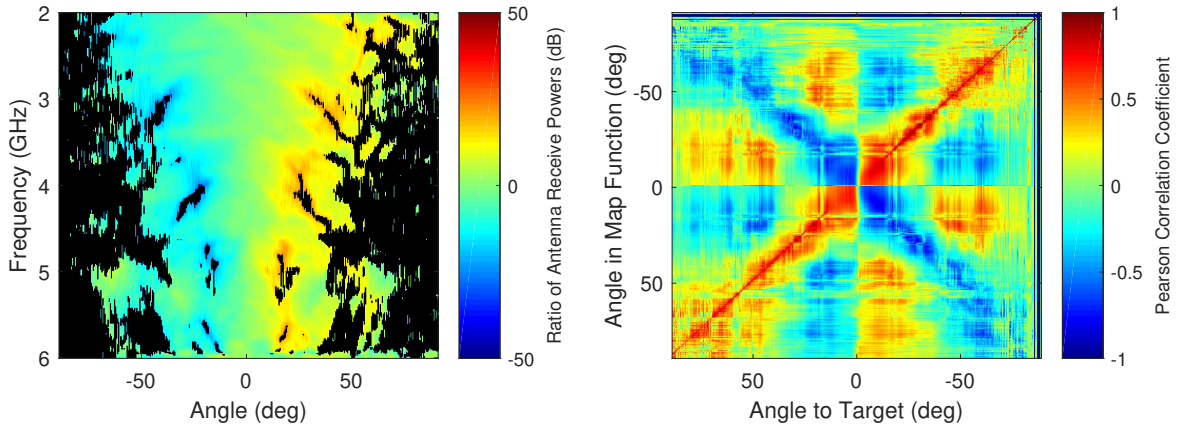


Figure 5.17: A comparison of SNRs on the joint pattern used for the power-based approach with and without signal magnitude filtering. The filtering excludes frequency components from the processing if the SNR falls below a threshold of 10 dB.



(a) (b)

Figure 5.18: Measurements with an SNR threshold of 10 dB. (a) Measured signal ratio, where black indicates areas where the SNR has fallen below the threshold. (b) Correlation surface.

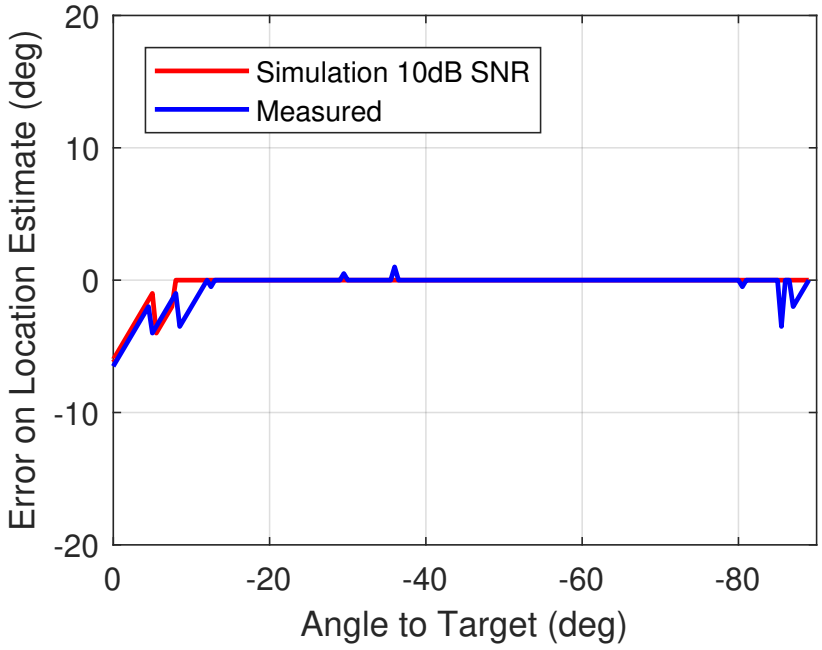


Figure 5.19: Simulated and measured errors of the power-based localization technique with an SNR threshold of 10 dB. In the $15 - 80^o$ region, the technique localizes the target to within 0.5^o (the angle bin size) in 128 out of 130 cases, with a standard deviation of 0.010^o.

5.5 Summary

The previous sections have laid out the theoretical basis for and the results of using a novel biologically-inspired radar technique based on the frequency-angle relationship of the received signals as a means of target localization. In this section, the processes underlying the results are discussed in more detail and the various advantages and drawbacks of this method are explored.

The most significant result is that using the techniques presented it is possible to achieve target localization over a wide range of angles using off-the-shelf horn antennas and no moving parts. These techniques rely on the variation of antenna gain as a function of frequency, with a pair of antennas effectively encoding space with unique markers that can be used to determine angular position. For successful target localization in angle, sufficient information must be encoded in the received signals and so the variation in the antenna beam patterns must also be significant. To achieve this significant variation, a wide band of frequencies was used, but this is not inherently a wideband technique, and the same results could be achieved with antennas that had the same variation in beam pattern over a much narrower bandwidth.

As discussed in Chapter 2, there is neuroscientific evidence that frequency-dependent binaural cues are used in nature [145, 44, 42]. Specifically, regions of neurons in the lateral superior olive (a region of the brain close to the start of the auditory pathway) respond selectively to ILDs as a function of frequency. This group of neurons pass the BTD cue along the auditory pathway to the inferior colliculus, and subsequently the superior colliculus where other cues are integrated and an auditory map of space is found. It is this set of neural processes that provide the strongest evidence for biological equivalents of the techniques presented in this chapter.

For both power- and phase-based approaches poor performance is seen in the boresight direction which is both expected and consistent with simulation results. This degraded performance occurred because the symmetry of the antenna patterns led to a flat profile within the map function (for both power- and phase-based approaches), which caused the

effects of noise to be more pronounced in the likelihood estimator. The poor performance in the boresight direction was compensated for by using an on-boresight detection approach, which identified when the measured signal profile fell below a threshold required for good localization performance. By switching to a phase monopulse approach in the boresight region, the overall performance was improved out to approximately 70^o where low SNR degrades the results. This performance increase demonstrates that this technique would likely be the most viable in conjunction with alternative techniques for target localization, and could be used to augment other techniques on existing systems.

The results for different bandwidths indicate the importance of using a band containing a frequency component which for the antenna used has a wide beamwidth with high gain. It is this low frequency which anchors the approach and determines the range of angular operation of the technique. It is also important to have a wide range of higher frequencies included in the band. A key finding of this research is that it is the variations in the beampatterns that provide the spectral features required to disambiguate between different angular positions. However, a limit must be imposed on the band of frequencies used, as a wider bandwidth permits more noise to enter the processing, degrading the overall performance. The imposition of a limit to the band is a trade off against the SNR requirements of the processing; if signals below an SNR threshold were filtered out of the processing, this band restriction could be relaxed.

SNR thresholding was carried out and shown to significantly improve the performance of the technique, localizing the target in the $15 - 80^o$ region almost perfectly, with a standard deviation of 0.010^o using a band of frequencies from 2-6 GHz. The application of a threshold increased the average SNR as the angle from the boresight increased. The threshold approach also demonstrates the feasibility of using any combination of frequencies within a band for the processing. The ability to select individual frequencies to consider in the correlation process opens up the possibility that this technique could conceivably be used with multiple narrowband radars operating at different frequencies, which is more feasible than using an ultra-wide bandwidth for real-world radar systems

that operate with stringent bandwidth restrictions.

This chapter has considered the single target case for frequency-diverse localization, this technique fails when two or more targets are present and are at a similar range. The reason for this breakdown is that the ratio of received signals, Equation (5.6), is no longer solely a ratio of antenna gains, and instead includes cross-terms from the result of division of different target returns. This is the same restriction that applies to conventional monopulse systems, which experience large errors if two targets are encountered at the same range. If the two targets are separable in range, the bio-inspired technique can still be used, operating on the returns of each target independently.

5.6 Conclusion

This chapter has developed a biologically-inspired radar system which uses the information encoded in the spectral content of a signal to localize a target. The wide band of frequencies provides an extra degree of freedom that the radar system can exploit for target localization. In addition to using a wide range of frequencies, human echolocators also use the degrees of freedom provided by movement of the head and body to build up a more complete perception of their environments.

The next chapter will explore the possibilities and challenges that adding movement to a radar platform presents. With additional degrees of freedom, a radar system should be able to leverage these to increase its understanding of its environment and to improve performance. As humans we understand our environments because of our cognitive abilities, and the robot in the next chapter will take inspiration from human cognitive structures in order to process and understand its environment.

Chapter 6

Cognitive Radar Architecture for Robot Navigation

Nosotros, de un vistazo, percibimos tres copas en una mesa; Funes, todos los vástagos y racimos y frutos que comprende una parra.

We, in a glance, perceive three wine glasses on the table; Funes saw all the shoots, clusters, and grapes of the vine.

— Jorge Luis Borges

Funes el memorioso, 1942

6.1 Introduction

This chapter is framed in terms of a radar-focused autonomous navigation task in which a robot is required to sense and avoid obstacles en route to a destination. Radar sensors are considered because their all-weather and day or night operational capabilities can make them a suitable sensing choice in certain unknown and adverse environments through which mobile robots may have to navigate.

Traditionally, radar has been widely used for collision avoidance in manned surface, ground, and air platforms. More recently, its use onboard ground [146] or aerial [147, 148, 149] robotic platforms has been considered. However, there are still major challenges to autonomous radar-only navigation, especially in confined spaces. A significant limiting factor is that radar systems generally have lower resolution and accuracy than other sensing modalities such as lidar. In addition, the complexity of radar scattering makes reliable information extraction more difficult.

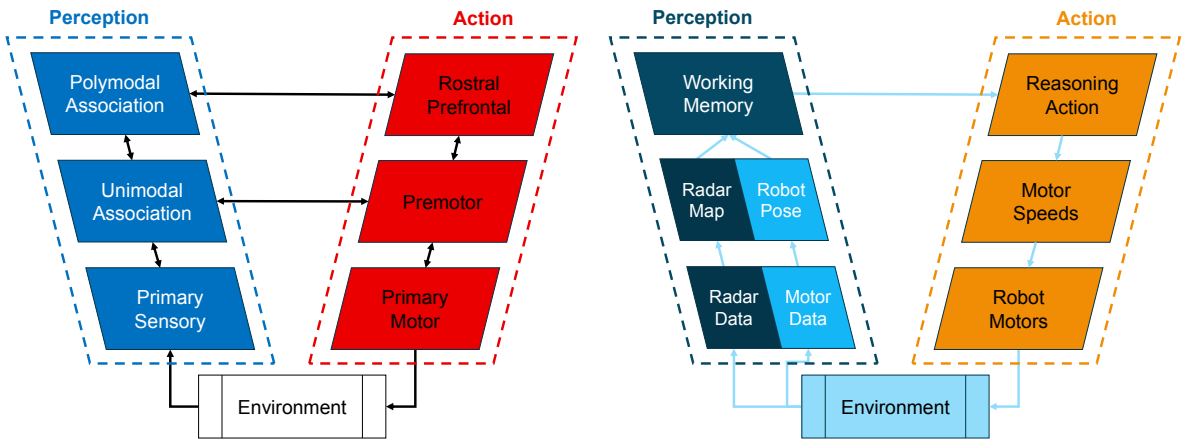
These challenges may be addressed by embedding intelligence within the radar sensor. This area is still in its infancy, but demands on improved autonomy are becoming more pressing all the time. Artificial radar intelligence may be manifested through reactive systems [150, 151], and can be inspired by natural principles, such as the echolocation of bats inspiring echoic flow techniques[76]. This work looks towards the creatures of the natural world and attempts to develop a system that captures some of the cognitive processes present in nature.

Creatures in the natural world use cognitive processes to operate on complex and variable sensory inputs while interacting with their environments. In 2006, the neuroscientist Joaquín Fuster introduced the idea of a 'cognit' to cognitive neuroscience. A cognit is a network of neurons ranging in size that represents a unit of memory or knowledge in the brain [153]. Figure 6.1a shows Fuster's hierarchical perception-action cycle, often referred to in literature on cognitive radar [154, 75, 56, 57, 155].

However, perception is only one of many functional elements of cognition. To list the functional elements of cognition exhaustively is a task beyond the scope of this work, but

commonly referenced elements are: Perception, Action, Memory, Attention, Language, Learning, and Reasoning [152]. These elements make it apparent that the perception-action cycle is the 'tip of the iceberg' of cognition and that the presence of many other functional elements alongside perception most likely enables higher-order cognitive behaviours.

Fuster's cognit and presentation of the perception-action cycle can be thought of as a useful roadmap for the development of cognitive systems. By following this roadmap, cognitive systems that directly mimic the functional structures present in Fuster's hierarchy can be built. Figure 6.1b shows how this chapter interprets Fuster's perception-action cycle for the mobile robot, relating the radar and motor data to raw perceptual input, and showing how this is propagated to a working memory which can be used for determining the course of action needed for the robot to complete its task.



(a) Natural perception-action cycle. (b) Robot perception-action cycle.

Figure 6.1: Schematics illustrating the functional components and hierarchy of processing in Fuster's perception-action cycle [153] and the robot's perception-action cycle.

This chapter takes initial steps into designing practical cognitive radar systems in the context of autonomous navigation. To do this, Fuster's hierarchy and the functional elements of cognition are used to guide the design of the system. The following sections will describe the theory and implementation of cognitive radar signal processing using the functional elements of cognition, explicitly including memory, for radar-guided autonomous navigation. Results of simulations and experiments are presented, in which

cognitive signal processing is used to perform a navigational task in the presence of obstacles. The role of the various functional elements of cognition are discussed in the context of these experiments, with a particular focus on working memory and how it relates to the radar data. Finally, possible future avenues of research into cognitive radar and robotics are presented.

As a note on terminology, this chapter refers to results of simulated and experimental trials under different conditions. 'Experimental' refers to practical experiments conducted in a laboratory, and 'simulated' refers to the output of an ideal simulation. A 'condition' is a particular set of starting conditions and processing parameters that may be tested repeatedly over multiple 'trials'.

Sections of this chapter are taken verbatim from novel works written entirely by the author [6, 7].

6.2 Memory

This chapter has a focus on the role of memory within autonomous cognitive radar systems, examining the impact that a working memory has on autonomous navigation performance, and so a brief overview of the relevant concepts is presented here.

Memory is generally conceptualized in three ways: sensory, working, and long-term. Sensory memory is the immediate perception by the senses that persists for a short period of time [156, 157]. The light trails that fireworks appear to leave across the sky are an example of the sensory memory of the eye. A person's working memory exists over a longer timescale than the sensory memory, and is the result of attentional processing of multiple sensory memories [158, 156, 159]. For example, if a light was flashed once in an otherwise darkened room, a person unfamiliar with the room would be able to navigate through it using the spatial information stored in their working memory from their visual sensory memory when the light was flashed. Long-term memory deals with the long term ability of creatures to remember information, such as the layout of a place they have

frequently visited [156, 160].

Memory is often discussed in the context of cognitive radar and varied approaches for implementing memory functions have been presented theoretically and in simulation [161, 162, 163], and experimental systems with capability to incorporate memory exist [103]. To the best of the authors' knowledge, this work presents the first results of an experimentally realized cognitive radar system that explicitly incorporates a short-term memory function for robotic navigation.

6.3 Perception and Attention

6.3.1 Primary Sensors

Before considering the cognitive radar signal processing and the control processes of the robot, it is useful to understand the physical sensors that generate the signals that will be used by the robot. The robot has two sources of external information: the radar and the motor encoders, which are described in the following sections. The physical hardware is equivalent to the primary sensory stage of Fuster's perception-action cycle. The robot has two sources of external information, the radar and the motor encoders, which are described in the following sections.

6.3.1.1 Robot

The robot, shown in Figure 6.2, is the Arlo Robot Platform from Parallax Inc. [164] which has a differential drive fixed-axle configuration with two driven wheels and two free coasting wheels for stabilization. The wheels are driven by brushed motors and their rotation is tracked with quadrature encoders fitted to each wheel block. The camera shown in Figure 6.2 is used to provide optical truth for the experiments, used for the post-trial analysis. The relevant parameters of the robot are provided in Table 6.1.

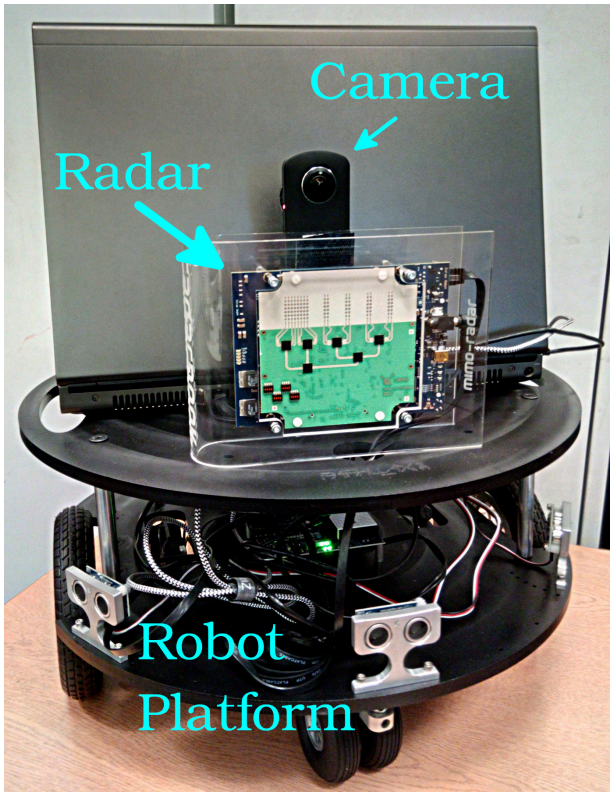


Figure 6.2: Photograph of the robot with the MIMO radar frontend.

Table 6.1: Arlo Robot Parameters.

Parameter	Value	Units
Wheelbase	39.2	cm
Drive Wheel Diameter	15.0	cm
Encoder Ticks per Revolution	144	-

6.3.1.2 Radar

The radar is the INRAS Radarbook with a 77GHz frontend which is a multiple-input multiple-output (MIMO) array in a 4-by-8 configuration resulting in an array of 29 virtual elements (due to overlap in the virtual array) [165, 166]. The radar is triggered by a robot control program, and generates a linear frequency modulated (LFM) upchirp which is sequentially transmitted from each of the 4 transmitters. Because of the MIMO configuration of the radar, the return signals are processed to form a virtual array using conventional MIMO techniques, indicated schematically in Figure 6.3. A more complete description of MIMO radar processing is presented in [167], but is omitted here for brevity.

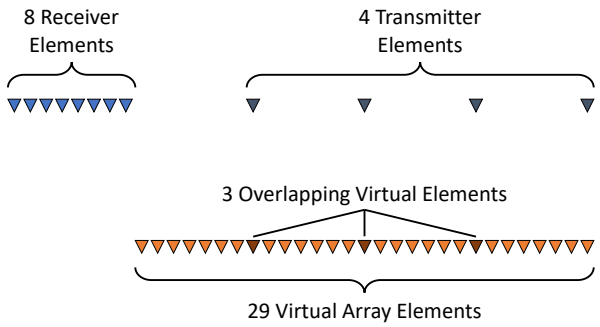


Figure 6.3: Diagram of the physical MIMO array (top) and the virtual MIMO array which this forms (bottom).

Table 6.2: INRAS Radarbook Parameters.

Parameter	Value	Units
Centre Frequency	77	GHz
FMCW Bandwidth	2	GHz
No. Tx Elements	4	-
No. Rx Elements	8	-
No. Virtual Elements	29	-
Sweep Time	60	μs
Sampling Frequency after downconversion	10	Msps
Tx Power	10	dBm

MIMO is a useful technique in the context of mobile robotics because it can have a small form-factor with low power requirements, and provides persistent scene observation over a wide angular range. The description of the radar beamforming in Section 6.3.2.1 refers to the array of virtual elements.

6.3.2 Unimodal Association

Unimodal association is the first layer of perception in Fuster's cycle above the primary sensory layer, and is characterized by the processing of a single sensory input. This section details how the separate signals of the motor rotation encoders and the radar system are processed independently to build up separate perceptions of the environment.

6.3.2.1 Radar Signal Processing

The radar signal processing is performed in two parts: (1) formation of a range-profile for a single (virtual) array element, and (2) the processing of several range profiles across the

array to create a range-angle map, the final output of the radar signal processing. The details of the signal processing are provided in Chapter 3, which gives an explanation of how to use a MIMO radar to form a virtual array, and how to form a range-angle map using an array of elements (virtual or otherwise).

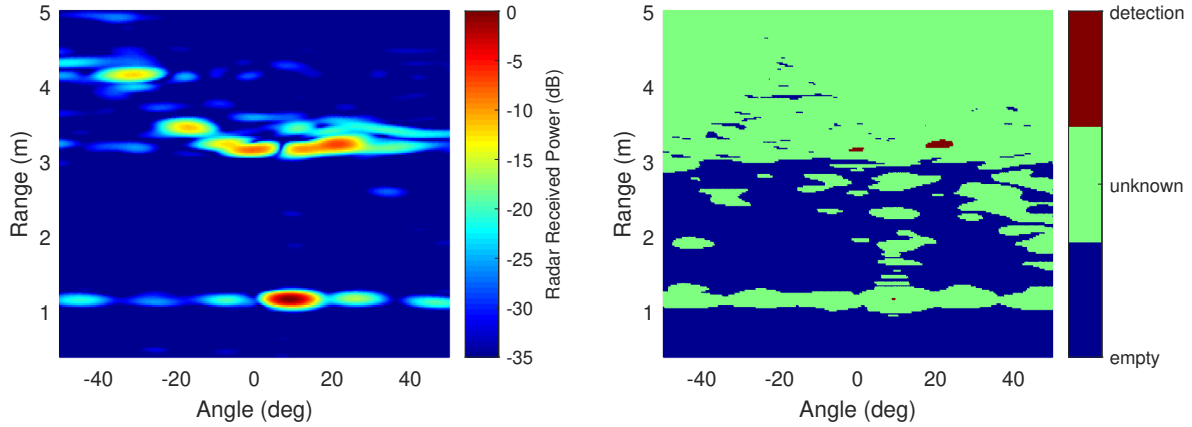
The radar sequentially transmits a pulse from each of its 4 transmit antennas and uses the signals received at the 8 receiving antennas to construct a virtual antenna array with 29 elements (due to element overlap in 3 positions). By performing a Fourier transform across the time-dimension and a Fourier transform across the array-dimension of the radar data, multiple beams are formed simultaneously, and a range-angle map is extracted. The range-angle map for a point target is given in Equation (6.1), which is derived in Chapter 3.

$$\begin{aligned} |m(f, f_a)| =& A\tau L \operatorname{sinc}(\pi\tau(f - 2kt_d)) \operatorname{sinc}(\pi L(f_a - f_c \sin\theta_d)) \\ |m(r, \theta)| \propto& \operatorname{sinc}\left(\frac{2\pi\beta}{c}(r - r_d)\right) \operatorname{sinc}(\pi L f_c(\sin\theta - \sin\theta_d)) \end{aligned} \tag{6.1}$$

When processing the radar data, a Kaiser window function was used for sidelobe suppression. Figure 6.4a shows an example range-angle map after beamforming of the radar data. An obstacle is present at 10^o and approximately 1m range and the wall of the measurement enclosure is present at approximately 3m. The field-of-view of the radar is restricted in software to be $\pm 50^o$ from the array boresight direction, as the azimuthal extent of the bins increases with azimuth angle. This sample measurement is taken from experimental trials detailed in Section 6.7.

6.3.2.2 Detection Thresholds

When propagating perception from an immediate sensory perception (the radar range-angle map) to a working memory, much of the information from the sensory perception is discarded and only the information required to perform the task at hand is kept. The retention of information can be considered as an attentional process whereby choices are



(a) An example range-angle map formed after beamforming of the raw radar data.

(b) An example determination map formed by application of the dual thresholds to the range-angle map.

Figure 6.4: Signal processing of a sample measurement, in polar coordinates centred on the radar. An obstacle is present at 10^o at approximately 1m range and the wall of the measurement enclosure is present at approximately 3m.

made about which information is important to parse [158, 159].

As the robot is trying to avoid obstacles and navigate through empty space, it is important to retain the regions of high power in the radar range-angle map, which are likely caused by the presence of an obstacle, and regions of low power, which are likely to be unobstructed. To determine which information to retain, the robot uses a pair of thresholds acting on the range-angle map to indicate whether a range-angle bin is occupied, vacant, or an intermediate state. The advantage of using a pair of thresholds is best seen in the context of a memory which can integrate multiple isolated perceptions as will be presented in Section 6.4.

In short, the two thresholds prevent the robot from making hard-decisions about the occupancy state of intermediate bins in the range-angle map. The main benefit of applying dual thresholds to the continuously varying range-angle map is that the resulting data is compressed to three states: occupied, unknown, and vacant, and the reduced data size allows the robot to process the information more quickly.

The upper and lower thresholds are given in dB relative to the direct signal from a transmit to a receive antenna element, in Equation (6.2) and Equation (6.3) respec-

tively, where P_u and P_l are the fixed upper and lower threshold values at a range of 1m respectively, and R is the round-trip distance to the target. The range is selected for mathematical convenience and corresponds to the order of magnitude of ranges that are important for the robot to accurately process. Points in the range-angle map with a power that exceeds the upper threshold are considered occupied and points that have a power that falls below the lower threshold are considered vacant.

To set these thresholds, it would be possible to carry out a full analysis of the statistics of the detection process and consider the desired detection rate and false alarm rate in the context of the dual thresholds. However, as this is preliminary research a detailed statistical analysis was beyond the scope of the work, and values were instead chosen empirically based on the results of experimental trials, and are detailed in Section 6.7.

$$t_+ = P_u + 20\log_{10}\left(\frac{1}{R}\right) \tag{6.2}$$

$$t_- = P_l + 20\log_{10}\left(\frac{1}{R}\right) \tag{6.3}$$

The locations of occupied and vacant points are stored in memory as a determination map which is used by the robot in subsequent processing. The determination map is the final stage of unimodal association for the radar data, and is in a polar coordinate reference frame that is centred on the radar sensor. The determination map stores the positions of occupied and vacant points. Any other points are declared as 'unknown' in the map (which avoids a hard decision about their state). It is also likely that multipath effects contributed to some of the unknown points on the map.

Figure 6.4b shows an example of the determination map for the range-angle map in Figure 6.4a. The main peak of the obstacle is detected, as well as two points on the wall of the enclosure. The area up to 3m is largely reported as empty, which was correct for this sample measurement. The obstacle sidelobes and much of the area beyond 3m were treated as having an unknown occupancy as the signals were above the lower threshold

at these points, but not large enough to cross the upper threshold. Some of the areas in the empty region below 3m that were marked as unknown are likely to originate from the enclosure walls on either side of the robot, but the scattering was not bright enough to cause the area to be marked as occupied.

6.3.2.3 Position Estimation

The perception the robot has of its wheel positioning is used to estimate its pose (a description of its position and orientation) which enables the robot to understand its relative position and angle to objects in the environment. The change in pose of the robot can then be determined by considering the equations of motion for a two-wheel, fixed-axle robot [168].

Encoders on the wheels record the number of ticks, n_l and n_r for the left and right wheels respectively, where a full revolution comprises of 144 ticks. The distance that each wheel travels may then be determined by multiplying the number of ticks by the distance travelled per tick, s.

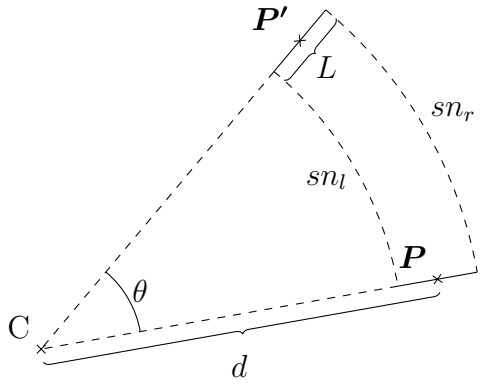


Figure 6.5: Geometry for determining robot motion.

The movement of the robot can be considered as a rotation of the starting pose, $\boldsymbol{P}$, about a point, C, as shown in Figure 6.5. To determine the final pose, $\boldsymbol{P'}$, the radius of rotation, d, and the angle of rotation, θ, must be found. Using the definition of the arc length of a circle allows equations for the arc lengths of the two wheels to be written as in Equation (6.4) and Equation (6.5), where L is the distance between the wheels. Solving

these gives Equation (6.6) for the angle of rotation and Equation (6.7) for the radius of rotation.

$$\theta\left(d-\frac{L}{2}\right)=sn_l \tag{6.4}$$

$$\theta\left(d+\frac{L}{2}\right)=sn_r \tag{6.5}$$

$$\theta=\frac{s}{L}\left(n_r-n_l\right) \tag{6.6}$$

$$d=L\frac{n_r+n_l}{n_r-n_l} \tag{6.7}$$

The initial pose, $\boldsymbol{P}$, comprises the Cartesian coordinates, x and y, and the heading, ϕ, of the robot as given in Equation (6.8). The relative pose of the robot to the centre of rotation, $\boldsymbol{V}$, is written in Equation (6.9). The rotation matrix, $\boldsymbol{M}$, which describes a rotation of a pose about the centre of rotation is given in Equation (6.10). The new pose, $\boldsymbol{P'}$, is then calculated by rotating the relative pose and adding the initial pose according to Equation (6.11).

$$\boldsymbol{P}=\begin{bmatrix}x\\y\\\phi\end{bmatrix} \tag{6.8}$$

$$\boldsymbol{V}=\begin{bmatrix}d\sin\phi\\d\cos\phi\\1\end{bmatrix} \tag{6.9}$$

$$\boldsymbol{M} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & \theta \end{bmatrix} \tag{6.10}$$

$$\boldsymbol{P'} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & \theta \end{bmatrix} \begin{bmatrix} d\sin\phi \\ d\cos\phi \\ 1 \end{bmatrix} + \begin{bmatrix} x \\ y \\ \phi \end{bmatrix} = \boldsymbol{MV} + \boldsymbol{P} \tag{6.11}$$

There are hardware limitations on this technique, for example due to the quantized nature of the motor encoder ticks as well as gradual wheel slippage, and although the errors that accumulate over the duration of the trials are small, they are of significance when considering the robustness of the working memory.

6.3.3 Polymodal Association

Polymodal Association is the section of Fuster's perception-action cycle which refers to the integration of multiple independent sensory perceptions to form a higher-level perception of a creature's surroundings. For the robot, the two independent sensory perceptions are the radar-based determination map (introduced in Section 6.3.2.2) and the motor-based estimate of its current pose (introduced in Section 6.3.2.3). These separate perceptions can be combined by performing a sensorimotor transform, which combines spatial and other sensory information [169, 170].

The reason for combining perceptions in this way is that the radar-based determination map of the environment is centred on the radar sensor and is in a polar coordinate system. Because the robot moves between measurements, the origin of this coordinate system also moves, which makes it difficult to compare different radar-based perceptions. Combination of sequential measurements is required to build up a working memory. If the determination map is projected into the fixed Cartesian world space using the robot's pose, it means that different radar measurements can be compared and combined into a

memory, as discussed in the next section.

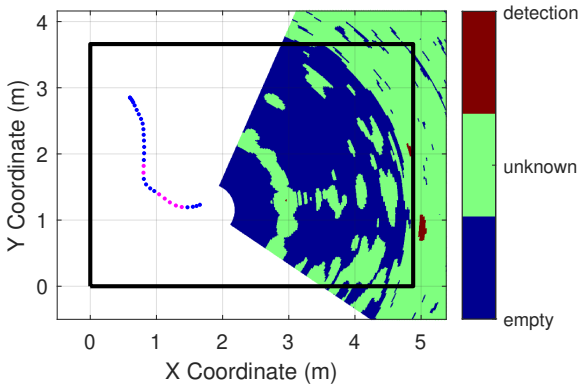
The coordinates are transformed from the radar polar coordinate frame, (R, θ), to the world frame, (X, Y), using Equation (6.12) which describes the sensorimotor transform, where x,y, and ϕ are taken from the robot pose relative to the origin located at one corner of the enclosure and r_0 is the radial offset of the radar from the centre of the robot in the forward direction. Figure 6.6a shows the sample determination map from Figure 6.4b after the sensorimotor transform, with the position (and historical track) of the robot shown in the Cartesian world space.

$$\begin{aligned} X &= x + r_0 \cos\phi + R\cos(\theta - \phi) \\ Y &= y - r_0 \sin\phi - R\sin(\theta - \phi) \end{aligned} \tag{6.12}$$

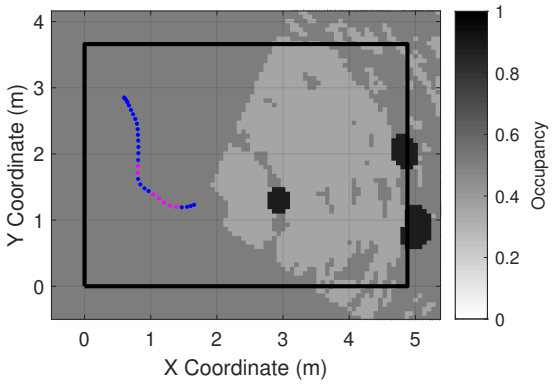
6.4 Memory

With the formation of an immediate sensory perception complete in the world frame, the robot is able to propagate this information up the cognitive hierarchy and build a working memory. The main reason for using a working memory is that it allows the robot to remember obstacles that it once detected that fall outside the field-of-view of the radar, and continue to avoid them. An additional benefit of the working memory is that it may reduce the robot's sensitivity to measurement errors, and this could be the subject of future statistical analysis.

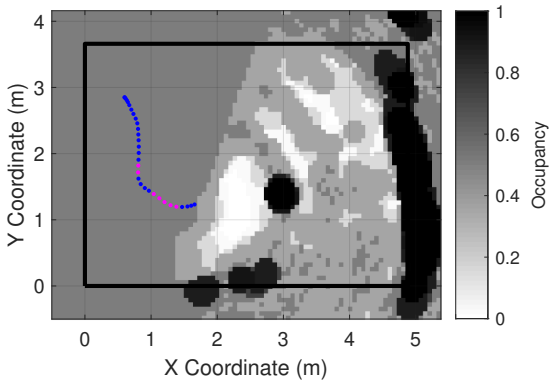
The radar-based determination map is registered to the fixed world space and so the robot is able to easily integrate multiple maps to form the working memory. The memory takes the form of an occupancy grid, which is populated with values derived from the radar-based determination map described in Section 6.3.2.2. The occupancy grid is an array which is mapped to the world space, with each grid cell containing a value that represents the robot's knowledge of occupancy at that point. A high occupancy value close



(a) Determination map registered to the world space after performing the sensorimotor transform.



(b) Working memory occupancy grid updated using the sample determination map.



(c) Working memory occupancy grid showing the result of integrating five determination maps.

Figure 6.6: Signal processing of a sample measurement. The dots indicate points at which a measurement was made, blue signifies avoidance behaviour and pink signifies goal finding behaviour. The black border represents the enclosure walls.

to 1 indicates that there is a high probability that the grid cell is occupied, while a low occupancy value close to 0 indicates that the grid cell is likely to be vacant. The occupancy values are between 0 and 1, and can be compared against thresholds to determine the presence or absence of an obstacle in a cell. The population of the occupancy grid can be time intensive when processing a large volume of data and so the attentional processing of the range-angle map into the smaller determination map allows the formation of working memory to occur quickly enough and for the robot to operate in real-time.

Equation (6.13) describes the reinforcement rules for generating the occupancy grid, m, where i and j are variables that index the grid and k_+ and k_- are rates of positive and negative reinforcement respectively. The reinforcement rules are designed so that the values in the grid cannot be larger than 1 or less than 0. This constraint is achieved by forcing the rates of reinforcement to be between 0 and 1. Occupied points in the determination map positively reinforce the corresponding grid cell, increasing its value, and vacant points in the determination map negatively reinforce the corresponding grid cell, decreasing its value.

$$m_{ij} = \begin{cases} m_{ij} + k_+ \left(1 - m_{ij}\right), & \text{if occupied} \\ m_{ij} = m_{ij} - k_- \left(m_{ij}\right), & \text{if empty} \\ m_{ij} = m_{ij}, & \text{otherwise} \end{cases} \tag{6.13}$$

Figure 6.6b shows an example of the working memory after using the single sample measurement to populate it. In the working memory, the obstacle before the robot is visible, as are two scattering centres from the far wall of the enclosure. The world exists beyond the field-of-view of the radar, but remains unpopulated and holds the default occupancy value of 0.5.

However, by integrating multiple determination maps into a single working memory map, a more complete representation of the world can be achieved by the robot. Integration of multiple measurements begins by using the most recent determination map (the n^{th} map) in conjunction with the update rules to update the memory map. Then the next most recent determination map (the $n - 1^{\text{th}}$ map) is used to update the memory map again. This process repeats until the desired number of measurements have been used to update the map. This number, N, represents the memory length of the working memory.

Figure 6.6c shows the result of integrating five measurements in the working memory. In this working memory, the full length of the back wall of the enclosure is visible, the wall to the right of the robot which is outside of its field-of-view is visible, and the location of

the immediate obstacle is clear. By building a more complete and consistent perception of the world, the robot is able to make more informed decisions about its actions. The length of the working memory is the number of determination maps that are integrated into the working memory, and will be seen to be of central importance.

The combination of the occupancy grid and the dual thresholds introduced in Section 6.3.2.2 reduces the effect of erroneous detections on the robot's performance. For example, if a particular region has been determined to be vacant for multiple measurements, the corresponding occupancy value will be low. A single spurious detection in the vacant region will not be sufficient to 'convince' the robot that there is an obstacle present. If however, the source of the detection is a real obstacle, it will be present over multiple measurements and the occupancy map will be positively reinforced, 'convincing' the robot that the space is occupied. Similarly, if an obstacle is present but not reflective enough to cause a detection at long range, the area in the determination map will be considered unknown and the occupancy map will retain the default state of 0.5. As soon as the robot is close enough for the obstacle's reflection to exceed the detection threshold, the occupancy grid will be positively reinforced and the robot will consider the space occupied. Of course, ensuring the signal to noise ratio is adequate through the radar design is as important as ever.

6.5 Reasoning and Action

With a fully-formed and integrated perception of its environment within the working memory, the robot must decide how to proceed with the task at hand. There are many different models of how the brain reasons, but for the purposes of this chapter it is sufficient to observe that reasoning is a cognitive function that satisfies the rostral prefrontal step in Fuster's perception-action cycle [171, 172, 173]. The steering scheme presented in this section is not necessarily suggested to be an accurate representation of any biological system or an ideal strategy. Rather, the scheme aims to provide a simple and effective

navigation strategy that enables the robot platform to successfully perform its task.

Ranges to obstacles in arbitrary directions from the robot are determined from the working memory. To do this, a vector at the required angle from the robot is formed in the memory, and the first occupied cell which lies along the vector is used to determine the range. A cell in the occupancy grid is considered occupied if it exceeds a threshold of 0.65 and vacant if it is below a threshold of 0.2. These values were set as the default values for the MATLAB occupancy grid class, and were found to be adequate.

The robot's movement is split into a speed component, S, and a rotation component, T. For a particular motion, the robot's wheels move at $S + T$ and $S - T$ for the left and right wheels respectively. The speed component is used to increase the time to collision in the forward direction which increases the number of observations the robot can make before colliding with an obstacle. The speed component requires minimal reasoning to determine, and is purely dependent on the distance from the robot to an obstacle in the forward direction, R_f. Equation (6.14) defines how the robot determines its linear speed, where D_s is the distance at which the robot's linear speed will drop to 0, and k_s is the coupling constant between range and speed.

$$S = k_s(R_f - D_s) \tag{6.14}$$

As the rotation component is responsible for both turning away from obstacles and turning towards the goal position, it requires more complex reasoning than the speed component, and so a strategy for deciding which type of behaviour to exhibit is important. Figure 6.7 provides a flowchart that describes the robot's decision-making process for the rotation component of its motion.

The robot begins by 'looking' in the direction of the goal (which may be outside of the current swath of the radar, but still present in the working memory) to determine if it can perceive an obstacle between it and the goal position. If it has line-of-sight to the goal (but has not reached it yet) and none of its other range measurements are less than the minimum allowed safe distance, r_s, the robot proceeds with its goal-oriented behaviour

and turns towards the goal position. k_θ is a coupling constant between angle to the goal, θ_g, and the speed of the robot's rotation.

If the robot does not have line-of-sight to the goal position, or if it is below the minimum safe distance, it performs obstacle avoidance by considering the relative difference between range measurements in the left and right beam, R_l and R_r respectively. The robot turns to reduce the difference between the ranges in the two lateral beams, where k_T is the coupling constant between relative range difference and rotational speed.

With the reasoning between goal-finding and obstacle avoidance complete, the robot's motion is converted to speeds to transmit to the motor controller, equivalent to the premotor step in Fuster's perception-action cycle, and then transmitted to the motor controller which actions the instruction. To complete the perception-action cycle, the motor controller can be considered the primary motor step in Fuster's cycle.

The action step concludes the robot's perception-action cycle, which it continues to repeat until it reaches the goal position. Throughout the cycle, the cognitive elements of Perception, Attention, Memory, and Reasoning are used to manipulate and interrogate raw sensory information.

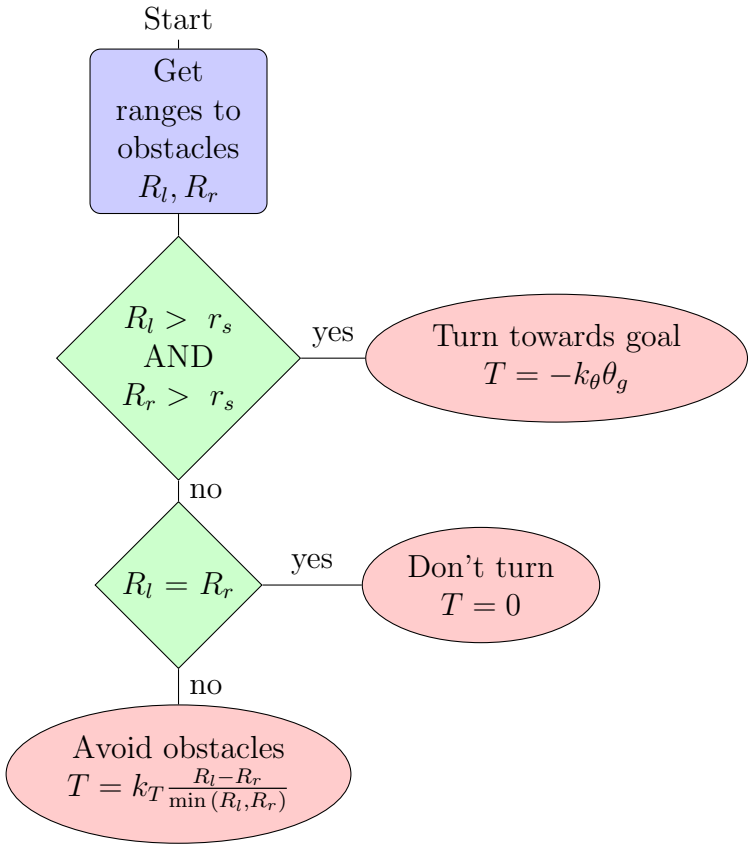
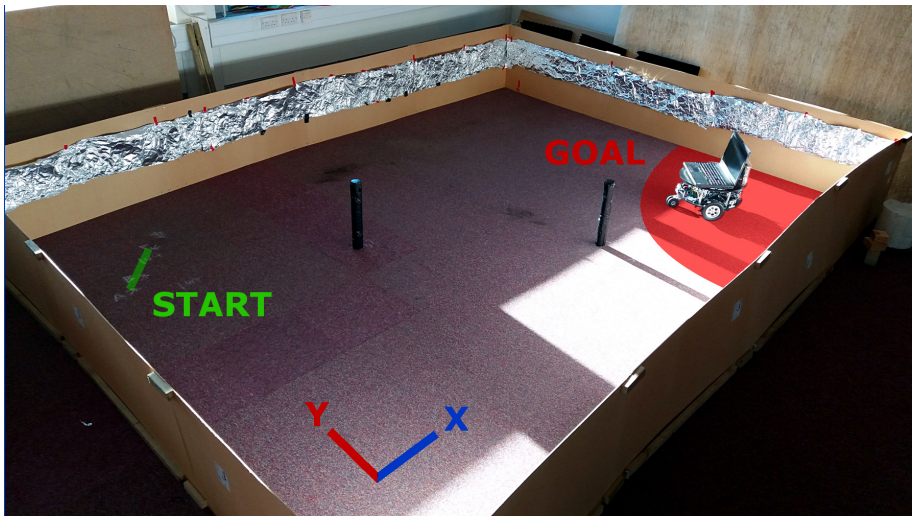


Figure 6.7: Flowchart to show the control rules of the robot for the rotation component of its movement.



(a) The enclosure used for the trials with obstacles and robot present.

(b) Obstacle close-up.

(c) The initial view of the robot.

Figure 6.8: Images to show the configuration of the robot enclosure including the obstacles used to obstruct the route of the robot.

6.6 Simulation

A simulation was written to test and validate the algorithm presented in the previous sections, and to investigate the effect of working memory on the robot's performance. The simulated trials were conducted with changes made to the memory length of the robot. The memory length refers to the number of sequential determination maps (originating from single radar pulses) that are integrated by the robot to form the working memory.

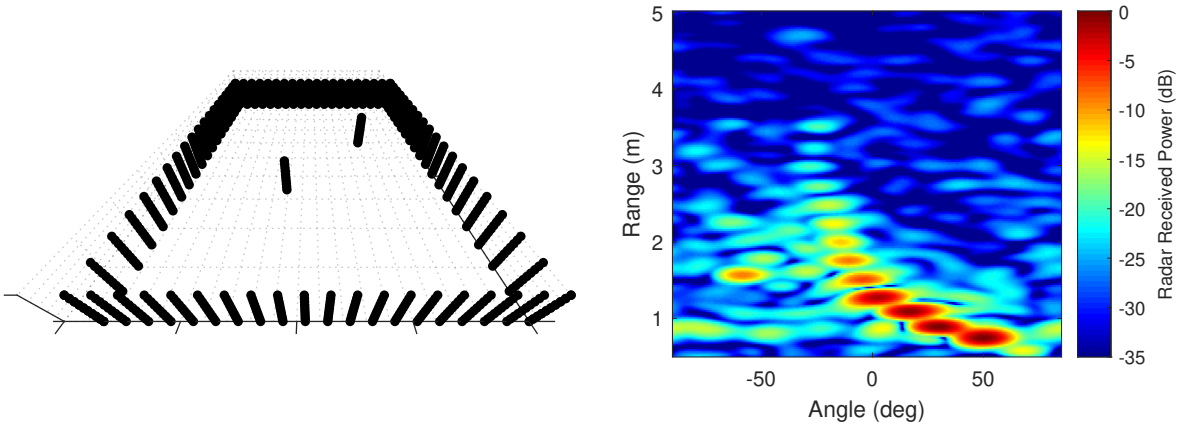
The simulation was written to be easily transferred to a laboratory experiment for direct comparison. The simulated robot had to navigate from some initial position to a goal position in virtual space avoiding two obstacles placed in its path. The trial was designed to require both obstacle avoidance and goal-finding behaviours from the robot in order to successfully complete the trial. Obstacles and walls were simulated with closely spaced point scatterers as shown in Figure 6.9a. The simulation was carried out in 2D, and the 3D perspective of the figure is for visualisation only. Multipath reflections and reflector occlusion effects were not included in the simulation.

A program emulating a 77 GHz MIMO radar (corresponding to the hardware that was available) was written for use in the simulation and an example output of this is shown in Figure 6.9b. Figure 6.9c shows the determination map for the simulated data, and Figure 6.9d shows the occupancy grid for the simulated robot during a trial, which had a cell size of 6.7 cm. The field-of-view of the radar was restricted to be $\pm 50^o$ from the boresight direction as the azimuth resolution of the radar becomes poor beyond this angle.

The power thresholds, reinforcement rates, and coupling constants used for the trial are given in Table 6.3. These values and the occupancy grid cell size were selected by iterating over multiple preliminary trials with different values and selecting those that gave the best observed performance.

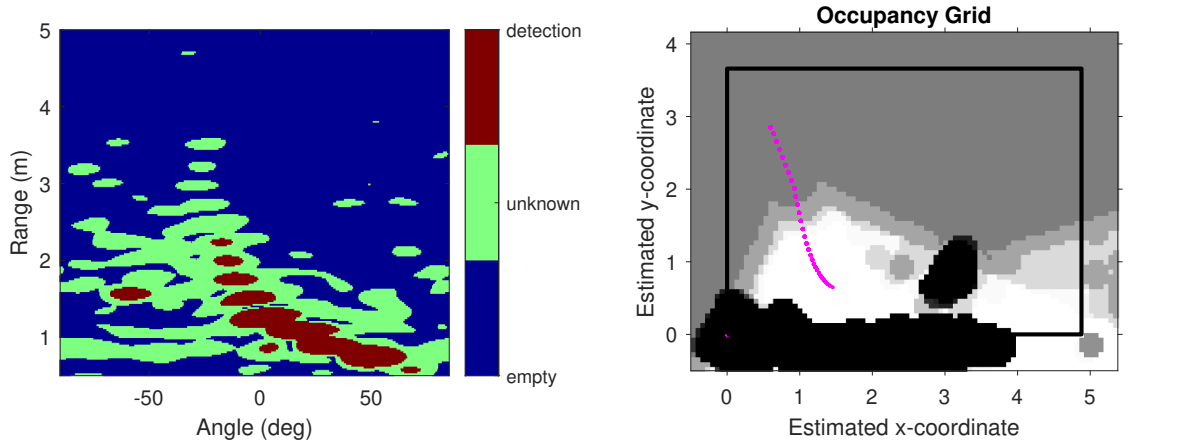
Table 6.3: Experiment and Simulation Parameters.

Parameter	Description	Value	Units
k_+	Positive reinforement rate	0.75	-
k_-	Negative reinforcement rate	0.3	-
P_u	Upper threshold at 1 m	-69	dB
P_l	Lower threshold at 1 m	-106	dB
k_s	Coupling of range and speed	20	ticks/s/m
k_θ	Coupling of angle to goal and rotation rate	20	ticks/s/radian
k_T	Coupling of range difference and rotation rate	5	ticks/s



(a) Simulated enclosure using multiple point scatterers in 3D space.

(b) Simulated 77GHz MIMO radar data.



(c) Detection map for the simulated measurement.

(d) Occupancy grid during a simulated trial.

Figure 6.9: Example plots from a simulated trial.

6.7 Experiment Design

As with the simulated trials, the goal of the experimental trials was to investigate the effect of the working memory on the robot's performance. As with the simulation, experimental trials were conducted with changes made to the memory length of the robot. The three conditions tested were: a memory length of 1 corresponding to no memory (since a single radar pulse is used to form a single range-angle map), a memory length of 15 for a short memory, and a memory length of 25 for a longer memory, all other steps of the process remained the same. Ten trials were conducted for each memory condition. Of course the amount of processing time taken to form the occupancy grid increases proportionally with memory length, however in all trials the robot navigated in real-time.

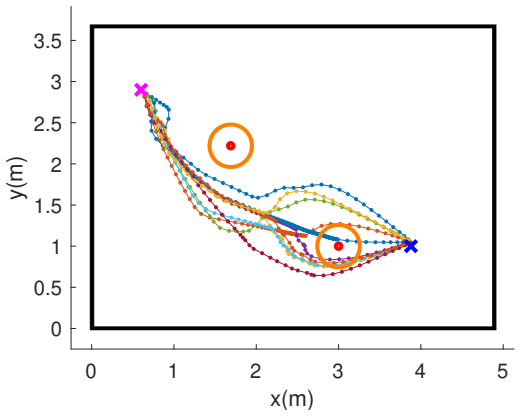
The experiment was designed with the same approach as the simulation; an enclosure measuring 4.88 m by 3.66 m was built and is shown in Figure 6.8a. The internal walls of the enclosure were covered with rough metal foil to provide diffuse scattering of the incident signals. The goal of the robot was to navigate from one corner of the enclosure to the other (left to right in Figure 6.8a). As indicated in Figure 6.8a, the Cartesian coordinate system is defined with the origin at the corner nearest the photographer and with the x dimension extending along the longer wall. The robot began the task at coordinate (0.6,2.85) with a heading of -45^o and was required to reach within 10 cm of the ending location, at coordinates (3.66,1.22).

To prevent the robot from being able to travel in a straight line between the start and end points, two metallic obstacles were placed within the enclosure mimicking the point scatterers present in the simulation. The obstacles were pillars that had a height of 46 cm which intersected the plane of the radar (35 cm above the ground). The pillars had a diameter of 6.6 cm and the surface was dented to present a radar cross section (RCS) that varied with viewing angle, making the task more realistic and challenging. A detailed image of one pillar is given in Figure 6.8b. The pillars were placed as shown in Figure 6.8a and Figure 6.8c to obstruct the route of the robot to the goal position. The placement of the pillars was such that the pillar nearest to the robot (the 'first' pillar)

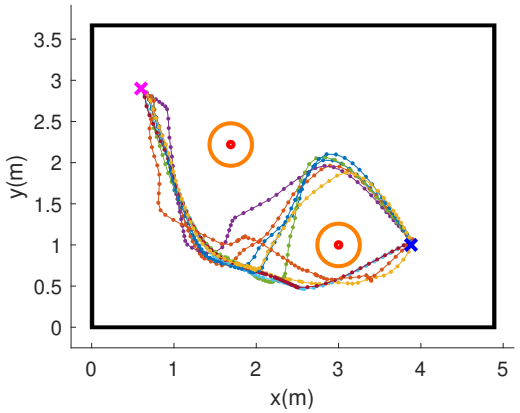
partially occluded the more distant pillar (the 'second' pillar) from part of the radar array.

The power thresholds, reinforcement rates, and coupling constants used for the trial are the same as the simulations and are given in Table 6.3. P_u and P_l are measured relative to the direct signal intensity from the transmit to the receive antenna elements. These values were selected by iterating over multiple preliminary trials with different values and selecting those that gave the best observed performance.

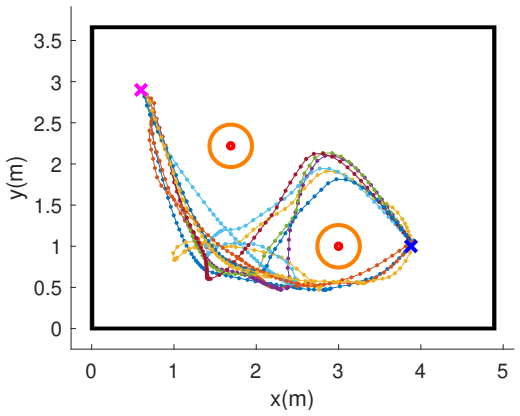
A trial was considered a success if the robot reached the goal position without touching either pillar or the enclosure walls. Conversely, a failed trial was recorded if the robot touched either the pillar or the walls. An intermediate state could occur if the robot reached the maximum allowed number of frames for the trial without reaching the goal, but this state did not occur.



(a) Memory Length 1.

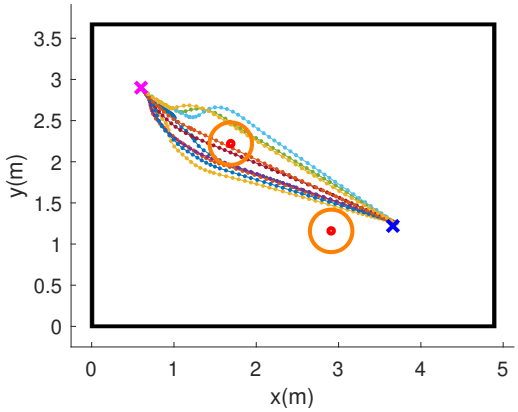


(b) Memory Length 15.

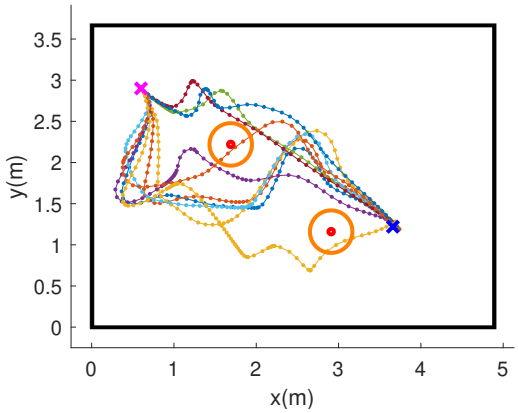


(c) Memory Length 25.

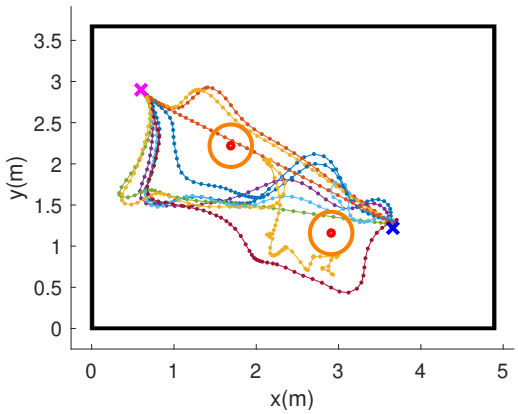
Figure 6.10: Simulated results for 10 trials under each trial condition, where memory length is the only difference between trials. The pink cross represents the starting point and the blue cross represents the end point. The small red circles represent the location of obstacles and the outer orange circles indicate the location of the object plus the radius of the robot such that a track passing through the outer circle indicates a collision.



(a) Memory Length 1.



(b) Memory Length 15.



(c) Memory Length 25.

Figure 6.11: Experimental results showing the robot ground track estimates for 10 trials under each trial condition, where memory length is the only difference between trials. The lines represent the robot track, with dots indicating the positions at which radar measurements were made. The pink cross represents the starting point and the blue cross represents the end point. The small red circles represent the location of obstacles and the outer orange circles indicate the location of the object plus the radius of the robot such that a track passing through the outer circle indicates a collision. The ground tracks are not a perfect representation of the robot's position, and collisions may appear to have occurred when they have not.

6.8 Results and Discussion

This is a new area of research and the trials conducted in this chapter are restricted in scope. As a result, these results are preliminary and there are many avenues of potential development. However, the aim of these trials is to generate sufficient evidence to evaluate the potential benefits of working memory in robotic guidance and to act as a stimulus for future efforts. A summary of the number of successful and failed trials for both simulation and experiment is given in Table 6.4, the mean trial durations are given in Table 6.5, and an aggregated view of the robot ground tracks for each condition is given in Figure 6.10 for the simulation and Figure 6.11 for the experiment.

Under all conditions, both simulated and experimental, the robot took different routes between trials. This is due to a combination of slight variations in initial conditions, including the pose at which the robot was placed; differences in the environment such as target orientation in the experimental trials; and noise effects causing bifurcation of behaviours often seen in complex systems.

The least successful trial condition was the no memory condition, in which the majority of trials were failed by the robot in both simulation and experiment. In the simulation, the first obstacle was successfully detected and avoided, but the robot collided with the second obstacle. In these failed trials, the robot began to take avoidance action around the second obstacle, however the robot turning away from the obstacle led to it leaving the field of view of the radar. With no memory of the obstacle, the robot turned back towards the obstacle, but was too close to detect the target again (as it fell below the minimum detection range of 0.5 m). As a result, the robot collided with the second obstacle. During

Table 6.4: Number of successful and failed trials for each condition.

	Simulation		Experiment	
Memory Length	Successes	Failures	Successes	Failures
1 (No Memory)	4	6	4	6
15 (Short Memory)	10	0	9	1
25 (Long Memory)	10	0	7	3

Table 6.5: Average number of measurements for each trial condition.

Memory Length	Simulation	Experiment
1 (No Memory)	59	59
15 (Short Memory)	70	66
25 (Long Memory)	80	65

the no memory experimental trials a similar behaviour was observed that led the robot to collide with the first obstacle. Another cause of failure in the experimental trials was a lack of consistent detection of the obstacles, which in several trials led to the obstacle being too close to the radar to be detected and subsequent collision with the robot. The lack of consistent detection contributed to collision with the first obstacle, rather than collisions with the second obstacle as seen in simulation.

In the short memory condition the robot exhibited more successful behaviour and was able to reach the goal without collision in all of the simulated trials, and in all but one of the experimental trials. In the simulated trials the robot followed a very similar initial route to the right of the first obstacle. On the approach to the second obstacle, the robot's route bifurcated and the robot succeeded in reaching the goal position in all cases without collision. In all but one of the experimental trials the robot detected and avoided the first obstacle. For the trials in which the robot proceeded to the left of the initial obstacle, it was then able to proceed almost directly to the goal without need for further avoidance behaviour. In trials where the robot passed to the right of the initial obstacle, the second obstacle was also avoided successfully. The ground tracks in Figure 6.11 indicate that two collisions occurred, whereas Table 6.4 states that only one occurred. This apparent discrepancy is due to the inaccuracy of the recorded position of the robot. The yellow track in Figure 6.11 did not result in a physical collision.

In the long memory condition, the robot succeeded in all of the simulated trials and in the majority of the experimental trials. In the simulation trials, very similar behaviour was observed as in the short memory condition, but the robot exhibited more sensitivity to measurement errors in the experimental trials. In one particular trial, denoted by a yellow track in Figure 6.11c, the robot followed an erratic path and collided with the wall

of the experiment enclosure. This erratic behaviour was a result of erroneous detections persisting in the memory for a long time, causing unnecessary avoidance manoeuvres.

The duration of the simulated trials was also notably affected by the memory length. As the memory length was increased the duration of the simulated trials increased significantly. This increase is a result of the more circuitous route taken by the robot, which is due to the persistence of objects in the working memory: a feature that significantly assists with obstacle avoidance but increases the duration of the trial. Ultimately however, this is due to a combination of erroneous detections, lack of consistent detections, and position error accumulation. Detection performance could be improved by carrying out a thorough statistical analysis of the detection thresholds and the relationship between this and the accumulation of detections in memory, and by using that analysis to determine the detection scheme. Positional errors could be decreased by using an external positioning system or a standard approach to ego-motion estimation. If these factors were addressed, the memory length could probably be increased, perhaps arbitrarily.

Memory in the cognitive functioning of the robot is essential for the robot to maintain a perception of obstacles outside of its field-of-view so that it can take appropriate obstacle avoidance actions. The implementation of memory described in this chapter is heavily influenced by the parameters chosen for the other cognitive processes of the robot. For example, if the radar measured more frequently, then a particular memory length would correspond to a shorter distance travelled by the robot than in the current implementation, and would have less accumulation of positional errors as a result.

However, it is still possible to come to some general conclusions about using a short-term memory: if the memory length is too long the robot is less able to recover from erroneous detections resulting from clutter, sidelobes, or multipath effects as they persist in the robot's working memory for a greater number of steps and can lead to collisions. Further, as the memory length increases, the cumulative positioning error within the integrated measurements could become significant which would further inhibit the robot's success.

6.9 Summary

This chapter has presented an example cognitive architecture for radar control of an autonomous robotic platform. The cognitive processing scheme contains several important elements of cognition including Perception, Memory, Attention, and Reasoning and relates these to functional elements of Fuster's perception action cycle.

Through simulation and experiment, this chapter has shown the structural benefits of including a working memory in a cognitive radar system. In general, memory allows a cognitive radar system to gradually develop its understanding of its environment, decreasing its overall sensitivity to erroneous measurements and enabling it to maintain perception outside of its immediate field-of-view. The stability of perception is an advantage for any cognitive system that is making rapid control decisions.

For the autonomous navigation task, the performance improvement with memory length was related to the ability of the robot to maintain a perception of an obstacle when it falls outside the radar's field-of-view. The persistence of obstacles in memory allowed for better navigation decisions to be made, which resulted in an increased number of successful trials. Some of the difficulties with using a working memory for autonomous navigation have also been explored, particularly the persistence of erroneous detections and accumulation of errors. For the robot the advantages of a working memory outweighed the disadvantages, but in some cases the accumulation of errors caused the robot to be unable to navigate effectively, leading to collisions.

6.10 Conclusion

This chapter has explored the possibilities of cognitive processing to leverage the degrees of freedom made available to a mobile radar platform. By taking inspiration from such a fundamental part of human experience, this chapter has shown how cognitive processing, particularly with regards to memory, can enhance the radar's quality of perception and significantly improve performance.

Chapter 7

Conclusion

All we have to believe with is our senses, the tools we use to perceive the world: our sight, our touch, our memory. If they lie to us, then nothing can be trusted. And even if we do not believe, then still we cannot travel in any other way than the road our senses show us; and we must walk that road to the end.

— Neil Gaiman

American Gods, 2001

7.1 Summary

This thesis has explored the processes of localization and echolocation in the natural world, and has made several central observations. Echolocators use wideband echolocation calls which have an extra degree of freedom that can encode information compared with narrowband calls. The mass of the head and the shape of the pinna attenuate different frequency components of a wideband sound signal by different amounts depending on the angle of incidence, and this can be described by the head-related transfer function (HRTF). Through the filtering process, the HRTF introduces binaural cues which can be used by the brain for sound source localization. The cognitive processes of the brain enable understanding of these cues, and the higher-level processes seen in mammalian echolocators enable complex behaviours and interactions with the environment.

By providing detailed analysis of human echolocator tongue clicks and their properties, this thesis has contributed to the understanding of human echolocation. This work found that the click signals have very wide beamwidths in both azimuth and elevation directions, almost uniformly illuminating the area before the echolocators. For the first time, the bandwidth of human echolocator tongue clicks was reported based on the analysis of a large database of tongue clicks and has contributed a new understanding of the diverse frequency components used by human echolocators. Using the descriptions of the different key components of human echolocator tongue clicks, an open access tool was developed and published to enable widespread synthesis of human echolocator tongue clicks.

Building on the central observations of echolocators in nature and research into human echolocator tongue click properties, this thesis has proposed and developed two distinct radar techniques.

The first technique is inspired by the wideband signals seen in nature's echolocators, binaural cues for localization, and the HRTF. This radar approach uses the information encoded in the spectral content of a signal to localize a target. The wide band of frequencies used by the radar provides an extra degree of freedom that the system can exploit for target localization. Power- and phase-based versions of this technique were developed,

with both approaches showing high angular accuracy over a wide angular range of operation. The phase-based approach outperformed the power-based approach in terms of accuracy, but is also expected to be less stable in noisy operating conditions.

The second radar technique presented in this thesis explored the possibilities of using cognitive processing to exploit the degrees of freedom made available to a mobile radar platform. By taking inspiration from the cognitive processes seen in mammalian echolocators, this work showed that a cognitive processing architecture can be used by a radar system. This research also investigated the effects of memory on a cognitive signal processing architecture and demonstrated that careful use of memory can significantly improve radar performance.

7.2 Core Contributions

Novel parametrization of human echolocator tongue clicks was carried out and resulted in the construction and release of open access synthetic human echolocator tongue clicks. The parametrization was carried out based on original analysis of a large database of human echolocator tongue clicks which was conducted in unprecedented detail. This analysis enhances understanding of the key echolocator click parameters and their variability.

A theoretical basis for a family of closely-related radar techniques has been developed, and these use the frequency diversity of the transmit waveform and spatial encoding by the receiving antennas to localize a target in angle. These techniques were verified and assessed through simulation and experiment, the results of which confirm the fundamental possibility that these techniques operate with high angular accuracy over a wide angular range using single measurements.

A novel autonomous cognitive radar system architecture for navigation was developed and tested using simulated and experimental results to assess its performance. The radar processing was designed to simulate and implement the functional blocks of cognition, which resulted in complex and variable emergent behaviours. To the author's best knowl-

edge this work is the first investigation into the effects of using memory in cognitive radar systems, which finds that some memory is useful for a cognitive radar, but that too much memory can hinder such a system.

7.3 Future Work

The exploration of biologically-inspired sensing techniques in this thesis has covered a wide range of approaches and has touched upon many interesting strands of research which could be developed further. This section sets out the most promising avenues for future research as the author sees them.

Future research into human echolocator tongue clicks could analyze the relationship between the centre frequencies of the components within the click spectrum. This research has the potential to develop synthetic tongue clicks that are both realistic and variable, enabling a wider variety of applications than the current model.

Future work on the frequency-diverse radar techniques could focus on integrating the power- and the phase-based approaches, which may improve the performance and robustness of angular localization. Integrating the two approaches should allow the correlation process to discriminate between closely spaced angles more reliably. The issues introduced by multiple closely-spaced targets merit more research as solving these would allow the radar to operate in more complex environments such as those with significant multipath and extended scatterers. Field trials would enable the $d << r_{tx}$ condition to be more convincingly satisfied and could provide useful insight into how this technique could be implemented in an operational environment, particularly with regard to the signal-to-noise ratio (SNR) thresholding approach which shows promising initial results.

Optimisation of the antenna patterns could provide a way to reduce the bandwidth of the technique to a practical bandwidth, and trials with multiple narrowband radars could allow the technique to become more practical in the wider world. Because the localization technique is independent of transmit signal (other than for meeting SNR

requirements), it is possible that this approach could be used passively, which would eliminate the requirement of generating a large bandwidth signal and provides another avenue of research.

Future work developed from the research into cognitive robotics could explore in more detail the effects of memory length on specific types of errors and how these errors accumulate and contribute to the overall performance of an autonomous cognitive radar platform. Specifically, a detailed statistical analysis of the detection process would enable better understanding of how detection parameter choice affects the robot's overall performance. This approach could be extended to include detailed analysis of the effects other parameter choices, including the occupancy grid resolution, motion coupling constants, and memory duration.

It would be especially interesting to explore the effects of including a long-term memory in the signal processing, which could be achieved by consolidating the observations made over the duration of a trial into a more complete map of the environment for use in a subsequent trial. This processing could be performed 'offline' and more computationally intensive techniques could be used to reduce the effects of errors. The use of a memory that persisted between trials could enable learning behaviour to be exhibited by the robot, and could lead to performance improvements as the robot increased its knowledge of the trial environment. Further, if the robot was given control of its own parameters, it may be possible to create a system that optimizes its parameters over multiple trials based on its interactions with the environment.

The treatment of attention in cognitive radar systems merits more research. While it is sometimes implemented as beamforming, attentional processing could also include different post-processing schemes such as the thresholding scheme presented in this thesis. The choice of processing scheme merits more research as this choice will affect the nature of a cognitive radar system and will impact the radar's performance.

7.4 Conclusion

This thesis set out to find new approaches to radar systems by exploring the paradigm of biologically-inspired design, and aimed to find effective ways of using the redundancy found in modern radar systems. By identifying frequency-diversity and platform motion as candidate degrees-of-freedom for utilization, this thesis has proposed and developed two effective biologically-inspired radar techniques that extract and process information from radar systems in new ways. The paradigm of biologically-inspired design is perfectly suited to research into radar, and the author hopes that researchers will continue to take inspiration from the natural world for many years to come.

Appendix A

Human Echolocator Click Database Counts

Table A.1: The number of sanitized click samples (meaning clicks after filtering, clipping rejection, and manual rejection) for each expert echolocator under each trial condition for the azimuth measurements. Each sample consists of two recordings, one measured at a reference microphone which remained static throughout the different trials, and one measured by the signal microphone at the indicated position.

Echolocator	**EE1**	**EE1**	**EE3**	**EE3**	**EE2**	**EE2**
Distance (cm)	**40**	**100**	**40**	**100**	**40**	**100**
Azimuth (deg)	**Counts**					
-90	43	47	30	28	45	48
-80	44	50	28	28	45	45
-70	42	42	30	20	46	45
-60	40	45	24	27	46	49
-50	40	41	26	28	46	54
-40	43	44	28	26	45	59
-30	42	44	27	28	48	53
-20	40	45	31	28	108	50
-10	84	45	20	28	92	47
0	84	50	52	23	110	56
10	43	42	29	28	45	49
20	43	42	30	26	62	50
30	42	41	30	26	59	50
40	42	41	29	28	57	48
50	46	42	30	28	57	51
60	44	45	30	30	49	52
70	44	43	29	26	53	46
80	40	40	30	28	45	52
90	44	40	31	27	44	49
100	43	42	29	28	45	51
110	45	42	30	26	49	57
120	40	44	28	31	43	48
130	43	42	28	28	58	51
140	40	34	25	27	56	58
150	40	41	27	28	50	53
160	41	42	28	26	44	58
170	55	40	27	25	74	52
180	43	43	30	26	56	60
Totals	1280	1199	816	756	1577	1441

Table A.2: The number of sanitized click samples (meaning clicks after filtering, clipping rejection, and manual rejection) for each expert echolocator under each trial condition for the elevation measurements. Each sample consists of two recordings, one measured at a reference microphone which remained static throughout the different trials, and one measured by the signal microphone at the indicated position.

Echolocator	**EE1**	**EE3**	**EE2**
Distance (cm)	**40**	**40**	**40**
Elevation (deg)	**Counts**		
-140	29	20	36
-150	26	20	40
-160	29	20	39
-170	34	22	50
-180	31	20	33
170	33	20	35
160	35	20	41
150	33	20	41
140	36	20	43
130	42	19	40
120	36	19	43
110	29	20	35
100	30	20	37
90	39	20	38
80	35	20	35
70	37	20	36
60	38	20	35
50	36	20	35
40	31	20	40
30	32	20	47
20	32	20	46
10	31	20	35
0	32	40	69
-10	30	20	35
-20	30	20	33
-30	30	20	35
-40	29	20	33
Totals	885	560	1065

Appendix B

Robot Control Hardware

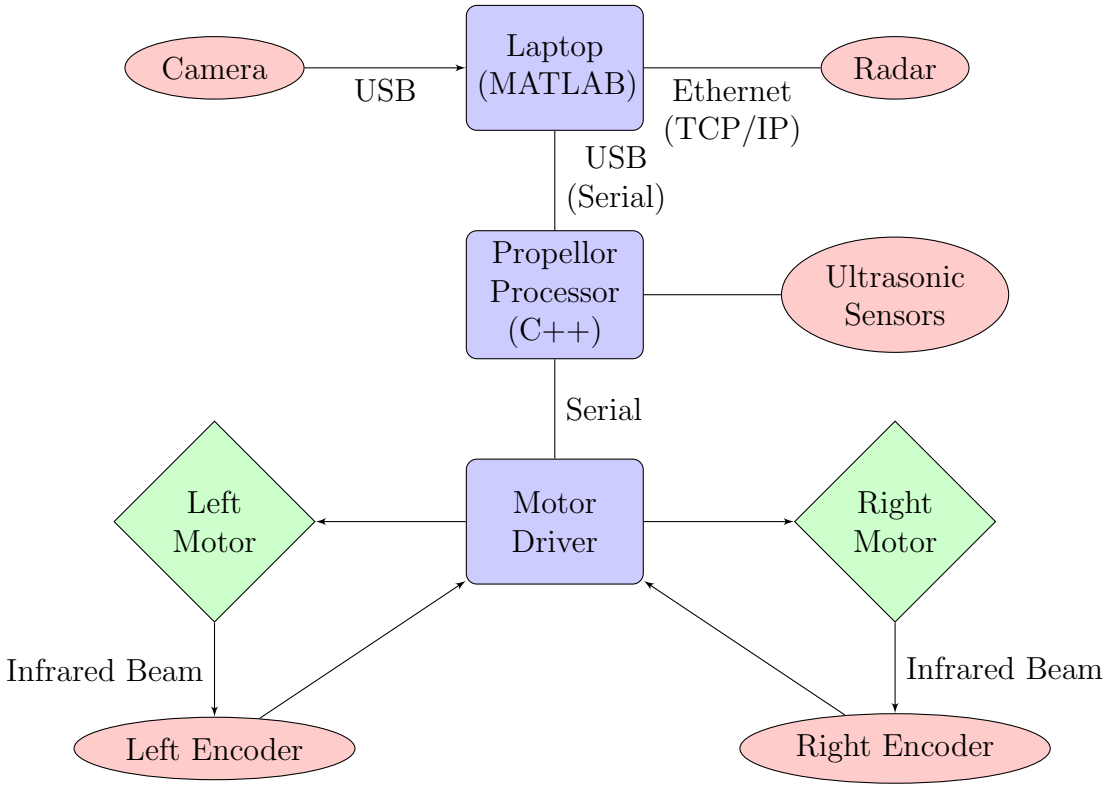


Figure B.1: Diagram to show the hardware components and connections of the robot.

This appendix details the hardware interfaces for the robot used in Chapter 6 and relates the flow of information to the MATLAB scripts which control the robot's functioning. An schematic overview of the connections for the robot is shown in Figure B.1.

B.1 Laptop Interfaces

B.1.1 Camera

The camera used for capturing the visual ground truth of the robot was a RICOH THETA S, which has a 360^o viewing angle [174] making it ideal for viewing the hemisphere before the robot. The following MATLAB code snippit indicates how the connection to the camera is made over a USB connection using the MATLAB Image Acquisition Toolbox.

```
vid = videoinput('winvideo', 4, 'RGB24_1280x640');
src = getselectedsource(vid);
vidSize =[320 0 640 640];
vid.ROIPosition = vidSize;
```

```
vid.FramesPerTrigger = 1;
hImage = image(zeros(vidSize(4),vidSize(3), vid.NumberOfBands);
preview(vid, hImage);
```

B.1.2 Radar

The INRAS Radarbook [165] communicates over ethernet, and is visible to MATLAB on the laptop as a TCP/IP object. The ethernet connection is handled by a MATLAB hardware class provided by INRAS.

B.1.3 Robot

The laptop communicates with the Propeller Activity Board WX [175] over a USB serial connection that is handled by a custom MATLAB class written for operating the robot. The serial connection has a Baud rate of 115200 and operates with 8 data bits.

B.2 Robot

B.2.1 Propeller

The Arlo robot package includes the Propeller Activity Board WX [175], which is a board based around the Propeller P8X32A microcontroller [176]. The board is designed to be the control unit of the robot, but as this control was moved to the laptop, the Propeller board provided a convenient way to communicate with the motor drivers and ultrasonic sensors.

The Propeller board is programmed in Propeller C and implements a serial forwarding program which connects the laptop serial connection to the motor driver serial connection. The connection can be escaped if data from the ultrasonic sensors, which are connected directly to the propeller, are requested.

B.2.2 Ultrasonic Sensors

Though not used in the material in this thesis, the robot is equipped with four ultrasonic PING))) sensors from Parallax [177]. The ultrasonic sensors are triggered by a short 5 V impulse from the Propeller board, and they return a 5 V signal with a width corresponding to the time delay of the sonar pulse, and hence range to an obstacle. The pulse width is parsed and returned as a range in centimeters.

B.2.3 Motor Driver Board

The motor driver board is the DHB-10 motor controller [178], which handles closed-loop control of the motors. The controller is connected to the Propeller board via a serial connection with the same properties as the laptop-Propeller serial connection.

The controller can drive each wheel independently, and tracks how far each wheel has rotated through use of the motor encoders which report the number of ticks travelled by each motor. The motor controller functionality is implemented in the MATLAB object class used on the laptop, so full use of the controllers capabilities is possible.

List of References

[1] L. Thaler, G. Reich, X. Zhang, D. Wang, G. Smith, M. Cherniakov, C. Baker, D. Kish, and M. Antoniou, "Mouth-clicks used by blind expert human echolocators - signal description and model based signal synthesis," *PLoS Computational Biology*, vol. 13, Aug. 2017.

[2] X. Zhang, G. M. Reich, M. Antoniou, M. Cherniakov, C. J. Baker, L. Thaler, D. Kish, and G. E. Smith, "Human echolocation: waveform analysis of tongue clicks," *Electronics Letters*, vol. 53, no. 9, pp. 580–582, 2017.

[3] G. M. Reich, M. Antoniou, and C. J. Baker, "Biologically-inspired wideband target localisation," *Sonar Navigation IET Radar*, vol. 12, no. 12, pp. 1410–1418, 2018.

[4] G. M. Reich, M. Antoniou, and C. J. Baker, "Bio-inspired techniques for target localization," in *2018 IEEE Radar Conference (RadarConf18)*, pp. 1239–1244, Apr. 2018.

[5] G. M. Reich, M. Antoniou, and C. J. Baker, "Frequency-dependent target localization," in *International Conference on Radar Systems (Radar 2017)*, pp. 1–6, Oct. 2017.

[6] G. M. Reich, M. Antoniou, and C. J. Baker, "Memory-Enhanced Cognitive Radar for Autonomous Navigation," 2019. Manuscript submitted for publication.

[7] G. M. Reich, M. Antoniou, and C. J. Baker, "MIMO Radar for Cognitive Robot Platform Control and Navigation," in *2019 International Radar Conference (RADAR)*, pp. 1–6, 2019. In Press.

[8] W. A. Yost, A. N. Popper, and R. R. Fay, eds., *Human Psychophysics*. No. v. 3 in Springer handbook of auditory research, New York: Springer-Verlag, 1993.

[9] J. Blauert, *Spatial hearing: the psychophysics of human sound localization*. Cambridge, Mass: MIT Press, rev. ed ed., 1997.

[10] J. H. Ginsberg, *Acoustics-A Textbook for Engineers and Physicists: Volume I: Fundamentals*. Springer International Publishing, 2018.

[11] J. H. Ginsberg, *Acoustics-A Textbook for Engineers and Physicists: Volume II: Applications*. Springer International Publishing, 2018.

[12] Lord Rayleigh, "On Our Perception of the Direction of a Source of Sound," *Proceedings of the Musical Association*, vol. 2, pp. 75–84, Jan. 1875.

[13] Lord Rayleigh, "Our Perception of the Direction of a Source of Sound," *Nature*, vol. 14, pp. 32–33, 1876.

[14] Lord Rayleigh, "Acoustical observations," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 3, pp. 456–464, June 1877.

[15] Lord Rayleigh, "Acoustical observations," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 13, pp. 340–347, May 1882.

[16] Lord Rayleigh, "On our perception of sound direction," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 13, pp. 214–232, Feb. 1907.

[17] L. J. Sivian and S. D. White, "On Minimum Audible Sound Fields," *The Journal of the Acoustical Society of America*, vol. 4, pp. 288–321, Apr. 1933.

[18] R. S. Woodworth, "Hearing," in *Experimental psychology / Robert S. Woodworth.*, pp. 501 – 538, Methuen, 1950.

[19] W. E. Feddersen, T. T. Sandel, D. C. Teas, and L. A. Jeffress, "Localization of High-Frequency Tones," *The Journal of the Acoustical Society of America*, vol. 29, pp. 988–991, Sept. 1957.

[20] G. F. Kuhn, "Model for the interaural time differences in the azimuthal plane," *The Journal of the Acoustical Society of America*, vol. 62, pp. 157–167, July 1977.

[21] D. Rowan and M. Gray, "Lateralization of high-frequency pure tones with interaural phase difference and bone conduction," *International Journal of Audiology*, vol. 47, pp. 404–411, Jan. 2008.

[22] D. W. Batteau, "Listening with the naked ear," in *The neuropsychology of spatially oriented behavior / edited by Sanford J. Freedman.* (S. J. Freedman, ed.), Dorsey series in psychology, pp. 109–33, Homewood Illinois: Dorsey Press, 1968.

[23] S. J. Freedman and H. G. Fisher, "The role of the pinna in auditory localization," in *The neuropsychology of spatially oriented behavior / edited by Sanford J. Freedman.* (S. J. Freedman, ed.), Dorsey series in psychology, pp. 135–152, Homewood Illinois: Dorsey Press, 1st ed., 1968.

[24] M. B. Gardner and R. S. Gardner, "Problem of localization in the median plane: effect of pinnae cavity occlusion," *The Journal of the Acoustical Society of America*, vol. 53, pp. 400–408, Feb. 1973.

[25] Y. Yorifuji, M. Morimoto, and Y. Ando, "Effect of training in sound localization in the median plane," *The Journal of the Acoustical Society of America*, vol. 57, pp. S37–S37, Apr. 1975.

[26] F. L. Wightman and D. J. Kistler, "Headphone simulation of free-field listening. I: Stimulus synthesis," *The Journal of the Acoustical Society of America*, vol. 85, pp. 858–867, Feb. 1989.

[27] V. R. Algazi, R. O. Duda, D. M. Thompson, and C. Avendano, "The CIPIC HRTF database," in *Applications of Signal Processing to Audio and Acoustics, 2001 IEEE Workshop on the*, pp. 99–102, 2001.

[28] P. Hofman and A. V. Opstal, "Binaural weighting of pinna cues in human sound localization," *Exp Brain Res*, vol. 148, pp. 458–470, Feb. 2003.

[29] R. S. Woodworth, *Experimental psychology / Robert S. Woodworth.* Methuen, 1950.

[30] C. F. Altmann, S. Terada, M. Kashino, K. Goto, T. Mima, H. Fukuyama, and S. Furukawa, "Independent or integrated processing of interaural time and level differences in human auditory cortex?," *Hear. Res.*, vol. 312, pp. 121–127, June 2014.

[31] N. H. Salminen, M. Takanen, O. Santala, J. Lamminsalo, A. Altoè, and V. Pulkki, "Integrated processing of spatial cues in human auditory cortex," *Hearing Research*, vol. 327, pp. 143–152, Sept. 2015.

[32] N. C. Higgins, S. A. McLaughlin, T. Rinne, and G. C. Stecker, "Evidence for cue-independent spatial representation in the human auditory cortex during active listening," *PNAS*, vol. 114, pp. E7602–E7611, Sept. 2017.

[33] J. R. Angell and W. Fite, "Contributions from the Psychological Laboratory of the University of Chicago: Further observations on the monaural localization of sound," *Psychological Review*, vol. 8, pp. 449–458, Sept. 1901.

[34] J. R. Angell and W. Fite, "From the Psychological Laboratory of the University of Chicago: The monaural localization of sound," *Psychological Review*, vol. 8, pp. 225–246, May 1901.

[35] J. Hebrank and D. Wright, "Spectral cues used in the localization of sound sources on the median plane," *The Journal of the Acoustical Society of America*, vol. 56, pp. 1829–1834, Dec. 1974.

[36] J. J. Rice, B. J. May, G. A. Spirou, and E. D. Young, "Pinna-based spectral cues for sound localization in cat," *Hearing Research*, vol. 58, pp. 132–152, Mar. 1992.

[37] H. Wallach, "The role of head movements and vestibular and visual cues in sound localization," *Journal of Experimental Psychology*, vol. 27, no. 4, pp. 339–368, 1940. Place: US Publisher: American Psychological Association.

[38] W. R. Thurlow and P. S. Runge, "Effect of Induced Head Movements on Localization of Direction of Sounds," *The Journal of the Acoustical Society of America*, vol. 42, pp. 480–488, Aug. 1967. Publisher: Acoustical Society of America.

[39] S. Perrett and W. Noble, "The contribution of head motion cues to localization of low-pass noise," *Percept Psychophys*, vol. 59, pp. 1018–1026, Oct. 1997.

[40] F. L. Wightman and D. J. Kistler, "Resolution of front–back ambiguity in spatial hearing by listener and source movement," *The Journal of the Acoustical Society of America*, vol. 105, pp. 2841–2853, Apr. 1999. Publisher: Acoustical Society of America.

[41] W. R. Thurlow, J. W. Mangels, and P. S. Runge, "Head Movements During Sound Localization," *The Journal of the Acoustical Society of America*, vol. 42, p. 489, July 2005. Publisher: Acoustical Society of America.

[42] A. Moiseff and M. Konishi, "Binaural characteristics of units in the owl's brainstem auditory pathway: precursors of restricted spatial receptive fields," *J. Neurosci.*, vol. 3, pp. 2553–2562, Dec. 1983.

[43] W. E. Sullivan and M. Konishi, "Neural Map of Interaural Phase Difference in the Owl's Brainstem," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 83, no. 21, pp. 8400–8404, 1986.

[44] E. I. Knudsen, "Neural derivation of sound source location in the barn owl. An example of a computational map," *Ann. N. Y. Acad. Sci.*, vol. 510, pp. 33–38, 1987.

[45] L. Kettler, H. Griebel, R. Ferger, and H. Wagner, "Combination of Interaural Level and Time Difference in Azimuthal Sound Localization in Owls," *eNeuro*, vol. 4, pp. ENEURO.0238–17.2017, Nov. 2017.

[46] L. M. Aitkin and R. L. Martin, "The representation of stimulus azimuth by high best-frequency azimuth-selective neurons in the central nucleus of the inferior colliculus of the cat," *Journal of Neurophysiology*, vol. 57, pp. 1185–1200, Apr. 1987.

[47] J. Zhang, K. T. Nakamoto, and L. M. Kitzes, "Binaural Interaction Revisited in the Cat Primary Auditory Cortex," *Journal of Neurophysiology*, vol. 91, pp. 101–117, Jan. 2004.

[48] D. C. Fitzpatrick, S. Kuwada, and R. Batra, "Neural Sensitivity to Interaural Time Differences: Beyond the Jeffress Model," *J. Neurosci.*, vol. 20, pp. 1605–1615, Feb. 2000.

[49] B. Tian, D. Reser, A. Durham, A. Kustov, and J. P. Rauschecker, "Functional Specialization in Rhesus Monkey Auditory Cortex," *Science*, vol. 292, pp. 290–293, Apr. 2001.

[50] J. C. Middlebrooks, I. A. Harrington, E. A. Macpherson, and G. C. Stecker, "3.46 - Sound Localization and the Auditory Cortex A2 - Gardner, Richard H. MaslandThomas D. AlbrightThomas D. AlbrightRichard H. MaslandPeter DallosDonata OertelStuart FiresteinGary K. BeauchampM. Catherine BushnellAllan I. BasbaumJon H. KaasEsther P.," in *The Senses: A Comprehensive Reference*, pp. 781–805, New York: Academic Press, 2008.

[51] N. Lessard, M. Paré, F. Lepore, and M. Lassonde, "Early-blind human subjects localize sound sources better than sighted subjects," *Nature*, vol. 395, pp. 278–280, Sept. 1998.

[52] M. Zwiers, A. Van Opstal, and J. Cruysberg, "Two-dimensional sound-localization behavior of early-blind humans," *Experimental Brain Research*, vol. 140, no. 2, pp. 206–222, 2001.

[53] J. Lewald, "Vertical sound localization in blind humans," *Neuropsychologia*, vol. 40, no. 12, pp. 1868–1872, 2002.

[54] C. F. Moss and A. Surlykke, "Probing the Natural Scene by Echolocation in Bats," *Front. Behav. Neurosci.*, vol. 4, 2010.

[55] M. Melcón, Y. Yovel, A. Denzinger, and H.-U. Schnitzler, "How greater mouse-eared bats deal with ambiguous echoic scenes," *Journal of Comparative Physiology A*, vol. 197, no. 5, pp. 505–514, 2011.

[56] K. L. Bell, C. J. Baker, G. E. Smith, J. T. Johnson, and M. Rangaswamy, "Cognitive Radar Framework for Target Detection and Tracking," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, pp. 1427–1439, Dec. 2015.

[57] G. E. Smith, Z. Cammenga, A. Mitchell, K. L. Bell, J. Johnson, M. Rangaswamy, and C. Baker, "Experiments with cognitive radar," *IEEE Aerospace and Electronic Systems Magazine*, vol. 31, pp. 34–46, Dec. 2016.

[58] C. V. Portfors and J. J. Wenstrup, "Delay-Tuned Neurons in the Inferior Colliculus of the Mustached Bat: Implications for Analyses of Target Distance," *Journal of Neurophysiology*, vol. 82, pp. 1326–1338, Sept. 1999.

[59] J. Tang, Z. Xiao, and N. Suga, "Bilateral Cortical Interaction: Modulation of Delay-Tuned Neurons in the Contralateral Auditory Cortex," *J. Neurosci.*, vol. 27, pp. 8405–8413, Jan. 2007.

[60] J. A. Simmons, "Bats use a neuronally implemented computational acoustic model to form sonar images," *Current Opinion in Neurobiology*, vol. 22, pp. 311–319, Apr. 2012.

[61] P. A. Saillant, J. A. Simmons, S. P. Dear, and T. A. McMullen, "A computational model of echo processing and acoustic imaging in frequency- modulated echolocating bats: The spectrogram correlation and transformation receiver," *The Journal of the Acoustical Society of America*, vol. 94, pp. 2691–2712, Nov. 1993.

[62] B. M. Siemers and H.-U. Schnitzler, "Natterer's bat (Myotis nattereri Kuhl, 1818) hawks for prey close to vegetation using echolocation signals of very broad bandwidth," *Behav Ecol Sociobiol*, vol. 47, pp. 400–412, May 2000.

[63] M. E. Bates, J. A. Simmons, and T. V. Zorikov, "Bats Use Echo Harmonic Structure to Distinguish Their Targets from Background Clutter," *Science*, vol. 333, pp. 627–630, July 2011.

[64] D. v. Helversen, M. W. Holderied, and O. v. Helversen, "Echoes of bat-pollinated bell-shaped flowers: conspicuous for nectar-feeding bats?," *Journal of Experimental Biology*, vol. 206, pp. 1025–1034, Mar. 2003.

[65] D. von Helversen and O. von Helversen, "Object recognition by echolocation: a nectar-feeding bat exploiting the flowers of a rain forest vine," *J Comp Physiol A*, vol. 189, pp. 327–336, May 2003.

[66] D. von Helversen, "Object classification by echolocation in nectar feeding bats: size-independent generalization of shape," *J Comp Physiol A*, vol. 190, pp. 515–521, July 2004.

[67] A. Balleri, K. Woodbridge, C. J. Baker, and M. W. Holderied, "Flower classification by bats: Radar comparisons," *IEEE Aerospace and Electronic Systems Magazine*, vol. 24, pp. 4–7, May 2009.

[68] A. Balleri, K. Woodbridge, C. J. Baker, and M. W. Holderied, "Classification of flowers by bats: comparison with the radar case," in *2009 International Waveform Diversity and Design Conference*, pp. 1–3, Feb. 2009.

[69] A. Balleri, H. D. Griffiths, K. Woodbridge, C. J. Baker, and M. W. Holderied, "Impact of flight trajectory on the detection and selection of flowers by nectar-feeding bats," in *2010 2nd International Workshop on Cognitive Information Processing*, pp. 350–353, June 2010.

[70] C. Chiu and C. F. Moss, "The role of the external ear in vertical sound localization in the free flying bat, Eptesicus fuscus," *The Journal of the Acoustical Society of America*, vol. 121, pp. 2227–2235, Apr. 2007.

[71] D. N. Lee, J. A. Simmons, P. A. Saillant, and F. Bouffard, "Steering by echolocation: a paradigm of ecological acoustics," *J Comp Physiol A*, vol. 176, pp. 347–354, May 1995.

[72] G. E. Smith and C. J. Baker, "Echoic flow for autonomous navigation," in *IET International Conference on Radar Systems (Radar 2012)*, pp. 1–6, Oct. 2012.

[73] G. E. Smith, C. J. Baker, and G. Li, "Echoic flow for radar target interception," in *2013 International Conference on Radar*, pp. 520–525, Sept. 2013.

[74] G. Smith, C. Baker, and G. Li, "Coupled echoic flow for cognitive radar sensing," in *2013 IEEE Radar Conference (RADAR)*, pp. 1–6, Apr. 2013.

[75] C. J. Baker, G. E. Smith, A. Balleri, M. Holderied, and H. D. Griffiths, "Biomimetic Echolocation With Application to Radar and Sonar Sensing," *Proceedings of the IEEE*, vol. 102, pp. 447–458, Apr. 2014.

[76] G. E. Smith, S. Alsaif, and C. J. Baker, "Echoic flow for cognitive radar guidance," in *2014 IEEE Radar Conference*, pp. 0490–0495, May 2014.

[77] T. A. Stroffregen and J. B. Pittenger, "Human Echolocation as a Basic Form of Perception and Action," *Ecological Psychology*, vol. 7, pp. 181–216, Sept. 1995.

[78] A. J. Kolarik, S. Cirstea, S. Pardhan, and B. C. J. Moore, "A summary of research investigating echolocation abilities of blind and sighted humans," *Hear. Res.*, vol. 310, pp. 60–68, Apr. 2014.

[79] L. Thaler and M. A. Goodale, "Echolocation in humans: an overview," *Wiley Interdiscip Rev Cogn Sci*, vol. 7, pp. 382–393, Nov. 2016.

[80] W. N. Kellogg, "Sonar System of the Blind: New research measures their accuracy in detecting the texture, size, and distance of objects "by ear."," *Science*, vol. 137, pp. 399–404, Aug. 1962.

[81] S. Teng, A. Puri, and D. Whitney, "Ultrafine spatial acuity of blind expert human echolocators," *Exp Brain Res*, vol. 216, pp. 483–488, Feb. 2012.

[82] J. L. Milne, M. Anello, M. A. Goodale, and L. Thaler, "A blind human expert echolocator shows size constancy for objects perceived by echoes," *Neurocase*, vol. 21, pp. 465–470, July 2015.

[83] S. Schörnich, A. Nagy, and L. Wiegrebe, "Discovering your inner bat: echo-acoustic target ranging in humans," *J. Assoc. Res. Otolaryngol.*, vol. 13, pp. 673–682, Oct. 2012.

[84] S. R. Arnott, L. Thaler, J. L. Milne, D. Kish, and M. A. Goodale, "Shape-specific activation of occipital cortex in an early blind echolocation expert," *Neuropsychologia*, vol. 51, pp. 938–949, Apr. 2013.

[85] D. Rowan, T. Papadopoulos, D. Edwards, H. Holmes, A. Hollingdale, L. Evans, and R. Allen, "Identification of the lateral position of a virtual object based on echoes by humans," *Hearing Research*, vol. 300, pp. 56–65, June 2013.

[86] J. L. Milne, M. A. Goodale, and L. Thaler, "The role of head movements in the discrimination of 2-D shape by blind echolocation experts," *Atten Percept Psychophys*, vol. 76, pp. 1828–1837, Aug. 2014.

[87] L. Wallmeier and L. Wiegrebe, "Ranging in Human Sonar: Effects of Additional Early Reflections and Exploratory Head Movements," *PLOS ONE*, vol. 9, p. e115363, Dec. 2014.

[88] L. Wallmeier and L. Wiegrebe, "Self-motion facilitates echo-acoustic orientation in humans," *R Soc Open Sci*, vol. 1, Nov. 2014.

[89] J. Reijniers and H. Peremans, "Biomimetic Sonar System Performing Spectrum-Based Localization," *IEEE Transactions on Robotics*, vol. 23, pp. 1151–1159, Dec. 2007.

[90] F. Guarato, J. Windmill, and A. Gachagan, "A beam based method for target localization: Inspiration from bats' directivity and binaural reception for ultrasonic sonar," *The Journal of the Acoustical Society of America*, vol. 133, pp. 4077–4086, June 2013.

[91] J. Ward, *The student's guide to cognitive neuroscience*. Hove: Psychology Press, 2. ed ed., 2010.

[92] L. A. Jeffress, "A place theory of sound localization," *Journal of Comparative*, vol. 41, pp. 35–39, Feb. 1948.

[93] K. Georgiev, A. Balleri, A. Stove, and M. Holderied, "Baseband version of the bat-inspired spectrogram correlation and transformation receiver," in *2016 IEEE Radar Conference (RadarConf)*, pp. 1–6, May 2016.

[94] K. Georgiev, A. Balleri, A. Stove, and M. Holderied, "Bio-inspired two target resolution at radio frequencies," in *2017 IEEE Radar Conference (RadarConf)*, pp. 0436–0440, May 2017.

[95] K. Georgiev, A. Balleri, A. Stove, and M. W. Holderied, "Bio-inspired processing of radar target echoes," *Sonar Navigation IET Radar*, vol. 12, no. 12, pp. 1402–1409, 2018.

[96] A. Balleri, A. Farina, and A. Benavoli, "Coordination of optimal guidance law and adaptive radiated waveform for interception and rendezvous problems," *IET Radar, Sonar & Navigation*, vol. 11, pp. 1132–1139, Feb. 2017.

[97] M. Athans, "On Optimal Allocation and Guidance Laws for Linear lnterception and Rendezvous Problems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-7, pp. 843–853, Sept. 1971.

[98] D. J. Kershaw and R. J. Evans, "Optimal waveform selection for tracking systems," *IEEE Transactions on Information Theory*, vol. 40, pp. 1536–1550, Sept. 1994.

[99] J. J. Gibson, "The Stimulus Variables for Visual Depth and Distance – The Active Observer," in *The perception of the visual world / by James J. Gibson ; [under the editorship of Leonard Carmichael].*, pp. 117–144, Boston: Houghton Mifflin Company, 1950.

[100] D. N. Lee and H. Kalmus, "The Optic Flow Field: The Foundation of Vision [and Discussion]," *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, vol. 290, no. 1038, pp. 169–179, 1980.

[101] D. N. Lee, "Guiding Movement by Coupling Taus," *Ecological Psychology*, vol. 10, pp. 221–250, Sept. 1998.

[102] K. Kugler, W. Greiter, H. Luksch, U. Firzlaff, and L. Wiegrebe, "Echo-acoustic flow affects flight in bats," *Journal of Experimental Biology*, vol. 219, pp. 1793–1797, June 2016.

[103] G. E. Smith, Z. Cammenga, A. Mitchell, K. L. Bell, M. Rangaswamy, J. T. Johnson, and C. J. Baker, "Experiments with cognitive radar," in *2015 IEEE 6th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pp. 293–296, Dec. 2015.

[104] A. E. Mitchell, G. E. Smith, K. L. Bell, and M. Rangaswamy, "Coordinate descent for cognitve radar adaptation," in *2016 CIE International Conference on Radar (RADAR)*, pp. 1–5, Oct. 2016.

[105] M. I. Skolnik, *Introduction to radar systems*. New York: McGraw-Hill, 2nd ed. ed., 1980.

[106] M. A. Richards, J. Scheer, W. A. Holm, and W. L. Melvin, eds., *Principles of modern radar*. Raleigh, NC: SciTech Pub, 2010.

[107] B. R. Mahafza, *Radar systems analysis and design using MATLAB*. 2016. OCLC: 1027156415.

[108] H. Griffiths, C. Baker, D. Adamy, and G. W. Stimson, *Stimson's introduction to airborne radar*. Edison, NJ: SciTech Publishing, third edition ed., 2014. OCLC: ocn868379889.

[109] Skolnik, "The Radar Equation," in *Introduction to Radar Systems*, p. 88, New Delhi: Tata McGraw-Hill, 3rd ed., 2001.

[110] B. R. Mahafza, "Sphere," in *Radar systems analysis and design using MATLAB*, pp. 495–497, 2016. OCLC: 1027156415.

[111] L. V. Blake and M. W. Long, *Antennas: fundamentals, design, measurement*. Raleigh, NC: SciTech Pub., 2009. OCLC: 746938702.

[112] Q-par Angus Ltd, "Miniature Single Polarisation 1 - 18 GHz Horn Antenna," Dec. 2008.

[113] T. N. Buell, C. Trahiotis, and L. R. Bernstein, "Lateralization of low-frequency tones: Relative potency of gating and ongoing interaural delays," *The Journal of the Acoustical Society of America*, vol. 90, pp. 3077–3085, Dec. 1991.

[114] T. M. Shackleton, J. M. Bowsher, and R. Meddis, "Lateralization of Very-short-Duration Tone Pulses of Low and High Frequencies," *The Quarterly Journal of Experimental Psychology Section A*, vol. 43, pp. 503–516, Aug. 1991.

[115] M. Rothacher, "Comparison of Absolute and Relative Antenna Phase Center Variations," *GPS Solutions*, vol. 4, pp. 55–60, Apr. 2001.

[116] D. N. Aloi, A. Rusek, and B. A. Oakley, "A relative technique for characterization of PCV error of large aperture antennas using GPS data," *IEEE Transactions on Instrumentation and Measurement*, vol. 54, pp. 1820–1832, 2005.

[117] R. Schmid, M. Rothacher, D. Thaller, and P. Steigenberger, "Absolute phase center corrections of satellite and receiver antennas," *GPS Solut*, vol. 9, pp. 283–293, Nov. 2005.

[118] P. Jarlemark, R. Emardson, J. Johansson, and G. Elgered, "Ground-Based GPS for Validation of Climate Models: The Impact of Satellite Antenna Phase Center Variations.," *IEEE Transactions on Geoscience & Remote Sensing*, vol. 48, pp. 3847–3855, Oct. 2010.

[119] K. Dawidowicz, "Phase center variations problem in GPS/GLONASS observations processing," *Environmental Engineering. Proceedings of the International Conference on Environmental Engineering. ICEE*, vol. 9, p. 1, 2014.

[120] D. K. Barton, *Monopulse radar*. Radars ; v.1, Dedham, Mass. ; London: Artech House : Adtech, 1974.

[121] S. M. Sherman and D. K. Barton, *Monopulse Principles and Techniques*. Norwood, UNITED STATES: Artech House, 2011.

[122] Rohde & Schwarz GmbH & Co. KG, "R&S®ZVA / R&S®ZVB / R&S® ZVT Vector Network Analyzers Operating Manual," 2011.

[123] L. G. Weiss, "Wavelets and wideband correlation processing," *IEEE Signal Processing Magazine*, vol. 11, pp. 13–32, Jan. 1994.

[124] J. G. Švec and G. Svante, "Tutorial and Guidelines on Measurement of Sound Pressure Level in Voice and Speech," *Journal of Speech, Language, and Hearing Research*, vol. 61, pp. 441–461, Mar. 2018.

[125] Scientific Committee on Emerging and Newly Identified Health Risks, "Potential health risks of exposure to noise from personal music players and mobile phones including a music playing function: Preliminary report," tech. rep., European Commission, Sept. 2008.

[126] K. G. Foote and J. A. Simmons, "Bat sonar and the role of frequency diversity," *The Journal of the Acoustical Society of America*, vol. 119, pp. 3318–3318, May 2006.

[127] S. D. Ibsen, W. W. L. Au, P. E. Nachtigall, and M. Breese, "Functional bandwidth of an echolocating Atlantic bottlenose dolphin (Tursiops truncatus)," *J. Acoust. Soc. Am.*, vol. 125, pp. 1214–1221, Feb. 2009.

[128] J. A. M. Rojas, J. A. Hermosilla, R. S. Montero, and P. L. L. Espí, "Physical Analysis of Several Organic Signals for Human Echolocation: Oral Vacuum Pulses," *Acta Acustica united with Acustica*, vol. 95, pp. 325–330, Mar. 2009.

[129] G. E. Smith and C. J. Baker, "Human echolocation waveform analysis," in *IET International Conference on Radar Systems (Radar 2012)*, pp. 1–5, Oct. 2012.

[130] DPA Microphones, "d:screet Miniature Microphones Users Manual," Oct. 2014.

[131] TASCAM, "DR-100MKII Linear PCM Recorder Owner's Manual," Dec. 2011.

[132] T. Parks and J. McClellan, "Chebyshev Approximation for Nonrecursive Digital Filters with Linear Phase," *IEEE Transactions on Circuit Theory*, vol. 19, pp. 189–194, Mar. 1972.

[133] S. A. Fulop and K. Fitz, "Algorithms for computing the time-corrected instantaneous frequency (reassigned) spectrogram, with applications," *The Journal of the Acoustical Society of America*, vol. 119, pp. 360–371, Jan. 2006.

[134] C. Condon, K. White, and A. Feng, "Bat Echolocation Chirp."

[135] B. G. Shinn-Cunningham, S. Santarelli, and N. Kopco, "Tori of confusion: Binaural localization cues for sources within reach of a listener," *The Journal of the Acoustical Society of America*, vol. 107, pp. 1627–1636, Mar. 2000.

[136] C. Oberzut and L. Olson, "Directionality and the head-shadow effect:," *The Hearing Journal*, vol. 56, p. 56, Apr. 2003.

[137] F. Schillebeeckx, F. D. Mey, D. Vanderelst, and H. Peremans, "Biomimetic Sonar: Binaural 3D Localization using Artificial Bat Pinnae," *The International Journal of Robotics Research*, vol. 30, pp. 975–987, July 2011.

[138] H. K. Dunn and D. W. Farnsworth, "Exploration of Pressure Field Around the Human Head During Speech," *The Journal of the Acoustical Society of America*, vol. 10, pp. 184–199, Jan. 1939.

[139] A. C. C. Warnock, W. T. Chu, and J.-C. Guy, "Directivity of human talkers," *Canadian Acoustics*, vol. 30, pp. 36–37, Sept. 2002.

[140] J. C. Middlebrooks and D. M. Green, "Sound Localization by Human Listeners," *Annual Review of Psychology*, vol. 42, no. 1, pp. 135–159, 1991.

[141] F. L. Wightman and D. J. Kistler, "Sound Localization," in *Human Psychophysics*, Springer Handbook of Auditory Research, pp. 155–192, Springer, New York, NY, 1993.

[142] S. R. Oldfield and S. P. A. Parker, "Acuity of Sound Localisation: A Topography of Auditory Space. III. Monaural Hearing Conditions," *Perception*, vol. 15, pp. 67–81, Feb. 1986.

[143] R. A. Butler, R. A. Humanski, and A. D. Musicant, "Binaural and monaural localization of sound in two-dimensional space," *Perception*, vol. 19, no. 2, pp. 241–256, 1990.

[144] S. Kingsley and S. Quegan, "Tracking Radar," in *Understanding Radar Systems*, p. 52, SciTech Publishing, 1999.

[145] E. I. Knudsen, S. du Lac, and S. D. Esterly, "Computational maps in the brain," *Annu. Rev. Neurosci.*, vol. 10, pp. 41–65, 1987.

[146] R. Rouveure, L. Jean, C. Debain, and R. Peuchot, "Robot Localization and Navigation with a Ground-based Microwave Radar," (Toulon), 2019.

[147] A. Moses, M. J. Rutherford, M. Kontitsis, and K. P. Valavanis, "UAV-borne X-band radar for collision avoidance," *Robotica*, vol. 32, pp. 97–114, Jan. 2014.

[148] S. Kemkemian, M. Nouvel-Fiani, P. Cornic, P. L. Bihan, and P. Garrec, "Radar systems for "Sense and Avoid" on UAV," in *2009 International Radar Conference "Surveillance for a Safer World" (RADAR 2009)*, pp. 1–6, Oct. 2009. ISSN: 1097-5764.

[149] Y. K. Kwag, M. S. Choi, C. H. Jung, and K. Y. Hwang, "Collision Avoidance Radar for UAV," in *2006 CIE International Conference on Radar*, pp. 1–4, Oct. 2006. ISSN: null.

[150] K. Ariyur, P. Lommel, and D. Enns, "Reactive inflight obstacle avoidance via radar feedback," in *Proceedings of the 2005, American Control Conference, 2005.*, pp. 2978–2982 vol. 4, June 2005. ISSN: 2378-5861.

[151] A. Viquerat, L. Blackhall, A. Reid, S. Sukkarieh, and G. Brooker, "Reactive Collision Avoidance for Unmanned Aerial Vehicles Using Doppler Radar," in *Field and Service Robotics: Results of the 6th International Conference* (C. Laugier and R. Siegwart, eds.), Springer Tracts in Advanced Robotics, pp. 245–254, Berlin, Heidelberg: Springer, 2008.

[152] K. Lamberts and R. Goldstone, *Handbook of Cognition*. 2005.

[153] J. M. Fuster, "The cognit: A network model of cortical representation," *International Journal of Psychophysiology*, vol. 60, pp. 125–132, May 2006.

[154] S. Haykin and J. M. Fuster, "On Cognitive Dynamic Systems: Cognitive Neuroscience and Engineering Learning From Each Other," *Proceedings of the IEEE*, vol. 102, pp. 608–628, Apr. 2014.

[155] M. J. Sagayaraj, V. Jithesh, J. B. Singh, D. Roshani, and K. G. Srinivasa, "A Hybrid Approach to Cognition in Radars," *Defence Science Journal; New Delhi*, vol. 68, pp. 183–189, Apr. 2018.

[156] I. Winkler and N. Cowan, "From Sensory to Long-Term Memory: Evidence from Auditory Memory Reactivation Studies," *Experimental Psychology*, vol. 52, no. 1, pp. 3–20, 2005.

[157] A. Baddeley, "The episodic buffer: a new component of working memory?," *Trends in Cognitive Sciences*, vol. 4, pp. 417–423, Nov. 2000.

[158] W. X. Schneider, "Visual-spatial working memory, attention, and scene representation: A neuro-cognitive theory," *Psychological Research Psychologische Forschung*, vol. 62, pp. 220–236, July 1999.

[159] A. M. Maxcey-Richard and A. Hollingworth, "The Strategic Retention of Task-Relevant Objects in Visual Working Memory," *Journal of Experimental Psychology*, vol. 39, pp. 760–772, May 2013.

[160] A. Hollingworth, "The Relationship Between Online Visual Representation of a Scene and Long-Term Scene Memory," *Journal of Experimental Psychology. Learning, Memory & Cognition*, vol. 31, pp. 396–411, May 2005.

[161] S. Haykin, A. Zia, Y. Xue, and I. Arasaratnam, "Control theoretic approach to tracking radar: First step towards cognition," *Digital Signal Processing*, vol. 21, pp. 576–585, Sept. 2011.

[162] S. S. Haykin, *Cognitive dynamic systems: perception–action cycle, radar, and radio.* Cambridge ; New York: Cambridge University Press, 2012.

[163] A. Charlish and F. Hoffmann, "Anticipation in cognitive radar using stochastic control," in *2015 IEEE Radar Conference (RadarCon)*, pp. 1692–1697, May 2015.

[164] Parallax Inc., "Arlo Complete Robot System | 28966 | Parallax Inc."

[165] INRAS GmbH, "INRAS Products - Radarbook," 2014.

[166] Inras GmbH, "MIMO-77-TX4RX8 Frontend," July 2015.

[167] M. S. Davis, "MIMO Radar," in *Principles of Modern Radar: Advanced Techniques* (W. L. Melvin and J. A. Scheer, eds.), pp. 119–145, Edison, NJ: SciTech Publishing, 2013.

[168] G. Dudek, "Forward kinematics for differential drive robots," in *Computational principles of mobile robotics / Gregory Dudek and Michael Jenkin.*, pp. 41–42, Cambridge: Cambridge University Press, 2000.

[169] H. G. Krapp, "Sensorimotor Transformation: From Visual Responses to Motor Commands," *Current Biology*, vol. 20, pp. R236–R239, Mar. 2010.

[170] K. Takiyama, "Sensorimotor transformation via sparse coding," *Scientific Reports*, vol. 5, p. 9648, Apr. 2015.

[171] J. M. Fuster, *The Prefrontal Cortex.* Elsevier, 2008.

[172] R. E. Passingham and S. P. Wise, *The Neurobiology of the Prefrontal Cortex: Anatomy, Evolution, and the Origin of Insight.* Oxford University Press, July 2012.

[173] M. Donoso, A. G. E. Collins, and E. Koechlin, "Foundations of human reasoning in the prefrontal cortex," *Science*, vol. 344, pp. 1481–1486, June 2014.

[174] Ricoh Company, Ltd, "Product | RICOH THETA S," Dec. 2017.

[175] Parallax Inc., "Propeller Activity Board WX (#32912)," Sept. 2017.

[176] Parallax Inc., "Propeller P8X32A Datasheet," 2011.

[177] Parallax Inc., "PING))) Ultrasonic Distance Sensor (#28015)," Feb. 2014.

[178] Parallax Inc., "DHB-10 Dual H-Bridge 10 Amp Motor Controller (#28231)," Apr. 2016.