CÉSAR SAULO BELLI ARROYO

**ESTIMAÇÃO DE IMAGENS ACÚSTICAS COM
ARRANJOS DE MICROFONES**

São Paulo
2015

CÉSAR SAULO BELLI ARROYO

**ESTIMAÇÃO DE IMAGENS ACÚSTICAS COM
ARRANJOS DE MICROFONES**

Dissertação apresentada à Escola Politécnica
da Universidade de São Paulo para obtenção
do Título de Mestre em Engenharia Elétrica.

São Paulo
2015

CÉSAR SAULO BELLI ARROYO

**ESTIMAÇÃO DE IMAGENS ACÚSTICAS COM
ARRANJOS DE MICROFONES**

Dissertação apresentada à Escola Politécnica
da Universidade de São Paulo para obtenção
do Título de Mestre em Engenharia Elétrica.

Área de Concentração:
Sistemas Eletrônicos

Orientador:
Prof. Dr. Vítor H. Nascimento

São Paulo
2015

Este exemplar foi revisado e corrigido em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, _____ de _____ de _____

Assinatura do autor: _____

Assinatura do orientador: _____

Catlogação-na-publicação

, César Saulo Belli Arroyo
ESTIMAÇÃO DE IMAGENS ACÚSTICAS COM ARRANJOS DE
MICROFONES / C. S. B. A. -- versão corr. -- São Paulo, 2015.
121 p.

Dissertação (Mestrado) - Escola Politécnica da Universidade de São
Paulo. Departamento de Engenharia de Sistemas Eletrônicos.

1.Processamento digital de sinais 2.Acústica 3.Estimação de parâmetros
4.Transformadas rápidas I.Universidade de São Paulo. Escola Politécnica.
Departamento de Engenharia de Sistemas Eletrônicos II.t.

Aos meus pais, e minha noiva Fiorella
Andrade.

AGRADECIMENTOS

Ao Prof. Dr. Vítor H. Nascimento pela grande oportunidade, imensurável apoio e por ter acreditado em mim fazendo possível este trabalho.

À Marinha de Guerra do Peru e aos Serviços Industriais da Marinha - SIMA, por ter me dado a grande oportunidade e apoio econômico para vir ao Brasil e poder seguir meus estudos de pós-graduação.

À Universidade de São Paulo e à Escola Politécnica; especialmente ao Departamento de Engenharia de Sistemas Eletrônicos, pelo apoio e grata disposição.

Aos colegas Dr. Bruno Masiero, Felipe Luiz F. O. Chamon, Felipe Durán V. G. dos Santos pelas contribuições feitas no transcorrer deste trabalho.

A todos que de alguma forma contribuíram e involutariamente foram omitos.

RESUMO

Nas últimas décadas, a poluição sonora tornou-se um grande problema para a sociedade. É por esta razão que a indústria tem aumentado seus esforços para reduzir a emissão de ruído. Para fazer isso, é importante localizar quais partes das fontes sonoras são as que emitem maior energia acústica. Conhecer os pontos de emissão é necessário para ter o controle das mesmas e assim poder reduzir o impacto acústico-ambiental.

Técnicas como “beamforming” e “Near-Field Acoustic Holography” (NAH) permitem a obtenção de imagens acústicas. Essas imagens são obtidas usando um arranjo de microfones localizado a uma distância relativa de uma fonte emissora de ruído. Uma vez adquiridos os dados experimentais pode-se obter a localização e magnitude dos principais pontos de emissão de ruído. Do mesmo modo, ajudam a localizar fontes aeroacústicas e vibro acústicas porque são ferramentas de propósito geral. Usualmente, estes tipos de fontes trabalham em diferentes faixas de frequência de emissão.

Recentemente, foi desenvolvida a transformada de Kronecker para arranjos de microfones, a qual fornece uma redução significativa do custo computacional quando aplicada a diversos métodos de reconstrução de imagens, desde que os microfones estejam distribuídos em um arranjo separável.

Este trabalho de mestrado propõe realizar medições com sinais reais, usando diversos algoritmos desenvolvidos anteriormente em uma tese de doutorado, quanto à qualidade do resultado obtido e à complexidade computacional, e o desenvolvimento de alternativas para tratamento de dados quando alguns microfones do arranjo apresentarem defeito.

Para reduzir o impacto de falhas em microfones e manter a condição de que o arranjo seja separável, foi desenvolvida uma alternativa para utilizar os algoritmos rápidos, eliminando-se apenas os microfones com defeito, de maneira que os resultados finais serão obtidos levando-se em conta todos os microfones do arranjo.

Palavras-Chave – Processamento do arranjo, Imagens acústicas, Formação de feixe, Arranjo de microfones.

ABSTRACT

In recent decades, noise pollution has become a major problem for society. It is for this reason that the industry has increased its efforts to reduce the emission of noise. To do this, it is important to find out which parts of the sound sources are emitting greater acoustic energy. Knowing the emission points is required to keep track of them and thus be able to reduce acoustic and environmental impact.

Techniques such as “beamforming” and “Near-Field Acoustic Holography” (NAH) allow obtaining acoustic images. These images are obtained using an array of microphones located at a relative distance from a noise source. Once acquired experimental data, one can obtain the location and the magnitude of the main points of noise emission. Similarly, we can find aeroacoustic and vibroacoustic sources, because they are general-purpose tools. Usually, these types of sources work in different frequency bands of the emission.

Recently, a Kronecker Array Transform (KAT) was developed to microphone arrays, which provides a significant reduction in computational cost when applied to various image reconstruction methods, provided that the microphones are distributed in a separable array.

This thesis proposes perform measurements with real signals using different algorithms previously developed in a dissertation about the quality of the obtained results and computational complexity, and the development of alternatives for data processing when some microphones of the array are defective.

To reduce the impact of failures in microphones and maintain the condition that the array is separable, one alternative has been developed to use the fast algorithms by removing only the defective microphones, so that the final results will be obtained taking into account every microphone of the array.

Keywords – Array processing, Acoustic images, Beamforming, Microphone array.

LISTA DE FIGURAS

1.1	Imagem acústica com método DAMAS2	1
1.2	Aeroporto de Congonhas, São Paulo - Brasil	2
1.3	Demonstração de uso do arranjo de sensores acústicos construídos pelos franceses na Primeira Guerra Mundial para detecção de aeronaves inimigas	3
1.4	Sea-Based X-Band Radar	4
1.5	Sonar CSU-3-22	5
1.6	de 27 antenas de rádio telescópio utilizado para análise de sinais eletromagnéticos em radioastronomia em Novo Mexico, USA	5
2.1	Sistema de coordenadas esféricas	14
2.2	Arranjo Linear Uniforme, ULA	16
2.3	Fonte Pontual em Campo Distante	17
3.1	Arranjo de microfones genérico	19
4.1	Discretização do espaço U	25
4.2	Exemplo de sinal original e sinal discretizado em janelas de K amostras	27
5.1	Delay and Sum Beamformer	30
5.2	Discretização do espaço com $M = 10 \times 10$ pontos do alto-falante	32
5.3	Relação arranjo de microfones com a discretização do espaço	32
5.4	Imagem depois de fazer o beamforming	32
5.5	Point - Spread Function de um arranjo uniforme linear $N = 8$ microfones	33
6.1	Arranjo separável	37
7.1	Arranjo de microfones com 64 elementos e com geometria separável	42
7.2	Placa de desenvolvimento Altera DE3	43
7.3	Alto-falante X-mini II	44

7.4	Estúdio do laboratório de Processamento de Sinais da Escola Politécnica da Universidade de São Paulo	45
7.5	Instalação usada para as medições com o arranjo de microfones	45
7.6	Laser STHT77340 CUBIX 12M da fabricante Stanley	46
7.7	Geometria separável do arranjo de microfones $N_x = N_y = 7$	47
7.8	Geometria separável do arranjo de microfones $N_x = 6, N_y = 7$	48
7.9	Esboço do estúdio com as posições do arranjo de microfones	49
7.10	Resultados com uma geometria de 49 microfones eliminando 15 elementos para manter a condição de separabilidade por causa dos 2 microfones inicialmente com defeito	51
7.11	Resultados com uma geometria de 42 microfones eliminando 22 elementos para manter a condição de separabilidade por causa dos 3 microfones com defeito . .	52
7.12	Comparação da resolução utilizando Delay and Sum beamformer	54
7.13	Comparação da resolução utilizando DAMAS2	55
7.14	Comparação da resolução utilizando o método de regularização ℓ_1	56
7.15	Resultados utilizando o método de regularização TV	57
7.16	Comparação dos resultados utilizando o método DAMAS2 modificando o número de iterações	59
7.17	Comparação dos resultados utilizando o método DAMAS2 ao longo de nove dias	60
7.18	Sinal de excitação <i>ruído branco</i>	62
7.19	Comparação da resolução utilizando regularização ℓ_1 , com uma frequência de 2 kHz e com calibração	63
7.20	Comparação da resolução utilizando regularização ℓ_1 , com uma frequência de 4 kHz e com calibração	64
7.21	Comparação da resolução utilizando regularização TV, com uma frequência de 2 kHz e com calibração	65
7.22	Comparação da resolução utilizando regularização TV, com uma frequência de 4 kHz e com calibração	66
7.23	Comparação da resolução utilizando DAMAS2, com 2048 amostras por janela .	67

7.24	Medição com um alto-falante na posição nº 1, utilizando a calibração do arranjo com frequência de 2 kHz	69
7.25	Medição com um alto-falante na posição nº 4, utilizando a calibração do arranjo com frequência de 2 kHz	70
7.26	Medição com um alto-falante na posição nº 1, utilizando a calibração do arranjo com frequência de 2 kHz e colocando espuma no solo	71
7.27	Medição com um alto-falante na posição nº 3, utilizando a calibração do arranjo com frequência de 2 kHz e colocando espuma no solo	72
7.28	Identificação lóbulos principais e laterais realizando uma medição com um alto-falante na posição nº 1, utilizando a calibração do arranjo com frequência de 2 kHz	74
7.29	Identificação lóbulos principais e laterais realizando uma medição com um alto-falante na posição nº1, utilizando a calibração do arranjo com frequência de 5 kHz	75
7.30	Resultados utilizando todos os métodos de reconstrução de imagens, com frequência de 2 kHz e com calibração do arranjo	79
7.31	Comparação de imagens acústicas utilizando 1 kHz, 2 kHz e 5 kHz com o método DAS	81
7.32	Comparação de imagens acústicas utilizando 1 kHz, 2 kHz e 5 kHz com o método DAMAS2	81
7.33	Comparação de imagens acústicas utilizando 1 kHz, 2 kHz e 5 kHz com o método de reconstrução covariance fitting utilizando regularização ℓ_1	82
7.34	Comparação de imagens acústicas utilizando 1 kHz, 2 kHz e 5 kHz com o método de reconstrução covariance fitting utilizando regularização TV	82
7.35	Um alto-falante no lado esquerdo com relação ao centro do arranjo de microfones, com uma frequência de 2 kHz e com calibração do arranjo	83
7.36	Um alto-falante no lado esquerdo com relação ao centro do arranjo de microfones, com uma frequência de 5 kHz e com calibração do arranjo	84
7.37	Um alto-falante no lado direito com relação ao centro do arranjo de microfones, com uma frequência de 2 kHz e com calibração do arranjo	85

7.38	Um alto-falante no lado direito com relação ao centro do arranjo de microfones, com uma frequência de 5 kHz e com calibração do arranjo	86
7.39	Um alto-falante no centro com relação ao arranjo de microfones com distância de 2.2 m, com uma frequência de 2 kHz e com calibração do arranjo	88
7.40	Um alto-falante no centro com relação ao arranjo de microfones com distância de 2.2 m, com uma frequência de 5 kHz e com calibração do arranjo	89
7.41	Dois alto-falantes no centro com relação ao arranjo de microfones com distância de 1.5 m e uma separação de 1.3 m, com sinais correlacionados, utilizando a geometria de 42 microfones	90
7.42	Dois alto-falantes no centro com relação ao arranjo de microfones com distância de 1.5 m e uma separação de 1.3 m, com sinais correlacionados, utilizando a geometria de 61 microfones	91
7.43	Sinais obtidas utilizando o ITA-toolbox	92
7.44	Dois alto-falantes com sinais não correlacionados (<i>ruído branco - ruído branco</i>), utilizando a geometria de 42 microfones	93
7.45	Dois alto-falantes com sinais não correlacionados (<i>ruído branco - ruído branco</i>), utilizando a geometria de 61 microfones	93
7.46	Dois alto-falantes com sinais não correlacionados (<i>ruído rosa - ruído rosa</i>), utilizando a geometria de 42 microfones	94
7.47	Dois alto-falantes com sinais não correlacionados (<i>ruído rosa - ruído rosa</i>), utilizando a geometria de 61 microfones	95
7.48	Dois alto-falantes com sinais não correlacionados (<i>ruído branco - ruído rosa</i>), utilizando a geometria de 42 microfones	95
7.49	Dois alto-falantes com sinais não correlacionados (<i>ruído branco - ruído rosa</i>), utilizando a geometria de 42 microfones	96

LISTA DE TABELAS

7.1	Tempo médio de cálculo para obtenção de imagens acústicas com diferentes algoritmos utilizando a KAT.	48
7.2	Tabela de tempos de reconstrução de imagens - Resolução.	58
7.3	Tabela com resultados dos valores pela simulação e dos 9 testes utilizando o método DAMAS2 com uma frequência de 2 kHz e com calibração	61
7.4	Tabela com resultados do custo computacional entre Ay e a KAT	78

LISTA DE SÍMBOLOS

Símbolos	Definição
τ	Atraso do sinal para um microfone
ϕ	Azimute
$*$	Conjugado complexo
$\mathbf{q}_m$	Coordenadas de uma fonte em coordenadas Cartesianas
$\mathbf{u}_m$	Coordenadas de uma fonte parametrizada no espaço U
$\mathbf{p}_n$	Coordenadas de um sensor em coordenadas Cartesianas
θ	Elevação
ω	Frequência angular em <i>rads/s</i>
$\mathbf{I}$	Intensidade sonora
$\mathbf{S}(\omega)$	Matriz de autocorrelação na frequência ω
$\mathbf{P}$	Matriz de permutação
$E\{\cdot\}$	Média temporal
M	Número de fontes incidendo no arranjo
N	Número de sensores no arranjo
$vec\{\cdot\}$	Operador de vetorização
σ^2	Potência do ruído branco
Y_m	Potência emitida por cada fonte
p	Pressão sonora em <i>Pa</i>
$\otimes$	Produto de Kronecker
T	Transposta de matriz ou vetor
H	Transposta Hermitiana de matriz ou vetor
$\mathbf{v}$	Velocidade da partícula
c	Velocidade do som em <i>m/s</i>
$\mathbf{r}$	Vetor de direção radial
$\mathbf{k}$	Vetor de onda
$\mathbf{a}$	Vetor de propagação
$\mathbf{v}(\omega)$	Vetor diretor de um arranjo
$\mathbf{w}$	Vetor número de pesos

SUMÁRIO

1	Introdução	1
1.1	Introdução	1
1.2	Objetivo do Estudo	8
1.3	Justificativa do Estudo	8
1.4	Metodologia do Estudo	9
1.5	Organização do Texto	9
2	Fundamentos Teóricos	10
2.1	Introdução	10
2.2	O som	10
2.3	Campo Acústico	11
2.4	Propagação de ondas sonoras	11
2.4.1	Velocidade do som	11
2.4.2	Equação da onda	11
2.4.3	Nível de pressão sonora	13
2.4.4	Intensidade Sonora	13
2.4.5	Coordenadas esféricas	14
2.5	Arranjos de Microfones	15
2.5.1	Arranjo Linear Uniforme (ULA)	15
2.6	Consideração de Campo Distante e Campo Próximo	16
3	Modelo dos sinais	18
3.1	Introdução	18
3.2	Resposta tempo - frequência	18

3.2.1	Obtenção do vetor $\hat{\mathbf{x}}(\omega)$	20
3.2.1.1	Com relação ao atraso $\tau_{n,m}$	20
3.2.1.2	Com relação ao vetor diretor (<i>array manifold vector</i>)	20
3.3	Reconstrução de Imagens	22
4	Estimação da matriz de autocorrelação	24
4.1	Introdução	24
4.2	Parametrização no espaço $\mathbf{U}$	24
4.3	Possível método de Solução	26
4.3.1	Cálculo do vetor $\tilde{\mathbf{x}}(\omega_k)$	26
4.3.2	Aproximação da matriz de autocorrelação $\mathbf{S}_{\mathbf{x}}(\omega_k)$	27
5	Beamformers	29
5.1	Introdução	29
5.2	Delay and Sum Beamforming	29
5.3	Point - Spread Function (PSF)	31
5.4	Beamformer MVDR	33
5.5	Beamformer MPDR	35
6	Transformada rápida para imagens acústicas	36
6.1	Introdução	36
6.2	Kronecker array transform (KAT)	37
6.3	Deconvolução de imagens acústicas	38
6.3.1	DAMAS	39
6.3.2	Covariance fitting	40
6.3.3	Regularização ℓ_1	40
6.3.4	Regularização TV	41

7	Resultados Finais e Equipamento utilizado	42
7.1	Introdução	42
7.2	Instalações para medição	44
7.3	Modificações	46
7.4	Resultados iniciais	47
7.4.1	Programa de simulação	50
7.4.2	Resultados utilizando a geometria de 49 e 42 microfones	50
7.4.3	Resultados modificando os parâmetros originais dos métodos	53
7.4.3.1	Resolução (modificando o valor da discretização do espaço) . . .	53
7.4.3.2	Modificação do número de iterações para o método DAMAS2 .	53
7.4.3.3	Modificação do número de iterações e do valor de sigma (σ) para a regularização ℓ_1	58
7.4.3.4	Modificação do número de iterações e o valor de (μ) para a regularização TV	61
7.4.3.5	Modificação do número de janelas (L) para realizar a estimação da matriz de autocorrelação $\mathbf{S}$	62
7.4.3.6	Medição com material de espuma no solo do estúdio	67
7.4.3.7	Medição com material de espuma no solo do estúdio e sem recobrimento na janela	68
7.4.3.8	Comparação dos artefatos dos resultados obtidos	71
7.5	Modificação dos métodos de reconstrução de imagens para eliminar só os microfones com defeito	73
7.6	Resultados finais utilizando os melhores parâmetros e a nova modificação nos método de reconstrução de imagens	80
7.6.1	Padrão: Um alto-falante no centro com relação ao arranjo de microfones	80
7.6.2	Padrão: Um alto-falante lado esquerdo com relação ao arranjo de microfones	83
7.6.3	Padrão: Um alto-falante lado direito com relação ao arranjo de microfones	85
7.6.4	Padrão: Um alto-falante na frente do arranjo com distância de 2,2 m . .	87

7.6.5	Padrão: Dois alto-falantes na frente do arranjo de microfones com sinais correlacionados	87
7.6.6	Padrão: Dois alto-falantes na frente do arranjo de microfones com sinais não correlacionados	90
8	Conclusões	97
8.1	Trabalhos Futuros	98
	Referências	100
	Apêndice A – Publicações relevantes	103

1 INTRODUÇÃO

1.1 Introdução

Imagens acústicas são amplamente utilizadas na indústria para a detecção de fontes de som. Permitem detectar a posição e a frequência de uma fonte de ruído. Este tipo de ferramenta permite trabalhar no controle de ruído em diferentes equipamentos industriais, na indústria mecânica e na otimização de projetos aerodinâmicos.

O objetivo é mapear a emissão acústica de uma máquina ou fonte de som, para obter informações sobre quais regiões irradiam mais ruído. Isto permitiria adotar medidas preventivas para reduzir a emissão de ruído.

Uma imagem acústica é um mapa da intensidade de som que incide sobre um arranjo de microfones, em função da direção de chegada. Como podemos ver na Figura 1.1, a intensidade sonora é representada por uma escala de cores. Neste exemplo, a cor vermelha representa maior intensidade e a cor azul intensidade menor.

Figura 1.1: Teste de um alto-falante (lado inferior esquerdo), frequência 500 Hz. Método DAMAS2, obtido com um arranjo de 49 microfones, como descrito na Seção 7.3.

Podemos dizer que as imagens acústicas são um método de diagnóstico da origem do ruído, o qual permite identificar de uma maneira gráfica e intuitiva as partes da máquina que emitem níveis de ruído elevado, com foco de possíveis alternativas de melhora. Identificar a fonte de ruído envolve a obtenção exata da posição e o conteúdo espectral dominante da fonte do ruído [1].

Alguns casos em que imagens acústicas são aplicadas são no projeto de aviões [2], turbinas de aviões e trens de pouso [3]. Outra aplicação onde podemos ver imagens acústicas é em aerodinâmica, por exemplo, para redução de ruído pela turbulência de geradores eólicos [4]. Aeronaves e veículos de qualquer tipo produzem ruído aerodinâmico que depende da velocidade, por exemplo, os trens modernos [5,6] superam os 350 km/h.

Devido ao aumento da população mundial, em muitos casos, aeroportos foram imersos no meio de cidades, gerando uma grande quantidade de poluição sonora, como é mostrado na Figura 1.2. Em geral, o ruído do tráfego também é um problema para os moradores que vivem perto das estradas ou grandes avenidas.

Figura 1.2: Aeroporto de Congonhas, São Paulo - Brasil.

As técnicas de processamento de sinais em arranjo de sensores foram inicialmente desenvolvidas com o uso de antenas de radar durante a Primeira Guerra Mundial (**World War I**). De acordo com Johnson e Dudgeon [7], a primeira utilização de um arranjo de sensores na identificação de fontes sonoras foi implementada, provavelmente, pelos franceses durante a **WWI** para localizar aeronaves inimigas. Na Figura 1.3 é apresentado um dispositivo constituído por dois arranjos, cada um apresentando seis sensores acústicos em forma de cornetas invertidas, distribuídos na forma de um hexágono, o que hoje seria comparável a um arranjo circular de seis microfones. Os sensores de cada arranjo eram ligados a um duto que funcionava como uma guia de onda acústica. Os dois dutos de igual comprimento eram levados aos ouvidos de um operador. O sinal que chegava no ouvido era um resultado médio do sinal dos seis sensores. Os dois arranjos eram separados por dois metros, o que melhorava a capacidade natural de

localização direcional. Modificando o ângulo do eixo dos dois arranjos ao mesmo tempo, a direção do som incidente sobre os arranjos poderia ser determinada. Este é um exemplo simplista de um Beamforming (formador de feixe) realizado pelos sensores em tempo real através do processamento analógico do sinal sonoro.

Figura 1.3: Demonstração de uso do arranjo de sensores acústicos construídos pelos franceses na Primeira Guerra Mundial para detecção de aeronaves inimigas. Retirado de [8].

As pesquisas sobre o uso do sonar aumentaram durante a Segunda Guerra Mundial (**World War II**), devido à necessidade de detectar submarinos [7]. Arranjos de hidrofones começaram a ser usados para melhorar a capacidade do SONAR (*Sound Navigation And Ranging Radio*) na localização de submarinos. Nesta aplicação, ondas acústicas são emitidas e suas reflexões são analisadas pelo arranjo. Esta classe de arranjo é definida como arranjos ativos. Aplicações passivas são utilizadas em submarinos, evitando a emissão de sinais que possam ser rastreados. Diversas aplicações submarinas são descritas por Urick [9].

Em aplicações militares, em particular na indústria de construção naval e de equipamento militar, um dos objetivos de projeto é não gerar muito ruído para evitar a detecção pelo inimigo. É por isso que, na indústria militar em particular são utilizados diferentes métodos para reduzir os problemas de ruído nos equipamentos industriais usando técnicas de processamento de sinais. Por exemplo, imagens acústicas são empregadas na construção da área de propulsão de um

navio para torná-lo menos barulhento. Imagens da pressão sonora na saída de turbinas de mísseis OTOMAT (principal míssil da Marinha do Peru) podem ser usadas para aumentar sua eficiência, confiabilidade e diagnosticar falhas.

Na indústria militar, sendo mais específico no âmbito naval, unidades de combate como os submarinos e navios de guerra requerem equipamento especial para defesa e segurança. O RADAR (*Radio Detection And Ranging*) e SONAR são equipamentos que têm arranjos de antenas e hidrofones respectivamente em seu interior [10–12].

O maior Radar militar do mundo pertence à Marinha dos Estados Unidos, como podemos ver na Figura 1.4. O Radar “*X-Band Sea-Based*” fica no topo de uma plataforma flutuante e tem a capacidade de se comunicar com mísseis interceptores dos EUA em diferentes áreas do país.

Figura 1.4: Sea-Based X-Band Radar, Marinha dos Estados Unidos. Retirado de [13]

A Marinha do Peru ¹ tem em seu poder naval seis submarinos. Cada um deles tem um Sonar tipo CSU-3-22, o qual é composto por um arranjo cilíndrico com 96 hidrofones como podemos ver na Figura 1.5. O hidrofone é um transdutor de som em eletricidade para ser utilizado em água ou em outro líquido, de forma análoga à utilização de um microfone no ar.

Kellerman [14] indica que, desde a década de 60, muitas aplicações foram exploradas na radioastronomia utilizando arranjos de antenas para análise de ondas eletromagnéticas radiadas por corpos celestes. Na Figura 1.6 é apresentada uma fotografia de um grande arranjo de antenas atualmente usado em radioastronomia pelo Observatório de Rádio Astronomía Norte Americano [15]. O arranjo, conhecido como **VLA** (sigla em inglês para *Very Large Array*), é constituído por um conjunto de 27 antenas de rádio e está localizado a 80 km da cidade de

¹O Serviço Industrial da Marinha (SIMA - PERU), principal estaleiro do Peru, me outorgou a oportunidade de poder seguir meus estudos de pós-graduação na Universidade de São Paulo - Brasil, tendo o grau militar de Primeiro Tenente e com uma segunda especialidade em Engenharia Eletrônica.

Figura 1.5: Arranjo cilíndrico de Hidrofones, Sonar CSU-3-22.

Socorro, no estado americano do Novo México. O arranjo apresenta geometria em forma de “Y” e pode ser organizado em quatro diferentes configurações: A, B, C e D. Cada configuração define uma distância entre as antenas no arranjo e as separações máximas entre duas antenas são respectivamente de: 36 km, 10 km, 3.6 km e 1 km. Estas diferentes configurações do arranjo permitem observar ondas eletromagnéticas em frequências entre 74 MHz e 50 GHz. A modificação entre estas configurações ocorre em períodos de quatro meses.

Figura 1.6: Arranjo de 27 antenas de rádio telescópio utilizado para análise de sinais eletromagnéticos em radioastronomia, em Novo Mexico, USA. Retirado de [15]

As imagens acústicas são obtidas na prática utilizando arranjos de microfones [16]. O arranjo de microfones é baseado na ideia de que os sinais recebidos em um conjunto de microfones, dispostos em posições conhecidas no espaço, podem ser usados para estimar as direções de diferentes frentes de onda incidindo simultaneamente no arranjo.

A principal limitação do arranjo é a *resolução espacial* que descreve a capacidade de discriminar as fontes sonoras adjacentes [17]. Pode ser corrigida pelo aumento do número de

microfones, e, até certo ponto, por técnicas mais modernas de processamento de sinais como é apresentado em [18].

Na indústria, arranjos de microfones podem ser posicionados em túneis de vento para determinar a distribuição de ruído sobre modelos devido ao fluxo de ar em alta velocidade [19, 20]. Estas medições são utilizadas na construção de automóveis, trens e aviões para torná-los mais silenciosos.

Uma das técnicas mais aplicadas para a detecção de fontes aeroacústicas é o método de beamforming, que é baseado no modelo de *campo distante* [21, 22]. Beamforming é uma técnica de filtragem espacial, que permite descartar sons vindos de algumas direções, e reforçar sons vindos de outras direções. Essa característica permite aumentar a relação sinal-ruído de um sinal desejado, através da atenuação de ruído e interferências vindas de outras direções [7, 17]. Novos métodos de processamento de sinais têm sido desenvolvidos ao longo do tempo para melhorar estes resultados e reduzir o custo computacional [18, 21, 22].

Antes dos anos 80, o processamento de arranjo de microfones era feito através de uma técnica chamada de atraso e soma (*Delay and Sum*) [23]. Avanços computacionais permitiram a evolução da técnica de atraso e soma para métodos mais poderosos para formação de feixe (beamforming). Estes métodos são amplamente aplicáveis, se a distância entre a fonte e o arranjo de microfones for de pelo menos alguns comprimentos de onda.

Em 1974, quase 60 anos depois do sistema utilizado na Primeira Guerra Mundial, a primeira publicação com a aplicação direta de um arranjo de microfones em medições acústicas, utilizando explicitamente o processamento de Beamforming, foi apresentada por Soderman e Noble [24]. Os autores desenvolveram um arranjo em linha para investigar o ruído irradiado por uma turbina instalada dentro de um túnel de vento, utilizando um sistema digital para processar atrasos no tempo dos sinais de cada microfone.

Em 1987, Brooks [25] utilizou um arranjo bidimensional de microfones para medir o ruído do modelo em escala de um helicóptero submetido ao fluxo de ar dentro de uma câmara anecóica para medições aeroacústicas. Foi usado pela primeira vez um conjunto de sensores de pressão e a técnica Delay and Sum (DAS) no domínio da frequência.

No caso de fontes vibroacústicas pode-se conseguir uma melhor resolução da imagem colocando o arranjo de microfones perto da fonte de ruído, de modo a ser possível captar ondas evanescentes da fonte. Para tratamento dos sinais, é utilizada uma técnica denominada “Near-Field Holography” [26]. Esta técnica permite a visualização de características da fonte que são muito menores do que o comprimento de onda acústica, mas requer que o arranjo seja colocado perto da fonte.

Mosher [27] descreve o procedimento de calibração de um arranjo de microfones. O procedimento consiste na obtenção de ajustes de fase para cada microfone e a calibração da resposta de magnitude e direcionalidade do arranjo completo, utilizando a medição de uma fonte pontual de referência.

Ao longo dos anos, foram sendo melhorados os algoritmos para imageamento acústico. Em 1974, Högbom [28] desenvolveu o algoritmo “CLEAN” para a deconvolução de imagens para radioastronomia, posteriormente adaptado para o uso com imagens acústicas. Em 1987 Henry Cox desenvolveu o algoritmo “*Robust Adaptive Beamforming*” [29], que é robusto contra erros no posicionamento e na calibração dos sensores. Os algoritmos aplicados mais recentemente à técnica beamforming são “DAMAS” [25], “LORE” [30] e “CLEAN-SC” [31], os quais conseguem detectar e separar as fontes com melhor resolução do que o beamforming convencional. Suprimir os lóbulos laterais no mapa acústico do beamforming é uma das chaves no processo de imagens acústicas.

Técnicas para reduzir lóbulos laterais têm sido desenvolvidas, onde procura-se quebrar a regularidade do arranjo de microfones com distribuições não-uniformes para os microfones. Isto aumenta a confiabilidade do arranjo e permite suprimir fontes fantasmas [16]. Existem diferentes geometrias de arranjos, por exemplo, aleatórias, ou espirais logarítmicas concêntricas [18].

Resumindo, diversos métodos para aumentar a qualidade das imagens sem aumentar o número de sensores foram propostos, os quais aplicam técnicas de deconvolução para eliminar o efeito da função de espalhamento do arranjo na imagem estimada. Por outro lado, a técnica de *compressive sensing* proposta por Tarik Yardibi em [32], é uma regularização de norma ℓ_1 à imagem estimada de modo a dar mais peso a imagens esparsas (o que funcionará melhor quando o campo acústico for bem aproximado por um pequeno número de fontes compactas). Os métodos como [25, 31], apresentam resultados razoáveis, mas têm um custo computacional elevado. O alto custo computacional destes algoritmos deve-se ao grande número de multiplicações matriz-vetor com dimensões elevadas.

Na tese de doutorado em [18] foi desenvolvido um algoritmo para reduzir o número de cálculos necessários para a obtenção das estimativas dos diferentes métodos mencionados. Esse novo algoritmo, denominado *Kronecker Array Transform*, ou *KAT*, possibilita a obtenção de estimativas de imagens acústicas a um custo computacional reduzido, com a restrição de que o arranjo de microfones utilizado seja separável, isto é, os microfones devem estar distribuídos em um reticulado cartesiano, não necessariamente uniforme. Diversas aplicações da KAT foram propostas e simuladas em [22].

Neste trabalho são realizados diferentes testes com os algoritmos que foram desenvolvidos em [18], usando dados reais e assumindo a condição de campo distante, com a finalidade de comparar alguns algoritmos para estimação de imagens acústicas, quanto à qualidade do resultado e à complexidade computacional.

Para reduzir o impacto de falhas em microfones e manter a condição de que o arranjo seja separável, foi desenvolvida uma alternativa para utilizar os algoritmos rápidos, eliminando-se apenas os microfones com defeito, de maneira que os resultados finais são obtidos levando-se em conta todos os microfones do arranjo (não incluindo os defeituosos).

1.2 Objetivo do Estudo

O presente trabalho tem como objetivo principal realizar a validação experimental e continuação aos projetos de pesquisa relacionados a imagens acústicas iniciados com [18], utilizando arranjos de microfones com geometria separável.

1.3 Justificativa do Estudo

Com esta pesquisa, que tem como finalidade primária dar continuidade aos trabalhos feitos na tese de doutorado [18], pretende-se levar os conhecimentos obtidos ao longo do tempo de estudo para apoiar ou melhorar os projetos atuais que estão sendo desenvolvidos pelo SIMA - PERU, tais como: DELFIN-SONAR ².

O SIMA - PERU S.A. é uma empresa estatal de direito privado, cujo único proprietário é o Estado Peruano. Foi criada em 14 de fevereiro de 1950 com a finalidade de promover e desenvolver a indústria naval, indústrias complementares e associadas. O SIMA - PERU é uma empresa dedicada a direcionar e executar projetos de médio e grande porte; ao desenvolvimento de projetos navais; à fabricação de projetos metal mecânicos; e sistemas de armas e eletrônica, contribuindo desta maneira à produção e manutenção do poder naval e marítimo, tanto para a defesa como para o desenvolvimento nacional. A capacidade que mostra a empresa lhe permite abarcar todas as especialidades da engenharia naval, realizando com sucesso, durante os últimos sessenta anos, os mais variados trabalhos de manutenção e modernização das unidades navais da Marinha de Guerra do Peru e das armadas de diferentes países, assim como de navios particulares de diferentes tipos e nacionalidades.

²Projeto nacional para criar um Sonar pela Marinha do Peru.

1.4 Metodologia do Estudo

As seguintes atividades fazem parte do desenvolvimento deste projeto:

1. Estudo bibliográfico das principais ferramentas usadas para a implementação de algoritmos e aquisição dos dados com o arranjo de microfones;
2. Implementação dos algoritmos em scripts do Matlab, e criação de um ambiente para simulação;
3. Simulações e gravações em ambiente controlado. Estas gravações foram realizadas no estúdio do Laboratório de Processamento de Sinais da Escola Politécnica da USP.

1.5 Organização do Texto

O presente trabalho está composto de oito capítulos. O segundo capítulo aborda os fundamentos teóricos relacionados à acústica, conhecimentos de arranjos de microfones e considerações entre campos próximo e distante.

No terceiro capítulo é apresentado um modelo matemático para descrever os sinais recebidos em um arranjo de microfones.

No quarto capítulo, apresenta-se mais detalhes sobre imagens acústicas, forma de discretizar o espaço, e um método para estimar a matriz de autocorrelação $\mathbf{S}$.

O quinto capítulo é um resumo sobre os métodos básicos de beamforming.

O sexto capítulo resume o desenvolvimento da transformada rápida KAT, apresentada na tese de doutorado [18].

O sétimo capítulo apresenta a descrição do hardware utilizado, materiais e resultados finais das medições realizadas no laboratório de processamento de sinais. É apresentada uma variação aos métodos de reconstrução de imagens para eliminar só os microfones com defeito.

Por último, no oitavo capítulo, apresentam-se as conclusões da dissertação e próximos trabalhos a realizar.

No Apêndice A são mencionadas duas publicações resultantes deste trabalho. O primeiro artigo foi aceito para o *International Symposium on Solid Mechanics, MecSol, B. Horizonte*, a se realizar em 25-27 de Maio de 2015 e o segundo artigo está em avaliação para o *XXXIII Simpósio Brasileiro de Telecomunicações, SBrT, Juiz de Fora, Minas Gerais*, a se realizar em 1-4 de Setembro de 2015.

2 FUNDAMENTOS TEÓRICOS

2.1 Introdução

Este capítulo descreve uma série de formulações básicas em acústica, que são fundamentais para uma melhor compreensão dos capítulos seguintes. O estudo da física acústica é essencial para entender as leis que governam a propagação do som e modelar a transmissão de ondas acústicas em campo aberto, fechado ou em ambientes semifechados. Para entendê-la melhor precisamos relembrar alguns conceitos, descritos a seguir.

2.2 O som

O som é uma onda mecânica que se propaga longitudinalmente e possui a intensidade e frequência necessária para ser percebida pelo ser humano. Entendemos como onda mecânica uma onda que precisa de meio físico, como o ar ou o solo, para se propagar.

As frequências audíveis pelo ouvido humano ficam entre 16 e 20000 Hz (20 kHz). Dentro desta faixa encontram-se a voz humana, instrumentos musicais, alto-falantes. Abaixo de 16 Hz temos os infrassons, produzidos por vibrações da água em grandes reservatórios, batidas do coração, etc. Acima de 20 kHz estão os ultrassons emitidos por alguns animais e insetos, sonares, aparelhos médicos e industriais.

Os dispositivos que produzem ondas sonoras são chamados de fontes sonoras. Entre os que mais se destacam estão aqueles compostos por:

1. Cordas vibrantes como violão, o piano, as cordas vocais.
2. Tubos sonoros como órgão, flauta, clarineta.
3. Membranas e placas vibrantes tais como o tambor ou alto falantes.

2.3 Campo Acústico

Pode-se definir o campo acústico como a variação de pressão sonora em nosso meio. A menor variação de pressão ambiente detectável pelo sistema auditivo, da ordem de 2×10^{-5} Pa, é chamada de *limiar da audição* [33]. Diferentes valores de pressão são atribuídos ao limiar da dor (maior valor da intensidade sonora suportável pelo ouvido), desde 20 até 200 Pa.

2.4 Propagação de ondas sonoras

Quando uma onda sonora se propaga, há mudanças na pressão e densidade do meio. Após a onda passar, o meio retorna ao seu estado inicial. Um campo sonoro pode ser considerado como a sobreposição de muitas ondas sonoras. A seguir são explicados conceitos relacionados com a propagação das ondas sonoras.

2.4.1 Velocidade do som

Em geral, a velocidade de propagação c no ar, é função da densidade ρ e do coeficiente de elasticidade E , sendo expressa por

$$c = \sqrt{\frac{E}{\rho}} \left[\frac{\text{m}}{\text{s}} \right], \quad (2.1)$$

onde ρ é dada em kg/m^3 e E em N/m^2 . A velocidade de propagação do som no ar depende da temperatura e pode ser calculada a partir da seguinte expressão

$$c = 331,4 \sqrt{1 + \frac{\Theta}{273}} \left[\frac{\text{m}}{\text{s}} \right], \quad (2.2)$$

onde Θ é a temperatura em graus Celsius [1]. A uma temperatura de 20°C a velocidade de propagação do som no ar será aproximadamente de 343 m/s.

2.4.2 Equação da onda

Uma onda é uma perturbação periódica no tempo de alguma grandeza física no espaço. São impulsos energéticos que se propagam através de um meio, porém o meio não acompanha essa propagação [34]. As ondas caracterizam-se também quanto à direção de propagação como: longitudinal, transversal e mista. Quando as vibrações ocorrem na mesma direção de propagação têm-se ondas longitudinais. Já nas ondas transversais as vibrações ocorrem perpendicularmente à direção de propagação. Por último, as ondas mistas constituem-se de vibrações

simultaneamente transversais e longitudinais.

A velocidade de propagação das ondas sonoras é dependente do meio em que se propagam, sendo maior na água que na atmosfera. Na atmosfera, a velocidade de propagação é independente da pressão atmosférica, variando somente em função da temperatura (T). A velocidade do som corresponde a cerca de 343 m/s na atmosfera (20°C), 520 m/s no mar e 1558 m/s no corpo humano.

No corpo humano, o ouvido percebe ondas sonoras com frequências compreendidas aproximadamente entre os 16 Hz (graves) e os 20 kHz (agudos). A frequência de uma onda é a grandeza que determina a altura tonal com que o som é percebido [33].

Os campos acústicos de interesse são sobreposições de soluções da chamada equação de onda. A propagação acústica segue o modelo resumido a seguir [34]. Podemos considerar um exemplo da equação de onda em acústica no caso unidimensional com uma fonte pontual harmônica. A equação de onda para pressão sonora neste caso é:

$$\frac{\partial^2 p(x, t)}{\partial x^2} = \frac{\partial^2 p(x, t)}{c^2 \partial t^2}, \quad (2.3)$$

com c a velocidade do som no ar em m/s, p a pressão sonora em Pa, e t o tempo em segundos.

Esta equação parcial diferencial tem soluções na forma:

$$p(x, t) = Ae^{j(\omega t - kx)} + Be^{j(\omega t + kx)}, \quad (2.4)$$

com A e B escolhidos para satisfazer condições de contorno específicas, em que $\omega = 2\pi f$ é a frequência angular e $k = \frac{2\pi}{\lambda}$ o número de onda ¹. Há então duas parcelas na solução que indicam a presença de ondas propagando-se em direções opostas.

Considerando três dimensões, o resultado vai se generalizar utilizando as coordenadas cartesianas $(x, y, z) \in \mathbb{R}^3$. É conveniente definir o *vetor número de onda*:

$$\mathbf{k} = k_x \vec{x} + k_y \vec{y} + k_z \vec{z}, \quad (2.5)$$

onde $(\vec{x}, \vec{y}, \vec{z})$ são vetores, e, tal que

$$k_x^2 + k_y^2 + k_z^2 = \frac{\omega^2}{c^2} = \left(\frac{2\pi}{\lambda}\right)^2. \quad (2.6)$$

Para uma onda plana se propagando em uma única direção dada pelo vetor de onda $\mathbf{k} =$

¹(λ) é o comprimento de onda, a distância de um pico de onda para o outro.

(k_x, k_y, k_z) , a pressão sonora é:

$$p(x, y, z, t) = Ae^{j(\omega t - k_x x - k_y y - k_z z)}, \quad (2.7)$$

como notação vetorial definimos $\mathbf{r} = x\vec{x} + y\vec{y} + z\vec{z}$. A Equação (2.7) fica:

$$p(\mathbf{r}, t) = Ae^{j(\omega t - \mathbf{k} \cdot \mathbf{r})}. \quad (2.8)$$

A solução obtida é utilizada no estudo da propagação esférica de uma fonte pontual. A onda esférica proveniente de uma fonte sonora pontual sofre uma atenuação proporcional à distância da fonte até um ponto qualquer [34]. Definindo $A = \frac{p}{r}$, com p representando a pressão sonora na superfície da fonte pontual e r a distância da fonte até um ponto qualquer. Consequentemente, a pressão sonora fica sendo:

$$p(\mathbf{r}, t) = \frac{p}{r} e^{j(\omega t - \mathbf{k} \cdot \mathbf{r})}, \quad (2.9)$$

com a Equação (2.9), calcula-se a propagação da pressão sonora de uma fonte pontual até um ponto (x, y, z) qualquer do espaço.

2.4.3 Nível de pressão sonora

Normalmente a pressão sonora (*sound pressure level, SPL*) não é expressa em pascais ($1 \text{ Pa} = 1 \text{ N/m}^2$), mas sim em uma escala logarítmica onde a unidade de medida é o decibel (dB). O nível de pressão sonora ou nível sonoro é a medida logarítmica da pressão sonora efetiva em relação ao valor de referência. O *SPL* é definido pela seguinte expressão

$$L = 10 \log_{10} \frac{p_{rms}^2}{p_0^2} = 20 \log_{10} \frac{p_{rms}}{p_0} \text{ [dB]}, \quad (2.10)$$

onde p_{rms} representa a pressão sonora efetiva ou RMS (*root mean square*), definida como

$$p_{rms} = \sqrt{\frac{1}{T} \int_0^T p^2(t) dt}, \quad (2.11)$$

e a referência de pressão sonora efetiva no ar, p_0 , é de $20 \mu\text{Pa}$ ($20 \times 10^{-6} \text{ Pa}$), o limiar da audição humana.

2.4.4 Intensidade Sonora

Outra grandeza importante é a intensidade sonora. É uma medida da energia transportada por uma onda de som. É definida como a quantidade de energia por segundo que passa através de uma seção plana de 1 m^2 perpendicular à direção de propagação.

O vetor de intensidade sonora é expresso como

$$\mathbf{I} = p\mathbf{v}, \quad (2.12)$$

onde p é a pressão sonora e $\mathbf{v}$ é a velocidade da partícula.

A intensidade sonora pode ser expressa com um nível de intensidade,

$$L_I = 10 \log_{10} \frac{|\mathbf{I}|}{I_0}, \quad (2.13)$$

onde $I_0 = 10^{-12} \text{ W/m}^2$ [35].

2.4.5 Coordenadas esféricas

Será necessário definir a relação entre coordenadas esféricas e cartesianas porque nos capítulos seguintes será mais confortável no momento de realizar os respectivos cálculos matemáticos. Como podemos ver na Figura 2.1,

$$\begin{aligned} x &= r \sin\theta \cos\phi, \\ y &= r \sin\theta \sin\phi, \\ z &= r \cos\theta, \end{aligned} \quad (2.14)$$

onde θ e ϕ são os ângulos de elevação e azimute respectivamente.

Figura 2.1: Sistema de coordenadas esféricas.

2.5 Arranjos de Microfones

Arranjos ou *arrays* de sensores são definidos como conjuntos ordenados de transdutores eletroacústicos ou microfones distribuídos espacialmente em uma determinada geometria. Os sinais recebidos nos sensores podem ser combinados de maneira a “direcionar” o arranjo para uma determinada região do espaço, o que permite uma captura máxima do sinal em uma direção específica e atenuar sinais procedentes de outras direções.

O processamento de um arranjo pode ser caracterizado como um processo de filtragem, onde os sinais amostrados são processados simultaneamente nos domínios do tempo e do espaço. Algumas aplicações atuais são: aquisição de imagens médicas por ultrassom, controle de tráfego aéreo, medição de ondas sísmicas, detecção e rastreamento de objetos subaquáticos, medições de campo acústico e aquisição de sinais de voz.

De maneira análoga, o uso de arranjos permite dar forma (gerar uma imagem, um objeto bidimensional) a uma distribuição de diversos sinais unidimensionais (as ondas sonoras) que são capturados pelos microfones individualmente.

2.5.1 Arranjo Linear Uniforme (ULA)

Do inglês (*Uniform Linear Array, ULA*), é o arranjo mais simples onde se consideram N microfones espaçados entre si de uma distância d , em linha reta. Como exemplo, podemos ver na Figura 2.2, um arranjo com dois microfones (x_0, x_1) . Escolhe-se um microfone como referência para o qual se assume a posição $\mathbf{p}_0 = [0, 0, 0] \in \mathbb{R}^3$. A posição do outro microfone é denotada por $\mathbf{p}_1 \in \mathbb{R}^3$.

Definimos o *vetor de propagação* $\mathbf{a}$, o qual é relacionado à direção de propagação de uma determinada onda:

$$\mathbf{a} = \begin{bmatrix} -\sin \theta \cos \phi \\ -\sin \theta \sin \phi \\ -\cos \theta \end{bmatrix}, \quad (2.15)$$

em que θ e ϕ são ângulos usados em coordenadas esféricas.

Considere, por exemplo, um sinal $f_0(t)$, chegando ao arranjo de uma direção com ângulo θ_0 com relação à posição dos microfones (ver Figura 2.2). Com esse ângulo e se os microfones estão separados por uma distância d , pode-se calcular a diferença de distância percorrida pela onda até incidir no segundo microfone, que será (assumindo $\mathbf{p}_0 = 0$)

$$\Delta = -d \cos(\theta) = \mathbf{a}^T \mathbf{p}_1, \quad (2.16)$$

o Δ é negativo porque nós escolhemos o microfone x_0 como referência e o sinal chega adiantado no microfone x_1 .

Figura 2.2: Arranjo Linear Uniforme, ULA.

2.6 Consideração de Campo Distante e Campo Próximo

Uma fonte de pequenas dimensões emite ondas esféricas. Longe o suficiente da fonte, a frente de onda pode ser aproximada por uma frente plana. A qualidade da aproximação depende da distância da fonte, do comprimento de onda, e das dimensões do arranjo.

Na situação de campo distante, presume-se que a onda se propaga como uma onda plana e a atenuação do sinal é a mesma para todos os microfones. A fonte é considerada no campo distante se $r > \frac{2L^2}{\lambda}$, em que r é a distância a partir da fonte para o arranjo, L é a maior dimensão do arranjo, e λ é o comprimento de onda [1].

Um exemplo de campo distante é mostrado na Figura 2.3, onde se considera uma fonte pontual no espaço com r maior que a distância total do arranjo L . A onda incide sobre um arranjo linear de N elementos uniformemente espaçados sobre o eixo x , com uma distância entre cada elemento d e o vetor número de onda $\mathbf{k}$.

Devemos definir que neste trabalho e para exemplos de simulações práticas a partir de agora será assumido campo distante, onde cada uma das fontes formará uma onda plana diferente.

Figura 2.3: Fonte Pontual em Campo Distante.

3 MODELO DOS SINAIS

3.1 Introdução

Neste capítulo, realizaremos a modelagem matemática do sistema para estimação de imagens acústicas, com base na tese de doutorado de [18]. A modelagem matemática permitirá expressar a intensidade sonora de ondas planas, incidindo de várias direções em um arranjo de microfones.

3.2 Resposta tempo - frequência

Consideremos um arranjo genérico com N microfones, conforme ilustrado na Figura 3.1, que amostram o campo nas coordenadas $\mathbf{p}_n \in \mathbb{R}^3$, com $n = 0, 1, \dots, N - 1$. Suporemos que o campo incidente no arranjo seja causado por M fontes $f_m(t)$, com $m = 0, 1, \dots, M - 1$, procedentes de direções diferentes no espaço.

Supondo conhecidas as fontes (as posições e sinal emitido), podemos estimar o sinal acústico incidente em determinado ponto. Devemos considerar que na prática as fontes não são conhecidas: a hipótese de fontes conhecidas será usada apenas para mostrar como os sinais nos microfones dependem das posições e dos sinais emitidos por cada fonte.

Considerando que estamos assumindo campo distante, cada uma das fontes formará uma onda plana diferente. Por isso, a forma de onda não se altera entre os elementos do arranjo, existe apenas um atraso entre o sinal recebido por cada um dos microfones.

Para um sinal emitido na frequência angular ω da direção $\mathbf{a}$, o vetor número de onda $\mathbf{k}$ é dado por:

$$\mathbf{k} = \frac{\omega}{c} \mathbf{a}, \quad (3.1)$$

onde $\omega = 2\pi f$ é a frequência angular, $\mathbf{a}$ foi descrito na Equação (2.15) e c é a velocidade do som.

O sinal recebido por cada sensor, $x_n(t)$, é a superposição dos sinais de cada fonte. Digamos

Figura 3.1: Arranjo de microfones genérico.

que temos M fontes pontuais escolhidas num campo distante, o sinal $x_n(t)$ pode ser escrito como:

$$x_n(t) = f_0(t - \tau_{n,0}) + f_1(t - \tau_{n,1}) + \cdots + f_m(t - \tau_{n,m}) + \cdots + f_{M-1}(t - \tau_{n,M-1}), \quad (3.2)$$

tendo a posição dos microfones $\mathbf{p}_n$ e utilizando o vetor de propagação $\mathbf{a}$, podemos obter $\tau_{n,m}$, o qual é o atraso (relativo à origem do sistema de coordenadas) com que chega o sinal da m -ésima fonte no n -ésimo microfone, dado por:

$$\tau_{n,m} = -\frac{\mathbf{a}_m^T \mathbf{p}_n}{c}, \quad (3.3)$$

sendo $\mathbf{a}_m^T$ a transposta do vetor.

Particularizando para arranjos de microfones lineares como foi visto na Subseção 2.5.1, e escolhendo a posição do primeiro microfone como origem ($\mathbf{p}_0 = 0$), o atraso relativo $\tau_{n,m}$ considerando-se uma distância simétrica entre cada microfone d e θ_m o ângulo de incidência da m -ésima fonte, fica

$$\begin{aligned} \tau_{0,m} &= 0, \\ \tau_{1,m} &= d \cos(\theta_m)/c, \\ &\vdots \\ \tau_{(N-1),m} &= (N-1)d \cos(\theta_m)/c. \end{aligned} \quad (3.4)$$

Agrupando os termos $x_n(t)$ em um novo vetor $\mathbf{x}(t)$ que tem dimensões $N \times 1$, cujas com-

ponentes são os sinais que chegam nos N microfones pelas M fontes, temos então:

$$\mathbf{x}(t) = \begin{bmatrix} x_0(t) \\ x_1(t) \\ \vdots \\ x_{N-1}(t) \end{bmatrix}. \quad (3.5)$$

3.2.1 Obtenção do vetor $\hat{\mathbf{x}}(\omega)$

3.2.1.1 Com relação ao atraso $\tau_{n,m}$

Por simplicidade, vamos considerar nesta seção sinais determinísticos. Podemos definir um vetor $\hat{\mathbf{x}}(\omega)$ dado pela Transformada de Fourier de (3.2):

$$\hat{x}_n(\omega) = \int_{-\infty}^{\infty} \sum_{m=0}^{M-1} f_m(t - \tau_{n,m}) e^{-j\omega t} dt. \quad (3.6)$$

Definindo

$$\hat{f}_m(\omega) = \int_{-\infty}^{\infty} f_m(t) e^{-j\omega t} dt, \quad (3.7)$$

$\hat{x}_n(\omega)$ pode ser reescrito como

$$\hat{x}_n(\omega) = \sum_{m=0}^{M-1} e^{-j\omega\tau_{n,m}} \hat{f}_m(\omega). \quad (3.8)$$

Agrupando os termos, obtemos:

$$\hat{\mathbf{x}}(\omega) = \begin{bmatrix} \hat{x}_0(\omega) \\ \hat{x}_1(\omega) \\ \vdots \\ \hat{x}_{N-1}(\omega) \end{bmatrix}. \quad (3.9)$$

3.2.1.2 Com relação ao vetor diretor (*array manifold vector*)

Modelando o espaço sobre o arranjo como uma semiesfera, definimos um vetor de direção, para cada possível fonte. Em outras palavras, para cada direção é associada uma fonte de onda plana.

Com essa definição podemos obter o *vetor diretor* $\mathbf{v}(\mathbf{k}_m)$, também conhecido como *array manifold vector*, o qual descreve a relação de atrasos em cada microfone. Portanto, o vetor diretor define completamente a geometria do arranjo de microfones e como ele responde a

diferentes números de onda,

$$\mathbf{v}(\mathbf{k}_m) = \begin{bmatrix} e^{-j\omega\tau_{0,m}} \\ e^{-j\omega\tau_{1,m}} \\ \vdots \\ e^{-j\omega\tau_{(N-1),m}} \end{bmatrix}, \quad (3.10)$$

sendo $\mathbf{k}_m$ o vetor número de onda definido em 3.1.

Com o vetor diretor, podemos definir uma única matriz $\mathbf{V}(\omega)$, que contém informações sobre a geometria dos microfones para o conjunto de direções do espaço, considerado em uma determinada frequência (ω). Esta matriz tem dimensões $N \times M$, temos então:

$$\mathbf{V}(\omega) = [\mathbf{v}(\mathbf{k}_0) \ \mathbf{v}(\mathbf{k}_1) \ \dots \ \mathbf{v}(\mathbf{k}_{M-1})]. \quad (3.11)$$

Podemos definir o vetor $\hat{\mathbf{x}}^{(m)}(\omega)$ como a componente de $\hat{\mathbf{x}}(\omega)$ devida à fonte $\hat{\mathbf{f}}_m(\omega)$:

$$\hat{\mathbf{x}}^{(m)}(\omega) = \hat{f}_m(\omega) \mathbf{v}(\mathbf{k}_m). \quad (3.12)$$

A Equação (3.8) fica então

$$\hat{\mathbf{x}}(\omega) = \hat{\mathbf{x}}^{(0)}(\omega) + \hat{\mathbf{x}}^{(1)}(\omega) + \dots + \hat{\mathbf{x}}^{(n)}(\omega) + \dots + \hat{\mathbf{x}}^{(N-1)}(\omega), \quad (3.13)$$

e usando (3.12),

$$\hat{\mathbf{x}}(\omega) = [\hat{f}_0(\omega) \mathbf{v}(\mathbf{k}_0) + \hat{f}_1(\omega) \mathbf{v}(\mathbf{k}_1) + \dots + \hat{f}_{(M-1)}(\omega) \mathbf{v}(\mathbf{k}_{M-1})], \quad (3.14)$$

ainda, agrupando os termos, obtemos:

$$\hat{\mathbf{x}}(\omega) = [\mathbf{v}(\mathbf{k}_0) \ \dots \ \mathbf{v}(\mathbf{k}_{M-1})] \begin{bmatrix} \hat{f}_0(\omega) \\ \vdots \\ \hat{f}_{M-1}(\omega) \end{bmatrix}, \quad (3.15)$$

simplificando mais as expressões, juntamos em um vetor os sinais que chegam ao arranjo emitido pelas M fontes no domínio da frequência:

$$\hat{\mathbf{f}}(\omega) = [\hat{f}_0(\omega) \ \dots \ \hat{f}_{M-1}(\omega)]^T, \quad (3.16)$$

obtemos a relação para todas as M fontes e os N microfones, com relação à frequência ω :

$$\hat{\mathbf{x}}(\omega) = \mathbf{V}(\omega) \hat{\mathbf{f}}(\omega). \quad (3.17)$$

3.3 Reconstrução de Imagens

As expressões obtidas na seção anterior são válidas para sinais determinísticos. Para processos estacionários as expressões não valem, pois $\hat{\mathbf{x}}(\omega)$ não é bem definido, só a sua densidade espectral de potência. Para prosseguir sem recorrer a derivações matemáticas extensas [36], vamos prosseguir usando a Transformada Discreta de Fourier (do inglês, *Discrete Fourier Transform, DFT*)

$$\tilde{\mathbf{x}}(\omega_k) = \sum_{n=0}^{K-1} \mathbf{x}(nt) e^{-j \frac{2\pi}{K} nk}, \quad (3.18)$$

$$\omega_k = \frac{2\pi}{K} k f_s, \quad 0 \leq k \leq \frac{K}{2}, \quad (3.19)$$

em que f_s é a frequência de amostragem, e similarmente para $\tilde{\mathbf{f}}(\omega_k)$. Para $\tilde{\mathbf{x}}(\omega_k)$ vale

$$\tilde{\mathbf{x}}(\omega_k) \approx \mathbf{V}(\omega_k) \tilde{\mathbf{f}}(\omega_k). \quad (3.20)$$

A partir da autocorrelação do sinal $\tilde{\mathbf{x}}(\omega_k)$, obtemos uma matriz de autocorrelação dada por:

$$\mathbf{S}_{\mathbf{x}}(\omega_k) = E\{\tilde{\mathbf{x}}(\omega_k) \tilde{\mathbf{x}}^H(\omega_k)\}, \quad (3.21)$$

em que $\{\cdot^H\}$ é o Hermitiano ou transposta conjugada da matriz (ou vetor). Substituindo a expressão (3.20) nesta última equação, obtemos:

$$\mathbf{S}_{\mathbf{x}}(\omega_k) \cong E\{\mathbf{V}(\omega_k) \tilde{\mathbf{f}}(\omega_k) \tilde{\mathbf{f}}^H(\omega_k) \mathbf{V}^H(\omega_k)\}, \quad (3.22)$$

podendo simplificar mais a expressão

$$\mathbf{S}_{\mathbf{x}}(\omega_k) = \mathbf{V}(\omega_k) E\{\tilde{\mathbf{f}}(\omega_k) \tilde{\mathbf{f}}^H(\omega_k)\} \mathbf{V}^H(\omega_k). \quad (3.23)$$

Para poder simplificar a expressão, admitiremos que as fontes pontuais são descorrelacionadas, conseqüentemente todos os termos fora da diagonal de $E\{\tilde{\mathbf{f}}(\omega_k) \tilde{\mathbf{f}}^H(\omega_k)\}$ se tornam nulos. Chamando de Y_m a potência emitida por cada fonte para $0 \leq m < M - 1$ (que na prática é desconhecida), isto é,

$$Y_m = E\{|\tilde{f}_m(\omega_k)|^2\}, \quad (3.24)$$

podemos definir o vetor $\mathbf{y}$:

$$\mathbf{y} = \begin{bmatrix} Y_0 & \dots & Y_{M-1} \end{bmatrix}^T, \quad (3.25)$$

o qual é o vetor que contém a imagem digital limpa que desejamos estimar. Ou seja, o que plotamos na imagem é uma estimativa da potência de uma onda plana.

Simplificando a matriz $\mathbf{S}_{\mathbf{x}}(\omega_k)$, obtemos:

$$\mathbf{S}_{\mathbf{x}}(\omega_k) \cong \sum_{m=0}^{M-1} Y_m \mathbf{v}(\mathbf{k}_m) \mathbf{v}^H(\mathbf{k}_m). \quad (3.26)$$

A hipótese de fontes descorrelacionadas é uma aproximação. Um dos objetivos deste trabalho é verificar, experimentalmente e por meio de simulações, o efeito desta hipótese na qualidade das imagens estimadas.

Finalmente, podemos obter o vetor $\mathbf{s} = \text{vec}(\mathbf{S}_{\mathbf{x}}(\omega_k))$, o qual é um vetor em que as colunas de $\mathbf{S}_{\mathbf{x}}(\omega_k)$ são empilhadas uma em cima da outra.

Podemos escrever:

$$\mathbf{s} \cong \mathbf{A} \mathbf{y}, \quad (3.27)$$

onde pode-se mostrar que a matriz $\mathbf{A}$, de dimensões $N^2 \times M$, depende apenas dos vetores $\mathbf{v}(\mathbf{k}_m)$, e é definida por:

$$\mathbf{A} = [\mathbf{v}^*(\mathbf{k}_0) \otimes \mathbf{v}(\mathbf{k}_0) \quad \mathbf{v}^*(\mathbf{k}_1) \otimes \mathbf{v}(\mathbf{k}_1) \cdots \mathbf{v}^*(\mathbf{k}_m) \otimes \mathbf{v}(\mathbf{k}_m)], \quad (3.28)$$

onde $\otimes$ é o produto de Kronecker, e $\mathbf{v}^*(\mathbf{k}_m)$ é o conjugado de $\mathbf{v}(\mathbf{k}_m)$.

Para entender o produto de Kronecker apresentamos um exemplo, onde consideramos duas matrizes A e B :

$$A = \begin{bmatrix} 1 & -2 \\ -1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 4 & -3 \\ 2 & 3 \end{bmatrix}, \quad (3.29)$$

$$A \otimes B = \begin{bmatrix} 1 \cdot 4 & 1 \cdot -3 & -2 \cdot 4 & -2 \cdot -3 \\ 1 \cdot 2 & 1 \cdot 3 & -2 \cdot 2 & -2 \cdot 3 \\ -1 \cdot 4 & -1 \cdot -3 & 0 \cdot 4 & 0 \cdot -3 \\ -1 \cdot 2 & -1 \cdot 3 & 0 \cdot 2 & 0 \cdot 3 \end{bmatrix} = \begin{bmatrix} 4 & -3 & -8 & 6 \\ 2 & 3 & -4 & -6 \\ -4 & 3 & 0 & 0 \\ -2 & -3 & 0 & 0 \end{bmatrix}. \quad (3.30)$$

Diversos métodos para realizar a reconstrução de imagens acústicas usam (3.27) como ponto de partida [18,22]. O vetor $\mathbf{s}$ poderá ser estimado diretamente a partir dos sinais nos microfones, e a tarefa dos algoritmos de reconstrução é inverter a relação (3.27).

4 ESTIMAÇÃO DA MATRIZ DE AUTOCORRELAÇÃO

4.1 Introdução

Em ampliação aos conceitos já apresentados, devemos saber que quando trabalhamos em aplicações reais com imagens acústicas realizaremos uma parametrização bidimensional do espaço com o qual vamos trabalhar. Ou seja, vamos supor que o campo incidente no arranjo é causado por M fontes procedentes de direções diferentes no espaço, que são escolhidas por nós no processo de discretização do espaço.

Para entender um pouco melhor, pode-se visualizar na Figura 4.1 uma semiesfera; nela suporemos que somente é possível existir alguma fonte sonora nos locais onde as arestas se cruzam na superfície. Isto é uma aproximação da realidade, em que a distribuição real, contínua, de fontes no espaço é aproximada por uma distribuição discreta com um determinado padrão, que pode em princípio ser escolhido de diferentes maneiras.

Em primeira aproximação, quanto mais pontos forem escolhidos no espaço (mais fontes possíveis forem admitidas), melhor a aproximação. Veremos depois que a escolha de um grande número de possíveis fontes acarreta diversos problemas, de complexidade computacional e de estimação, dado o número limitado de medidas disponível em um arranjo de microfones.

4.2 Parametrização no espaço $\mathbf{U}$

Uma das formas de poder parametrizar o espaço em forma bidimensional pode ser da seguinte forma; com a finalidade de obter uma correspondência conjunta entre $\mathbf{a}$ (que é uma direção específica de propagação) e as coordenadas de um pixel da imagem acústica, definimos $\mathbf{u} = -\mathbf{a}$, como nossa direção de visada. Quer dizer, para cada direção é associado um versor $\mathbf{u}$, que corresponde à direção que o arranjo está observando. Parametrizada em coordenadas

Figura 4.1: Discretização do espaço U.

esféricas, obtemos de (2.15),

$$\mathbf{u} = \begin{bmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{bmatrix}. \quad (4.1)$$

Definindo:

$$\begin{aligned} u_x(\theta, \phi) &= \sin\theta \cos\phi, \\ u_y(\theta, \phi) &= \sin\theta \sin\phi, \end{aligned} \quad (4.2)$$

obtemos

$$\mathbf{u} = \begin{bmatrix} u_x \\ u_y \\ \sqrt{1 - u_x^2 - u_y^2} \end{bmatrix}, \quad (4.3)$$

para $u_x^2 + u_y^2 \leq 1$. Com esta parametrização podemos representar qualquer direção de chegada definindo as coordenadas $(u_x, u_y) \in [-1, 1]^2 \triangleq U$. Como descrito em [18], a discretização das direções em termos de u_x e u_y (em vez de diretamente sobre os ângulos θ e ϕ) permite o uso de algoritmos para reduzir o número de cálculos necessários para a estimação de campos acústicos.

Usando (3.3), e mudando o vetor $\mathbf{a}$ para o novo vetor $\mathbf{u}$, obtemos

$$\tau_{n,m} = -\frac{\mathbf{a}^T \mathbf{p}_n}{c} = \frac{\mathbf{u}^T \mathbf{p}_n}{c}, \quad (4.4)$$

substituindo (4.4) em (3.10), definimos novamente o vetor diretor que descreve a relação de

atrasos em cada microfone em função da frequência ω e do vetor $\mathbf{u}$,

$$\mathbf{v}(\omega : \mathbf{u}) = \begin{bmatrix} e^{-j\omega \mathbf{u}^T \mathbf{p}_0/c} \\ e^{-j\omega \mathbf{u}^T \mathbf{p}_1/c} \\ \vdots \\ e^{-j\omega \mathbf{u}^T \mathbf{p}_{(N-1)}/c} \end{bmatrix}, \quad (4.5)$$

com esta nomenclatura das direções do espaço e os conceitos apresentados nos capítulos posteriores, podemos descrever um método de estimar a matriz de autocorrelação $\mathbf{S}_{\mathbf{x}}(\omega_k)$ em função do conjunto de fontes discretas escolhido.

4.3 Possível método de Solução

Podemos dizer que a equação definida em (3.27) é a base para a estimação do vetor $\mathbf{y}$, pois o vetor $\mathbf{s}$ repetido aqui novamente como,

$$\mathbf{s} = \text{vec}(\mathbf{S}_{\mathbf{x}}(\omega_k)) = \text{vec}(E\{\tilde{\mathbf{x}}(\omega_k)\tilde{\mathbf{x}}^H(\omega_k)\}), \quad (4.6)$$

pode ser aproximado diretamente a partir dos sinais $x_n(t)$ dos microfones, e a função dos algoritmos de reconstrução é inverter a relação (3.27). Dividiremos o procedimento em duas partes para poder compreender melhor e obter uma aproximação do vetor $\mathbf{s}$:

4.3.1 Cálculo do vetor $\tilde{\mathbf{x}}(\omega_k)$

Primeiro vamos calcular o vetor $\tilde{\mathbf{x}}(\omega_k)$, definido em (3.20). Como exemplo, suponhamos que temos um campo de onda genérico, que será modelado pela superposição de M fontes de onda plana localizadas nas direções $\mathbf{u}_m$ do espaço total, seja $m = 0, 1, \dots, M-1$.

Os sinais recebidos em cada um dos N microfones assim como foi definido em (3.2), serão segmentados em janelas de K amostras cada uma. Podemos ver na Figura 4.2 (A), por exemplo, um sinal senoidal contínuo em função do tempo excluindo o ruído externo e a Figura 4.2 (B) apresenta a discretização do sinal que neste exemplo é dividida em seções de $K = 20$ amostras cada uma.

As mencionadas seções na Figura 4.2 (B) serão definidas agora como janelas, às quais aplicamos a Transformada Discreta de Fourier (TDF) a fim de obter aproximações de $\hat{\mathbf{x}}(\omega_k)$, em que a frequência ω_k , pertencente ao conjunto

$$\omega_k = \frac{k \times 2\pi}{K} \times f_s, \quad (4.7)$$

Figura 4.2: Exemplo de sinal original e sinal discretizado em janelas de K amostras.

$0 < k < \frac{K}{2}$, é a frequência particular que gostaríamos de observar (f_s é a frequência de amostragem).

Com esta frequência ω_k forma-se um vetor de N elementos, com um elemento por microfone, para cada uma das janelas. Cada vetor vai representar o componente da frequência escolhida anteriormente. Seja L a quantidade de janelas escolhidas no domínio da frequência. O novo vetor $\tilde{\mathbf{x}}_l(\omega_k)$, que tem dimensões $N \times 1$, sendo $l = 0, 1, \dots, L - 1$, é portanto dado por

$$\tilde{\mathbf{x}}_l(\omega_k) = \begin{bmatrix} \tilde{x}_{0,l}(\omega_k) \\ \tilde{x}_{1,l}(\omega_k) \\ \vdots \\ \tilde{x}_{(N-1),l}(\omega_k) \end{bmatrix}, \quad (4.8)$$

o qual sera uma aproximação de $\hat{\mathbf{x}}(\omega_k)$, definido na Subseção 3.2.1:

$$\tilde{\mathbf{x}}_l(\omega_k) \cong \hat{\mathbf{x}}(\omega_k). \quad (4.9)$$

4.3.2 Aproximação da matriz de autocorrelação $\mathbf{S}_{\mathbf{x}}(\omega_k)$

O segundo procedimento que temos que fazer é encontrar uma aproximação da matriz de autocorrelação $\mathbf{S}_{\mathbf{x}}(\omega_k)$ definida em (3.21). A partir das L janelas definidas no domínio da

frequência, podemos estimar esta autocorrelação como:

$$\hat{\mathbf{S}}(\omega_k) = \frac{1}{L} \sum_{l=0}^{L-1} \tilde{\mathbf{x}}_l(\omega_k) \tilde{\mathbf{x}}_l^H(\omega_k), \quad (4.10)$$

podendo dizer que

$$\hat{\mathbf{S}}(\omega_k) \cong \mathbf{S}_{\mathbf{x}}(\omega_k). \quad (4.11)$$

Para processos estacionários, como médias e correlações são constantes, quanto maior for o valor de L teremos uma melhor aproximação da matriz de autocorrelação $\mathbf{S}_{\mathbf{x}}(\omega_k)$. Em processos não estacionários, a escolha de um valor muito grande para L também resultaria em erro elevado. Nesse caso, é necessária a escolha de um valor adequado de L , nem muito grande, nem muito pequeno. Por isso, o procedimento dependerá dos valores escolhidos de K e L .

Em seguida, obtemos o vetor $\hat{\mathbf{s}} = \text{vec}(\hat{\mathbf{S}}(\omega_k))$. Supondo que $\hat{\mathbf{s}}$ seja uma aproximação razoável para $\mathbf{s}$, obtemos de (3.27)

$$\hat{\mathbf{s}} \cong \mathbf{A}\mathbf{y}, \quad (4.12)$$

que, dependendo das dimensões de $\mathbf{A}$, pode ser resolvido para calcular $\mathbf{y}$.

5 BEAMFORMERS

5.1 Introdução

Neste capítulo são resumidas algumas técnicas para conformação de feixe (beamforming)¹ [37]. Os métodos que são expostos a seguir supõem que os processos envolvidos sejam estacionários para o intervalo de medida. Na prática, poucos processos práticos podem ser considerados absolutamente estacionários. Serão apresentados fundamentos de beamforming, e o procedimento para obter uma imagem acústica.

5.2 Delay and Sum Beamforming

O Delay and Sum Beamformer² (DSB) é uma das técnicas mais simples de processamento de sinais com arranjos de microfones, mas pode-se dizer que tem bons resultados. A ideia principal do método é obter uma amostragem do som por meio de um arranjo de microfones, em seguida, os sinais serão atrasados para cada um dos microfones e ao final serão adicionados em conjunto [38].

Realizaremos uma explicação mais intuitiva para entendê-lo. A Figura 5.1 ilustra o caso de uma fonte sonora pontual qualquer f_m localizada em frente do arranjo de N microfones, que emite uma onda que se propaga com um vetor de onda $\mathbf{k}_m$.

O som da fonte viaja para cada microfone por caminhos diferentes (etapa 1). Os sinais captados pelos microfones serão semelhantes em forma de onda, mas apresentam diferentes atrasos proporcionais às distâncias percorridas (etapa 2). Os atrasos podem ser determinados a partir da velocidade do som e a distância entre os microfones com a fonte. Na Figura 5.1, Δ/c corresponde ao atraso $\tau_{n,m}$, como foi visto em (3.3).

¹O termo Beamforming pode ser traduzido como “formação de feixe”. Entretanto, devido à extensa utilização da palavra na língua inglesa, o nome original será mantido neste texto se referindo à técnica de processamento de sinais capaz de formar um feixe espacial seletivo em um conjunto de transdutores.

²O delay and sum beamformer é conhecido como o beamformer clássico (DAS) na literatura.

Figura 5.1: Delay and Sum Beamformer. Retirado de [39].

Definimos um vetor de pesos $\mathbf{w}$ (com relação ao vetor diretor que se deseja *olhar* $\mathbf{v}(\mathbf{k}_m)$, definido em (3.10)),

$$\mathbf{w} = \frac{\mathbf{v}(\mathbf{k}_m)^*}{N}. \quad (5.1)$$

O passo seguinte (3) está em mudar o sinal de cada microfone para obter uma forma ordenada dos picos do sinal. Como resultado, os componentes do sinal em todos os canais estarão em uma mesma fase.

Na etapa de processamento final (4) os sinais de todos os canais são somados e, finalmente, o sinal de soma é normalizado ao número de microfones. Obtendo a saída para este exemplo,

$$z(t) = \mathbf{w}^H \mathbf{x}(t), \quad (5.2)$$

sendo $\mathbf{x}(t)$ o vetor cujas componentes são os sinais que chegam no arranjo, definido em (3.5).

A estimativa da imagem na direção m será:

$$\begin{aligned} \check{Y}_m &= E\{|\mathbf{w}^H \mathbf{x}(t)|^2\}, \\ \check{Y}_m &= \mathbf{v}^H(\mathbf{k}_m) E\{\mathbf{x}(t) \mathbf{x}^H(t)\} \mathbf{v}(\mathbf{k}_m), \\ \check{Y}_m &= \mathbf{v}^H(\mathbf{k}_m) \mathbf{S}_x \mathbf{v}(\mathbf{k}_m). \end{aligned} \quad (5.3)$$

No DAS tradicional primeiro é calculado $z(t)$ e em seguida é calculada a potência média. Neste procedimento primeiro é calculada uma aproximação da matriz de autocorrelação $\mathbf{S}_x$ e depois é multiplicado pelo vetor diretor.

Tomando o vec de (5.3), obtemos:

$$\begin{aligned} vec(\check{Y}_m) &= vec(\mathbf{v}^H(\mathbf{k}_m)\mathbf{S}_x\mathbf{v}(\mathbf{k}_m)), \\ &= (\mathbf{v}^T(\mathbf{k}_m) \otimes \mathbf{v}^H(\mathbf{k}_m))vec(\mathbf{S}_x), \\ vec(\check{Y}_m) &= \mathbf{a}_m^H \mathbf{s}, \end{aligned} \quad (5.4)$$

em que $\mathbf{a}_m$ é a m -ésima coluna de $\mathbf{A}$.

Isto ocorre porque $vec(\check{Y}_m) = \check{Y}_m$ (pois $\check{Y}_m$ é um escalar), e $vec(\mathbf{XYZ}) = (\mathbf{Z}^T \otimes \mathbf{X})vec(\mathbf{Y})$. Colecionando os valores $\check{Y}_m$ em um vetor $\hat{\mathbf{y}}$, resulta

$$\hat{\mathbf{y}} = \mathbf{A}^H \mathbf{s}. \quad (5.5)$$

Comparando (5.5) com (3.27), vê-se que beamforming só resulta em uma estimativa perfeita da imagem se $\mathbf{A}$ for ortogonal, o que não ocorre na prática.

Para poder entender o procedimento de uma melhor forma, apresentaremos um exemplo.

Vamos considerar que temos dois alto-falantes, e discretizamos o espaço em M pontos, ver Figura 5.2. Na frente dos alto-falantes vamos ter um arranjo de N microfones, nas posições $\mathbf{p}_n \in \mathbb{R}^3$, sendo $n = 0, 1, \dots, N - 1$. O arranjo vai estimar ponto por ponto a potência de sinal vindo de cada uma das regiões definidas na Figura 5.2 obtendo uma escala de cores não uniforme.

Deseja-se estimar a intensidade sonora de cada um desses pontos no espaço. Primeiro escolhemos um tempo ou frequência específica, em seguida aplica-se o filtro espacial (atrasos e soma) e se plota o ponto do espaço estimado com uma escala de cores pré-definida, ver Figura 5.3. Para este exemplo, a maior intensidade de som estará vindo da membrana de cima, como podemos ver na Figura 5.4.

5.3 Point - Spread Function (PSF)

A *Point - Spread Function* (PSF) descreve a resposta de um sistema de imagem para uma fonte pontual. De uma forma geral é a resposta ao impulso que expressa a distribuição de intensidade da imagem estimada, supondo que uma única fonte pontual está presente. É obtida por

$$PSF(\mathbf{k}_m) = |\mathbf{w}^H \mathbf{v}(\mathbf{k}_m)|^2. \quad (5.6)$$

A Figura 5.5 ilustra o comportamento da PSF para diferentes direções de um arranjo uniforme linear com $N = 8$ microfones, o valor entre o espaçamento dos sensores e o comprimento

Figura 5.2: Discretização do espaço com $M = 10 \times 10$ pontos do alto-falante. Retirado de [39].

Figura 5.3: Relação arranjo de microfones com a discretização do espaço. Retirado de [39].

Figura 5.4: Imagem depois de fazer o beamforming. Retirado de [39].

de onda é de $\frac{d}{\lambda} = 0.5$ e o $\theta_{\mathbf{k}} = 0$ (fonte localizada a frente do arranjo no eixo y).

Figura 5.5: Point - Spread Function de um arranjo uniforme linear $N = 8$ microfones.

5.4 Beamformer MVDR

O beamformer *minimum variance distortionless response* (MVDR) foi proposto na literatura em [40] e também é conhecido como beamformer *CAPON*. A seguir, apresentaremos uma descrição do método, vamos supor que temos um sinal (de uma fonte) de banda estreita e analítico

$$f_m(t) = \alpha_m e^{j\omega t}, \quad (5.7)$$

onde α_m é a amplitude do sinal que incide no arranjo. O sinal no microfone n devido à fonte m é:

$$x_n(t) = \alpha_m e^{j\omega(t - \tau_{n,m})}, \quad (5.8)$$

ou também,

$$x_n(t) = \alpha_m e^{-j\mathbf{k}_m^T \mathbf{p}_n} e^{j\omega t}. \quad (5.9)$$

Arranjando de uma forma vetorial, obtemos:

$$\mathbf{x}(t) = \begin{bmatrix} x_0(t) \\ x_1(t) \\ \vdots \\ x_{N-1}(t) \end{bmatrix} = \begin{bmatrix} 1 \\ e^{-j\mathbf{k}_m^T \mathbf{p}_1} \\ \vdots \\ e^{-j\mathbf{k}_m^T \mathbf{p}_{N-1}} \end{bmatrix} \alpha_m e^{j\omega t}, \quad (5.10)$$

o vetor $\mathbf{x}(t)$ vai ser a multiplicação do vetor diretor $\mathbf{v}(\mathbf{k}_m)$ para uma determinada direção pelo sinal da fonte $\alpha_m e^{j\omega t}$.

Supondo agora que haja ruído, temos que

$$\mathbf{x}(t) = \mathbf{v}(\mathbf{k}_m)\alpha_m e^{j\omega t} + \mathbf{r}(t), \quad (5.11)$$

sendo $\mathbf{r}(t)$ o ruído. Define-se a matriz de autocorrelação para o ruído $\mathbf{S}_n = E\{\mathbf{r}(t)\mathbf{r}^H(t)\}$.

Será introduzido o conceito de *distortionless constraint*, o qual é utilizado para garantir uma resposta unitária na direção de interesse e assim poder obter o ganho do arranjo [41]. Definindo

$$\mathbf{w}^H \mathbf{v}(\mathbf{k}_m) = 1, \quad (5.12)$$

garante-se que qualquer sinal analítico na frequência ω que se propaga na direção de $\mathbf{k}_m$ vai passar através do filtro com ganho unitário.

Nesse caso, pode-se procurar estimar o sinal de saída de interesse $z(t) = \mathbf{w}^H \mathbf{x}(t)$, minimizando a variância da estimativa:

$$\min_{\mathbf{w}^H \mathbf{v}(\mathbf{k}_m)=1} E\{|z(t) - \alpha_m e^{j\omega t}|^2\}, \quad (5.13)$$

para este caso,

$$z(t) - \alpha_m e^{j\omega t} = \mathbf{w}^H \mathbf{r}(t), \quad (5.14)$$

portanto deve-se minimizar,

$$\min_{\mathbf{w}^H \mathbf{v}(\mathbf{k}_m)=1} E\{|\mathbf{w}^H \mathbf{r}(t)|^2\}, \quad (5.15)$$

ou ainda

$$\min_{\mathbf{w}^H \mathbf{v}(\mathbf{k}_m)=1} E\{\mathbf{w}^H \mathbf{r}(t)\mathbf{r}^H(t)\mathbf{w}\} = \min_{\mathbf{w}^H \mathbf{v}(\mathbf{k}_m)=1} \mathbf{w}^H \mathbf{S}_n \mathbf{w}. \quad (5.16)$$

Usando multiplicadores de Lagrange

$$\min_{\mathbf{w}, \lambda} \mathbf{w}^H \mathbf{S}_n \mathbf{w} + \lambda^* (\mathbf{w}^H \mathbf{v}(\mathbf{k}_m) - 1) + \lambda (\mathbf{v}^H(\mathbf{k}_m) \mathbf{w} - 1), \quad (5.17)$$

derivando

$$\frac{\partial}{\partial \mathbf{w}^H} \Rightarrow \mathbf{S}_n \mathbf{w} + \lambda^* \mathbf{v}(\mathbf{k}_m) = 0, \quad (5.18)$$

$$\frac{\partial}{\partial \lambda^*} \Rightarrow \mathbf{w}^H \mathbf{v}(\mathbf{k}_m) - 1 = 0, \quad (5.19)$$

concluimos de (5.18) e (5.19) que:

$$\frac{\partial}{\partial \mathbf{w}^H} \Rightarrow \mathbf{S}_n \mathbf{w} + \lambda^* \mathbf{v}(\mathbf{k}_m) \Rightarrow \mathbf{w} = -\lambda^* \mathbf{S}_n^{-1} \mathbf{v}(\mathbf{k}_m), \quad (5.20)$$

$$\frac{\partial}{\partial \lambda^*} \Rightarrow \mathbf{w}^H \mathbf{v}(\mathbf{k}_m) - 1 \Rightarrow -\lambda \mathbf{v}^H(\mathbf{k}_m) \mathbf{S}_n^{-1} \mathbf{v}(\mathbf{k}_m) = 1, \quad (5.21)$$

de (5.21), obtemos

$$\lambda = \frac{-1}{\mathbf{v}^H(\mathbf{k}_m)\mathbf{S}_n^{-1}\mathbf{v}(\mathbf{k}_m)}, \quad (5.22)$$

então, o $\mathbf{w}$ ótimo é dado por,

$$\mathbf{w}_{ot} = \frac{\mathbf{S}_n^{-1}\mathbf{v}(\mathbf{k}_m)}{\mathbf{v}^H(\mathbf{k}_m)\mathbf{S}_n^{-1}\mathbf{v}(\mathbf{k}_m)}. \quad (5.23)$$

Se o ruído for branco, o $\mathbf{w}_{ot}$ é o beamforming clássico,

$$\mathbf{S}_n = \sigma_n^2 \mathbf{I}, \quad (5.24)$$

então,

$$\mathbf{w}_{ot} = \frac{\sigma_n^{-2}}{\sigma_n^{-2}} \frac{\mathbf{v}(\mathbf{k}_m)}{\mathbf{v}^H(\mathbf{k}_m)\mathbf{v}(\mathbf{k}_m)}, \quad (5.25)$$

$$\mathbf{w}_{ot} = \frac{1}{N} \mathbf{v}(\mathbf{k}_m). \quad (5.26)$$

5.5 Beamformer MPDR

Do inglês, *minimum power distortionless response* ou MPDR. Podemos dizer que se não for possível estimar a matriz de autocorrelação $\mathbf{S}_n$, usa-se este método, em que é minimizada a potência do sinal recebido

$$\min_{\mathbf{w}^H \mathbf{v}(\mathbf{k}_m)=1} E|z(t)|^2, \quad (5.27)$$

e lembrando que $\mathbf{S}_x(\omega) = E\{\hat{\mathbf{x}}(\omega)\hat{\mathbf{x}}^H(\omega)\}$, a minimização fica:

$$\min_{\mathbf{w}^H \mathbf{v}(\mathbf{k}_m)=1} E\{\mathbf{w}^H \hat{\mathbf{x}}(\omega) \hat{\mathbf{x}}^H(\omega) \mathbf{w}\} = \min_{\mathbf{w}^H \mathbf{v}(\mathbf{k}_m)=1} \mathbf{w}^H \mathbf{S}_x \mathbf{w}, \quad (5.28)$$

e obtemos

$$\mathbf{w}_{MPDR} = \frac{\mathbf{S}_x^{-1}\mathbf{v}(\mathbf{k}_m)}{\mathbf{v}^H(\mathbf{k}_m)\mathbf{S}_x^{-1}\mathbf{v}(\mathbf{k}_m)}. \quad (5.29)$$

6 TRANSFORMADA RÁPIDA PARA IMAGENS ACÚSTICAS

6.1 Introdução

Na tese de doutorado de [18], foi desenvolvido um método para realizar a aceleração dos diversos métodos que realizam a desconvolução de imagens acústicas com a finalidade de reduzir os custos computacionais e tempos de processamento.

Podemos ver na Figura 6.1 um esquema de um arranjo separável. Quer dizer, as posições dos microfones estarão em um reticulado plano, não necessariamente uniforme. O nome “separável” é usado no estudo de transformadas de Fourier de duas dimensões. A transformada de duas dimensões $F(\omega_1, \omega_2)$ da função $f(x, y)$ é separável se puder ser escrita na forma $F(\omega_1, \omega_2) = F_x(\omega_1) \cdot F_y(\omega_2)$. Será visto na Seção (6.2) que a transformada de Kronecker para arranjos (KAT) é baseada em uma decomposição semelhante, do vetor diretor $\mathbf{v}(\mathbf{u})$ por um produto de Kronecker $\mathbf{v}(\mathbf{u}) = \mathbf{v}_x(\mathbf{u}_x) \otimes \mathbf{v}_y(\mathbf{u}_y)$. Isso pode ser feito somente se os microfones estiverem num “grid” retangular, o que denominamos “arranjo separável”.

Considerando campo distante, o vetor diretor para a direção visada $\mathbf{u}$ é dado por (4.5), repetido abaixo:

$$\mathbf{v}(\omega : \mathbf{u}) = \begin{bmatrix} e^{-j\omega \mathbf{u}^T \mathbf{p}_0/c} \\ e^{-j\omega \mathbf{u}^T \mathbf{p}_1/c} \\ \vdots \\ e^{-j\omega \mathbf{u}^T \mathbf{p}_{(N-1)}/c} \end{bmatrix}. \quad (6.1)$$

a forma do vetor diretor será usada na próxima seção para deduzir a KAT.

Figura 6.1: Arranjo separável.

6.2 Kronecker array transform (KAT)

Nesta seção, será descrito o procedimento da KAT desenvolvido na tese de doutorado de [18]. No caso de escolhermos uma direção $\mathbf{u}_m$ que temos interesse de estimar, conseguimos escrever:

$$\mathbf{v}(\mathbf{u}_m) = \mathbf{v}_x(\mathbf{u}_m) \otimes \mathbf{v}_y(\mathbf{u}_m), \quad (6.2)$$

onde $\otimes$ é o produto de Kronecker. Para verificar (6.2), suponha que o arranjo esteja no plano xy de modo que, a posição dos microfones seja $\mathbf{p}_n = \begin{bmatrix} x_n \\ y_n \\ 0 \end{bmatrix} \in \mathbb{R}^3$.

Na hora de multiplicar a posição dos microfones pelo vetor número de onda $\mathbf{k}$, obtemos

$$[k_x \vec{x} + k_y \vec{y} + k_z \vec{z}] \begin{bmatrix} x_n \\ y_n \\ 0 \end{bmatrix} = k_x x + k_y y, \quad (6.3)$$

e resulta:

$$e^{-j\mathbf{k}_i^T \mathbf{p}_n} = e^{-jk_x x_n} e^{-jk_y y_n}. \quad (6.4)$$

Podemos dizer que por causa da propriedade (6.4) e se o arranjo for separável, o vetor diretor total $\mathbf{v}(\mathbf{u}_m)$ é o produto de Kronecker do vetor diretor da direção x ($\mathbf{v}_x(\mathbf{u}_m)$) vezes o vetor diretor da direção y ($\mathbf{v}_y(\mathbf{u}_m)$). Por conseguinte, teremos um arranjo linear na direção de

x e um arranjo linear na direção y com N_x microfones na direção horizontal e N_y microfones na direção vertical, sendo o número total de microfones $N = N_x \times N_y$ para este caso.

Quando fazemos isso, cada coluna da matriz $\mathbf{A}$ da Equação (3.28) será:

$$\mathbf{v}^*(\mathbf{u}_m) \otimes \mathbf{v}(\mathbf{u}_m) = [\mathbf{v}_x^*(\mathbf{u}_m) \otimes \mathbf{v}_y^*(\mathbf{u}_m)] \otimes [\mathbf{v}_x(\mathbf{u}_m) \otimes \mathbf{v}_y(\mathbf{u}_m)], \quad (6.5)$$

ou seja,

$$\mathbf{A} = [(\mathbf{v}_x^*(\mathbf{u}_0) \otimes \mathbf{v}_y^*(\mathbf{u}_0)) \otimes (\mathbf{v}_x(\mathbf{u}_0) \otimes \mathbf{v}_y(\mathbf{u}_0)) \cdots (\mathbf{v}_x^*(\mathbf{u}_{M-1}) \otimes \mathbf{v}_y^*(\mathbf{u}_{M-1})) \otimes (\mathbf{v}_x(\mathbf{u}_{M-1}) \otimes \mathbf{v}_y(\mathbf{u}_{M-1}))]. \quad (6.6)$$

Pode-se definir uma matriz $\mathbf{V}_x$ de dimensões $N_x^2 \times M_x$ e uma matriz $\mathbf{V}_y$ de dimensões $N_y^2 \times M_y$, as quais vão depender de $\mathbf{v}_x(\mathbf{u}_m)$ e $\mathbf{v}_y(\mathbf{u}_m)$ respectivamente [34].

A partir dessas matrizes, considere o produto

$$\mathbf{Z} = \mathbf{V}_y \mathbf{Y} \mathbf{V}_x^T, \quad (6.7)$$

para uma matriz $\mathbf{Y}$ com dimensões $M_y \times M_x$, e seja o vetor $\mathbf{y} = \text{vec}(\mathbf{Y})$. Então, se $\mathbf{s} = \mathbf{A}\mathbf{y}$, e $\mathbf{Z} = \mathbf{V}_y \mathbf{Y} \mathbf{V}_x^T$, vale $\mathbf{s} = \mathbf{P} \text{vec}(\mathbf{Z})$, em que $\mathbf{P}$ é uma matriz de permutação.

A vantagem da KAT reside em que o produto direto de $\mathbf{A}\mathbf{y}$ requer aproximadamente $\frac{1}{2}M_x M_y N_x^2 N_y^2$ operações e o produto através de (6.7) da KAT requer $\frac{1}{2}N_y^2 M_x M_y + N_x^2 N_y^2 M_x$ operações. Se, por exemplo, temos um arranjo de microfones 8×8 e quisermos uma imagem 16×16 , nós ganharíamos usando a KAT um total de 7.1 vezes menos operações.

6.3 Deconvolução de imagens acústicas

As imagens acústicas produzidas por beamformers como o delay and sum apresentam muita distorção, pois a PSF cumpre uma função similar ao de filtros passa-baixas. Esta característica é indesejável na prática, pois reduz a resolução das imagens reconstruídas e impede a determinação de níveis de potência na presença de múltiplas fontes.

Para poder melhorar a qualidade de resolução das imagens, têm sido propostas diversas técnicas de desconvolução como CLEAN [28, 31], DAMAS [25, 32], DAMAS2 [34, 35] e LORE [30]. Elas utilizam como entrada a imagem produzida por beamforming (DSB) e a PSF do arranjo, retornando uma aproximação mais confiável da distribuição de fontes verdadeira.

No Capítulo 5, foi desenvolvido um método para obter uma aproximação da matriz de autocorrelação $\mathbf{S}_x(\omega_k)$. Nesta seção, para simplificar a notação, ocasionalmente ignoraremos

o argumento ω_k (sob a hipótese que o processamento é feito em banda estreita) e o subíndice $\{\mathbf{x}\}$. Assim, $\mathbf{S}_{\mathbf{x}}(\omega_k)$ será denotada por $\mathbf{S}$.

A seguir, apresentaremos a técnica de deconvolução para imagens acústicas DAMAS2.

6.3.1 DAMAS

Um dos métodos para efetuar desconvolução de imagens acústicas é conhecido como DAMAS (*deconvolution approach for the mapping of acoustic sources*) [25, 32]. É um algoritmo que recupera os níveis de fontes reais dos resultados do *delay and sum beamformer*, definindo um problema inverso; o qual poderá ser representado como um sistema de equações lineares. Posteriormente melhorado em [34, 35] e nomeado DAMAS2.

DAMAS2 é um método que apresenta melhoras significativas com relação ao custo computacional. Este método assume que a imagem obtida pelo delay and sum beamformer é igual à imagem limpa convoluída com a PSF do beamformer, considerando a aproximação de campo distante.

Sejam $\check{\mathbf{Y}}$ a imagem obtida por delay and sum beamformer, $\mathbf{P}_{\mathbf{SF}}$ a PSF, $\mathbf{Y}$ a imagem limpa e $\hat{\mathbf{Y}}^{(k)}$ a imagem reconstruída na iteração k . Definindo,

$$\check{\mathbf{Y}} = \mathbf{P}_{\mathbf{SF}} * \mathbf{Y}, \quad (6.8)$$

onde $(*)$ representa convolução 2D.

DAMAS2 obtém uma aproximação de $\mathbf{Y}$ iterando

$$\hat{\mathbf{Y}}^{(k+1)} = \max \left\{ \hat{\mathbf{Y}}^{(k)} + \frac{1}{a} \left[\check{\mathbf{Y}} - \left(\mathbf{P}_{\mathbf{SF}} * \hat{\mathbf{Y}}^{(k)} \right) \right], 0 \right\}, \quad (6.9)$$

onde $\max \{\cdot, \cdot\}$ retorna o máximo ponto a ponto, $a = \sum_{i,j} [\mathbf{P}]_{i,j}$, $\hat{\mathbf{Y}}^{(0)} = 0$ e a convolução é implementada usando uma FFT bidimensional e *zero-padding*¹.

DAMAS2 atualiza todas as variáveis simultaneamente, não depende da ordem dos pixels e tende a produzir resultados muito mais limpos.

¹Refere-se à adição de zeros a um sinal no domínio do tempo para aumentar seu comprimento.

6.3.2 Covariance fitting

Em [32], é proposta uma técnica de aproximação da matriz de covariância (*covariance matrix fitting*) partindo de (3.27), e aqui repetida como:

$$\mathbf{s} = \mathbf{A}\mathbf{y}. \quad (6.10)$$

A matriz $\mathbf{A}$ em geral não é quadrada com um número de incógnitas (dimensão de $\mathbf{y}$) maior que o número de equações (dimensão de $\mathbf{s}$). Para se obter uma solução que faça sentido em um problema desse tipo, é necessário introduzir algum tipo de regularização, que introduza alguma informação prévia sobre como se espera que a solução desejada $\mathbf{y}$ deva ser. Por exemplo, regularização ℓ_1 é adequada quando se sabe que $\mathbf{y}$ é esparso, e regularização TV (do inglês, *Total Variation*) é adequada quando a imagem $\mathbf{Y}$ é suave.

Os autores de [32] propõem resolver (6.10) por meio de uma minimização

$$\min_{\hat{\mathbf{Y}}, \sigma^2} \|\text{vec}\{\mathbf{S}\} - \mathbf{A}\text{vec}\{\hat{\mathbf{Y}}\} - \sigma^2 \text{vec}\{\mathbf{I}\}\|_2^2, \quad (6.11)$$

tal que $\hat{\mathbf{Y}} \geq 0$, $\sigma^2 \geq 0$, e $\|\text{vec}(\hat{\mathbf{Y}})\|_1 \leq 1$, onde σ^2 é a variância do ruído de medida em cada microfone. Repare que (6.11) é obtida supondo que o ruído (de medida) seja independente em cada microfone, mas sempre com a mesma variância. Neste caso, a matriz de autocorrelação do ruído será

$$\mathbf{S}_n = \sigma^2 \mathbf{I} = \begin{bmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \sigma^2 \end{bmatrix}. \quad (6.12)$$

Devemos ter em consideração três restrições para o problema: $\hat{\mathbf{Y}} \geq 0$, $\sigma^2 \geq 0$, e $\|\text{vec}(\hat{\mathbf{Y}})\|_1 \leq \lambda$. As duas primeiras restrições são devidas ao fato de que tanto a imagem quanto a variância devem ser positivas, e a última é a regularização para forçar a solução a ser esparsa.

6.3.3 Regularização ℓ_1

A Equação (6.11) pode ser resolvida através do método de regularização ℓ_1 como proposto em [32], o que permite dar uma solução estável para o sistema de equações lineares. Esta regularização somente é razoável se a distribuição de fontes for de fato esparsa, o que pode não ser o caso. Ainda o procedimento vai continuar tendo um custo computacional alto porque continuamos trabalhando com a matriz $\mathbf{A}$. Esta proposta supera o método DAMAS em relação

à precisão de reconstrução de imagens para o caso de imagens esparsas.

Com a finalidade de obter uma solução rápida, foi proposta em [18] a reformulação de (6.11) para tentar eliminar o ruído; a nova expressão é dada por

$$\min_{\hat{\mathbf{Y}}} \|\hat{\mathbf{Y}}\|_1 \quad \text{sujeito a} \quad \|\text{vec}\{\mathbf{S}\} - \mathbf{Avec}\{\hat{\mathbf{Y}}\}\|_2 \leq \sigma. \quad (6.13)$$

Para fornecer a solução de (6.13), o autor de [18] utilizou o algoritmo SPGL1 de [42] para solucionar o problema de otimização, acelerado pela KAT para o cálculo do produto $\mathbf{Avec}\{\hat{\mathbf{Y}}\}$.

6.3.4 Regularização TV

Esta formulação foi proposta pela primeira vez para deconvolução de imagens em [43]. Em situações quando as imagens acústicas não sejam esparsas, é proposta a reconstrução de imagens acústicas com a regularização TV.

Dado $\mathbf{Y} \in \mathbb{C}^{M_x \times M_y}$, define-se sua variação isotrópica total como

$$\|\mathbf{Y}\|_{\text{BV}} = \sum_{i,j} \sqrt{[\nabla_x \mathbf{Y}]_{i,j}^2 + [\nabla_y \mathbf{Y}]_{i,j}^2}, \quad (6.14)$$

onde ∇_x e ∇_y são operadores de primeira diferença ao longo das dimensões x e y com limites periódicos, para $0 \leq i < M_x$ e $0 \leq j < M_y$. $\|\cdot\|_{\text{BV}}$ é chamado de *bounded variation* (variação limitada).

O autor de [18] propôs realizar a solução de

$$\min_{\hat{\mathbf{Y}}} \|\hat{\mathbf{Y}}\|_{\text{BV}} + \mu \|\text{vec}\{\mathbf{S}\} - \mathbf{Avec}\{\hat{\mathbf{Y}}\}\|_2^2, \quad (6.15)$$

sujeito a $\hat{\mathbf{Y}}_{i,j} \geq 0$, utilizando o método TVAL3 de [44] para resolver o problema de minimização, também acelerado pela KAT. A primeira parte de (6.15) leva a solução a reduzir a quantidade de oscilações que apresenta a imagem, e a segunda parte garante um bom ajuste entre a imagem reconstruída e os dados medidos.

7 RESULTADOS FINAIS E EQUIPAMENTO UTILIZADO

7.1 Introdução

Neste trabalho, foi utilizado o arranjo de microfones montado em [18], com o qual conseguimos fazer diferentes medições no laboratório de acústica da Universidade de São Paulo. Foram feitas diversas correções e modificações no software do arranjo de microfones com ajuda de alunos de mestrado e pós-doutorando.

O arranjo utilizado está montado em uma placa de circuito impresso (ver Figura 7.1). É composto por 64 microfones MEMS (micro electrical-mechanical system), os quais possuem conversores A/D sigma-delta de 1 bit integrados às cápsulas. Cada microfone tem como saída apenas um sinal digital de 1 bit à taxa de 2.4 MHz.

Figura 7.1: Arranjo de microfones com 64 elementos e com geometria separável ($N_x = N_y = 8$). Em (a) observa-se a placa de circuito impresso original fabricada com os 64 microfones MEMS. Em (b) observa-se a geometria do arranjo de microfones original na qual os círculos representam as posições dos microfones. Em (c) observa-se a tabela com as coordenadas de cada microfone.

Cada um dos microfones MEMS é fabricado com processos de litografia, deposição e corrosão

semelhantes aos que são usados para produção de circuitos integrados. Os processos utilizados permitem a produção de microfones com tolerâncias melhores do que as presentes em cápsulas tradicionais de eletreto, o que em princípio deveria reduzir significativamente os problemas de casamento e calibração existentes em arranjos tradicionais. Infelizmente, devido ao processo de soldagem essa expectativa não se realizou. Cada microfone produz um único sinal digital, por isso um arranjo de N elementos produz N linhas digitais a 2.4 MHz.

Foi utilizada uma FPGA¹ Altera Stratix III para realizar a aquisição de dados (ver Figura 7.2). Esta controla um módulo de 1 GB de memória DDR3, que é usado como buffer para armazenar os sinais adquiridos pelo arranjo. A mesma FPGA também está conectada a uma interface USB 2.0, que é utilizada para transferir os dados da memória DDR3 para um computador, onde é executado o algoritmo de reconstrução de imagens.

Cada FPGA pode adquirir sinais de até 256 microfones. Arranjos maiores podem ser construídos utilizando mais FPGAs, bastando que todos os microfones estejam sincronizados por um mesmo CLOCK de 2.4 MHz. Como os conversores A/D já estão integrados aos microfones, o custo dos arranjos com estes sensores torna-se consideravelmente inferior ao dos arranjos tradicionais.

Figura 7.2: Placa de desenvolvimento Altera DE3.

Com a finalidade de gerar um sinal de excitação para realizar as medições no laboratório, foram utilizados dois alto-falantes miniaturizados modelo X-mini II adquiridos para realizar este projeto, ver Figura 7.3. Têm uma potência de 1.9 W e faixa de frequência de 100 Hz a 20 kHz, com bateria recarregável (12 horas de reprodução).

¹Do inglês, Field Programmable Gate Array.

Figura 7.3: Alto-falante X-mini II.

No 4º trimestre de 2014, foi efetuada a calibração do arranjo de microfones pelo Dr. Bruno Masiero. A calibração de um microfone consiste em determinar a sua resposta em frequência com relação a um microfone de referência. A calibração dos microfones foi feita de forma independente, colocando cada microfone em uma mesma posição relativa a uma fonte e a um microfone de referência. Este procedimento teve várias complicações porque os microfones estão soldados na placa. A calibração do arranjo de microfones permite obter uma melhor qualidade na reconstrução de imagens acústicas, como será apresentado nos resultados finais.

Para realizar o pós-processamento, foi utilizado o *ITA-Toolbox* [45] do programa MATLAB, o qual fornece facilidades de medição acústica, com diferentes aplicações disponíveis dentro do mesmo toolbox. Este toolbox apresenta uma plataforma onde os usuários têm acesso ao código fonte da maioria das funções que lhes permite seguir e compreender todas as partes do processamento de sinais. Com esta plataforma, as tarefas de processamento de sinais podem ser totalmente automatizadas, o que permite que os usuários realizem tarefas como, por exemplo, medir diretividade.

7.2 Instalações para medição

Foi requerida uma instalação física adequada onde as medições pudessem ser realizadas. O estúdio do laboratório de Processamento de Sinais da EPUSP foi utilizado para todas as medições apresentadas nesta dissertação; a Figura 7.4 apresenta as dimensões do laboratório. O estúdio é uma sala pré-fabricada propriamente revestida com um material de espuma com espessura de 5 cm. Pode-se dizer que a sala é uma câmara semi-anecóica, já em seu interior os ecos e ruídos do exterior são reduzidos, mas não eliminados por completo. Esta sala tem um chão sólido recoberto só com carpete da mesma cor que a espuma colocada nas paredes da sala, e uma das paredes tem uma janela de vidro de 1 m² a qual foi recoberta com o

mesmo material de espuma de maneira improvisada durante as medições, ver Figura 7.5 (a, b). Portanto, é importante observar que esta estrutura não permitirá um isolamento completo de ruídos externos nem eliminação de todas as reflexões da sala.

Os alto-falantes e o arranjo de microfones foram posicionados no interior do estúdio usando-se pedestais de microfones já disponíveis no estúdio, ver Figura 7.5 (c). Para realizar a fixação da placa foi utilizado um sistema de ventosa preso a um pedestal de microfone e foi utilizado um adaptador para poder segurar o alto-falante.

Figura 7.4: Estúdio do laboratório de Processamento de Sinais da Escola Politécnica da Universidade de São Paulo.

(a)

(b)

(c)

Figura 7.5: Instalação usada para as medições com o arranjo de microfones. Em (a) observa-se o estúdio do laboratório de Processamento de Sinais EPUSP. Em (b) observa-se a janela recoberta do estúdio. Em (c) observa-se a posição do arranjo de microfones e de um alto-falante, embaixo da placa encontra-se posicionada a FPGA.

Para a conexão com o computador foi necessário adquirir cabos de extensão para os alto-falantes e uma extensão com múltiplas tomadas USB para conectar a placa FPGA ao computador (necessita de dois cabos USB). Para o posicionamento correto dos alto-falantes e do arranjo de microfones foi adquirido também um laser STHT77340 CUBIX 12M da fabricante Stanley, ver Figura 7.6.

Figura 7.6: Laser STHT77340 CUBIX 12M da fabricante Stanley. Em (a) observa-se a imagem do laser. Em (b) observa-se a imagem de como foi utilizado o equipamento.

As medições foram controladas e pós-processadas por um PC com processador Intel Core i7-3770 rodando o programa **MATLAB**[®] que já estava disponível no laboratório. Foi utilizada uma placa de áudio Edirol Firewire, a qual foi utilizada para produzir os sinais de excitação enviados para os alto-falantes.

7.3 Modificações

Na etapa inicial deste projeto, foram feitos alguns testes de aquisição de dados com a placa eletrônica com tempos de 5 a 10 segundos, onde comprovamos no início que dois microfones (nº 6 e nº 16) só conseguiam pegar ruído branco, e possivelmente estavam danificados. Com estes testes pudemos comprovar que nosso arranjo de microfones com 64 ($N_x = N_y = 8$) elementos teria que mudar em princípio para 49 ($N_x = N_y = 7$) elementos (total 15 microfones eliminados), já que perderíamos uma linha e coluna completa para poder continuar trabalhando com uma geometria separável, como apresentado na Figura 7.7.

Após alguns meses, um novo microfone apresentou defeito (nº 52), e a geometria do arranjo teria de ser reduzida de 49 ($N_x = N_y = 7$) microfones para 42 ($N_x = 6$, $N_y = 7$). No total teríamos 3 elementos quebrados e para manter a condição de geometria separável teríamos que eliminar todas as linhas e colunas correspondentes para aqueles elementos com defeito (total 22 microfones eliminados). Esta nova geometria é apresentada na Figura 7.8.

Para reduzir o impacto de falhas em microfones, foi desenvolvida uma alternativa para utilizar os algoritmos rápidos, eliminando-se apenas os microfones com defeito, de maneira que

Figura 7.7: Geometria separável do arranjo de microfones $N_x = N_y = 7$. Em (a) observa-se a geometria do arranjo de microfones com dois elementos danificados (nº 6 e nº 16), são destacadas, assim como a linha e coluna eliminadas no pós-processamento. Em (b) observa-se a tabela com as coordenadas de cada microfone.

os últimos resultados foram obtidos levando-se em conta 61 microfones.

7.4 Resultados iniciais

A seguir, apresentamos simulações feitas no computador utilizando um algoritmo proposto nesta dissertação, e medições feitas utilizando o arranjo de microfones com e sem calibração, utilizando as instalações apresentadas na Seção 7.2. São comparados resultados obtidos usando Delay and Sum beamformer, DAMAS2, e covariance fitting com reconstrução regularizada ℓ_1 com o solver SPGL1 [42] e reconstrução regularizada TV com o solver TVAL3 [44]. Todos os métodos foram acelerados com a transformada rápida de Kronecker (KAT), como resumido no Capítulo 6.

O autor de [18] utilizou alguns parâmetros iniciais para desenvolver cada algoritmo, como por exemplo, para resolver a Equação (6.13) utilizou 200 iterações, com $\sigma = 0.01 \|S\|_F^2$. Para obter a solução da Equação (6.15) foram utilizadas 200 iterações com $\mu = 2^{10}$. DAMAS2 usou 1000 iterações. Todos estes parâmetros são modificados no longo da seção para poder realizar comparações de quais são os mais adequados para cada um dos algoritmos e poder

² $\|\cdot\|_F$ é a norma de Frobenius.

Figura 7.8: Geometria separável do arranjo de microfones $N_x = 6$, $N_y = 7$. Em (a) observa-se a geometria do arranjo de microfones com três elementos danificados (nº 6, nº 16 e nº 52), são destacadas, assim como a linha e coluna eliminadas no pós-processamento. Em (b) observa-se a tabela com as coordenadas de cada microfone.

obter uma reconstrução de imagem com boa resolução e melhor qualidade. A média do tempo de reconstrução por cada método está apresentada na Tabela 7.1. Podemos observar que o Delay and Sum beamformer é o método mais rápido comparado com os demais algoritmos.

Algoritmos	Tempo [s]
DAS	0.13
DAMAS2	1.58
Regularização ℓ_1	1.96
Regularização TV	2.46

Tabela 7.1: Tempo médio de cálculo para obtenção de imagens acústicas com diferentes algoritmos utilizando a KAT.

É utilizada uma discretização do espaço U com $M = M_x \times M_y = 256^2$ pontos (e correspondentemente pixels). Note-se que como a KAT requer uma distribuição separável de pontos, alguns pontos calculados estarão fora da área visível (isto é, para os pontos com $u_x^2 + u_y^2 > 1$). O ideal é que as imagens resultantes devem ter apenas valores pequenos (azuis) fora da região visível, isto constitui um teste preliminar da qualidade dos algoritmos de reconstrução. Nas próximas figuras o eixo vertical representa $u_x \in [-1, 1]$ e o eixo horizontal representa $u_y \in [-1, 1]$. O limite da região visível ($u_x^2 + u_y^2 = 1$) é delimitado com um círculo em linha preta.

Para poder realizar a estimação da matriz de autocorrelação $\mathbf{S}$, inicialmente foram escolhidas 65 janelas (L) e 2048 amostras (K), como foi descrito no Capítulo 4.

Para todas as configurações foi usado um sinal de excitação tipo *ruído branco gaussiano*, o qual foi gerado utilizando a função *randn* do Matlab, chamada através da função *ita-generate* do pacote ITA-Toolbox [45] do MATLAB com os seguintes parâmetros: um desvio padrão igual a 1, duração de 2^{17} amostras e com uma taxa de amostragem de 48 kHz.

A seguir são apresentados resultados para diferentes configurações de fontes sonoras. Também comparamos os resultados com e sem calibração para verificar a influência desta nos algoritmos de reconstrução. A disposição dos alto-falantes e do arranjo de microfones no estúdio está especificada na Figura 7.9.

Figura 7.9: Esboço do estúdio com as posições do arranjo de microfones (à direita) e algumas posições onde os alto-falantes foram posicionados (à esquerda).

Esta seção estará dividida em 3 subseções, na primeira subseção, será descrito o programa para efetuar a simulação das imagens acústicas proposto nesta dissertação. Na segunda subseção, são apresentados resultados utilizando a geometria de 49 e 42 microfones eliminando 15 e 22 elementos correspondentes para cada geometria. Na terceira subseção, são apresentados resultados modificando os parâmetros originais para comprovar quais valores vão dar os melhores resultados.

7.4.1 Programa de simulação

O programa de simulação desenvolvido nesta dissertação, utilizou um sinal tipo *ruído branco* com as mesmas condições do sinal utilizado para realizar as medições no estúdio (como foi descrito no início deste capítulo).

O programa permite escolher a geometria dos sensores utilizados, seja o caso para 61, 49 ou 42 microfones. As unidades estão em metros, sendo que os valores padrão servem para visualizar um arranjo de 35 cm por 35 cm como o projetado em [18]. É utilizada uma discretização do espaço U com $M = 256^2$, da mesma forma como foi descrito no início deste capítulo.

Logo, escolhemos uma frequência (ω_k) em particular que gostaríamos de observar (como foi descrito no Capítulo 4), e a partir dos sinais $\mathbf{f}(t)$ (vetor que representa o sinal vindo das M direções no domínio do tempo), estimamos o valor de $\hat{\mathbf{f}}(\omega_k)$ realizando a Transformada Discreta de Fourier e efetuamos o produto matricial $\mathbf{V}(\omega_k)\hat{\mathbf{f}}(\omega_k)$ para obter $\hat{\mathbf{x}}(\omega_k)$ como foi descrito no Capítulo 4. Desta forma, a relação entre os atrasos dos sinais que chegam nos microfones é levada em conta automaticamente pela matriz $\mathbf{V}(\omega_k)$. Obtemos a matriz de autocorrelação $\mathbf{S} = E\{\hat{\mathbf{x}}(\omega_k)\hat{\mathbf{x}}^H(\omega_k)\}$, como descrito na Equação (4.10).

Do mesmo modo, obtemos as duas matrizes $\mathbf{V}_x$ e $\mathbf{V}_y$ para poder acelerar os métodos de reconstrução de imagens com a KAT.

Com o programa de simulação não foi possível obter resultados com o método *covariance fitting* com reconstrução regularizada TV. Devido ao tempo de pesquisa, não foi possível consertar o algoritmo desenvolvido para realizar as simulações, é por isso que deixaremos esta correção para projetos que continuem com esta linha de investigação com imagens acústicas.

7.4.2 Resultados utilizando a geometria de 49 e 42 microfones

Posicionou-se um alto-falante na posição nº 1, a uma distância de 1,5 m do centro do arranjo de microfones com uma altura de 1,4 m. Os resultados da simulação e medição com o arranjo de microfones para uma geometria de 49 e 42 elementos com e sem calibração utilizando uma frequência da 1 kHz são apresentados nas Figuras 7.10 e 7.11.

Na primeira coluna apresentamos os resultados da simulação, que serviram como referência para observar a qualidade das imagens reconstruídas a partir das medições.

Os resultados mostram a importância da calibração, que torna a imagem calibrada mais similar à imagem simulada comparado com o resultado sem calibração. A reconstrução com regularização ℓ_1 contém muitos artefatos enquanto que a reconstrução usando DAMAS2 e

regularização TV apresentam alguns artefatos no limite da área visível.

Fica claro, comparando a Figura 7.10 com 7.11, que à medida que perdemos mais microfones a reconstrução de imagem vai ficando mais distorcida para o caso de DAS, mas para este exemplo, os métodos de DAMAS2, regularização ℓ_1 e TV obtêm um melhor resultado, sendo semelhantes aos resultados das simulações.

Figura 7.10: Resultados com uma geometria de 49 microfones eliminando 15 elementos para manter a condição de separabilidade por causa dos 2 microfones inicialmente com defeito. Coluna (a): Simulações. Coluna (b): Medições obtidas sem calibração. Coluna (c): Medições obtidas com calibração. 1ª linha: Delay and Sum beamformer. 2ª linha: DAMAS2. 3ª linha: regularização ℓ_1 . 4ª linha: regularização TV.

Figura 7.11: Resultados com uma geometria de 42 microfones eliminando 22 elementos para manter a condição de separabilidade por causa dos 3 microfones com defeito. Coluna (a): Simulações. Coluna (b): Medições obtidas sem calibração. Coluna (c): Medições obtidas com calibração. 1ª linha: Delay and Sum beamformer. 2ª linha: DAMAS2. 3ª linha: regularização ℓ_1 . 4ª linha: regularização TV.

7.4.3 Resultados modificando os parâmetros originais dos métodos

Posicionou-se um alto-falante na posição nº 1, à distância de 1,5 m do centro do arranjo de microfones com uma altura de 1,4 m. Foram alterados os diferentes parâmetros para cada um dos algoritmos, com a finalidade de identificar os mais convenientes para obter um melhor resultado. São apresentados resultados modificando o valor da discretização do espaço, a partir do valor nominal usado em [18], de $M = M_x \times M_y = 256^2$. Da mesma forma, o valor nominal para o parâmetro do método DAMAS2 são 1000 iterações. A regularização ℓ_1 trabalha com 200 iterações e $\sigma = 0.01 \|S\|_F$. A regularização TV trabalha com 200 iterações com $\mu = 10^3$.

Para todos os resultados seguintes obtidos com simulação e medição utilizando o arranjo de microfones, foi utilizada a geometria de 42 elementos com e sem calibração para uma frequência de 1 kHz e 2 kHz. Na primeira coluna, apresentamos os resultados da simulação, que serviram como referência para observar a qualidade das imagens reconstruídas a partir das medições. Para realizar a estimação da matriz de autocorrelação $\mathbf{S}$ foram escolhidas novamente 65 janelas (L) e 2048 amostras (K). Os métodos utilizados são Delay and Sum beamformer, DAMAS2, regularização ℓ_1 e TV, todos foram acelerados pela KAT.

7.4.3.1 Resolução (modificando o valor da discretização do espaço)

São apresentados resultados modificando os valores de discretização do espaço U para $M = M_x \times M_y = 576^2$, $M = M_x \times M_y = 256^2$, $M = M_x \times M_y = 64^2$ e $M = M_x \times M_y = 16^2$. Os resultados obtidos nas Figuras 7.12, 7.13, 7.14 e 7.15 são para Delay and Sum beamformer, DAMAS2, regularização ℓ_1 e TV, todos para uma frequência de 1 kHz.

A Figura 7.15 só apresenta resultados obtidos com medições no estúdio e com calibração, ainda não conhecemos o por que não foi possível obter algum resultado com o programa de simulação para o método com regularização TV.

Podemos concluir que o melhor resultado é dado com as discretizações utilizando $M = 576^2$ e $M = 256^2$, mas o tempo de processamento para o primeiro caso é maior comparado com o segundo caso (original), como podemos ver na Tabela 7.2. É por isso que utilizaremos a discretização original até o final da tese.

7.4.3.2 Modificação do número de iterações para o método DAMAS2

São apresentados resultados modificando o número de iterações para o método DAMAS2. O valor utilizado em [18] para o número de iterações é 1000. Para poder comparar os resultados obtidos foi utilizada uma frequência de 1 kHz e 100, 500 e 5000 iterações.

Figura 7.12: Comparação da resolução utilizando Delay and Sum beamformer. Coluna (a): Simulação. Coluna (b): Medição com 1 kHz, com calibração. 1ª linha: $M = 576 \times 576$. 2ª linha: $M = 256 \times 256$. 3ª linha: $M = 64 \times 64$. 4ª linha: $M = 16 \times 16$.

Figura 7.13: Comparação da resolução utilizando DAMAS2. Coluna (a): Simulação. Coluna (b): Medição com 1 kHz, com calibração. 1ª linha: $M = 576 \times 576$. 2ª linha: $M = 256 \times 256$. 3ª linha: $M = 64 \times 64$. 4ª linha: $M = 16 \times 16$.

Figura 7.14: Comparação da resolução utilizando o método de regularização ℓ_1 . Coluna (a): Simulação. Coluna (b): Medição com 1 kHz, com calibração. 1ª linha: $M = 576 \times 576$. 2ª linha: $M = 256 \times 256$. 3ª linha: $M = 64 \times 64$. 4ª linha: $M = 16 \times 16$.

Figura 7.15: Resultados utilizando o método de regularização TV. Coluna (a): Medição com 1 kHz, com calibração. 1ª linha: $M = 576 \times 576$. 2ª linha: $M = 256 \times 256$. 3ª linha: $M = 64 \times 64$. 4ª linha: $M = 16 \times 16$.

Resolução - Tempo do processamento [s]				
	M = 576 ²	M = 256 ²	M = 64 ²	M = 16 ²
Delay and Sum	0.1309	0,065	0.063	0.058
Regularização TV	18.593	3.421	0.202	0.112
Regularização ℓ_1	11.674	2.331	0.359	0.184
DAMAS2	5.442	1.235	0.156	0.069

Tabela 7.2: Tabela de tempos de reconstrução de imagens - Resolução.

Podemos observar na Figura 7.16, que com um número de 5000 iterações obtemos uma melhor qualidade que com 1000, só que o tempo de processamento é de aproximadamente 5 vezes maior. Observamos que para todos os casos os resultados com calibração apresentam uma considerável diferença nas imagens, e novamente aparece na parte inferior uma possível reflexão ou algum problema com artefatos próprios do algoritmo.

Para estudar a variância de imagens obtidas sob as mesmas condições, na Figura 7.17 são apresentadas nove medições feitas reiniciando todo o equipamento e software, com o alto falante e o arranjo de microfones na mesma posição e utilizando o mesmo sinal de excitação tipo *ruído branco* descrito no início deste capítulo. Foi utilizado o método de reconstrução DAMAS2 com uma frequência de 2 kHz, pois este método apresenta um melhor resultado de reconstrução, assim como o método utilizando regularização TV. Verifica-se que as imagens obtidas não são iguais, vão aparecendo alguns pequenos fantasmas de maior ou menor intensidade nas diferentes imagens.

Procurou-se quantificar a variabilidade entre as imagens calculando-se a diferença entre cada imagem e a obtida com simulação, como é apresentado na Tabela 7.3. Observa-se na Coluna 3, os resultados obtidos efetuando a diferença dos valores de cada imagem produto de cada medição com os valores pela simulação.

7.4.3.3 Modificação do número de iterações e do valor de sigma (σ) para a regularização ℓ_1

São apresentados a seguir resultados modificando o número de iterações e valores para sigma (σ) na regularização ℓ_1 . Este método tem como parâmetro original 200 iterações e $\sigma = 0.01$ em seu processamento e, para poder comparar os resultados obtidos, foi utilizada uma frequência de 2 kHz.

O parâmetro (σ) pode ser estimado, já que é relativo ao nível de ruído na sala. A Figura 7.18 mostra no lado esquerdo o sinal de excitação *ruído branco* que estamos utilizando para fazer todas as medições como foi descrito no início do capítulo, e no lado direito a medição que

Figura 7.16: Comparação dos resultados utilizando o método DAMAS2. Coluna (a): Simulação. Coluna (b): Medição sem calibração. Coluna (c): Medição com calibração. 1ª linha: 1000 iterações (original). 2ª linha: 100 iterações. 3ª linha: 500 iterações. 4ª linha: 5000 iterações.

Figura 7.17: Comparação dos resultados utilizando o método DAMAS2 por um período de nove dias, utilizando uma frequência de 2 kHz. Lado esquerdo: Figura (a): Simulação. Lado direito: Coluna (b, c, d): Medições com calibração.

chega em todos os microfones do arranjo. No lado superior direito, observa-se no fundo o sinal do microfone nº 52 que está com defeito representado com a cor vermelha. Observa-se que o sinal de excitação que estamos utilizando tem 2^{15} amostras iguais a zero no início e no final. Durante a medição observa-se que nesse mesmo intervalo de tempo existe um ruído dentro da sala, é este o nível de ruído que podemos estimar.

Primeiro, escolhemos um microfone como referência e logo é calculado o valor médio da sinal nesse intervalo de tempo (só ruído ambiente), que será denotado como $\mathbf{E}$ [46, 47], utilizando

$$\mathbf{E} = \frac{1}{n} \sum_{i=0}^{n-1} f_i, \quad (7.1)$$

onde f_i é o sinal no intervalo de tempo escolhido e n é o número total de amostras do sinal. Depois é calculado o desvio padrão (σ), o qual representa a média do desvio do sinal de acordo com a sua potência, utilizando

$$\sigma^2 = \frac{1}{n} \sum_{i=0}^{n-1} (f_i - \mathbf{E})^2, \quad (7.2)$$

para nosso exemplo, utilizando o sinal de excitação tipo *ruído branco* obtemos o valor do desvio padrão estimado no sinal medido de $\sigma = 0.009$. Os resultados utilizando este valor estimado são apresentados nas Figuras 7.19 e 7.20.

DAMAS2 - 2 kHz - 9 Testes		
Provas feitas	Tempo [s]	$\sum \mathbf{Y}_{\text{teste}} - \mathbf{Y}_{\text{simulação}} ^2$
Simulação	1.55	—
Teste 1	1.63	98.27
Teste 2	1.56	90.71
Teste 3	1.89	87.74
Teste 4	1.81	92.56
Teste 5	1.57	95.28
Teste 6	1.61	85.19
Teste 7	1.86	87.57
Teste 8	1.57	86.18
Teste 9	1.48	88.61

Tabela 7.3: Tabela com resultados dos valores pela simulação e dos 9 testes utilizando o método DAMAS2 com uma frequência de 2 kHz e com calibração. 1ª Coluna: Provas feitas. 2ª Coluna: Tempo de reconstrução dos resultados. 3ª Coluna: Variância dos resultados (diferença dos valores de cada imagem com o valor da imagem da simulação).

Na Figura 7.19 são apresentados resultados modificando o número de iterações por 50, 100, 500 e 1000, e os valores para sigma $\sigma = 0.01$ (original). Observa-se que com um número menor de iterações o algoritmo produz uma melhor resolução da imagem e o tempo do processamento é reduzido em quase um terço, por exemplo, para 200 iterações o tempo necessário é de 1,41 s e para 100 iterações, de 0,48 s.

Na Figura 7.20, observa-se que à medida que se aumenta a frequência, fica um pouco mais complicado poder identificar a fonte. Por exemplo, quando utilizamos uma frequência de 4 kHz com 500 e 1000 iterações já é difícil identificar a posição da fonte comparando com os resultados apresentados na Figura 7.19.

Observa-se do mesmo modo, que utilizando o valor de $\sigma = 0.1$ obtemos um melhor resultado. Portanto, utilizaremos para a regularização ℓ_1 os valores de 50 iterações e $\sigma = 0.1$ até o final da dissertação.

7.4.3.4 Modificação do número de iterações e o valor de (μ) para a regularização TV

São apresentados resultados modificando o número de iterações e valores para μ na regularização TV. Este método tem como parâmetro original 200 iterações e $\mu = 2^{10}$ em (6.15), para poder comparar os resultados obtidos foi utilizada uma frequência de 2 kHz. Podemos observar nas Figuras 7.21 e 7.22 os resultados obtidos modificando o número de iterações para 100, 500 e 1000, e os valores para $\mu = 2^{10}$ usado em [18].

Figura 7.18: Sinal de excitação *ruído branco*. Lado esquerdo: Sinal original. Lado superior direito: Medição nos 64 microfones. Lado inferior direito: Medição nos 64 microfones em escala de decibéis.

Observamos na Figura 7.21, que modificar o número de iterações produz algumas diferenças na resolução da imagem com relação aos pequenos fantasmas produto dos artefatos do algoritmo, e variação do tempo de processamento em aproximadamente 1 s.

Na Figura 7.22 observa-se que, trabalhando com uma frequência maior, os resultados terão um melhor resultado em comparação aos resultados mostrados na Figura 7.21.

No momento de modificar os valores para μ , o autor de [18] recomenda utilizar como valor mínimo $\mu = 2^3$ e como máximo $\mu = 2^{10}$. É por isso, que nas Figura 7.21 e 7.22 percebem-se os mesmos resultados para valores de μ acima de $\mu = 2^{10}$ e uma considerável diferença entre valores inferiores. Por conseguinte, continuaremos trabalhando com os parâmetros originais para este método.

7.4.3.5 Modificação do número de janelas (L) para realizar a estimação da matriz de autocorrelação S

Apresentamos resultados modificando o número janelas para estimar a matriz de autocorrelação **S**. Cada janela tem 2048 amostras (K). Na Figura 7.23, observamos resultados para 65, 32 e 1 janela (L), sendo o primeiro valor o máximo número de janelas que escolhemos para realizar todas as medições utilizando um sinal de excitação *ruído branco* como foi descrito no início do capítulo. Foi utilizado o método DAMAS2 com uma frequência de 2 kHz. Inicialmente

Figura 7.19: Comparação da resolução utilizando regularização ℓ_1 , com uma frequência de 2 kHz e com calibração. 1ª Coluna: sigma ($\sigma = 0.01$) (original). 2ª Coluna: sigma ($\sigma = 0.1$). 3ª Coluna: sigma ($\sigma = 0.001$). 4ª Coluna: sigma ($\sigma = 0.009$) valor estimado para nosso exemplo. 1ª linha: 50 iterações. 2ª linha: 100 iterações. 3ª linha: 200 iterações (original). 4ª linha: 500 iterações. 5ª linha: 1000 iterações.

Figura 7.20: Comparação da resolução utilizando regularização ℓ_1 , com uma frequência de 4 kHz e com calibração. 1ª Coluna: sigma ($\sigma = 0.01$) (original). 2ª Coluna: sigma ($\sigma = 0.1$). 3ª Coluna: sigma ($\sigma = 0.001$). 4ª Coluna: sigma ($\sigma = 0.009$) valor estimado para nosso exemplo. 1ª linha: 50 iterações. 2ª linha: 100 iterações. 3ª linha: 200 iterações (original). 4ª linha: 500 iterações. 5ª linha: 1000 iterações.

Figura 7.21: Comparação da resolução utilizando regularização TV, com uma frequência de 2 kHz e com calibração. 1ª Coluna: ($\mu = 2^{10}$) (original). 2ª Coluna: ($\mu = 2^3$). 3ª Coluna: ($\mu = 10^6$). 1ª linha: 100 iterações. 2ª linha: 200 iterações (original). 3ª linha: 500 iterações. 4ª linha: 1000 iterações.

Figura 7.22: Comparação da resolução utilizando regularização TV, com uma frequência de 4 kHz e com calibração. 1ª Coluna: ($\mu = 2^{10}$) (original). 2ª Coluna: ($\mu = 2^3$). 3ª Coluna: ($\mu = 10^6$). 1ª linha: 100 iterações. 2ª linha: 200 iterações (original). 3ª linha: 500 iterações. 4ª linha: 1000 iterações.

observamos que os resultados com simulação não apresentam diferença. Podemos observar uma pequena diferença entre os resultados obtidos com uma e 65 janelas escolhidas, como por exemplo, para o resultado obtido com 1 janela os pequenos fantasmas que vão aparecendo são um pouco maiores com relação aos obtidos com 65 janelas. A variação do tempo de processamento segundo os resultados obtidos é de aproximadamente 1.64 s para 1 só janela e 1.69 s para 65 janelas. Para dar continuidade aos resultados apresentados até o momento continuaremos trabalhando com o número de janelas ($L = 65$).

Figura 7.23: Comparação da resolução utilizando DAMAS2, com 2048 amostras por janela. Coluna (a): Simulação. Coluna (b): Medição com calibração. 1ª linha: número máximo de janelas ($L=65$) para este exemplo. 2ª linha: metade do número de janelas ($L=32$). 3ª linha: primeira janela ($L=1$).

7.4.3.6 Medição com material de espuma no solo do estúdio

Foram feitas duas medições nas posições nº 1 e nº 4 (ver Figura 7.9) com distâncias de 1.5 m e 2.2 m respectivamente com relação ao centro do arranjo de microfones e uma altura de 1.4 m. Foi utilizado o mesmo sinal de excitação *ruído branco* como foi descrito no início do capítulo. Foram utilizados todos os métodos de reconstrução de imagens com uma frequência de 2 kHz. O solo do estúdio foi recoberto parcialmente (distância entre o alto-falante e o arranjo de microfones) com material de espuma para poder descartar uma possível reflexão, a qual aparece claramente na parte inferior dos resultados utilizando os métodos de Delay and

Sum, DAMAS2, regularização ℓ_1 e TV. A descrição do estúdio foi feita na Seção 7.2.

Os resultados apresentados nas Figuras 7.24 e 7.25 mostram no lado esquerdo medições sem colocar espuma no solo, e no lado direito colocando material de espuma de maneira improvisada durante a medição.

Na Figura 7.24 observa-se uma melhora considerável com relação aos resultados obtidos sem colocar espuma no solo do estúdio. Novamente os métodos de DAMAS2 e regularização TV apresentam os melhores resultados e podemos observar que o reflexo na parte inferior desapareceu, mas ainda vão aparecendo alguns fantasmas nas imagens.

A Figura 7.25 mostra os resultados para o microfone na posição nº 4, observa-se que devido à distância e ao recobrimento parcial do estúdio continua aparecendo um pequeno reflexo na parte inferior. Podemos ver que os resultados utilizando espuma para os métodos DAMAS2 e regularização TV apresentam uma redução considerável na intensidade da reflexão que aparece no lado inferior da imagem. O melhor resultado nesta figura é utilizando o método com regularização ℓ_1 , o qual apresenta uma boa resolução na imagem.

Dos resultados obtidos nas Figuras 7.24 e 7.25, podemos dizer que se tivéssemos uma sala que cumpra todas as condições de uma câmara anecoica (evitar os ecos e ruídos internos dentro da sala) poderíamos obter uma melhor qualidade nos resultados. Deixaremos este ponto para possíveis trabalhos futuros.

7.4.3.7 Medição com material de espuma no solo do estúdio e sem recobrimento na janela

Foram feitas duas medições nas posições nº 1 e nº 3 (ver Figura 7.9) com distância de 1,5 m com relação ao centro do arranjo de microfones e uma altura de 1,4 m. Foi utilizado o mesmo sinal de excitação *ruído branco* como foi descrito no início do capítulo. Foram utilizados todos os métodos de reconstrução de imagens com uma frequência de 2 kHz. O solo do estúdio foi recoberto parcialmente (distância entre o alto-falante e o arranjo de microfones) com material de espuma para poder descartar uma possível reflexão e foi retirada a espuma colocada na janela do lado direito da sala (descrição do estúdio foi feita na seção 7.2).

Os resultados apresentados nas Figuras 7.26 e 7.27 mostram no lado esquerdo medições sem colocar espuma na janela do estúdio, e no lado direito resultados colocando material de espuma de maneira improvisada durante a medição.

Dos resultados apresentados nas Figuras 7.26 e 7.27 observa-se que, sem colocar espuma na janela do estúdio a imagem está com uma pequena inclinação para o lado esquerdo para

Figura 7.24: Medição com um alto-falante na posição nº 1, utilizando a calibração do arranjo com frequência de 2 kHz. Coluna (a): Medição sem colocar espuma. Coluna (b): Medição utilizando espuma. 1ª linha: Delay and Sum beamformer. 2ª linha: DAMAS2. 3ª linha: regularização ℓ_1 . 4ª linha: regularização TV.

Figura 7.25: Medição com um alto-falante na posição nº 4, utilizando a calibração do arranjo com frequência de 2 kHz. Coluna (a): Medição sem colocar espuma. Coluna (b): Medição utilizando espuma. 1ª linha: Delay and Sum beamformer. 2ª linha: DAMAS2. 3ª linha: regularização ℓ_1 . 4ª linha: regularização TV.

todos os métodos de reconstrução. Já quando é colocado o material de espuma os resultados apresentam uma melhor resolução.

Figura 7.26: Medição com um alto-falante na posição nº 1, utilizando a calibração do arranjo com frequência de 2 kHz e colocando espuma no solo. Coluna (a): Medição sem espuma na janela. Coluna (b): Medição com espuma na janela. 1ª linha: Delay and Sum beamformer. 2ª linha: DAMAS2. 3ª linha: regularização ℓ_1 . 4ª linha: regularização TV.

7.4.3.8 Comparação dos artefatos dos resultados obtidos

É feita uma comparação entre os resultados obtidos colocando um alto-falante na posição nº 1 (ver Figura 7.9), sem colocar espuma no solo do estúdio e com o arranjo calibrado. Foi utilizado o mesmo sinal de excitação *ruído branco* como foi descrito no início do capítulo. Foram utilizados todos os métodos de reconstrução de imagens com uma frequência de 2 kHz e 5 kHz.

Observa-se na Figura 7.28 resultados utilizando uma frequência de 2 kHz, onde os melhores resultados são apresentados utilizando os métodos DAMAS2, regularização TV. Neste exemplo, o método com regularização ℓ_1 , não apresenta muitos artefatos como os exemplos

Figura 7.27: Medição com um alto-falante na posição nº 3, utilizando a calibração do arranjo com frequência de 2 kHz e colocando espuma no solo. Coluna (a): Medição sem espuma na janela. Coluna (b): Medição com espuma na janela. 1ª linha: Delay and Sum beamformer. 2ª linha: DAMAS2. 3ª linha: regularização ℓ_1 . 4ª linha: regularização TV.

anteriormente apresentados, já que foi utilizado os parâmetros de $\sigma = 0.1$ e 50 iterações. Comparando os lóbulos principais, observa-se que o método DAS apresenta uma baixa qualidade de resolução, e percebe-se as diferenças entre os diferentes lóbulos laterais nos diferentes métodos de reconstrução.

Podemos observar na Figura 7.29 que trabalhando com uma frequência de 5 kHz, obtém-se uma melhor resolução quanto ao lóbulo principal, mas um incremento considerável quanto aos lóbulos laterais.

7.5 Modificação dos métodos de reconstrução de imagens para eliminar só os microfones com defeito

No Capítulo 6, foi apresentado o método KAT desenvolvido em [18], o qual permite acelerar diversos métodos que realizam a desconvolução de imagens acústicas com a finalidade de reduzir os custos computacionais e tempos de processamento sob a condição de que os elementos (microfones) estejam localizados sobre um reticulado plano para que o arranjo seja separável. Se um dos microfones do arranjo apresenta defeito o arranjo deixará de ter uma geometria separável, violando o principal requisito da KAT.

É proposta a seguir uma variação aos métodos de Delay and Sum beamformer, DAMAS2, regularização ℓ_1 e TV apresentados nos Capítulos 5 e 6. Para poder realizar esta variação será acrescentada uma matriz de pesos na formulação dos algoritmos de reconstrução de imagens para eliminar apenas os elementos (microfones) que estão com defeito no processamento, e não as linhas e colunas inteiras, como foi nos resultados apresentados anteriormente. Esta nova matriz de pesos será representada como $\mathbf{A}_{\text{masc}}$.

Todas as variações feitas são para modificar as contas do tipo $\mathbf{A}\hat{\mathbf{y}}$ e $\mathbf{A}^H\hat{\mathbf{s}}$ dentro dos algoritmos de reconstrução de imagens, sendo $\mathbf{A}$ uma matriz que depende dos vetores diretores como foi definido no Capítulo 3, $\hat{\mathbf{s}}$ é o vetor da matriz de autocorrelação estimada como foi visto no Capítulo 4 e $\hat{\mathbf{y}}$ é o vetor que contém os valores da imagem como foi definido no Capítulo 5.

Por exemplo, para um arranjo separável com cinco elementos temos que sua matriz de

Figura 7.28: Identificação lóbulos principais e laterais realizando uma medição com um alto-falante na posição n° 1, utilizando a calibração do arranjo com frequência de 2 kHz sem colocar espuma no solo. Coluna (a): Medição 2 dimensões com escala em decibéis. Coluna (b): Medição 3 dimensões com escala logarítmica. 1^a linha: Delay and Sum beamformer. 2^a linha: DAMAS2. 3^a linha: regularização ℓ_1 . 4^a linha: regularização TV.

Figura 7.29: Identificação lóbulos principais e laterais realizando uma medição com um alto-falante na posição n°1, utilizando a calibração do arranjo com frequência de 5 kHz sem colocar espuma no solo. Coluna (a): Medição 2 dimensões com escala em decibéis. Coluna (b): Medição 3 dimensões com escala logarítmica. 1^a linha: Delay and Sum beamformer. 2^a linha: DAMAS2. 3^a linha: regularização ℓ_1 . 4^a linha: regularização TV.

autocorrelação $\mathbf{S}^{(5)}$ é dada pelos seguintes valores

$$\mathbf{S}^{(5)} = \mathbf{E} \left\{ \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} \begin{bmatrix} a_1 & a_2 & a_3 & a_4 & a_5 \end{bmatrix}^* \right\} = \mathbf{E} \left\{ \begin{bmatrix} a_1 a_1^* & a_1 a_2^* & a_1 a_3^* & a_1 a_4^* & a_1 a_5^* \\ a_2 a_1^* & a_2 a_2^* & a_2 a_3^* & a_2 a_4^* & a_2 a_5^* \\ a_3 a_1^* & a_3 a_2^* & a_3 a_3^* & a_3 a_4^* & a_3 a_5^* \\ a_4 a_1^* & a_4 a_2^* & a_4 a_3^* & a_4 a_4^* & a_4 a_5^* \\ a_5 a_1^* & a_5 a_2^* & a_5 a_3^* & a_5 a_4^* & a_5 a_5^* \end{bmatrix} \right\}, \quad (7.3)$$

e se o elemento nº 2 estivesse com defeito, este último não deve ser utilizado nos cálculos. No entanto, se simplesmente eliminamos este microfone com defeito dos cálculos já não será possível acelerar os métodos de reconstrução utilizando a KAT. Veremos a seguir como eliminar o efeito dos microfones com defeito nos diversos algoritmos, de forma a ainda ser possível usar a KAT.

Analisando para o caso do Delay and Sum beamformer, da Equação (5.3) do Capítulo 5, lembramos que a estimativa da imagem é $\check{Y}_m = \mathbf{v}^H(\mathbf{u}_m) \mathbf{S} \mathbf{v}(\mathbf{u}_m)$, onde $\mathbf{v}(\mathbf{u}_m)$ é o vetor diretor para a direção de visada $\mathbf{u}_m$ que temos interesse de estimar.

Tomando o $vec(\check{Y}_m)$, verifica-se que

$$vec \left(\mathbf{v}^H(\mathbf{u}_m) \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{S}^{(5)} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{v}(\mathbf{u}_m) \right), \quad (7.4)$$

aplicando o produto de Kronecker, tal que $vec(\mathbf{XYZ}) = (\mathbf{Z}^T \otimes \mathbf{X})vec(\mathbf{Y})$, reorganizamos

$$\left\{ \left(\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{v}(\mathbf{u}_m) \right)^T \otimes \left(\mathbf{v}^H(\mathbf{u}_m) \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \right) \right\} vec(\mathbf{S}^{(5)}), \quad (7.5)$$

e utilizando a propriedade do produto de Kronecker $(\mathbf{AC}) \otimes (\mathbf{BD}) = (\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D})$, rees-

crevemos a Equação (7.5) como

$$\left(\mathbf{v}^T(\mathbf{u}_m) \otimes \mathbf{v}^H(\mathbf{u}_m) \right) \left(\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \right) \text{vec}(\mathbf{S}^{(5)}). \quad (7.6)$$

Comparando o resultado obtido na última equação com a Equação (5.4) do Capítulo 5, verifica-se que o primeiro parênteses é uma das linhas da matriz $\mathbf{A}_{(5)}^H$, e a multiplicação $\mathbf{A}_{(5)}^H \left[(\mathbf{a}_{\text{masc}}^{(5)} \otimes \mathbf{a}_{\text{masc}}^{(5)}) \hat{\mathbf{s}}^{(5)} \right]$ pode ser calculada de maneira eficiente com a KAT.

Concluindo, para o caso de Delay and Sum beamformer, efetuamos

$$\check{\mathbf{y}} = \mathbf{A}^H \mathbf{A}_{\text{masc}} \hat{\mathbf{s}}, \quad (7.7)$$

onde $\mathbf{A}_{\text{masc}} = (\mathbf{a}_{\text{masc}} \otimes \mathbf{a}_{\text{masc}})$ e $\mathbf{a}_{\text{masc}}$ é uma matriz diagonal com $\mathbf{a}_{\text{masc}_{ii}} = 1$ se o microfone i estiver funcionando e $\mathbf{a}_{\text{masc}_{ii}} = 0$ se o microfone i estiver com defeito.

Para o caso de DAMAS2, a modificação será em duas posições, a primeira será igual ao caso do Delay and Sum beamformer para obter a imagem $\check{\mathbf{Y}}$ e a segunda variação é na Equação (6.9) para obter a convolução de $\mathbf{P}_{\text{SF}} * \hat{\mathbf{Y}}^{(k)}$, mostrada novamente

$$\hat{\mathbf{Y}}^{(k+1)} = \max \left\{ \hat{\mathbf{Y}}^{(k)} + \frac{1}{a} \left[\check{\mathbf{Y}} - (\mathbf{P}_{\text{SF}} * \hat{\mathbf{Y}}^{(k)}) \right], 0 \right\}, \quad (7.8)$$

que o autor de [18] resolve utilizando,

$$\hat{\mathbf{y}}^{(k+1)} = \max \left\{ \hat{\mathbf{y}}^{(k)} + \frac{1}{a} [\check{\mathbf{y}} - (\mathbf{A}^H \mathbf{A} \hat{\mathbf{y}}^{(k)})], 0 \right\}, \quad (7.9)$$

para dar solução ao produto de $\mathbf{A}^H \mathbf{A} \hat{\mathbf{y}}^{(k)}$ o autor utilizou as matrizes $\mathbf{V}_x$ e $\mathbf{V}_y$ (com dimensões $N \times M$) definidas no Capítulo 6, e é aqui onde serão efetuadas as modificações já que estas contém as informações sobre a geometria dos microfones para o conjunto de direções do espaço. São zerados todos os valores de cada linha dos microfones com defeito, utilizando o comando do programa Matlab $\mathbf{V}_x(n^\circ, :) = 0$ e $\mathbf{V}_y(n^\circ, :) = 0$, onde (n°) corresponde ao número de cada linha do microfone com defeito que para nosso caso são $(n^\circ 6, n^\circ 16 \text{ e } n^\circ 52)$, e $(:)$ é utilizado para selecionar todos os componentes dessa mesma linha selecionada.

No caso de regularização ℓ_1 e TV, a modificação será efetuada multiplicando o $\text{vec}(\mathbf{a}_{\text{masc}})$ nas Equações (6.13) e (6.15), alterando a restrição $\left\| \text{vec}\{\mathbf{S}\} - \mathbf{A} \text{vec}\{\hat{\mathbf{Y}}\} \right\|_2$ para

$$\left\| \text{vec}(\mathbf{a}_{\text{masc}}) \odot (\text{vec}\{\mathbf{S}\} - \mathbf{A} \text{vec}\{\hat{\mathbf{Y}}\}) \right\|_2, \quad (7.10)$$

onde $\odot$ será utilizado para denotar o produto ponto-a-ponto, do mesmo modo, também podem ser acelerados utilizando a KAT.

Lembrando que $\mathbf{s} = \mathbf{A}\mathbf{y} = \mathbf{P}vec(\mathbf{Z})$, onde $\mathbf{P}$ é uma matriz de permutação e $\mathbf{Z} = \mathbf{V}_x \mathbf{Y} \mathbf{V}_y^T$ é a KAT, será a seguir determinado qual é o máximo número de microfones com defeito para o qual ainda é útil usar a KAT em comparação com o cálculo da imagem acústica diretamente pela multiplicação matriz-vetor. Podemos observar na Tabela 7.4, que trabalhando com a KAT vamos reduzir o número de contas consideravelmente que utilizando $\mathbf{A}\mathbf{y}$ e assim poder obter um menor custo computacional. Por exemplo, trabalhando com um arranjo de $N = 64$ e $M = 256^2$ ganhamos um total de 42,66 vezes menos operações utilizando a KAT. Por outro lado se trabalhamos com dez microfones ganhamos 1,03 vezes menos operações.

Na tabela, o número de operações mostrado para o cálculo usando produto matriz-vetor assume que só fazemos os cálculos da imagem que aparece dentro da circunferência delimitada por $u_x^2 + u_y^2 = 1$ e descartamos os valores fora desta região. Isto resulta em um ganho de 1,27 vezes menos operações que calculando os resultado para todas as $M = 256^2$ direções com o produto matriz-vetor $\mathbf{A}\mathbf{y}$. Então se a razão mostrada na quarta coluna for menor que um, podemos dizer que não será útil utilizar a KAT, como será o caso para valores de $N \leq 9$.

	Método	Matriz - Vetor	KAT (64 microfones)	Comparação
Nº	Operações	$\mathbf{A}\mathbf{y}(u_x^2 + u_y^2 \leq 1)$	$\frac{1}{2}N_y^2 M_x M_y + N_x^2 N_y^2 M_x$	$\frac{\mathbf{A}\mathbf{y}}{\text{KAT}}$
1	$N = 64, M = 256^2$	134203664	3145728	42,66
2	$N = 49, M = 256^2$	78661904	3145728	25,01
3	$N = 36, M = 256^2$	42453264	3145728	13,49
4	$N = 10, M = 256^2$	3262736	3145728	1,03
5	$N = 9, M = 256^2$	2640144	3145728	0,84
6	$N = 4, M = 256^2$	510224	3145728	0,16

Tabela 7.4: Tabela com resultados do custo computacional entre $\mathbf{A}\mathbf{y}$ e a KAT. 1ª Coluna: Número total de microfones (N) e espaço discretizado (M). 2ª Coluna: Número de operações para $\mathbf{A}\mathbf{y}$ (cálculos da imagem que aparece dentro da circunferência delimitada). 3ª Coluna: Número de operações para a KAT com 64 microfones. 4ª Coluna: Comparação.

Para poder comprovar o funcionamento desta modificação dentro dos algoritmos de reconstrução de imagens, são comparados resultados obtidos com uma medição utilizando a geometria com 42 microfones (22 elementos eliminados) e 61 microfones (3 elementos eliminados) do arranjo. O espaço U é discretizado em $M = 256 \times 256$ direções. Foi utilizado um sinal de excitação tipo *ruído branco* (descrito no início do Capítulo) posicionado a uma distância de 1,5 m com relação ao centro do arranjo de microfones e uma altura de 1,4 m. O solo do estúdio foi recoberto parcialmente com espuma para realizar esta medição como foi descrito na subsubseção (7.4.3.6). Para realizar a estimação da matriz de autocorrelação $\mathbf{S}$ foram escolhidas 65 janelas (L) e 2048 amostras (K), com uma frequência de 2 kHz. Foram utilizados todos os métodos de

reconstrução de imagens, com a calibração do arranjo de microfones e acelerado utilizando a KAT.

Na Figura 7.30, observa-se dos resultados que utilizando 42 microfones com 22 elementos eliminados (coluna do lado esquerdo), a imagem vai se tornando um pouco turva. A coluna do lado direito utiliza os 61 microfones, só eliminando 3 elementos com defeito (caso real que acontece com nosso arranjo de microfones), e observa-se nos resultados uma melhor qualidade de imagem. Da mesma forma, para todos os casos continuam aparecendo alguns pequenos fantasmas possivelmente próprios dos artefatos dos algoritmos. A reflexão do solo foi reduzida, mas não eliminada por completo já que o recobrimento da sala não é na totalidade.

Figura 7.30: Resultados utilizando todos os métodos de reconstrução de imagens, com frequência de 2 kHz e com calibração do arranjo. Coluna (a): Medições com 42 microfones (22 elementos eliminados). Coluna (b): Medições com 61 microfones (3 elementos eliminados). 1ª linha: Delay and Sum beamformer. 2ª linha: DAMAS2. 3ª linha: regularização ℓ_1 . 4ª linha: regularização TV.

7.6 Resultados finais utilizando os melhores parâmetros e a nova modificação nos método de reconstrução de imagens

Realizaremos uma comparação das imagens obtidas pelas medições utilizando o arranjo com 42 microfones (22 elementos eliminados) e com 61 microfones (3 elementos eliminados) devidamente calibrado, com uma discretização do espaço U ($M = 256^2$). Para realizar a estimação da matriz de autocorrelação $\mathbf{S}$ foram escolhidas 65 janelas (L) e 2048 amostras (K), trabalhando com frequências de 1 kHz, 2 kHz e 5 kHz nos diferentes resultados. Os métodos a comparar são Delay and Sum beamformer, DAMAS2, regularização ℓ_1 e TV, todos foram acelerados pela KAT. Nos resultados que são apresentados nesta subseção não foi colocada espuma no solo do estúdio onde foram feitas as medições, pois não contamos com material suficiente para poder colocar em toda a sala.

7.6.1 Padrão: Um alto-falante no centro com relação ao arranjo de microfones

Posicionou-se um alto-falante na posição nº 1, à distância de 1,5 m com relação ao centro do arranjo de microfones e uma altura de 1,4 m. Para realizar a comparação das Figuras 7.31, 7.32, 7.33 e 7.34, observa-se nas primeiras linhas os resultados obtidos utilizando o arranjo com 42 microfones (22 elementos eliminados) e nas segundas linhas resultados obtidos com 61 microfones (3 elementos eliminados) utilizando a variação proposta em cada método de reconstrução.

Observa-se para todos os casos que utilizando o arranjo de microfones completo eliminando só 3 elementos obtém-se uma melhor imagem com relação aos resultados obtidos utilizando só 42 microfones do arranjo. É possível observar que o método DAS não apresenta uma boa qualidade de imagem. Podemos comprovar que os métodos de reconstrução DAMAS2 e regularização TV apresentam um melhor resultado para poder reconhecer a localização da fonte principal utilizando uma frequência de 2 kHz e 5 kHz. Com a regularização ℓ_1 podemos comprovar que vão aparecendo alguns fantasmas fora do centro da imagem, possivelmente sendo produto dos artefatos próprios dos algoritmos.

Figura 7.31: Comparação de imagens acústicas utilizando 1 kHz, 2 kHz e 5 kHz (cada coluna) com o método de reconstrução Delay and Sum beamformer e com calibração do arranjo. 1ª linha: Medição utilizando 42 microfones (22 elementos eliminados). 2ª linha: Medição utilizando 61 microfones (3 elementos eliminados).

Figura 7.32: Comparação de imagens acústicas utilizando 1 kHz, 2 kHz e 5 kHz (cada coluna) com o método de reconstrução DAMAS2 e com calibração do arranjo. 1ª linha: Medição utilizando 42 microfones (22 elementos eliminados). 2ª linha: Medição utilizando 61 microfones (3 elementos eliminados).

Figura 7.33: Comparação de imagens acústicas utilizando 1 kHz, 2 kHz e 5 kHz (cada coluna) com o método de reconstrução covariance fitting utilizando regularização ℓ_1 , com nova variação dentro do algoritmo (SPGL1), 100 iterações e $\sigma = 0.1$ e com calibração do arranjo. 1ª linha: Medição utilizando 42 microfones (22 elementos eliminados). 2ª linha: Medição utilizando 61 microfones (3 elementos eliminados).

Figura 7.34: Comparação de imagens acústicas utilizando 1 kHz, 2 kHz e 5 kHz (cada coluna) com o método de reconstrução covariance fitting utilizando regularização TV, 200 iterações e $\mu = 2^{10}$ e com calibração do arranjo. 1ª linha: Medição utilizando 42 microfones (22 elementos eliminados). 2ª linha: Medição utilizando 61 microfones (3 elementos eliminados).

7.6.2 Padrão: Um alto-falante lado esquerdo com relação ao arranjo de microfones

Posicionou-se um alto-falante na posição nº 2, ou seja, 1,5 m na frente e 0.64 m para a esquerda e com uma altura de 1,4 m. Para realizar a comparação das Figuras 7.35, 7.36, observa-se nas primeiras colunas os resultados obtidos utilizando o arranjo com 42 microfones (22 elementos eliminados) e nas segundas colunas resultados obtidos com 61 microfones (3 elementos eliminados) utilizando a variação proposta em cada método de reconstrução.

Observa-se novamente que para todos os casos o melhor resultado é utilizando a variação proposta eliminando só 3 elementos. Nas Figuras 7.35 e 7.36, trabalhamos com 2 kHz e 5 kHz, observamos que os resultados significativamente melhores são apresentados utilizando os métodos de DAMAS2 e regularização TV em comparação com o DAS.

Figura 7.35: Um alto-falante no lado esquerdo com relação ao centro do arranjo de microfones (posição nº 2), com uma frequência de 2 kHz e com calibração do arranjo. Coluna (a): Medição com 42 microfones (22 elementos eliminados). Coluna (b): Medição com 61 microfones (3 elementos eliminados), utilizando a nova variação nos métodos. 1ª linha: Delay and Sum beamformer. 2ª linha: DAMAS2. 3ª linha: Regularização ℓ_1 . 4ª linha: Regularização TV.

Figura 7.36: Um alto-falante no lado esquerdo com relação ao centro do arranjo de microfones (posição nº 2), com uma frequência de 5 kHz e com calibração do arranjo. Coluna (a): Medição com 42 microfones (22 elementos eliminados). Coluna (b): Medição com 61 microfones (3 elementos eliminados), utilizando a nova variação nos métodos. 1ª linha: Delay and Sum beamformer. 2ª linha: DAMAS2. 3ª linha: Regularização ℓ_1 . 4ª linha: Regularização TV.

7.6.3 Padrão: Um alto-falante lado direito com relação ao arranjo de microfones

Posicionou-se um alto-falante na posição nº 3, ou seja, 1,5 m na frente e 0.64 m para a direita e com uma altura de 1,4 m. Para realizar a comparação das Figuras 7.37, 7.38, observa-se nas primeiras colunas os resultados obtidos utilizando o arranjo com 42 microfones (22 elementos eliminados) e nas segundas colunas resultados obtidos com 61 microfones (3 elementos eliminados) utilizando a variação proposta em cada método de reconstrução.

Observa-se novamente que para todos os casos o melhor resultado é utilizando a variação proposta eliminando só 3 elementos. DAMAS2 e regularização TV apresentam os melhores resultados para poder identificar a posição da fonte com o sinal de excitação.

Figura 7.37: Um alto-falante no lado direito com relação ao centro do arranjo de microfones (posição nº 3), com uma frequência de 2 kHz e com calibração do arranjo. Coluna (a): Medição com 42 microfones (22 elementos eliminados). Coluna (b): Medição com 61 microfones (3 elementos eliminados), utilizando a nova variação no método. 1ª linha: Delay and Sum beamformer. 2ª linha: DAMAS2. 3ª linha: Regularização ℓ_1 . 4ª linha: Regularização TV.

Figura 7.38: Um alto-falante no lado direito com relação ao centro do arranjo de microfones (posição nº 3), com uma frequência de 5 kHz e com calibração do arranjo. Coluna (a): Medição com 42 microfones (22 elementos eliminados). Coluna (b): Medição com 61 microfones (3 elementos eliminados), utilizando a nova variação no método. 1ª linha: Delay and Sum beamformer. 2ª linha: DAMAS2. 3ª linha: Regularização ℓ_1 . 4ª linha: Regularização TV.

7.6.4 Padrão: Um alto-falante na frente do arranjo com distância de 2,2 m

Posicionou-se um alto-falante na posição nº 4, ou seja, com 2,2 m na frente e uma altura de 1,4 m. Para realizar a comparação das Figuras 7.39, 7.40, observa-se nas primeiras colunas os resultados obtidos utilizando o arranjo com 42 microfones (22 elementos eliminados) e nas segundas colunas resultados obtidos com 61 microfones (3 elementos eliminados) utilizando a variação proposta em cada método de reconstrução.

Os melhores resultados são apresentados utilizando a variação proposta nos algoritmos. DAMAS2 e regularização TV mostram os melhores resultados. Observa-se que quanto maior for a distância de separação do arranjo com o alto-falante, a reflexão que aparece no lado inferior dos resultados vai ser maior. Do mesmo modo, os resultados obtidos com o método DAS e DAMAS2 apresentam outra possível reflexão no lado direito central na imagem, possivelmente por causa da janela que não está completamente recoberta com espuma. O problema das reflexões poderia-se resolver com um estúdio que cumpra as condições de uma câmara completamente anecoica, como foi comprovado nos resultados apresentados nas Subsubseções (7.4.3.6, 7.4.3.7).

7.6.5 Padrão: Dois alto-falantes na frente do arranjo de microfones com sinais correlacionados

Posicionaram-se dois alto-falantes na posição nº 2 e nº 3, ou seja, 1,5 m na frente, 0,64 m para a esquerda e 0,64 m para a direita do centro do arranjo de microfones e com uma altura de 1,4 m. Para realizar a medição com sinais correlacionados, foi utilizado o sinal de excitação *ruído branco* como foi descrito no início do capítulo, este sinal vai ser o mesmo que vai sair pelos dois alto-falantes. Os resultados apresentados são para as geometrias com 42 microfones (22 elementos eliminados) e com 61 microfones (3 elementos eliminados), esta última utilizando a variação proposta em cada método de reconstrução. Para realizar uma melhor comparação dos resultados utilizamos frequências de 500 Hz, 1 kHz, 2 kHz e 5 kHz.

Na Figura 7.41, são apresentados resultados utilizando uma geometria com 42 microfones. Observa-se que utilizando frequências baixas não é possível identificar a presença de duas fontes, mas utilizando frequências a partir de 2 kHz já é possível identificar as duas fontes. O método DAS com a usual baixa qualidade de reconstrução.

Na Figura 7.42 são apresentados resultados obtidos com 61 microfones. Observa-se a mesma situação que para os resultados apresentados na Figura 7.41. Utilizando a variação proposta nos algoritmos de reconstrução para frequências a partir de 2 kHz, observa-se uma melhora

Figura 7.39: Um alto-falante no centro com relação ao arranjo de microfones com distância de 2.2 m (posição nº 4), com uma frequência de 2 kHz e com calibração do arranjo. Coluna (a): Medição com 42 microfones (22 elementos eliminados). Coluna (b): Medição com 61 microfones (3 elementos eliminados), utilizando a nova variação no método. 1ª linha: Delay and Sum beamformer. 2ª linha: DAMAS2. 3ª linha: Regularização ℓ_1 . 4ª linha: Regularização TV.

Figura 7.40: Um alto-falante no centro com relação ao arranjo de microfones com distância de 2.2 m (posição nº 4), com uma frequência de 5 kHz e com calibração do arranjo. Coluna (a): Medição com 42 microfones (22 elementos eliminados). Coluna (b): Medição com 61 microfones (3 elementos eliminados), utilizando a nova variação no método. 1^a linha: Delay and Sum beamformer. 2^a linha: DAMAS2. 3^a linha: Regularização ℓ_1 . 4^a linha: Regularização TV.

considerável nos resultados DAMAS2 e com regularização TV comparados com os obtidos na Figura 7.41.

Figura 7.41: Dois alto-falantes no centro com relação ao arranjo de microfones com distância de 1.5 m e uma separação de 1.3 m (posição nº 2 e nº 3) com sinais correlacionados, utilizando a geometria de 42 microfones (22 elementos eliminados) e com calibração do arranjo. 1ª Coluna: Delay and Sum beamformer. 2ª Coluna: DAMAS2. 3ª Coluna: Regularização ℓ_1 . 4ª Coluna: Regularização TV. 1ª linha: Freq. 500 Hz. 2ª linha: Freq. 1 kHz. 3ª linha: Freq. 2 kHz. 4ª linha: Freq. 5 kHz.

7.6.6 Padrão: Dois alto-falantes na frente do arranjo de microfones com sinais não correlacionados

Posicionaram-se dois alto-falantes na posição nº 2 e nº 3 (ver Figura 7.9). Para realizar a medição com sinais não correlacionados foram utilizadas dois sinais diferentes, foi utilizada a função *ita-generate* do pacote ITA-Toolbox [45] da biblioteca de MATLAB. Utilizamos um sinal de excitação tipo *ruído branco gaussiano* como foi definido no início do capítulo e um sinal tipo *ruído rosa gaussiano* que tem uma sequência pseudoaleatória, com um desvio padrão igual a 1, duração de 2^{17} amostras e com uma taxa de amostragem de 48 kHz.

Na Figura 7.43, são amostrados os dois sinais de excitação utilizadas, na escala de tempo e frequência, observa-se que o sinal tipo *ruído rosa* tem um decaimento de aproximadamente -30 dB com relação ao *ruído branco*. Cada um dos sinais é independente para cada um dos alto-falantes. Os resultados apresentados são para as geometrias com 42 microfones (22 elementos

Figura 7.42: Dois alto-falantes no centro com relação ao arranjo de microfones com distância de 1.5 m e uma separação de 1.3 m (posição nº 2 e nº 3) com sinais correlacionados, utilizando a geometria de 61 microfones (3 elementos eliminados) e com calibração do arranjo. 1ª Coluna: Delay and Sum beamformer. 2ª Coluna: DAMAS2. 3ª Coluna: Regularização ℓ_1 . 4ª Coluna: Regularização TV. 1ª linha: Freq. 500 Hz. 2ª linha: Freq. 1 kHz. 3ª linha: Freq. 2 kHz. 4ª linha: Freq. 5 kHz.

eliminados) e com 61 microfones (3 elementos eliminados), esta última utilizando a variação proposta em cada método de reconstrução.

Figura 7.43: Sinais obtidas utilizando o ITA-toolbox [45]. Coluna (a): observa-se o sinal ruído branco em escala de tempo e frequência. Coluna (b): observa-se o sinal ruído rosa em escala de tempo e frequência.

Serão apresentados resultados trabalhando com três tipos de combinações utilizando os sinais apresentados na Figura 7.43. A primeira combinação será *ruído branco* com *ruído branco*, a segunda é *ruído rosa* com *ruído rosa* e a última será a combinação de *ruído branco* com *ruído rosa*.

Observa-se nas Figuras 7.44 e 7.45 os resultados obtidos com sinais não correlacionados tipo *ruído branco* com *ruído branco*, cada figura apresenta resultados utilizando as geometrias com 42 e 61 microfones respectivamente. Os resultados utilizando uma frequência de 500 Hz permitem identificar a presença de alguma fonte no centro do arranjo, já trabalhando com frequência a partir de 1 kHz pode-se identificar que existem duas fontes. Com frequências com 2 kHz e 5 kHz obtém-se um melhor resultado e é possível identificar claramente as duas fontes.

Utilizando a modificação nos métodos, observa-se nos resultados apresentados na Figura 7.45 que com DAMAS2 e regularização TV com uma frequência de 5 kHz obtém-se um melhor resultado. Podemos dizer que utilizando sinais não correlacionados os métodos de reconstrução de imagens apresentam melhores resultados comparando com os resultados obtidos com sinais correlacionados.

Nas Figuras 7.46 e 7.47 são apresentados resultados com sinais não correlacionados tipo

Figura 7.44: Dois alto-falantes na posição nº 2 e nº 3 com sinais não correlacionados (*ruído branco - ruído branco*), utilizando a geometria de 42 microfones (22 elementos eliminados) e com calibração do arranjo. 1ª Coluna: Delay and Sum beamformer. 2ª Coluna: DAMAS2. 3ª Coluna: Regularização ℓ_1 . 4ª Coluna: Regularização TV. 1ª linha: Freq. 500 Hz . 2ª linha: Freq. 1 kHz. 3ª linha: Freq. 2 kHz. 4ª linha: Freq. 5 kHz.

Figura 7.45: Dois alto-falantes na posição nº 2 e nº 3 com sinais não correlacionados (*ruído branco - ruído branco*), utilizando a geometria de 61 microfones (3 elementos eliminados) e com calibração do arranjo. 1ª Coluna: Delay and Sum beamformer. 2ª Coluna: DAMAS2. 3ª Coluna: Regularização ℓ_1 . 4ª Coluna: Regularização TV. 1ª linha: Freq. 500 Hz . 2ª linha: Freq. 1 kHz. 3ª linha: Freq. 2 kHz. 4ª linha: Freq. 5 kHz.

ruído rosa com *ruído rosa*, cada figura apresenta resultados utilizando as geometrias com 42 e 61 microfones respectivamente. Da mesma forma, os resultados a partir de uma frequência menor como 500 Hz permite identificar a presença de alguma fonte no centro do arranjo comparando com os resultados obtidos com sinais correlacionados, já trabalhando com frequências maiores observa-se a presença de duas fontes. Utilizando a modificação dos métodos e trabalhando com frequências de 2 kHz e 5 kHz os métodos DAMAS2 e regularização TV apresentam os melhores resultados.

Figura 7.46: Dois alto-falantes na posição nº 2 e nº 3 com sinais não correlacionados (*ruído rosa* - *ruído rosa*), utilizando a geometria de 42 microfones (22 elementos eliminados) e com calibração do arranjo. 1ª Coluna: Delay and Sum beamformer. 2ª Coluna: DAMAS2. 3ª Coluna: Regularização ℓ_1 . 4ª Coluna: Regularização TV. 1ª linha: Freq. 500 Hz . 2ª linha: Freq. 1 kHz. 3ª linha: Freq. 2 kHz. 4ª linha: Freq. 5 kHz.

Nas Figuras 7.48 e 7.49 são apresentados resultados com sinais não correlacionados tipo *ruído branco* com *ruído rosa*, utilizando as geometrias com 42 e 61 microfones respectivamente. Observa-se as mesmas características dos resultados anteriores, e de igual forma, trabalhando com sinais não correlacionados os resultados vão ser melhores que trabalhando com sinais correlacionados.

Figura 7.47: Dois alto-falantes na posição nº 2 e nº 3 com sinais não correlacionados (*ruído rosa* - *ruído rosa*), utilizando a geometria de 61 microfones (3 elementos eliminados) e com calibração do arranjo. 1ª Coluna: Delay and Sum beamformer. 2ª Coluna: DAMAS2. 3ª Coluna: Regularização ℓ_1 . 4ª Coluna: Regularização TV. 1ª linha: Freq. 500 Hz . 2ª linha: Freq. 1 kHz. 3ª linha: Freq. 2 kHz. 4ª linha: Freq. 5 kHz.

Figura 7.48: Dois alto-falantes na posição nº 2 e nº 3 com sinais não correlacionados (*ruído branco* - *ruído rosa*), utilizando a geometria de 42 microfones (22 elementos eliminados) e com calibração do arranjo. 1ª Coluna: Delay and Sum beamformer. 2ª Coluna: DAMAS2. 3ª Coluna: Regularização ℓ_1 . 4ª Coluna: Regularização TV. 1ª linha: Freq. 500 Hz . 2ª linha: Freq. 1 kHz. 3ª linha: Freq. 2 kHz. 4ª linha: Freq. 5 kHz.

Figura 7.49: Dois alto-falantes na posição nº 2 e nº 3 com sinais não correlacionados (*ruído branco - ruído rosa*), utilizando a geometria de 61 microfones (3 elementos eliminados) e com calibração do arranjo. 1ª Coluna: Delay and Sum beamformer. 2ª Coluna: DAMAS2. 3ª Coluna: Regularização ℓ_1 . 4ª Coluna: Regularização TV. 1ª linha: Freq. 500 Hz . 2ª linha: Freq. 1 kHz. 3ª linha: Freq. 2 kHz. 4ª linha: Freq. 5 kHz.

8 CONCLUSÕES

Nesta dissertação de mestrado são apresentadas medições feitas utilizando o arranjo de microfones com e sem calibração, utilizando as instalações apresentadas na Seção 7.2. São comparados resultados obtidos usando Delay and Sum beamformer, DAMAS2, e covariance fitting com reconstrução regularizada ℓ_1 com o solver SPGL1 [42] e reconstrução regularizada TV com o solver TVAL3 [44]. Todos os métodos foram acelerados com a transformada rápida de Kronecker (KAT).

São comparados resultados obtidos com uma medição utilizando a geometria com 42 microfones (22 elementos eliminados) e 61 microfones (3 elementos eliminados) do arranjo.

Dos resultados apresentados, concluímos que:

1. O método Delay and Sum beamformer é o mais rápido comparado com os demais algoritmos, mas apresenta uma baixa qualidade de resolução, e percebe-se as diferenças entre os diferentes lóbulos laterais nos diferentes métodos de reconstrução.
2. Comparando os resultados utilizando as geometrias de 49 e 42 microfones, fica claro que à medida que perdemos mais microfones a reconstrução de imagem vai ficando mais distorcida.
3. Observa-se para o método DAMAS2 que com um número de 5000 iterações obtemos uma melhor qualidade que com 1000 no resultado, só que o tempo de processamento é de aproximadamente 5 vezes maior. Do mesmo modo, observamos na Figura 7.17 a variância das imagens obtidas sob as mesmas condições.
4. Para a reconstrução regularizada ℓ_1 o parâmetro (σ) pode ser estimado, já que é relativo ao nível de ruído na sala, e utilizando um número menor de iterações o algoritmo produz uma melhor resolução da imagem e o tempo do processamento é reduzido em quase um terço. Mesmo assim, é o método que apresenta menor confiabilidade devido a que seus resultados apresentam uma constante variância com relação aos artefatos do algoritmo.

5. Para realizar a estimação da matriz de autocorrelação $\mathbf{S}$, os resultados obtidos para 1 e 65 janelas não apresentaram muita diferença no tempo de processamento, por isso, nas medições foram utilizadas o valor de $L = 65$.
6. Dos resultados obtidos colocando espuma no solo e na janela do estúdio, podemos dizer que se tivéssemos uma sala que cumpra todas as condições de uma câmara anecóica (evitar os ecos e ruídos internos dentro da sala) poderíamos obter uma melhor qualidade nos resultados.
7. Utilizando a modificação dos métodos de reconstrução de imagens, observa-se dos resultados que utilizando 42 microfones a imagem vai se tornando um pouco turva. Com 61 microfones, observa-se nos resultados uma melhor qualidade de imagem.
8. Comprovamos que os métodos de reconstrução DAMAS2 e regularização TV apresentam um melhor resultado para poder reconhecer a localização da fonte principal utilizando uma frequência de 2 kHz e 5 kHz. Com a regularização ℓ_1 podemos comprovar que vão aparecendo alguns fantasmas fora do centro da imagem, possivelmente sendo produto dos artefatos próprios dos algoritmos.
9. Utilizando duas fontes não correlacionadas, os resultados utilizando uma frequência de 500 Hz permitem identificar a presença de alguma fonte no centro do arranjo, já trabalhando com frequência a partir de 1 kHz pode-se identificar que existem duas fontes. Com frequências com 2 kHz e 5 kHz obtém-se um melhor resultado e permite identificar claramente as duas fontes.

8.1 Trabalhos Futuros

Dos resultados obtidos, podemos concluir que será necessário realizar futuras medições numa câmara anecóica para poder descartar completamente as reflexões e ecos dentro da sala, com a finalidade de poder obter possivelmente melhores resultados na reconstrução de imagens.

Realizar a montagem de um novo arranjo de microfones na Universidade de São Paulo, para portanto, ser possível continuar reduzindo o custo de fabricação de um arranjo de microfones e fazer a comparação dos diferentes algoritmos com uma nova geometria de posição dos microfones. E ao mesmo tempo tentar consertar os três microfones que estão com defeito na placa desenvolvida em [18].

Por sua vez, os mesmos estudos e testes no trabalho, apresentaram uma grande contribuição aos conhecimentos a ser desenvolvidos e a implementação em projetos futuros para a construção

de equipamento militar e manutenção dos mesmos na Marinha do Peru. O departamento de Serviço de Armas e Eletrônica (SAE) da Marinha, na atualidade está trabalhando em parceria com a Universidade Católica do Peru (PUCP) e a Universidade Nacional de Engenharia (UNI) na construção de um sonar propriamente no Peru. Mas representa um alto custo para a Marinha o recrutamento de pessoas de outras universidades privadas e nacionais, de modo que a Marinha envia oficiais e técnicos para realizar cursos de pós-graduação fora do país, a fim de que o pessoal esteja totalmente preparado intelectualmente para executar trabalhos específicos em cada uma de suas especialidades.

REFERÊNCIAS

- [1] ZIOMEK, L. *Fundamentals of acoustic field theory and space-time signal processing*. [S.l.]: CRC press, 1994.
- [2] SILLER, H. et al. Fly-over source localisation on a boeing 747-400. In: *Berlin Beamforming Conference (BeBeC)*. [S.l.: s.n.], 2010.
- [3] REMILLIEUX, M. C. *Aeroacoustic study of a model-scale landing gear in a semi-anechoic wind-tunnel*. Tese (Doutorado) — Virginia Polytechnic Institute and State University, 2007.
- [4] OERLEMANS, S. *Detection of aeroacoustic sound sources on aircraft and wind turbines*. Tese (Doutorado) — University of Twente, 2009.
- [5] LAUTERBACH, A. et al. Microphone array measurements on high-speed trains in wind tunnels. In: *Berlin Beamforming Conference, BeBeC*. [S.l.: s.n.], 2010.
- [6] KITAGAWA, T. *An investigation into inconsistencies between theoretical predictions and microphone array measurements of railway rolling noise*. Tese (Doutorado) — University of Southampton, 2007.
- [7] MICHEL, U. et al. History of acoustic beamforming. In: *Berlin Beamforming Conference*. [S.l.: s.n.], 2006. p. 20–21.
- [8] FONSECA, W. D. et al. *Desenvolvimento e aplicação de sistema para obtenção de imagens acústicas pelo método do beamforming para fontes em movimento*. Dissertação (Mestrado) — Universidade Federal de Santa Catarina, 2009.
- [9] URICK, R. J. *Principles of underwater sound for engineers*. [S.l.]: Tata McGraw-Hill Education, 1967.
- [10] HAFIZOVIC, I.; NILSEN, C.; KJOLERBAKKEN, M. Acoustic tracking of aircraft using a circular microphone array sensor. *IEEE International Symposium on Phased Array Systems and Technology (ARRAY)*, p. 1025–1032, 2010.
- [11] SOLOVIEV, A. et al. Sonar measurements in ship wakes simultaneous with terrasar-x overpasses. *IEEE Transactions on Geoscience and Remote Sensing*, v. 48, n. 2, p. 841–851, 2010.
- [12] CLAPP, M. A.; ETIENNE-CUMMINGS, R. Sonar echo-location in 2-d using mini-microphone array and spatiotemporal frequency filtering. *IEEE International Symposium on Circuits and Systems*, p. 4775–4778, 2005.
- [13] RAYTHEON, C. *Raytheon Company*. Maio, 2015. Disponível em: <<http://www.raytheon.com/capabilities/products/xbr/>>.

- [14] KELLERMANN, K.; MORAN, J. The development of high-resolution imaging in radio astronomy 1. *Annual Review of Astronomy and Astrophysics*, USA, v. 39, n. 1, p. 457–509, 2001.
- [15] FOUNDATION, N. S. *National Radio Astronomy Observatory*. Abril, 2015. Disponível em: <<http://www.vla.nrao.edu/.htm>>.
- [16] JOHNSON, D. H.; DUDGEON, D. E. *Array signal processing: concepts and techniques*. [S.l.]: Simon & Schuster, 1992.
- [17] CADUDA, F. G. *Método de alta resolução em imageamento acústico*. Dissertação (Mestrado) — Universidade de São Paulo, 2011.
- [18] RIBEIRO, F. P. *Arrays de microfones para medida de campos acústicos*. Tese (Doutorado) — Universidade de São Paulo, 2012.
- [19] HORNE K. JAMES, T. A. P. S. N. B. W.; JAEGER, S. Aeroacoustic study of 26 percent scale semispan model of a boeing 777 wing in the NASA Ames 40 - 80, foot wind tunnel. *International Journal of Aeroacoustics*, Multi-Science, n. 10, 2005.
- [20] HUMPHREYS, W. M.; BROOKS, T. F. Noise spectra and directivity for a scale-model landing gear. *International Journal of Aeroacoustics*, Multi-Science, v. 8, n. 5, p. 409–443, 2009.
- [21] RIBEIRO, F. P.; NASCIMENTO, V. H. Fast transforms for acoustic imaging. Part I: Theory. *IEEE Transactions on Image Processing*, IEEE, v. 20, n. 8, p. 2229–2240, 2011.
- [22] RIBEIRO, F. P.; NASCIMENTO, V. H. Fast transforms for acoustic imaging. Part II: Applications. *IEEE Transactions on Image Processing*, IEEE, v. 20, n. 8, p. 2241–2247, 2011.
- [23] BRANDSTEIN, M.; WARD, D. *Microphone arrays: signal processing techniques and applications*. [S.l.]: Springer Science & Business Media, 2001.
- [24] SODERMAN, P. T.; NOBLE, S. C. Directional microphone array for acoustic studies of wind tunnel models. *Journal of Aircraft*, v. 12, n. 3, p. 168–173, 1975.
- [25] BROOKS, T. F.; HUMPHREYS, W. M. A deconvolution approach for the mapping of acoustic sources (DAMAS) determined from phased microphone arrays. *Journal of Sound and Vibration*, Elsevier, v. 294, n. 4, p. 856–879, 2006.
- [26] WILLIAMS, E. G. *Fourier acoustics: sound radiation and nearfield acoustical holography*. [S.l.]: Academic Press, 1999.
- [27] MOSHER, M. et al. Calibration of microphone arrays for phased array processing. *AIAA Paper*, v. 1678, p. 1997, 1997.
- [28] CORNWELL, T. Hogbom’s CLEAN algorithm. Impact on astronomy. *Astronomy & Astrophysics*, EDP Sciences, v. 500, n. 1, p. 65–66, 2009.
- [29] COX, H.; ZESKIND, R. M.; OWEN, M. M. Robust adaptive beamforming. *IEEE Transactions on, Acoustics, Speech, and Signal Processing*, IEEE, v. 35, n. 10, p. 1365–1376, 1987.

- [30] RAVETTA, P. A. *LORE approach for phased array measurements and noise control of landing gears*. Tese (Doutorado) — Mechanical Engineering Department, Virginia Tech, 2005.
- [31] SIJTSMA, P. Clean based on spatial source coherence. *International Journal of Aeroacoustics*, Multi-Science, v. 6, n. 4, p. 357–374, 2007.
- [32] YARDIBI, T. et al. Sparsity constrained deconvolution approaches for acoustic source mapping. *The Journal of the Acoustical Society of America*, Acoustical Society of America, v. 123, n. 5, p. 2631–2642, 2008.
- [33] BISTAFA, S. R. *Acústica aplicada ao controle do ruído*. [S.l.]: Edgard Blücher, 2008.
- [34] KINSLER, L. E. et al. Fundamentals of acoustics. *Fundamentals of Acoustics, 4th Edition, by Lawrence E. Kinsler, Austin R. Frey, Alan B. Coppens, James V. Sanders, pp. 560. ISBN 0-471-84789-5. Wiley-VCH, December*, v. 1, 1999.
- [35] KUTTRUFF, H. *Room acoustics*. [S.l.]: CRC Press, 2009.
- [36] TREES, H. L. V. *Detection, estimation, and modulation theory*. [S.l.]: John Wiley & Sons, 2004.
- [37] VEEN, B. D. V.; BUCKLEY, K. M. Beamforming: A versatile approach to spatial filtering. *IEEE ASSP Magazine*, v. 5, n. 2, p. 4–24, 1988.
- [38] CIGADA, A.; RIPAMONTI, F.; VANALI, M. The delay & sum algorithm applied to microphone array measurements: numerical analysis and experimental validation. *Mechanical systems and signal processing*, Elsevier, v. 21, n. 6, p. 2645–2664, 2007.
- [39] FONSECA, W. D. et al. *Beamforming considerando difração acústica em superfícies cilíndricas*. Tese (Doutorado) — Universidade Federal de Santa Catarina, 2013.
- [40] CAPON, J. High-resolution frequency-wavenumber spectrum analysis. *Proceedings of the IEEE*, IEEE, v. 57, n. 8, p. 1408–1418, 1969.
- [41] EHRENFRIED, K.; KOOP, L. Comparison of iterative deconvolution algorithms for the mapping of acoustic sources. *AIAA journal*, v. 45, n. 7, p. 1584–1595, 2007.
- [42] BERG, E. V. D.; FRIEDLANDER, M. P. Probing the pareto frontier for basis pursuit solutions. *SIAM Journal on Scientific Computing*, SIAM, v. 31, n. 2, p. 890–912, 2008.
- [43] RUDIN, L. I.; OSHER, S.; FATEMI, E. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, Elsevier, v. 60, n. 1, p. 259–268, 1992.
- [44] LI, C. *An efficient algorithm for total variation regularization with applications to the single pixel camera and compressive sensing*. Tese (Doutorado) — Rice University, Texas, 2009.
- [45] MASIERO, B.; SCHARRER, R.; VORLÄNDER, M. ITA-Toolbox—An open source MATLAB toolbox for acousticians. 38TH German Annual Conference on Acoustics, 2012.
- [46] PROAKIS, J. G. *Digital signal processing: principles, algorithms, and application-3/E*. [S.l.]: PRENTICE-HALL INTERNATIONAL, INC., 1996.
- [47] OJEDA, L. R. *Probabilidad y estadística básica para ingenieros*. [S.l.]: ICM ESPOL, 2014.

APÊNDICE A – PUBLICAÇÕES RELEVANTES

1. MASIERO, B. S.; BELLI, C. A.; NASCIMENTO, V. H. Fast acoustic imaging with separable arrays. *International Symposium on Solid Mechanics, MecSol, B. Horizonte*. A se realizar em 25-27 de Maio de 2015.
2. BELLI, C. A.; MASIERO, B. S.; NASCIMENTO, V. H. Estimação de Imagens Acústicas com Arranjos de Microfones. *XXXIII Simpósio Brasileiro de Telecomunicações, SBrT, Juiz de Fora, Minas Gerais*. A se realizar em 1-4 de Setembro de 2015.