

# University of Wollongong - Research Online

## Thesis Collection

Title: Developing a subband model for blind signal separation in an acoustic environment

Author: Iain Trent Russell

Year: 2005

Repository DOI:

### Copyright Warning

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following: This work is copyright. Apart from any use permitted under the Copyright Act 1968, no part of this work may be reproduced by any process, nor may any other exclusive right be exercised, without the permission of the author. Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material.

Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

**Unless otherwise indicated, the views expressed in this thesis are those of the author and do not necessarily represent the views of the University of Wollongong.**

Research Online is the open access repository for the University of Wollongong. For further information contact the UOW Library: [research-pubs@uow.edu.au](mailto:research-pubs@uow.edu.au)

2005

## Developing a subband model for blind signal separation in an acoustic environment

Iain Trent Russell  
*University of Wollongong*

Follow this and additional works at: <https://ro.uow.edu.au/theses>

### University of Wollongong

#### Copyright Warning

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following: This work is copyright. Apart from any use permitted under the Copyright Act 1968, no part of this work may be reproduced by any process, nor may any other exclusive right be exercised, without the permission of the author. Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material.

Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

Unless otherwise indicated, the views expressed in this thesis are those of the author and do not necessarily represent the views of the University of Wollongong.

### Recommended Citation

Russell, Iain Trent, Developing a subband model for blind signal separation in an acoustic environment, PhD thesis, School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, 2005. <http://ro.uow.edu.au/theses/553>

## **NOTE**

This online version of the thesis may have different page formatting and pagination from the paper copy held in the University of Wollongong Library.

## **UNIVERSITY OF WOLLONGONG**

### **COPYRIGHT WARNING**

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site. You are reminded of the following:

Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material. Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

# Developing A Subband Model for Blind Signal Separation in an Acoustic Environment

A thesis submitted in fulfilment of the  
requirements for the award of the degree

Doctor of Philosophy

from

THE UNIVERSITY OF WOLLONGONG

by

Iain Trent Russell  
Bachelor of Telecommunications Engineering (Honours Class I)

SCHOOL OF ELECTRICAL, COMPUTER  
AND TELECOMMUNICATIONS ENGINEERING  
2005

# Statement of Originality

This is to certify that the work described in this thesis is entirely my own, except where due reference is made in the text.

No work in this thesis has been submitted for a degree to any other university or institution.

Signed

A handwritten signature in black ink, appearing to read 'Iain Russell', written in a cursive style.

Iain Russell

22nd August, 2005

*Dedicated to my family*

# Acknowledgments

I would like to thank my supervisors, Dr. Jiangtao Xi, Prof. Joe Chicharo, and Prof. Alfred Mertins for their academic advice and continual support throughout the PhD. Gratitude is extended to Alfred and Jiangtao for making it an easy transition between supervisors over the course of the first two years of the project. I am also very grateful to Alfred for approaching me and providing the opportunity to do a PhD, providing the necessary scholarship and initial insight to the overall thesis topic.

Special thanks goes to Gaurav Srivastava and Dr. Mehran Abolhasan who as work colleagues always created an enjoyable atmosphere in the office and always fostered good discussions on topics ranging from women, politics, engineering, and religion.

Many thanks go to the members of TITR and the Whisper lab who made the last three years of hard work a pleasureable experience! The insight from different areas of research often provided a different way of approaching many problems specific to the area of BSS and was greatly appreciated. Special thanks also goes to Dr. Jason Lukasiak for his academic advice from time to time.

Finally I would like to express my deepest thanks to my parents, and family, for their encouragement, support and understanding over the last few years.

# Author's Publications

Much of the work in this thesis has been published or has been submitted for publication as academic papers. These papers are:

1. Alfred Mertins and Iain Russell, "An extended ACDC algorithm for the blind estimation of convolutive mixing systems," in *Proceedings of Seventh International Symposium on Signal Processing and its Applications (ISSPA 2003)*, Paris, France, July 2003, vol. 2, pp. 527-530.
2. Iain Russell, Alfred Mertins, and Jiangtao Xi, "Time domain optimization techniques for blind separation of non-stationary convolutive mixed signals," in *Proceedings of 9th IASTED International Conference on Signal and Image Processing (SIP 2003)*, Honolulu, Hawaii, USA, August 2003, pp. 440-445.
3. Iain Russell, Jiangtao Xi, Alfred Mertins, and Joe Chicharo, "Blind separation of nonstationary convolutively mixed signals in the time domain," in *Proceedings of 7th International Symposium on DSP for Communication Systems (DSPCS03)*, Coolangatta, Qld, Australia, December 2003, pp. 93-98.
4. Iain Russell, Jiangtao Xi, Alfred Mertins, and Joe Chicharo, "Blind source separation of nonstationary convolutively mixed signals in the subband domain," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2004)*, Montreal, Canada, May 2004, pp. V-481-V-484.



5. Iain Russell, Jiangtao Xi, Alfred Mertins, and Joe Chicharo, "Integration of DFT and cosine-modulated filter banks with blind separation of convolutively mixed nonstationary sources," in *Proceedings of 3rd Sensor Array and Multichannel Signal Processing Workshop (SAM2004)*, Barcelona, Spain, July 2004, pp. CDRom.
6. Iain Russell, Jiangtao Xi, and Alfred Mertins, "Time Domain Blind Separation of Nonstationary convolutively mixed signals," in *Signal Processing for Telecommunications and Multimedia*, Vol. 27, Springer, New York, 2004, pp. 15-29.
7. Iain Russell, Jiangtao Xi, and Alfred Mertins, "Global optimization of uninitialized convolutive blind signal separation problems in the time domain," *Proceedings of 3rd Workshop on the Internet, Telecommunications, and Signal Processing (WITSP 2004)*, Adelaide, Australia, December 2004, pp. CDRom.
8. Iain Russell, Jiangtao Xi, Alfred Mertins, and Joe Chicharo, "Uninitialized sub-band blind signal separation of nonstationary convolutively mixed signals in acoustics using global optimization," Submitted to *IEEE Transactions on Speech and Audio Processing*.

# Abstract

The focus of this thesis is to develop a framework for solving convolutively mixed blind signal separation problems in the subband domain. Current methods generally employ a discrete Fourier transform (DFT) to change the time domain convolutive model into many instantaneous multiplicative models to save on computations and convergence time. The motivation for approaching the problem from the subband domain is that there is an upper bound on the quality of separation for frequency domain methods where the mixing is done in a reverberant environment and there is a high number of unknown variables to solve for. This is shown with reference to the works in (S. Araki, S. Makino, T. Nishikawa, and H. Saruwatari, 2001; M. Ikram, and D. Morgan, 2000; R. Mukai, S. Araki, H. Sawada, and S. Makino, 2004). The model is developed throughout the thesis in a series of stages. Firstly we investigate modelling the convolutive Blind Signal Separation (BSS) problem completely in the time domain. The benefit of this is that by not performing any transforms we eliminate the local frequency permutation problem that is inherent in all convolutive BSS problems. To solve the permutation problem requires additional computational overhead. There is a tradeoff however according to how complex the mixing/demixing system is. The longer the reverberation time of an acoustic environment, the more unknown variables must be solved. The savings of performing multiplication in the frequency domain as opposed to convolution in the time domain must be compared to the savings of not doing the transform operator twice, as well as ensuring the local

permutation problem is solved.

Two new algorithms that avoid the local permutation problem are proposed and investigated. The first uses an alternating least squares approach (ALS) while the second uses joint diagonalization of output correlation matrices of the recovered signals. Where it is plausible to assume that we have some sort of a priori information that provides a good initial starting point for the unknown demixing system, then we only need to consider some type of local optimization procedure to solve the unknown demixing system. Two local optimization procedures investigated include the steepest gradient descent and Newton methods. Both types of local solvers were compared and the merits and disadvantages of each are specified in regards to the convolutive BSS time domain algorithm proposed. Where small convolutive mixing systems exist, such as in wireless communication mixing systems that assume a two ray model, the computational overhead that is increased by doing convolution in the time domain is offset more by the savings of not having to solve the local permutation problem and execute the transform operation.

In some cases, information pertaining to problem is unavailable. Geometric source separation assumes that there is some additional knowledge about the layout of the sensors with spatial reference to the source positions. This allows an angle of incidence of the sound wave impinging on the sensor array to either be known directly or calculated using various beamforming techniques. If we cannot assume to know such information, then multivariate complex problems with a high number of parameters become harder to solve for without getting spurious results from ill-convergence to local multiminima as opposed to the preferred global minima which corresponds to the desired demixing system that will allow signal separation. To avoid this, we integrate one of the proposed time domain convolutive BSS algorithms with a global optimization routine that is catered to suit the BSS convolutive problem model. A

branch and bound algorithm that uses division by hyper-rectangles is used to solve the uninitialized optimization BSS problem. With the validity of the proposed BSS time domain convolutive algorithm and the global optimization approach being justified, attention will then be focused on integrating these contributions into a model which uses subband decomposition before performing signal separation.

Various methods of subband decomposition are considered including using a uniform FIR analysis/synthesis filter bank based on DFT modulation as well as cosine modulation. The prototype window used is based on an extended lapped transform and was chosen due to the computational benefits of using lapped transforms. A framework for developing such a subband model is made with the main aspects of the model being the BSS algorithm and optimization approaches used, the way in which the observed signals from a multiple-input-multiple-output (MIMO) mixing system are decomposed via a filter bank, and the way in which the local permutation problem is overcome. In our work we propose a new subband detection, correction, and sorting routine for separated but arbitrarily permuted subbands over the entire spectrum.

Finally, a general and systematic approach for obtaining experimental measurements for generating the impulse response of an acoustic environment such as a typical office room, as well as the inverting MIMO system using wiener-hopf and optimal filtering theory is presented to allow full availability of information for the problem modelled in a practical environment as opposed to synthetic testing methods which are also examined.

# Contents

<b>1 Preliminaries</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.2 Problem Statement . . . . .	3
1.3 Blind Signal Separation: A Historical Review . . . . .	8
1.4 Motivations and Applications . . . . .	10
1.5 Approach and Contributions of this Thesis . . . . .	12
1.5.1 Literature Review, Pre-Requisites and Outstanding Issues . .	12
1.5.2 Acoustic Modelling in Reverberant Environments . . . . .	13
1.5.3 Time and Frequency Domain Convolutional BSS Algorithms .	13
1.5.4 Uninitialized BSS with Global Optimization . . . . .	14
1.5.5 Subband BSS Model . . . . .	15
1.6 Summary of Contributions in Order of Presentation . . . . .	15
<b>2 Literature Review, Pre-requisites, and Outstanding Issues</b>	<b>18</b>
2.1 Introduction . . . . .	18
2.2 General criteria for Signal Separability . . . . .	19
2.3 Source Modelling . . . . .	20
2.4 Mixing Systems . . . . .	21
2.5 Measuring Performance . . . . .	25

2.6	Stationary Instantaneous BSS . . . . .	28
2.6.1	Higher Order Statistics . . . . .	29
2.6.2	Estimation Theory . . . . .	30
2.6.3	Information Theory . . . . .	32
2.6.4	Global Scaling and Permutation Ambiguities . . . . .	34
2.7	Stationary Convolutional BSS . . . . .	35
2.7.1	Frequency Domain Methods . . . . .	36
2.7.2	Local Scaling and Permutation Ambiguities . . . . .	37
2.8	Non-stationary Sources . . . . .	38
2.9	Subband Decomposition . . . . .	39
2.9.1	Fundamental Concepts . . . . .	39
2.9.2	Maximally Decimated Filter Banks and QMF . . . . .	42
2.9.3	Block Transforms . . . . .	45
2.9.4	Lapped Transforms . . . . .	48
2.9.5	Extended Lapped Transform . . . . .	50
2.9.6	Conclusion . . . . .	53
2.10	Identifying Areas for Contribution . . . . .	54
<b>3</b>	<b>Acoustic Modelling</b>	<b>66</b>
3.1	Introduction . . . . .	66
3.2	Acquiring Room Response . . . . .	67
3.3	Inverting the Room . . . . .	71
3.4	Conclusions . . . . .	76
<b>4</b>	<b>Time and Frequency Domain Convolutional BSS models</b>	<b>77</b>
4.1	Introduction . . . . .	77

4.2	ALS Approach . . . . .	79
4.2.1	Algorithmic Model . . . . .	81
4.2.2	Simulation Results . . . . .	86
4.3	Fullband TDBSS Approach . . . . .	89
4.3.1	General Overview . . . . .	89
4.3.2	Review of Instantaneous BSS . . . . .	91
4.3.3	Proposed Convolutional TDBSS . . . . .	93
4.3.4	Simulation Results . . . . .	98
4.4	Conclusions . . . . .	106
<b>5</b>	<b>Uninitialized BSS with Global Optimization</b>	<b>107</b>
5.1	Introduction . . . . .	107
5.2	DiRect Algorithm . . . . .	111
5.2.1	Convex Hulls . . . . .	117
5.2.2	Clustering . . . . .	117
5.2.3	Sequential Quadratic Programming . . . . .	119
5.3	Simulation Results . . . . .	121
5.4	Benefits for Small to Medium Scale Systems . . . . .	124
5.5	Conclusions . . . . .	126
<b>6</b>	<b>Subband BSS Model</b>	<b>127</b>
6.1	Introduction . . . . .	127
6.2	Frequency Domain Methods and Limitations . . . . .	130
6.3	Subband Model . . . . .	133
6.3.1	Cosine Modulated FB . . . . .	134
6.3.2	DFT Modulated FB . . . . .	135

6.4	Integration of TDBSS into Subband Model . . . . .	136
6.5	Comparing Subband and Frequency Domain BSS Models . . . . .	138
6.6	Subband Coupling Metric . . . . .	140
6.7	Dyadic Sorting Routine . . . . .	142
6.8	Simulation Results . . . . .	144
6.8.1	Benchmark ICA'99 Dataset . . . . .	144
6.8.2	Synthetic Testing . . . . .	145
6.8.3	Real Testing . . . . .	151
6.9	Conclusions . . . . .	153
<b>7</b>	<b>Conclusions and Suggestions for Further Research</b>	<b>154</b>
7.1	Conclusions . . . . .	154
7.2	Suggestions for Future Research . . . . .	157
<b>A</b>	<b>Proof of Closed Form Analytical Expressions for Gradient and Hessian</b>	<b>169</b>
A.1	Proof of Closed Form Analytical Expressions for Gradient and Hessian	169



# List of Figures

1.1	General Mixing/Demixing system BSS model . . . . .	4
1.2	$\tilde{M}$ -channel uniform filter QMF bank (P. Vaidyanathan, 1993). . . .	5
1.3	BSS model of a two-input-two-output (TITO) system. . . . .	6
1.4	BSS model of a TITO system using subband decomposition. . . . .	7
2.1	Multipath Propagation in a room. . . . .	23
2.2	Multirate building blocks . . . . .	40
2.3	Decimation by factor $R = 2$ in the time and frequency domain . . .	40
2.4	Expanding by factor $R = 2$ in the time and frequency domain . . . .	41
2.5	Frequency responses of an $\tilde{M}$ th band filter $H(z)$ . . . . .	44
2.6	Polyphase structure of $\tilde{M}$ -channel QMF maximally decimated filter bank . . . . .	45
2.7	Basis functions of block transforms and lapped transforms . . . . .	48
2.8	Westner conference room photo (A. Westner, and V. Bove, 1999). .	54
2.9	Westner conference room layout (A. Westner, and V. Bove, 1999). .	57
3.1	A section of a typical MLS signal . . . . .	69
3.2	Typical polar patterns for Shure SM57 (Shure Incorporated, 2004) .	70
3.3	Digidesign Digi001 8 channel analogue I/O with 48kHz sampling. .	70
3.4	Geometrical properties of room layout and direction of arrival of sources impinging on array manifold. . . . .	71

3.5	Measured room impulse responses with reverberation time of 200ms i.e. $P=1600$ , for down-sampled rate of 8kHz for a TITO system. . .	72
3.6	System identification of the inverse FIR filter for a SISO system. . .	73
3.7	MIMO system identification for demixing FIR system using MMSE. . .	74
3.8	Wiener solution to TITO FIR demixing system with reverberation time of 250ms i.e. $Q=2000$ , delay of for down-sampled rate of 8kHz. . .	75
3.9	Global TITO FIR system, i.e. cascaded mixing and demixing FIR systems. . . . .	76
4.1	Value of objective function on a logarithmic scale versus the number of full iterations. . . . .	87
4.2	Source power spectral densities (psd's) and their estimates. Legend: — psd1; - · - psd2; o estimate for psd1; x estimate for psd2. . .	88
4.3	Convergence of differing optimization methods for instantaneous BSS. . .	100
4.4	Convergence of gradient descent and Newton algorithms for a first order TITO FIR demixing system over 10 trials. . . . .	102
4.5	Convergence of Newton algorithms for first and third order TITO FIR demixing systems over 10 trials. . . . .	104
4.6	(a) and (b) are the two original signals, (c) and (d) are the convolatively mixed signals, (e) and (f) are the permuted separated results. . .	105
5.1	An iteration of the Direct algorithm for $n = 2$ dimensions. i.e. rectangles (D. R. Jones, C. D. Perttunen, and B. E. Stuckman, 1992) . .	116
5.2	An iteration of the Direct algorithm for $n = 3$ dimensions. i.e. rectangular prism (D. R. Jones, C. D. Perttunen, and B. E. Stuckman, 1992) . . . . .	116
5.3	Hyper-rectangles on lower right convex hull (D. R. Jones, DIRECT, 2001) . . . . .	119
5.4	Comparison of global and local optimization routines, glcCluster and snopt . . . . .	123
5.5	(a) Above shows 'flops' using glcCluster and the method from (L. Parra and C. Spence, 2000) (b) Below shows is a magnification of (a). . .	125

6.1	General subband MIMO BSS model with oversampling factor $\frac{\tilde{M}}{R}$ . . .	135
6.2	(a) Shows the impulse responses for the analysis filters for the first two subbands and (b) shows the magnitude frequency responses of the first three subbands for an ELT CM FIR FB when $\tilde{M} = 8$ . . . .	136
6.3	Separation performance using three different BSS techniques for two TIMIT speech segments recorded with two cardioid microphones in a reverberant office environment. . . . .	140
6.4	Example of dyadic permutation sorting algorithm from (K. Rahbar, and J. Reilly, 2001) for separated subbands when total number of subbands $\tilde{M} = 8$ . . . . .	144
6.5	Virtual room synthetic mixing environment at 8kHz sampling frequency generated with the simroommix.m function. . . . .	146
6.6	Virtual room synthetic mixing impulse responses at 8kHz sampling frequency and reverberation time of 130 ms. . . . .	147
6.7	Wiener solution to TITO FIR synthetic demixing system with reverberation time of 200ms i.e. $Q=1600$ , delay=32ms. . . . .	148
6.8	Ideal global TITO FIR system for virtual room. . . . .	150
6.9	First two separated adjacent subbands for $p = 1, 2$ when $\tilde{M} = 256, R = 128$ . . . . .	151

# List of Tables

4.1	Closed form analytical expressions for the gradient and Hessian of the cost function and row-normilization constraint. . . . .	97
4.2	Gradient descent subband BSS algorithm for the joint-diagonalization task with a weighted constraint. . . . .	98
4.3	Newton-type subband BSS algorithm for the joint-diagonalization task with a weighted constraint. . . . .	98
5.1	Steps for algorithm "DIRECT" (M. Björkman and K. Holmström, 1999) . . . . .	112
5.2	Parameters for algorithm "DIRECT" (M. Bjorkman and K. Holmström, 1999) . . . . .	113
5.3	Steps for algorithm "DIRECT" (D. R. Jones, C. D. Perttunen, and B. E. Stuckman, 1992) . . . . .	114
5.4	Steps for algorithm "Conhull" (M. Björkman and K. Holmström, 1999)	118
6.1	Separation performance comparison using SIR . . . . .	151

# List of Abbreviations

ABF	Adaptive Beamformers
AC	Alternating Columns
AIR	Acoustical Impulse Response
ALS	Alternating Least Squares
ANC	Adaptive Noise Cancellation
AR	Auto Regressive
ARMA	Auto Regressive Moving Average
BSS	Blind Signal Separation
CM	Cosine Modulated
CNV	Central Nervous System
DC	Diagonal Centers
DFT	Discrete Fourier Transform
DIRECT	Dividing Rectangles
DOA	Direction of Arrival
DSP	Digital Signal Processing
EEG	Electro-encephalographic
ELT	Extended Lapped Transform
EVA	Eigenvector solution for Blind Equalization
FB	Filter Bank
FIR	Finite Impulse response

---

FLOPS	Floating point Operations
GSM	Global System for MObile Communication
HOS	Higher Order Statistic
ICA	Independent Component Analysis
IID	Independent and Identically Distributed
IIR	Infinite Impulse Response
JADE	Joint Approximate Diagonalization of Eigenmatrices
KLD	Kullback-Liebler Divergence
LOT	Lapped Orthogonal Transform
LTl	Linear Time Invariant
MA	Moving Average
MAP	Maximum a Posteriori
MEG	Magneto-encephalographic
MI	Mutual Information
MIMO	Multiple Input Multiple Output
ML	Maximum Likelihood
MLS	Maximum Length Sequence
MLT	Modulated Lapped Transform
MRI	Magnetic Resonance Imaging
MSE	Minimum Squared Error
NP	Nonlinear Programming
NRR	Noise Reduction Ratio
PCA	Principle Component Analysis
PCI	Peripheral Component Interconnect
PI	Performance Index
PR	Perfect Reconstruction

QAM	Quadrature Amplitude Modulation
QMF	Quadrature Mirror Filter
QP	Quadratic Programming
QPSK	Quadrature Phase Shift Keying
RHS	Right Hand Side
SBSS	Subband Blind Signal Separation
SGD	Steepest Gradient Descent
SIR	Signal to Interference Ratio
SISO	Single Input Single Output
SNOPT	Sparse Nonlinear Optimization
SNR	Signal to Noise Ratio
SOS	Second Order Statistic
SQP	Sequential Quadratic Programming
SSARS	Source Separation Algorithm with Reference System
STFT	Short Time Fourier Transform
TD	Time Domain
TITO	Two Input Two Output