Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.
ACOUSTIC SOURCE LOCALISATION

A THESIS PRESENTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
DEGREE OF
MASTER OF SCIENCE
IN
MATHEMATICS
AT MASSEY UNIVERSITY, PALMERSTON NORTH,
NEW ZEALAND.

Alexander Lyndon White

2019
Contents

Abstract iv
Acknowledgements v

1 Introduction 1
1.1 Motivation 1
1.2 Objectives 2
1.3 Results 2
1.4 Outline 3

2 Literature Review 4
2.1 Acoustic Source Localisation 4
2.1.1 The Generalised Cross-Correlation 5
2.1.2 The Multiple Signal Classification (MUSIC) Algorithm 6
2.1.3 The Min-Norm Algorithm 7
2.1.4 ESPRIT 7
2.1.5 Beamforming 8
2.1.6 Machine Learning 8
2.2 History of the Polynomial Eigenvalue Decomposition 9

3 Generalised Cross-Correlation (GCC) 12
3.1 The Fourier Transform 13
3.2 Cross-Correlation 14
3.3 Generalised Cross-Correlation 15
3.3.1 GCC Processors 17

4 Polynomial Matrices 22
4.1 Overview 22
4.1.1 Description 25
4.1.2 Implications 25
4.2 Existence and Uniqueness of Parahermitian Matrix EVDs 26
4.2.1 Analyticity 26
4.2.2 EVD on the Unit Circle 27
4.2.3 Continuity of Eigenvalues 28
4.2.4 Existence of Parahermitian Polynomial Matrix Inverses 28
4.3 Calculation of the PEVD 29
4.3.1 The SBR2 Algorithm 29
4.3.2 Example Use of SBR2 ................................................. 31
4.3.3 Notes on Practical Implementation .......................... 33
4.3.4 Extensions to SBR2 ................................................ 34

5 The MUSIC Algorithm .................................................. 35
5.1 Multiple Signal Classification (MUSIC) ...................... 35
  5.1.1 Limitations ......................................................... 39
  5.1.2 Extensions to the MUSIC Algorithm .................. 40
5.2 Broadband MUSIC .................................................... 41
  5.2.1 The MUSIC Spectrum ......................................... 42
  5.2.2 Limitations ......................................................... 44
  5.2.3 Equivalence of PS(S)-MUSIC to AF-MUSIC .......... 45

6 Experiment .............................................................. 47
  6.1 Aim & Overview ...................................................... 47
  6.2 Audio File ............................................................ 48
  6.3 Method ............................................................... 48

7 Results ................................................................. 51
  7.1 GCC Performance .................................................. 51
    7.1.1 False Positives ............................................... 51
    7.1.2 Choice of Processor ........................................ 53
  7.2 AF-MUSIC Performance ......................................... 55
  7.3 Remarks on Computational Cost .............................. 55
  7.4 Array Radius ........................................................ 56
  7.5 Future Work ........................................................ 58

Appendices ............................................................... 59
A Figures ................................................................. 60
  A.1 SBR2 Visualisation ................................................ 60

B Image Generation ..................................................... 65
  B.1 Graphical Representation ...................................... 65
    B.1.1 GCC Heatmap Generation ................................ 66
    B.1.2 Radial Plots ................................................... 67
    B.1.3 AF-MUSIC Image Generation ............................ 67

C Code Demonstration .................................................. 69
  C.1 Simulating and Loading Data .................................. 69
  C.2 Localisation ........................................................ 71
    C.2.1 GCC-Based Methods ........................................ 71
    C.2.2 MUSIC-Based Methods ..................................... 72
  C.2.3 Example Using Real Data .................................. 74
  C.3 Runtime Comparison .............................................. 77

Bibliography ........................................................... 77
Abstract

Many New Zealand native bird species are under threat, and as such conservationists are interested in obtaining accurate estimates of population density in order to closely monitor the changes in abundance of these species over time. One method of estimating the presence and abundance of birdlife in an area is using acoustic recorders; currently, omnidirectional microphones are used, which provide no estimate of the direction of arrival of the call. An estimate of the direction from which each sound came from would help to discern one individual calling multiple times, from multiple birds calling in succession – thus providing more accurate information to models of population density. The estimation of this direction-of-arrival (or DOA) for each source is known as acoustic source localisation, and is the subject of this work.

This thesis contains a discussion and application of two families of algorithm for acoustic source localisation: those based on the Generalised Cross-Correlation (GCC) algorithm, which applies weightings to the calculation of the cross-correlation of two signals; and those based on the Multiple Signal Classification (MUSIC) algorithm, which provides an estimate of source direction based on subspaces generated by the covariance matrix of the data. As the MUSIC algorithm was originally described for narrowband signals – an assumption not applicable to birdsong – we discuss several adaptations of MUSIC to the broadband scenario; one such adaptation requiring the use of polynomial matrices, which are described herein.

An experiment was conducted during this work to determine the effect that the distance between the microphones in a microphone array has on the ability of that array to localise various acoustic signals, including the New Zealand native North Island Brown Kiwi, *Apteryx mantelli*. It was found that both GCC and MUSIC benefit from larger inter-array spacings, and that a variant of the MUSIC algorithm known as autofocusing MUSIC (or AF-MUSIC) provided the most precise DOA estimates.

Though native birdlife was the motivator for the research, none of the methods described within this thesis are necessarily bound only to work on recordings of birdsong; indeed, any multichannel audio which satisfies the necessary assumptions for each algorithm would be suitable.

As well as a description of the algorithms, an implementation of GCC, MUSIC, and AF-MUSIC was produced in the Python 3 programming language, and is available at https://github.com/alexW335/Locator
Acknowledgements

I would like to thank my supervisors, Dr. Richard Brown and Prof. Stephen Marsland, for their encouragement, patience, and advice throughout my degree. I’d like to thank my father for providing me with the theme “quantum electrodynamics” for my poster on the letter ‘Q’ in primary school, and starting me on a long journey of looking up things I don’t understand solely because they sound interesting. I’d like to thank my mother for her unwavering support in whatever I do, my brother for generally being rad 24/7, and I’d like to thank Kelsey for putting up with me for quite some time now; I couldn’t have done it without you.

In addition, I am exceptionally grateful to have received funding from the National Science Challenge on Science for Technological Innovation, and the Ministry of Business, Innovation and Employment to enable me to have this opportunity.