

LUANA TORQUETE LARA

APLICAÇÃO DE WAVELETS E TÉCNICAS CLÁSSICAS DE
PROCESSAMENTO DE SINAIS À CARACTERIZAÇÃO DE
MATERIAIS E FONTES ACÚSTICAS

Belo Horizonte
07 de julho de 2017

DISSERTAÇÃO DE MESTRADO Nº 990

**APLICAÇÃO DE WAVELETS E TÉCNICAS CLÁSSICAS DE
PROCESSAMENTO DE SINAIS À CARACTERIZAÇÃO DE MATERIAIS E
FONTES ACÚSTICAS**

Luana Torquete Lara

DATA DA DEFESA: 07/07/2017

Universidade Federal de Minas Gerais

Escola de Engenharia

Programa de Pós-Graduação em Engenharia Elétrica

**APLICAÇÃO DE WAVELETS E TÉCNICAS CLÁSSICAS DE
PROCESSAMENTO DE SINAIS À CARACTERIZAÇÃO DE
MATERIAIS E FONTES ACÚSTICAS**

Luana Torquete Lara

Dissertação de Mestrado submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Escola de Engenharia da Universidade Federal de Minas Gerais, como requisito para obtenção do Título de Mestre em Engenharia Elétrica.

Orientador: Prof. Wallace do Couto Boaventura

Belo Horizonte - MG

Julho de 2017

L318a

Lara, Luana Torquete.

Aplicação de wavelets e técnicas clássicas de processamento de sinais à caracterização de materiais e fontes acústicas [manuscrito] / Luana Torquete Lara. - 2017.
xvi, 77 f., enc.: il.

Orientador: Wallace do Couto Boaventura.

Dissertação (mestrado) Universidade Federal de Minas Gerais, Escola de Engenharia.

Bibliografia: f. 73-77.

1. Engenharia elétrica - Teses. 2. Wavelets (Matemática) - Teses. 3. Sinal-ruído (Acústica) - Teses. I. Boaventura, Wallace do Couto. II. Universidade Federal de Minas Gerais. Escola de Engenharia. III. Título.

CDU: 621.3(043)

**"Aplicação de Wavelets e Técnicas Clássicas de
Processamento de Sinais à Caracterização de
Materiais e Fontes Acústicas"**

Luana Torquete Lara

Dissertação de Mestrado submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Escola de Engenharia da Universidade Federal de Minas Gerais, como requisito para obtenção do grau de Mestre em Engenharia Elétrica.

Aprovada em 07 de julho de 2017.

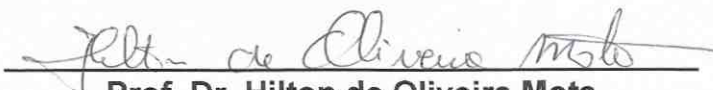
Por:



Prof. Dr. Wallace do Couto Boaventura
DEE (UFMG) - Orientador



Prof. Dr. Maurílio Nunes Vieira
DELT (UFMG)



Prof. Dr. Hilton de Oliveira Mota
DEE (UFMG)

Resumo

Este trabalho estuda técnicas de processamento de sinais aplicadas a sinais acústicos relacionados a dois problemas. O primeiro refere-se à remoção de ruído em sinais que simulam medições em tubo de impedância para posterior cálculo do coeficiente de absorção sonora. Já o segundo problema refere-se à localização de fontes sonoras. O tubo de impedância é um sistema utilizado para estimar a curva de absorção de materiais acústicos, e o método mais utilizado para isso é o dos dois microfones. Tal método consiste em obter uma função de resposta em frequência a partir de sinais de pressão sonora medidos por dois microfones presentes no tubo de impedância, a qual permite obter o coeficiente de absorção. Porém, o ruído de medição degrada a qualidade da estimação dessa curva, principalmente quando a relação sinal ruído é baixa. Assim, propõe-se a utilização de um método de *denoising* baseado em wavelets. Um programa de simulação numérica do tubo de impedância foi utilizado como caixa preta, sendo que uma varredura senoidal exponencial é fornecida ao programa como sinal de excitação do alto-falante, e o programa retorna os sinais dos dois microfones. Estes são contaminados com ruídos Gaussianos brancos e em seguida são submetidos ao processo de remoção de ruído via wavelets. Os sinais filtrados são então processados e a curva de absorção sonora do material é estimada. Os desempenhos de várias famílias de wavelets são investigados. A comparação da curva de absorção estimada com a "exata" mostra que as wavelets são eficientes em baixas frequências. Quanto ao segundo problema, um espaço bidimensional e outro tridimensional são simulados e a localização da fonte sonora é estimada por mínimos quadrados utilizando a diferença de tempo de chegada da onda sonora, TDOA, entre vários microfones dispostos no recinto. Para determinar o valor de TDOA, são utilizadas a correlação cruzada e a correlação cruzada generalizada com transformação de fase. Além disso, é analisada a influência das reflexões e do ruído presente no espaço nessa estimação. Para isso, aos sinais simulados dos microfones são incorporados ruídos gaussianos brancos. Vários sinais de teste são comparados e para todos eles é possível localizar a posição da fonte

sonora.

Palavra-chaves: wavelet, remoção de ruído, absorção acústica, tubo de impedância, localização de fonte acústicas e diferença de tempo de chegada.

Abstract

This work studies signal processing techniques applied to acoustic signals related to two problems. The first refers to the noise removal in signals simulating impedance tube measurements for further calculation of the sound absorption coefficient. The second problem is the sound sources location. The impedance tube is a system used to estimate the absorption curve of acoustic materials, and the method most used is the method of the two microphones. Such method consists of obtaining a frequency response function from sound pressure signals measured by two microphones mounted in the impedance tube, which allows obtaining the absorption coefficient. However, the measurement noise degrades the estimated curve, especially if the signal-to-noise ratio is low. To tackle this problem, a wavelet-based denoising method is proposed. A numerical impedance tube simulation program is used as a black box and an exponential sine sweep is fed to the program as a loudspeaker excitation signal. Then, the program computes the signals from the two microphones. These are further contaminated with white Gaussian noise and subjected to noise removal via wavelets. The filtered signals are then processed and the sound absorption curve of the material is estimated. The performance of several wavelet families is investigated. A comparison of the estimated and "exact" absorption curves shows that wavelets are efficient in noise removal at low frequencies. For the second problem, a two-dimensional and a three-dimensional space are simulated and the sound source location is estimated in the least-squares sense by using the time difference of arrival of the sound wave, TDOA, between several microphones arranged in the enclosure. To evaluate the TDOA, cross-correlation and generalized cross-correlation with phase transformation are used. In addition, the influence of the wall reflections and the background noise in this estimation is investigated. To this end, the simulated microphone signals are corrupted by white Gaussian noise. Several test signals are compared and for all of them it is possible to locate the

position of the sound source.

Keywords: wavelet, denoising, acoustic absorption, impedance tube, acoustic source localization and time difference of arrival.

Dedico meu título de mestre ao Alex, aos meu pais e à minha irmã.

Agradecimentos

Aos meus pais e minha irmã pelo apoio pessoal ao longo do meu mestrado.

Ao Alex, pelo apoio pessoal e profissional e pelo incentivo incondicional.

A meu orientador, Prof. Wallace Boaventura, por aceitar ser meu orientador e compartilhar comigo seu conhecimento.

Aos membros da banca, por se dedicarem a ler e a avaliar meu trabalho.

Aos professores, colegas do Laboratório de Compatibilidade Eletromagnética, e amigos do mestrado, por todas as contribuições dadas ao longo dessa jornada.

À Nossa Senhora do Carmo por me guiar e presentear com tantas pessoas queridas!

Lista de Símbolos e Abreviações

Letras Latinas

a	-	Parâmetro de escala
a_0^m	-	Discretização do parâmetro a
a_l	-	Coefficientes reais da expansão de sinais
$a_{j,k}$	-	Coefficientes wavelets
B	-	Densidade de fluxo magnético
b	-	Parâmetro de deslocamento
b_0	-	Discretização do parâmetro b
C_ψ	-	Constante de admissibilidade
c	-	Velocidade do som [m/s]
c_l	-	Constante de amortecimento viscoso do alto-falante excitador
c_s	-	Constante de amortecimento viscoso do alto-falante absorvedor
c_j	-	Coefficiente de aproximação
d	-	Diâmetro do tubo de impedância
$d_{j,k}$	-	Coefficientes de detalhes
d_{rf}	-	Distância entre o microfone de referência e a fonte
d_{if}	-	Distância entre o i -ésimo microfone e a fonte
F_{img}	-	Fonte imagem
F_{re}	-	Fonte real
f_e	-	Força eletromagnética
f_l	-	Frequência inferior do tubo de impedância
f_s	-	Frequência de amostragem
f_u	-	Frequência superior do tubo de impedância
g_{mi}	-	Resposta ao impulso acústico entre a fonte e o i -ésimo microfone
H	-	Resposta em frequência
$H_i(\omega)$	-	Transformada de Fourier do sinal y_i filtrado
$H_r(\omega)$	-	Transformada de Fourier do sinal y_r filtrado
h	-	Coefficientes do filtro passa-baixas
h_1	-	Coefficientes do filtro passa-altas

Im	-	Parte imaginária
$i = \sqrt{-1}$	-	Formador da parte imaginária de números complexos
j	-	Parâmetro de escala
k	-	Parâmetro de deslocamento
K	-	Constante para o tubo de impedância
k_l	-	Rigidez da suspensão do alto-falante excitador
k_s	-	Rigidez da suspensão do alto-falante absorvedor
L	-	Comprimento do tubo de impedância
L^2	-	Espaço Hilbert
l	-	Distância entre o microfone 1 e a amostra de teste no tubo de impedância
l_e	-	Comprimento da bobina do alto-falante
m_l	-	Massa do alto-falante excitador
m_s	-	Massa do alto-falante absorvedor
m_v	-	Momento de esvaecimento (<i>vanishing</i>)
M_i	-	h microfone
N_m	-	Número total de microfones
$P(z, t)$	-	Pressão acústica
$P(\omega)$	-	Transformada de Fourier da pressão sonora
$P_i(\omega)$	-	Transformada de Fourier da pressão incidente
$P_r(\omega)$	-	Transformada de Fourier da pressão refletida
p_i	-	Pressão incidente
p_r	-	Pressão refletida
$p(l, t)$	-	Pressão sonora no ponto l no instante de tempo t
$p(l + s, t)$	-	Pressão sonora no ponto $l + s$ no instante de tempo t
R	-	Coeficiente de reflexão sonora
\Re	-	Parte Real
R_e	-	Resistência da bobina
$R_{y_r y_i}^{cc}(\xi)$	-	FCC entre os sinais y_r e y_i
$R_{y_r y_i}^{ccg}(\xi)$	-	FCCG entre os sinais y_r e y_i
S	-	Área da seção transversal do tubo de impedância
s	-	Distância entre os microfones
T	-	Tempo de duração da varredura exponencial
T_h	-	Regra de <i>thresholding hard</i>
T_s	-	Regra de <i>thresholding soft</i>
t	-	Tempo
$U_{in}(\omega)$	-	Transformada de Fourier da velocidade de partícula
$V(\omega)$	-	Transformada de Fourier do sinal de tensão do alto-falante excitador
$v(t)$	-	Sinal de tensão do alto-falante

\hat{v}	- Amplitude da varredura exponencial
x	- Coordenada espacial
$x_s(t)$	- Sinal emitido pela fonte sonora
Z_s	- Impedância acústica característica
z	- Coordenada espacial
$Y_r(\omega)$	- Transformada de Fourier do sinal do microfone de referência
$Y_i(\omega)$	- Transformada de Fourier do sinal do i -ésimo microfone
y_1	- Sinal de pressão sonora medido pelo microfone 1
y_2	- Sinal de pressão sonora medido pelo microfone 2
y_i	- Sinal medido pelo i microfone
$\tilde{y}_i(t)$	- Envelope extraído do sinal y_i
$\hat{y}_i(t)$	- Transformada Hilbert do sinal y_i
$W_{a,b}$	- Transformada contínua de wavelets
$W_{m,n}$	- Transformada discreta de wavelets

Letras Gregas

α	- Coeficiente de absorção sonora
β	- Constante de atenuação da onda sonora no espaço
Γ_{ri}	- Função de ponderação dos sinais y_r e y_i
ζ_l	- Deslocamento resultante do diafragma do alto-falante excitador
ζ_s	- Deslocamento resultante do diafragma do alto-falante absorvedor
$\dot{\zeta}_1$	- Velocidade da massa m_l
Θ	- Matriz com as posições da fonte e dos microfones
$\tilde{\Theta}$	- Matriz com as posições da fonte e dos microfones estimada
Θ_f	- Vetor com as coordenadas espaciais da fonte sonora
Θ_{mi}	- Vetor com as coordenadas espaciais do i -ésimo microfone
Θ_{mr}	- Vetor com as coordenadas espaciais do microfone de referência
κ	- Número de onda
$\bar{\kappa}$	- Número de onda complexo
λ	- Limiar para remoção de ruído
$\nu(t)$	- Ruído presente no ambiente
ξ	- Atrasos da função de correlação
Ξ	- Constante de atenuação no tubo de impedância
ρ	- Densidade do meio gasoso
ρ_e	- Densidade das reflexões presentes no espaço
σ	- Desvio padrão
τ	- Tempo de atraso da onda sonora
τ_{12}	- TDOA entre os sinais dos microfones 1 e 2

$\phi(t)$	-	Função escala
$\phi_{\text{Haar}}(t)$	-	Função escala de Haar
$\psi(t)$	-	Wavelet mãe
$\psi_{a,b}$	-	Expansão wavelet em tempo contínuo
$\psi_{j,k}$	-	Expansão wavelet
$\psi_{m,n}$	-	Expansão wavelet em tempo discreto
$\Psi(\omega)$	-	Transformada de Fourier da função wavelet
ψ_{Haar}	-	Wavelet de Haar
ψ_{Morlet}	-	Wavelet de Morlet
ψ_{Mexh}	-	Wavelet chapéu mexicano
ω_0	-	Frequência central da wavelet
ω_1	-	Frequência inicial da varredura exponencial
ω_2	-	Frequência final da varredura exponencial

Abreviações

CC	-	Correlação Cruzada
CCG	-	Correlação Cruzada Generalizada
DFT	-	Transformada Discreta de Fourier
FCC	-	Função de Correlação Cruzada
FCCG	-	Função de Correlação Generalizada
FIR	-	Resposta so Impulso Finita
FRF	-	Função Resposta em Frequência
ITD	-	Diferença de Tempo Interaural (<i>Interaural Time Difference</i>)
LS	-	Mínimos Quadrados sem Restrição (<i>Least-Square</i>)
MR	-	Multiresolução
MUSIC	-	Classificação Múltiplo Sinal (<i>Multiple Signal Classification</i>)
PHAT	-	Transformação de Fase (<i>Phase Transform</i>)
PSMG	-	Pulso Senoidal Modulado por uma Gaussiana
SA	-	Sinal Aleatório
SD	-	Sinal com decaimento exponencial
SNR	-	Relação Sinal-Ruído (<i>Signal-to-Noise Ratio</i>)
TCW	-	Transformada Contínua de Wavelets
TDE	-	Estimação do Tempo de Atraso (<i>Time Delay Estimation</i>)
TDOA	-	Diferença de Tempo de Chegada (<i>Time Difference Of Arrival</i>)
TDW	-	Transformada Discreta de Wavelets
TOA	-	Tempo de Chegada (<i>Time Of Arrival</i>)
TW	-	Transformada de Wavelets
VE	-	Varrredura Exponencial

Lista de Figuras

1.1	Tubo de impedância e o método dos dois microfones.	2
2.1	Representação do tubo de impedância com dois microfones.	6
2.2	Modelo do tubo de impedância.	10
2.3	Varredura exponencial com $\omega_1 = 2\pi 100$ rad/s, $\omega_2 = 2\pi 1000$ rad/s e $T = 0.5$ s.	12
2.4	Espectro da varredura exponencial com $\omega_1 = 2\pi 100$ rad/s, $\omega_2 = 2\pi 1000$ rad/s e $T = 0.5$ s	13
2.5	Wavelet de Haar (à esquerda) e escala de Haar (à direita).	18
2.6	Construção da wavelet de Morlet (à esquerda) e parte real e imaginária da wavelet de Morlet (à direita).	19
2.7	Wavelet chapéu mexicano.	20
2.8	Wavelets Daubechies (à esquerda) e funções escala Daubechies (à direita).	20
2.9	Obtenção dos coeficientes de resolução mais baixa a partir dos de alta resolução (à esquerda) e dos coeficientes de resolução mais alta a partir dos de baixa resolução (à direita).	25
2.10	Implementação da transformada de wavelets discreta através de um banco de filtros digitais.	25
2.11	Reconstrução do sinal original através do banco de filtros.	26
2.12	Espectrograma representativo da varredura logarítmica com frequências iniciais e finais de 30 Hz e 6 kHz.	28
2.13	Sinais dos microfones 1 (à esquerda) e 2 (à direita).	29
2.14	Diferença da pressão sonora medida entre os microfones 1 e 2.	30
2.15	Sinais original e filtrado (db45 com 11 níveis, $\sigma = 0,006$, limiar = 0,0005%) do microfone 2 (à esquerda) e um zoom nesses sinais (à direita).	30
2.16	Coefficiente de absorção exato.	31
2.17	Função wavelet db45 (à esquerda) e função escala db45 (à direita).	32
2.18	Função wavelet db20 (à esquerda) e função escala db20 (à direita).	32
2.19	Função wavelet coif5 (à esquerda) e função escala coif5 (à direita).	32
2.20	Função wavelet sym20 (à esquerda) e função escala sym20 (à direita).	33

2.21	Curvas de absorção sonora estimadas (à esquerda) e diferença entre os sinais original e filtrado no microfone 2 (à direita) para db45 com 11 níveis, $\sigma = 0,006$, limiar = 0,0005%.	33
2.22	Curvas de absorção sonora estimadas (à esquerda) e diferença entre os sinais original e filtrado no microfone 2 (à direita) para db45 com 11 níveis, $\sigma = 0,006$, limiar = 0,001%.	33
2.23	Curvas de absorção sonora estimadas (à esquerda) e diferença entre os sinais original e filtrado no microfone 2 (à direita) para db20 com 13 níveis, $\sigma = 0,006$, limiar = 0,0005%.	34
2.24	Curvas de absorção sonora estimadas (à esquerda) e diferença entre os sinais original e filtrado no microfone 2 (à direita) para sym20 com 10 níveis, $\sigma = 0,006$, limiar = 0,0005%.	34
2.25	Curvas de absorção sonora estimadas (à esquerda) e diferença entre os sinais original e filtrado no microfone 2 (à direita) para coif5 com 9 níveis, $\sigma = 0,006$, limiar = 0,0005%.	34
2.26	Curvas de absorção estimadas (à esquerda) e diferença entre os sinais original e filtrado no microfone 2 (à direita) para db45 com 11 níveis, $\sigma = 0,002$, limiar = 0,0005%.	35
3.1	Representação de um ambiente reverberante com uma fonte sonora e um microfone.	37
3.2	Representação do método da fonte imagem.	39
3.3	Simulação de um espaço bidimensional com quatro microfones considerando três reflexões e ruído ambiente (à esquerda) e correlação cruzada entre o sinal de entrada e o microfone 2 (à direita).	43
3.4	Resultado da correlação cruzada generalizada utilizando transformação de fase (CCG-PHAT) entre o sinal emitido pela fonte sonora e o microfone 2.	45
3.5	Pulso senoidal com modulação gaussiana e seu envelope.	46
3.6	Representação de um ambiente de duas dimensões com uma fonte sonora e dois microfones.	48
3.7	Posições da fonte sonora e dos microfones em \mathbb{R}^2	51
3.8	Método das fontes imagens em \mathbb{R}^2 considerando duas reflexões de primeira ordem e uma de segunda ordem.	52
3.9	Diagrama das simulações realizadas para estimação da posição da fonte sonora no espaço.	54
3.10	Sinal emitido pela fonte sonora, PSMG, e sinais de resposta simulados para os 4 microfones em \mathbb{R}^2 sem considerar as reflexões à esquerda, e considerando as reflexões à direita.	56
3.11	Sinais dos microfones sem considerar as reflexões e com um ruído de $\sigma = 0,2$	57

3.12	Erros relativos de todos os algoritmos testados para os sinais dos microfones com ruído de $\sigma = 0,2$ e sem a presença de reflexões quando o sinal de entrada é o PSMG à esquerda, e os melhores resultados à direita.	58
3.13	Erros de todos os algoritmos testados para sinais dos microfones com ruído de $\sigma = 0,2$ e $\beta = 0,4$ quando o sinal de entrada é o PSMG à esquerda, e os melhores resultados à direita.	58
3.14	Sinal emitido pela fonte sonora, sinal aleatório, e sinais de resposta simulados para os 4 microfones em \mathbb{R}^2 sem considerar as reflexões locais.	59
3.15	Sinal aleatório da fonte sonora e sinais dos microfones com $\sigma = 0,2$ e $\beta = 0,4$	60
3.16	Diferença entre sinais dos microfones com ruídos, $\sigma = 0,2$ e $\beta = 0,4$, e os filtrados pela wavelet.	61
3.17	Sinal emitido pela fonte sonora, decaimento exponencial, e sinais de resposta simulados para os 4 microfones em \mathbb{R}^2 sem considerar as reflexões locais.	61
3.18	Sinais dos microfones com ruído de $\sigma = 0,2$ e $\beta = 0,4$ quando o sinal da fonte sonora é um decaimento exponencial.	62
3.19	Melhores resultados de erros na estimação da posição da fonte sonora quando os sinais dos microfones possuem $\sigma = 0,2$ e $\beta = 0,4$ e o sinal de teste é o SD.	63
3.20	Sinal emitido pela fonte sonora, varredura exponencial, e sinais de resposta simulados para os 4 microfones com $\beta = 0$ (à esquerda) e $\beta = 0,4$ (à direita).	63
3.21	Resultados para todos os algoritmos testados quando o sinal da fonte é a varredura exponencial e os sinais dos microfones possuem $\sigma = 0,2$ e $\beta = 0,4$ (à esquerda) e melhores resultados (à direita)	64
3.22	Sinal da fonte sonora, varredura exponencial, e sinais dos microfones com $\beta = 0,4$ e sem ruído em \mathbb{R}^3	66
3.23	Resultados de alguns dos algoritmos testados para os sinais dos microfones com $\sigma = 0,2$ à esquerda e $\sigma = 0,4$ à direita.	67
3.24	Sinal emitido pela fonte, SD, e sinais medidos pelos microfones no espaço tridimensional.	68
3.25	Resultados para alguns dos algoritmos testados quando os sinais dos microfones possuem ruído com $\sigma = 0,2$, para o sinal de decaimento exponencial como sinal de teste.	68

Lista de Tabelas

3.1	Posições das fontes sonoras, real e imaginárias, e dos microfones em \mathbb{R}^2	52
3.2	Posição das fontes sonoras, real e imaginária, e dos microfones em \mathbb{R}^3	53
3.3	Erros (%) relativos da estimação da posição da fonte sonora quando o sinal da fonte sonora é o PSMG considerando alguns valores do coeficiente de atenuação (β).	57
3.4	Erros (%) relativos da estimação da posição da fonte sonora quando o sinal da fonte sonora é aleatório e os sinais dos microfones não possuem ruídos.	59
3.5	Erros (%) relativos da estimação da posição da fonte sonora quando o sinal da fonte é o sinal de decaimento exponencial com frequência de 100 Hz e não há ruído nos sinais dos microfones.	62
3.6	Erros (%) relativos da estimação da posição da fonte sonora para a varredura exponencial como sinal de teste e sinais dos microfones sem ruído.	64
3.7	Erros (%) relativos da estimação da posição da fonte sonora para todos os sinais de teste no espaço tridimensional.	65
3.8	Erros (%) relativos da estimação da posição da fonte sonora para os algoritmos que utilizam decomposição em wavelets no espaço tridimensional.	66
3.9	Erros (%) relativos da estimação da posição da fonte sonora para todos os sinais de teste no espaço tridimensional considerando as reflexões do ambiente.	66
3.10	Menores erros (%) da estimação da posição da fonte e seus respectivos algoritmos para simulação do espaço bidimensional.	70

Sumário

1	Introdução	1
1.1	Motivação e objetivos	1
1.2	Principais contribuições deste trabalho	3
1.3	Organização do texto	4
2	Medição de absorção sonora de materiais acústicos	5
2.1	Caracterização de materiais acústicos	5
2.1.1	Sinal de teste	11
2.2	Fundamentos de wavelets	13
2.2.1	Wavelets	14
2.2.2	Famílias de wavelets	17
2.2.3	Transformada contínua de wavelets	21
2.2.4	Transformada discreta de wavelets	22
2.2.5	Algoritmo de Mallat	23
2.3	Wavelets na remoção de ruído	26
2.4	Metodologia	27
2.5	Resultados	29
2.6	Conclusões	35
3	Localização de fontes sonoras	36
3.1	Fundamentos de localização de fontes sonoras	36
3.1.1	Modelo de um ambiente reverberante	37
3.2	Algoritmos de TDOA	40
3.2.1	Método da correlação cruzada	41
3.2.2	Método da correlação cruzada generalizada	43
3.2.3	Transformada wavelets na estimação de TDOA	45
3.3	Localização de fontes sonoras	47
3.4	Metodologia	51
3.5	Resultados	55
3.5.1	Pulso senoidal com modulação gaussiana	56

3.5.2	Sinal Aleatório	59
3.5.3	Sinal com decaimento exponencial	61
3.5.4	Varredura exponencial	63
3.5.5	Simulações no espaço tridimensional	65
3.6	Conclusões	68
4	Considerações finais	71
4.1	Conclusões gerais	71
4.2	Sugestões de trabalhos futuros	72
4.3	Publicações	72
	Referências Bibliográficas	73

Capítulo 1

Introdução

1.1 Motivação e objetivos

Existem várias ferramentas de processamento de sinais aplicadas a problemas relacionados a sinais acústicos. A wavelet é uma dessas ferramentas e pode ser aplicada para remover ruído incorporados aos sinais. Assim, esse trabalho tem como um dos objetivos minimizar ou até mesmo remover o ruído presente em sinais simulados em tubos de impedância para posterior cálculo da curva de absorção sonora de um material. Outro objetivo desse trabalho é comparar os métodos mais clássicos de estimação da posição da fonte sonora.

O tubo de impedância é um sistema acústico bastante conhecido para avaliar as propriedades de reflexão e absorção de materiais acústicos. Ele consiste de um tubo contendo em uma extremidade um alto-falante excitador e na outra um material a ser testado. E um método bastante utilizado é o método dos dois microfones, Fig. 1.1, no qual um sinal de teste excita o alto-falante e a pressão sonora em dois pontos diferentes do tubo é medida através de dois microfones. A partir desses sinais é calculado a curva de absorção sonora do material. Porém, um dos problemas na aplicação desse método é a presença de ruído nos sinais medidos, fazendo com que a curva de absorção seja alterada de forma errada. Assim, esse trabalho propõe aplicar as transformadas de wavelets nos sinais dos microfones para remover o ruído antes de estimar tal curva.

Wavelets são ondas de curta duração e com energia finita concentrada em uma determinada região do sinal [12]. Elas tiveram um rápido desenvolvimento na década de 1980 com esforços compartilhados por pesquisadores de diversas áreas e têm sido utilizadas como base para representação de funções e como técnica para análise tempo-frequência [18]. Sua representação matemática é dada por operações de deslocamento no tempo e de escala a partir de uma função primária denominada de wavelet mãe, formando um conjunto de funções com suporte compacto, ou seja, a energia do sinal é

concentrada em uma região do sinal.

A transformada de wavelets possui a capacidade de detectar fenômenos não estacionários tais como surtos, descontinuidades e impulsos momentâneos. Ela tem sido proposta como uma ferramenta flexível para a decomposição multiresolução (MR) de sinais de tempo contínuo, e em seguida, estuda cada componente com uma resolução correspondente à sua escala [17], sendo possível ter intervalos maiores para adquirir informações mais precisas em baixas frequências e intervalos menores quando a precisão deve ser sobre altas frequências. Assim, a transformada wavelet é melhor do que a transformada breve de Fourier para análise de fenômenos de alta frequência, ou de curta duração, como transientes presentes em sinais. Para exemplificar essa aplicação, têm-se o trabalho de Ribeiro [38] que mostra a transformada wavelet como uma ferramenta de análise para distorções harmônicas não estacionárias em sistemas de energia. Santos et al. [40] aplicam essa transformada para avaliar a qualidade de energia e Wikinson et al. [52] utilizam a transformada discreta de wavelets para análise de transientes em sistemas de potência. Além dessas aplicações, pode-se citar a utilização de tal teoria em sistemas de comunicação [4], em problemas de compressão de sinais [51] e de imagens [46], em remoção de ruído [50, 42, 41, 33], e em processamentos de sinais de voz [22].

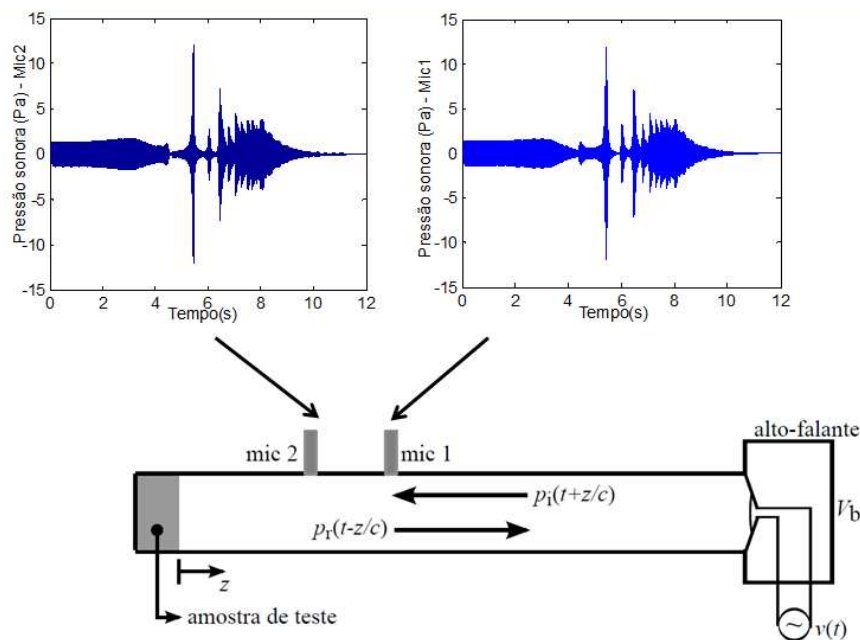


Figura 1.1: Tubo de impedância e o método dos dois microfones.

A localização de fontes acústicas é a determinação das coordenadas da fonte sonora em relação a um ponto do espaço. Ela é uma ferramenta relevante em diferentes aplicações, por exemplo na garantia da qualidade sonora, localização de fontes de ruído criada por uma máquina, no aprimoramento da fala, em geofísica, em sistemas

de vigilância ou até mesmo para localizar uma pessoa através do som emitido por ela e outras aplicações.

A base da localização de fontes sonoras está no princípio da propagação de ondas acústicas e recepção dessas ondas por um arranjo de microfones em diferentes pontos no espaço. Uma das formas, amplamente estudadas e bem compreendidas na literatura, de identificar a fonte sonora é através da estimativa de tempo de atraso (*Time Delay Estimation* - TDE), isto é, do tempo que a onda precisa para alcançar os microfones [5]. A estimação de TDE pode ser muito útil, por exemplo, para uma sala de teleconferência automática, em que uma câmera deve apontar para o falante. Nesse caso, o ruído na sala, além da reverberação, pode também ser um problema se a relação sinal-ruído se tornar menor que 20 dB, mas isso raramente ocorre em salas de teleconferência [6].

Existem várias abordagens para estimação do TDE, por exemplo, que incluem a determinação da função de transferência entre canais dos microfones de forma adaptativa [37], ou as repostas ao impulso entre a fonte e os receptores [6]. Mas os métodos mais comuns são baseados em correlação cruzada e correlação cruzada generalizada. Na correlação cruzada generalizada a estimativa de atraso é obtida como o intervalo de tempo que maximiza a correlação cruzada entre as versões filtradas dos sinais que estão sendo processados. Porém um problema nessa técnica é a presença de ruído, embora o maior problema na estimação de TDE seja a presença da reverberação [6]. Assim, o presente trabalho também estuda esses métodos de estimação de TDE e avalia a influência do ruído e das reflexões no processo de estimação da fonte sonora.

1.2 Principais contribuições deste trabalho

Como contribuição desse trabalho e possível solução para o primeiro problema, propõe-se a utilização da transformada discreta de wavelets para minimizar o ruído dos sinais. Assim, é realizado um estudo sobre a utilização dessa ferramenta no processo de *denoising*.

Quanto ao segundo problema, é realizado um estudo comparativo entre os métodos, presentes na literatura, mais comuns na utilização de estimação da posição da fonte sonora. Assim, são simulados os espaços bidimensional e tridimensional e aos sinais que seriam medidos pelos microfones, presentes nesses espaços, são incorporados ruídos gaussianos brancos, para posterior tentativa de minimização desse ruído via wavelets.

1.3 Organização do texto

O presente trabalho apresenta dois objetivos correlatos que são introduzidos e em seguida, são apresentados a fundamentação teórica, simulações e resultados do primeiro objetivo e posteriormente do segundo. De tal forma, os capítulos são divididos em:

- Capítulo 1: Introdução: apresentação da motivação da escolha do tema desse trabalho, objetivos e contribuições desse estudo;
- Capítulo 2: Medição de absorção sonora de materiais acústicos: descrição do sistema de tubo de impedância, seu funcionamento e utilização para cálculo do coeficiente de absorção sonora; apresentação dos fundamentos de wavelets, suas aplicações, transformadas e principalmente o algoritmo para utilizá-la na remoção de ruído; metodologia utilizada nas simulações referentes à remoção de ruído no sistema de tubo de impedâncias, resultados obtidos e conclusões.
- Capítulo 3: Localização de fontes sonoras: estudo comparativo entre alguns dos métodos mais utilizados na literatura para estimar a localização da fonte sonora; metodologia utilizada nas simulações referentes à estimação da posição da fonte sonora, resultados e conclusões.
- Capítulo 6: Considerações finais: Conclusões gerais, sugestões de trabalhos futuros e publicações realizadas.
- Referências Bibliográficas.

Capítulo 2

Medição de absorção sonora de materiais acústicos

Os tubos de impedância acústica são equipamentos amplamente utilizados para medir os coeficientes de reflexão e absorção de materiais acústicos sob incidência normal de ondas sonoras, bem como a impedância da superfície do material. Um método bastante conhecido para obter esses coeficientes utilizando o tubo de impedância é o método dos dois microfones. Porém, um dos problemas na aplicação desse método é a presença de ruído nos sinais medidos pelos microfones. Como solução desse problema, esse trabalho propõe utilizar wavelets para minimizar tal ruído. Deste modo, esse capítulo apresenta uma breve revisão bibliográfica sobre os fundamentos da caracterização de materiais acústicos e os fundamentos de wavelets para que o leitor possa compreender a aplicação da técnica de *denoising* via wavelets a este problema. Além disso, este capítulo apresenta a metodologia das simulações realizadas, resultados e conclusões obtidos.

2.1 Caracterização de materiais acústicos

O tubo de impedância consiste em um tubo reto de seção transversal constante, geralmente circular ou retangular, sendo que uma fonte sonora é montada em uma de suas extremidades, e uma amostra do material a ser testado é montada na outra extremidade, conforme mostra a Fig. 2.1.

Usualmente, a fonte sonora é um alto-falante montado em uma caixa de volume V_b e alimentado com um sinal de teste de tensão $v(t)$, e assim, serão produzidas ondas acústicas planas dentro do tubo. A técnica mais usual de medição consiste em utilizar um sinal de banda larga como sinal de tensão do alto-falante, e captar as respostas através de dois microfones montados na parede do tubo e espaçados entre si por uma

distância s [1, 2]. Através do processamento dos sinais de pressão sonora medidos, pode-se estimar a curva de absorção sonora do material analisado.

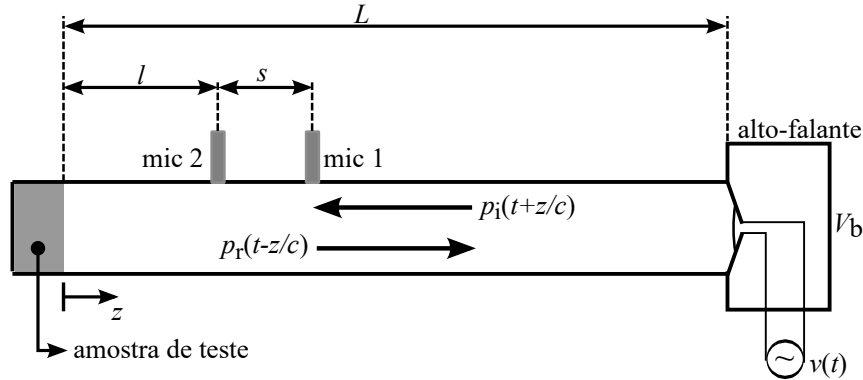


Figura 2.1: Representação do tubo de impedância com dois microfones.

A propagação de ondas sonoras planas dentro de um duto rígido de seção transversal constante é governada pela equação diferencial parcial [30],

$$\frac{\partial^2 p(z, t)}{\partial z^2} = \frac{1}{c^2} \frac{\partial^2 p(z, t)}{\partial t^2}, \quad (2.1)$$

em que $p(z, t)$ é a pressão acústica, c é a velocidade do som, z é a coordenada espacial ao longo do eixo do duto (ver Fig. 2.1) e t é o tempo. Esta equação é válida para ondas de pequena amplitude e negligencia quaisquer processos dissipativos que possam ocorrer no domínio acústico ($0 < z < L$, onde L é o comprimento do tubo), tais como vibrações e perdas viscosas e térmicas na parede interna do duto. Além disso, a condição de onda plana requer que [2]

$$d < K \frac{c}{f_u}, \quad (2.2)$$

no qual f_u é a frequência superior de operação do tubo. Para um tubo circular, d corresponde ao seu diâmetro e ASTM E1050 [2] utiliza $K = 0,586$, enquanto que a ISO 10534-2 [1] considera $K = 0,58$; para um tubo retangular, d é a maior dimensão da sua seção e a ASTM E1050 [2] utiliza $K = 0,500$, já a ISO 10534-2 [1] considera $K = 0,5$.

A Eq. (2.2) é usada para definir o tamanho da seção transversal de um tubo de impedância para evitar ondas não planas, como exigido para medições de tubo de impedância. Observe que a seção transversal deve ser reduzida à medida que a frequência máxima de interesse aumenta. De fato, ondas não planas sempre surgem na vizinhança da fonte sonora e da amostra de teste. Portanto, os microfones não devem ser colocados próximos às extremidades do tubo. Recomenda-se ter uma distância mínima de três diâmetros do tubo entre o alto-falante e o microfone 1 (mic 1 na Fig. 2.1), e

dois diâmetros do tubo entre o material de teste e seu microfone mais próximo (mic 2 Fig. 2.1) [2], isto é, $l > 2d$.

Se o tubo de impedância for projetado de modo que as suposições descritas acima sejam cumpridas, a Eq. (2.1) é um modelo matemático adequado. Sua solução é a soma de uma onda plana incidente viajando na direção $-z$, $p_i(t + z/c)$, e uma onda plana refletida viajando na direção $+z$, $p_r(t - z/c)$, isto é,

$$p(z, t) = p_i(t + z/c) + p_r(t - z/c). \quad (2.3)$$

Sejam $P_i(z, \omega)$ e $P_r(z, \omega)$ as transformações de Fourier para um dado z de $p_i(t + z/c)$ e $p_r(t - z/c)$, respectivamente, onde ω é a frequência angular. O coeficiente de reflexão da amostra de ensaio é definido como

$$R(\omega) \equiv \frac{P_r(0, \omega)}{P_i(0, \omega)}, \quad (2.4)$$

sendo um número complexo que expressa as diferenças de amplitude e fase entre as ondas incidentes e refletidas na superfície do espécime. O coeficiente de absorção acústica, $\alpha(\omega)$, é definido como a relação entre a potência acústica absorvida pelo espécime e a potência acústica da onda incidente. Dado que a potência absorvida é proporcional a $|P_i(0, \omega)|^2 - |P_r(0, \omega)|^2$, e a potência incidente é proporcional a $|P_i(0, \omega)|^2$ (a constante de proporcionalidade é a mesma em ambos os casos), tem-se que [2]

$$\alpha(\omega) = 1 - |R(\omega)|^2. \quad (2.5)$$

Além dos coeficientes de reflexão e de absorção, a impedância acústica específica da amostra, $Z_s(\omega)$, também é importante, sendo definida por

$$Z_s(\omega) \equiv \frac{P(0, \omega)}{U_{in}(\omega)}, \quad (2.6)$$

no qual $P(0, \omega)$ e $U_{in}(\omega)$ são as Transformadas de Fourier, respectivamente, da pressão sonora na superfície da amostra e da velocidade de partícula fluida entrando na amostra (sentido negativo de z).

As relações,

$$\begin{cases} P(0, \omega) &= P_r(0, \omega) + P_i(0, \omega), \\ U_{in}(\omega) &= -[P_r(0, \omega) - P_i(0, \omega)]/(\rho c), \end{cases} \quad (2.7)$$

juntamente com as definições de $R(\omega)$ e $Z_s(\omega)$, resultam em

$$\frac{Z_s(\omega)}{\rho c} = \frac{(1 + R(\omega))}{(1 - R(\omega))}. \quad (2.8)$$

sendo ρ a densidade do meio gasoso, neste caso, o ar.

As equações (2.5) e (2.8) mostram que $\alpha(\omega)$ e $Z_s(\omega)$ podem ser obtidos de $R(\omega)$ com facilidade. No entanto, a Eq. (2.4) não fornece um meio para a avaliação experimental de $R(\omega)$ porque $P_r(0, \omega)$ e $P_i(0, \omega)$ não podem ser medidos diretamente. Portanto, uma expressão que forneça $R(\omega)$ a partir dos sinais de pressão sonora medidos pelos dois microfones deve ser derivada. Primeiro, a aplicação a transformada de Fourier à Eq. (2.3), conduz a

$$P(z, \omega) = P_i(0, \omega)e^{i\kappa z} + P_r(0, \omega)e^{-i\kappa z}, \quad (2.9)$$

sendo $\kappa = \omega/c$ o número de onda, $i \equiv \sqrt{-1}$, e a propriedade de deslocamento no tempo da transformada de Fourier foi usada. Seja $z = l$ a posição do microfone 2, e s seja o espaçamento entre os microfones, tal que $z = l + s$ é a posição do microfone 1, como mostrado na Fig. 2.1. Consequentemente, $P(l, \omega)$ e $P(l + s, \omega)$ são as transformadas de Fourier dos sinais de pressão medidos pelos microfones 2 e 1, respectivamente. Então, a Eq. (2.9) pode ser usada para relacionar $P(l, \omega)$ e $P(l + s, \omega)$ com $P_r(0, \omega)$ e $P_i(0, \omega)$, e assim obter $R(\omega)$. Após algumas manipulações algébricas, obtém-se

$$R(\omega) = \left[\frac{H(\omega) - e^{-i\kappa s}}{e^{-i\kappa s} - H(\omega)} \right] e^{i2\kappa(l+s)}, \quad (2.10)$$

no qual,

$$H(\omega) \equiv \frac{P(l, \omega)}{P(l + s, \omega)} \quad (2.11)$$

é a função resposta em frequência (FRF) de um sistema linear invariante no tempo cuja entrada é $p(l + s, t)$, e a saída é $p(l, t)$, ou seja, é a FRF entre os microfones 2 e 1. O espaçamento entre os microfones desempenha um papel importante na acurácia das medições. A ASTM E1050 [2] recomenda

$$\frac{0,01c}{f_l} < s < \frac{0,8c}{2f_u} \quad (2.12)$$

e a ISO 10534-2 [1] recomenda

$$\frac{0,05c}{f_l} < s < \frac{0,9c}{2f_u}, \quad (2.13)$$

em que f_l é a menor frequência de operação do tubo. Na prática, $P(l, \omega)$ e $P(l+s, \omega)$ são estimados através da transformada discreta de Fourier (DFT) dos sinais amostrados de tempo contínuo, $p(l, t)$ e $p(l+s, t)$, ou seja, da DFT dos sinais medidos pelos microfones 1 e 2. Portanto, deve-se ter cuidado para reduzir artefatos indesejáveis devido ao aliasing, tanto no domínio do tempo como no da frequência. Isso inclui a filtragem anti-aliasing passa-baixas antes da amostragem, bem como uma escolha adequada da frequência de amostragem e da duração do sinal.

Não é necessário medir os dois sinais simultaneamente. De fato, um único canal de medição pode ser usado tomando duas medições, alterando a posição do microfone. Neste caso, devem ser obtidas duas FRFs:

$$\begin{cases} H_1(\omega) \equiv P(l+s, \omega)/V(\omega), \\ H_2(\omega) \equiv P(l, \omega)/V(\omega), \end{cases} \quad (2.14)$$

sendo $V(\omega)$ a transformada de Fourier do sinal de tensão $v(t)$. Assim, tem-se que

$$H(\omega) = \frac{H_2(\omega)}{H_1(\omega)}. \quad (2.15)$$

Se forem utilizados dois canais de medição, as diferenças entre as respostas dos dois microfones levam a erros adicionais. Esta incompatibilidade pode ser compensada por um fator de correção derivado de duas medições de calibração de uma amostra altamente absorvente, nos quais as posições dos microfones são intercambiadas. As normas ASTM E1050 [2] e ISO 10534-2 [1] fornecem uma descrição detalhada de tal procedimento.

A propagação da onda sonora em tubos está sujeita à atenuação, que se deve principalmente a perdas viscosas e térmicas dentro da camada limite na parede interna do tubo. Este efeito pode ser levado em conta no domínio da frequência substituindo κ por um número de onda complexo, $\bar{\kappa} = \kappa - i\Xi$, na Eq. (2.10), na qual a constante de atenuação Ξ é [2]

$$\Xi(\omega) = \frac{0,002203}{cd} \sqrt{\left(\frac{\omega}{2\pi}\right)}. \quad (2.16)$$

Embora tais perdas sejam fenômenos tridimensionais, o comportamento geral pode ser modelado no domínio do tempo adicionando um termo de dissipação à Eq. (2.1). No entanto, esse procedimento torna o modelo significativamente mais complexo e aumenta consideravelmente o custo computacional. E como a atenuação do tubo é desprezível se l não exceder cerca de três diâmetros do tubo [2], neste trabalho, as perdas térmicas e viscosas não são consideradas no modelo de domínio do tempo adotado para as simulações realizadas. Para mais detalhes, o leitor pode reportar-se as Refs. [16, 47].

A amostra utilizada para teste não precisa ser um material específico, podendo ser um alto-falante que irá atuar de forma passiva. Assim, quando a onda plana incidir sobre o cone do alto-falante ele irá se mover fazendo com que a bobina acoplada a ele se mova no interior de um imã gerando uma corrente induzida. Logo, parte da energia incidente é absorvida e parte é refletida. Neste caso, o alto-falante absorvedor e o alto-falante que age como uma fonte sonora podem ser modelados como osciladores mecânicos com um grau de liberdade, conforme mostra a Fig. 2.2.

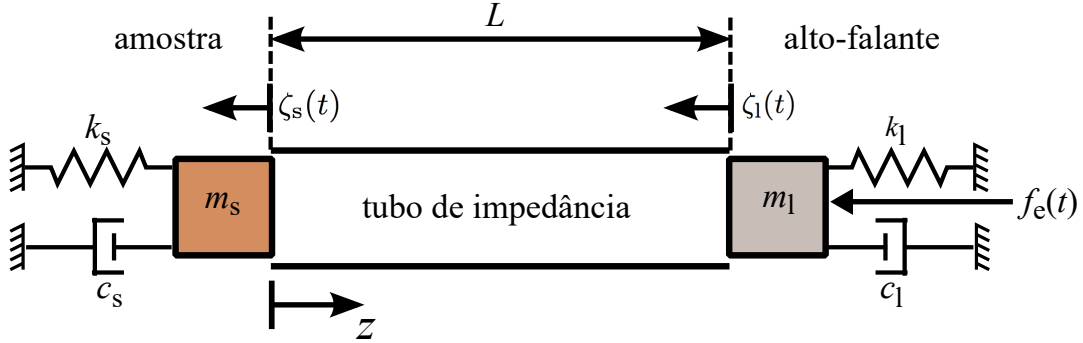


Figura 2.2: Modelo do tubo de impedância.

Na Fig. 2.2, os parâmetros e funções com subscrito l referem-se ao alto-falante excitador e os com subscrito s referem-se ao alto-falante absorvedor. Assim, m_l e m_s são a massa suspensa do alto-falante (membrana e bobina móvel), k_l e k_s são a rigidez da suspensão do alto-falante, c_l e c_s são a constante de amortecimento viscoso e $\zeta_l(t)$ e $\zeta_s(t)$ são o deslocamento resultante do diafragma do alto-falante. A função f_e é a força de origem eletromagnética, devido à aplicação da tensão elétrica nos terminais do alto-falante que age sobre a membrana deste é dada por [43]

$$f_e(t) = \frac{Bl_e}{R_e}[v(t) - Bl_e\dot{\zeta}_l(t)]. \quad (2.17)$$

Na equação acima, $\dot{\zeta}_l$ é a velocidade da massa m_l , B é a densidade de fluxo magnético devido ao ímã permanente do alto-falante, l_e é o comprimento da bobina e R_e é a resistência da bobina.

Considerando essa modelagem a impedância acústica específica do alto-falante absorvedor pode ser calculada por

$$Z_s(\omega) = \frac{[c_s + i(m_s\omega - k_s/\omega)]}{S}, \quad (2.18)$$

sendo S a área da seção transversal do tubo.

Um dos problemas do tubo de impedância é a presença de ruído e para minimizá-lo, este trabalho propõe a utilização de uma técnica de *denoising* via wavelets, que é uma ferramenta matemática que pode ser aplicada em processamento de sinais. Os

fundamentos de wavelets, bem como toda a descrição técnica necessária para remoção de ruído utilizando esta ferramenta estão descritos na Seção 2.2.

2.1.1 Sinal de teste

Há uma série de sinais de teste, $v(t)$, que podem ser usados para obter uma estimativa adequada de $H(\omega)$. Eles estão bem cobertos na literatura (ver, por exemplo, [23, 35]). Um exemplo de sinal de teste é a varredura senoidal cujas frequências instantâneas variam no tempo, sendo que essa variação pode ser linear ou exponencial.

A varredura linear possui espectro branco e a energia do sinal é distribuída no tempo de acordo com a fase do espectro, que por sua vez é proporcional ao atraso de grupo. Sua frequência aumenta a uma taxa constante por unidade de tempo. Já na varredura exponencial, a taxa de variação da frequência é exponencial, e assim apresenta melhor relação sinal/ruído nas baixas frequências e seu espectro é rosa [23].

A varredura senoidal é descrita matematicamente por,

$$v(t) = \hat{v}\text{sen}(f(t)), \quad (2.19)$$

em que \hat{v} é a amplitude da tensão elétrica nos terminais do alto-falante.

A frequência instantânea é dada pela derivada no tempo do argumento da função seno. Então, se $f(t) = \omega t$, sendo ω constante, a frequência instantânea também será constante. Porém, se a frequência varia linearmente a partir de uma frequência angular inicial, ω_1 (em $t = 0$) com término em ω_2 (em $t = T$) durante um intervalo de tempo T , obtém-se:

$$\omega(t) = \frac{d(f(t))}{dt} = \omega_1 + \left(\frac{\omega_2 - \omega_1}{T}\right)t. \quad (2.20)$$

Integrando esta Eq. (2.20), obtém-se

$$f(t) = \omega_1 t + \left(\frac{\omega_2 - \omega_1}{T}\right)\frac{t^2}{2} + \gamma, \quad (2.21)$$

em que $\gamma = 0$ para que $f(0) = 0$.

Considerando esse mesmo procedimento para a uma varredura exponencial, com início em ω_1 e término em ω_2 durante um intervalo de tempo T , tem-se

$$\omega(t) = \frac{d(f(t))}{dt} = \omega_1 e^{\ln(\omega_2/\omega_1)t/T}. \quad (2.22)$$

Integrando essa equação, chega-se a

$$f(t) = \frac{\omega_1 T}{\ln\left(\frac{\omega_2}{\omega_1}\right)} e^{\ln(\omega_2/\omega_1)t/T} + \gamma, \quad (2.23)$$

em que

$$\gamma = \frac{-\omega_1 T}{\ln\left(\frac{\omega_2}{\omega_1}\right)} \quad \text{para } f(0) = 0. \quad (2.24)$$

Finalmente, a partir das Eqs. 2.19 e 2.23, tem-se [23]

$$v(t) = \hat{v} \text{sen} \left(\frac{T\omega_1}{\ln\left(\frac{\omega_2}{\omega_1}\right)} \left(e^{\frac{t}{T} \ln\left(\frac{\omega_2}{\omega_1}\right)} - 1 \right) \right). \quad (2.25)$$

Uma varredura exponencial está ilustrada na Fig. 2.3 com $\omega_1 = 2\pi 100$ rad/s, $\omega_2 = 2\pi 1000$ rad/s, $T = 0,5$ s, frequência de amostragem, f_s , igual a 40 kHz.

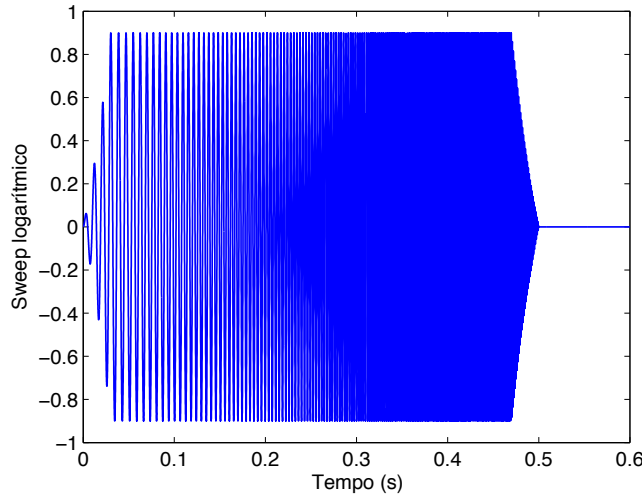


Figura 2.3: Varredura exponencial com $\omega_1 = 2\pi 100$ rad/s, $\omega_2 = 2\pi 1000$ rad/s e $T = 0,5$ s.

Vale ressaltar que o início e término abruptos da varredura introduzem transientes que podem degradar os resultados de medições em torno das frequências ω_1 e ω_2 . Estes efeitos indesejáveis podem ser reduzidos suavizando as extremidades do sinal. Assim, para obtenção da varredura e seu espectro mostrados, respectivamente, nas Figs. 2.3 e 2.4, um crescimento exponencial de amplitude (de 0 até \hat{v}) durante 0,03 s foi utilizado no início do sinal, bem como um decaimento exponencial de duração 0,03 s no final do sinal. Este procedimento também reduz as oscilações do espectro dessa varredura próximo às frequências ω_1 e ω_2 .

Esse sinal possui as características do ruído rosa (ocorre frequentemente na natureza), ou seja, o espectro de frequências bem como a densidade espectral de potência são inversamente proporcionais à sua frequência. Assim a sua energia possui um decaimento correspondente a 3dB/oitava, Fig. 2.4. Tal característica torna relevante a utilização da varredura exponencial em acústica.

Para uma dada aplicação, ω_1 , ω_2 , T e \hat{v} devem ser cuidadosamente escolhidos a fim de garantir uma elevada relação sinal ruído, do inglês *signal-to-noise ratio* SNR, suficiente em toda a faixa de frequência desejada. A frequência ω_1 deve ser menor do que a frequência mais baixa de interesse, e ω_2 deve ser maior do que a frequência máxima de interesse. Além disso, aumentando a duração da varredura, T , e/ou \hat{v} melhora a SNR. Altos valores de \hat{v} podem saturar o alto-falante e/ou produzir níveis sonoros muito altos no interior do tubo, levando a uma resposta não-linear. Na verdade, uma das vantagens da varredura exponencial em relação a outros sinais de teste é que, ao utilizá-la, é possível separar as componentes linear e não-linear da resposta do sistema somente utilizando processamento de sinais [23, 35].

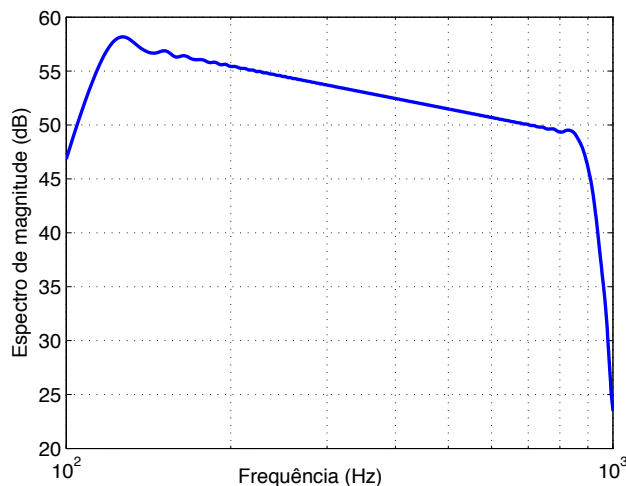


Figura 2.4: Espectro da varredura exponencial com $\omega_1 = 2\pi 100$ rad/s, $\omega_2 = 2\pi 1000$ rad/s e $T = 0.5$ s

2.2 Fundamentos de wavelets

As wavelets são ondas de curta duração localizadas em um determinado instante de tempo e com crescimento e decaimento rápido em um determinado período. Sua denominação vem do inglês, o termo em francês é ondelettes e em português ondinhas, ondículas, ôndulas, ondeletas, onduletas ou ondaleta. No entanto, historicamente o termo original, quando usado, era *wavelets of constant shape*, pois o termo wavelet já

possuía outro significado na comunidade geofísica [13]. As wavelets foram primeiramente aplicadas em geofísica para analisar dados de levantamentos sísmicos, os quais são usados na exploração de petróleo e minerais para obter "imagens" de camadas em rocha subterrânea [7].

Jean Morlet (1983) utilizou produtos de funções Gaussianas transladadas, dilatadas e contraídas para uma análise tempo-frequência de sinais sísmicos na geofísica [13]. Mais tarde, tais funções receberam o nome de wavelet de Morlet. Esse procedimento de análise foi formalizado com os esforços de matemáticos, como Alex Grossman e Yves Meyer, sendo denominado, atualmente, de transformada contínua de wavelets [13].

As transformadas de wavelets, TW, particularmente as que utilizam a wavelet ortonormal e biorthogonal com suporte finito, surgiram como uma nova ferramenta matemática para a multiresolução e decomposição de sinais de tempo contínuo com aplicações potenciais em visão computacional, codificação de sinal, processamento de imagens e de sinais e remoção de ruído. Elas são, geralmente, comparadas à transformada breve de Fourier, porém as TW são melhores. Isso porque na transformada breve de Fourier, independente da frequência, as janelas de tempo possuem a mesma largura, enquanto na TW as janelas no tempo são ajustáveis à frequência.

As funções de base wavelet são geradas por escala e deslocamento de uma wavelet denominada wavelet mãe, e estão bem localizadas nos domínios de tempo e da frequência, permitindo uma análise multiresolução.

2.2.1 Wavelets

Um sinal, $f(t)$, pode ser expresso como uma decomposição linear dada por [12]

$$f(t) = \sum_l a_l \psi_l(t), \quad (2.26)$$

onde l é um índice inteiro, a_l são os coeficientes reais da expansão e ψ_l são um conjunto de funções reais de t chamadas de conjunto de expansão. Se a expansão for única, o conjunto é chamado de "base", e se a base for ortonormal, ou seja, ortogonal,

$$\langle \psi_k(t), \psi_l(t) \rangle = \int \psi_k \psi_l(t) dt = 0, \quad k \neq l, \quad (2.27)$$

e normalizada,

$$\langle \psi_k(t), \psi_l(t) \rangle = \int \psi_k \psi_l(t) dt = 1, \quad k = l, \quad (2.28)$$

então os coeficientes podem ser calculados pelo produto interno,

$$a_l = \langle f(t), \psi_l(t) \rangle = \int f(t) \psi_l(t) dt. \quad (2.29)$$

Para decompor um sinal usando como base uma função wavelet, um sistema dependente de dois parâmetros é construído tal que a Eq. (2.26) torna-se [12]

$$f(t) = \sum_k \sum_j a_{j,k} \psi_{j,k}(t), \quad (2.30)$$

onde j e k são índices inteiros do somatório e $\psi_{j,k}(t)$ são funções de expansão wavelets. O conjunto dos coeficientes dessa expansão é chamado de “Transformada Discreta de Wavelets” (TDW) de $f(t)$ e a Eq. (2.30) é a transformada inversa [12]. A TDW será discutida com mais detalhes na Seção 2.2.4.

Definição: A função wavelet, então, $\psi(t) \in L^2(\mathbb{R})$ é chamada wavelet ortonormal se ela pode ser usada para definir uma base de Hilbert, que é um sistema ortonormal completo, para o espaço de Hilbert de funções integráveis quadradas, $\in L^2(\mathbb{R})$, (funções com energia finita). Ela possui média zero,

$$\int_{-\infty}^{\infty} \psi(t) dt = 0. \quad (2.31)$$

A função é normalizada $\|\psi(t)\| = 1$ e centrada próxima de $t = 0$ [34]. E para que uma função seja definida como tal, é preciso satisfazer os seguintes itens [34]:

- **Teorema 1: Calderón, Grossmann and Morlet (Condição de admissibilidade):** Seja $\psi \in L^2(\mathbb{R})$ uma função real,

$$C_\psi = 2\pi \int_0^{+\infty} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < +\infty, \quad (2.32)$$

em que C_ψ é uma constante que depende da wavelet escolhida, sendo denominada de condição de admissibilidade e $\Psi(\omega)$ representa a transformada de Fourier da wavelet mãe $\psi(t)$. Para garantir que essa integral seja finita é preciso que $\Psi(0) = 0$, nível DC seja zero, justificando porque a função wavelet possui média zero, Eq. (2.31). Caso essa condição seja verdadeira e $\Psi(\omega)$ seja continuamente diferenciável,

$$\int_{-\infty}^{\infty} (1 + |t|) |\psi(t)| dt < +\infty, \quad (2.33)$$

a condição de admissibilidade é satisfeita, garantindo que a transformada wavelet tenha inversa.

- **Condição de energia unitária:** A função wavelet deve ter energia unitária,

$$\int_{-\infty}^{\infty} |\psi(t)|^2 dt = 1. \quad (2.34)$$

Isso garante que a função wavelet possua suporte compacto, com um decaimento rápido de amplitude (do inglês, *e-folding time*), conhecido como suporte efetivo, garantindo a localização espacial.

O Teorema 1 foi provado, não exatamente aplicado ao conceito atual de wavelet (wavelets não apareceram como tal), pela primeira vez em 1964 pelo matemático Calderón. Porém, sem conhecer os trabalhos de Calderón, Grossmann e Morlet provaram a mesma fórmula para processamento de sinais [34].

A expansão wavelets,

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) dt, \quad (2.35)$$

é um conjunto de funções wavelets bases em tempo contínuo gerados a partir de escala e deslocamento de uma wavelet denominada de wavelet mãe, $\psi(t)$, no qual a e $b \in \mathbb{R}$ são variáveis contínuas. O fator de normalização, $|a|^{-1/2}$ é incluído na Eq. (2.35) para que $\|\psi_{a,b}\| = \|\psi\|$ [24]. Além disso, a é um parâmetro de escala, com efeito de dilatação quando for maior do que 1 e de contração quando $0 < a < 1$ e b é um parâmetro de deslocamento ou translação no tempo. A compressão de um sinal no domínio do tempo causa uma expansão do espectro desse sinal no domínio das frequências, simultaneamente, causando um deslocamento para a região de altas frequências. Já a expansão no tempo, comprime o espectro e o desloca para a região de baixas frequências. Essa compressão e expansão permite varrer o espectro de frequências a partir da wavelet mãe. Próximo as baixas frequências o espectro da wavelet torna-se mais comprimido, resultando em um aumento da resolução em baixas frequências. Já em altas frequências o espectro é expandido, resultando em diminuição da resolução.

Considerando que $\psi(t)$ é centrada no tempo zero e na frequência ω_0 , então a sua versão dilatada e translada, Eq. (2.35), é centrada no ponto b e na frequência ω_0/a . Assim, as características da função são refletidas para novos pontos, $(t, \omega_0) = (b, \omega_0/a)$. Como o parâmetro a sofre variações,

$$\psi_{a,0}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t}{a}\right) dt, \quad (2.36)$$

a função englobará várias frequências e assim valores elevados do parâmetro de escala, a , correspondem a frequências baixas ou a uma escala grande, enquanto que os valores pequenos de a representam as altas frequências ou uma escala fina. E portanto, a alteração no parâmetro b permite deslocar o centro de localização do tempo e assim cada $\psi_{a,b}(t)$ está localizada em torno de $t = b$ [17].

Uma característica importante nas wavelets é o número de momentos de esvaecimento (*vanishing*). E uma função ψ tem m_v desses momentos se

$$\int_{-\infty}^{\infty} t^p \psi(t) dt = 0, \quad \text{para } 0 \leq p < m_v. \quad (2.37)$$

Isso significa que ψ é ortogonal para qualquer polinômio de grau $m_v - 1$. Além disso, esse número pode ajudar na escolha da wavelet utilizada para processar o sinal, pois se f tem poucas singularidades, ou seja, pontos descontínuos isolados e é muito regular entre as singularidades, deve-se escolher uma wavelet com muitos momentos de vanishing para produzir um grande número de pequenos coeficientes wavelets, $\langle f, \psi_{j,k} \rangle$. Se a densidade das singularidades aumenta, pode ser melhor diminuir o tamanho de seu suporte ao custo de reduzir o número desses momentos. De fato, as wavelets que se sobrepõem às singularidades criam coeficientes de alta amplitude [34]. Assim, os momentos de *vanishing* são importantes para definir as wavelets utilizadas para detectar singularidades em sinais, sendo que esse método pode ser usado, por exemplo, para detectar rachaduras em um material.

2.2.2 Famílias de wavelets

Existem várias famílias de wavelets, tais como Daubechies, Coifelets, Symlets, Haar e outras. Essas famílias são subdivididas em classes distinguidas pelo número de coeficientes e pelo nível de interação. Na maioria das vezes, elas são classificadas também pelo número de momentos nulos. Uma descrição mais aprofundada sobre esses momentos pode ser encontrado em [12, 34]. Valores elevados de coeficientes correspondem a um grau maior de suavidade e também permitem uma representação mais exata de polinômios de ordem mais elevada [27]. As famílias de wavelets possuem uma função wavelet denominada wavelet mãe, $\psi(t)$, e uma função escala, $\phi(t)$, denominada de wavelet pai.

Após a introdução dos conceitos e definições sobre o que é uma função wavelet, a seguir são apresentadas algumas famílias de wavelets para que o leitor possa compreender melhor esta ferramenta matemática.

2.2.2.1 Wavelets de Haar

A wavelet de Haar, Fig. 2.5 à esquerda, é a mais simples possível das wavelets e foi criada por Haar em 1910, sendo

$$\psi_{\text{Haar}}(t) = \begin{cases} 1, & \text{se } 0 \leq t < 1/2 \\ -1, & 1/2 \leq t < 1 \\ 0, & \text{caso contrário} \end{cases} \quad (2.38)$$

Enquanto que a função escala de Haar, Fig. 2.5 à direita, é definida por

$$\phi_{\text{Haar}}(t) = \begin{cases} 1, & \text{se } 0 \leq t < 1 \\ 0, & \text{caso contrário} \end{cases} \quad (2.39)$$

Logo, a wavelet de Haar é a função [7].

$$\psi_{\text{Haar}}(t) = \phi(2t) - \phi(2t - 1). \quad (2.40)$$

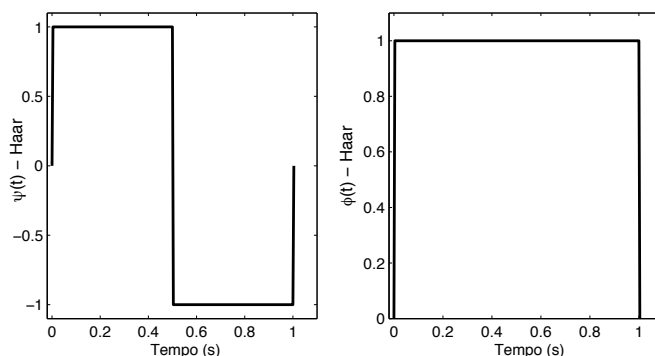


Figura 2.5: Wavelet de Haar (à esquerda) e escala de Haar (à direita).

2.2.2.2 Wavelets de Morlet

A wavelet de Morlet, Fig. 2.6 à direita, é uma exponencial modulada por uma gaussiana, sendo da seguinte forma [48]

$$\psi_{\text{Morlet}}(t) = \pi^{-1/4} e^{i\omega_0 t} e^{-0.5t^2}, \quad (2.41)$$

em que $i = \sqrt{-1}$, ω_0 é uma frequência central que determina a forma do sinal modulador da equação, t é o vetor tempo. Geralmente, ω_0 é escolhido de modo que o segundo

máximo de $\Re[\psi(t)]$, $t > 0$, seja metade do primeiro em $t = 0$. Logo [48],

$$\omega_0 = \pi \sqrt{\frac{2}{\ln 2}} = 5,536. \quad (2.42)$$

Nota-se que essa wavelet não satisfaz a condição de admissibilidade, pois $\Psi(\omega) \neq 0$ quando $\omega = 0$, mas esse valor é desprezível ($\approx 7 \cdot 10^{-7}$), e assim não há qualquer problema na prática. Há um fator para corrigir isso, mas é um valor muito pequeno. A wavelet de Morlet é aproximadamente um filtro de banda de oitava.

A Fig. 2.6 à esquerda representa a construção da wavelet de Morlet, sendo a curva em azul a parte imaginária da exponencial com frequência de $\omega = 6$ rad/s, a curva verde o envelope gaussiano, e por fim a parte imaginária da wavelet de Molet em vermelho, com $\omega_0 = 6$ rad/s. Nota-se que o ponto máximo da wavelet possui amplitude menor do que a curva azul para o mesmo ponto devido ao fator $\pi^{-1/4}$ na Eq. 2.41.

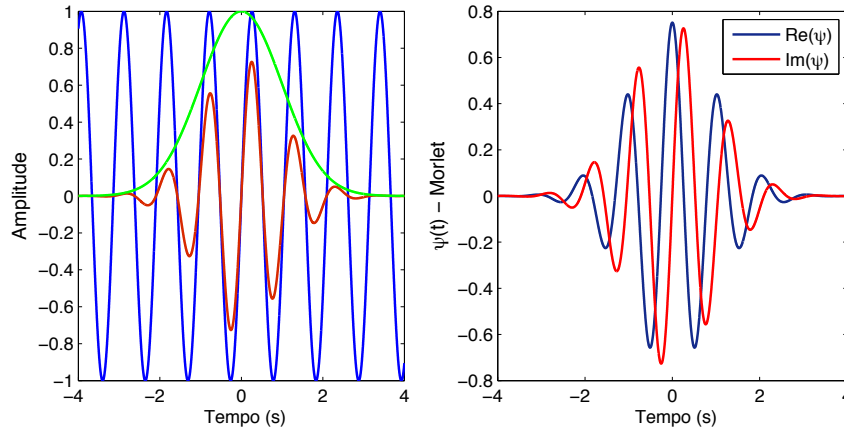


Figura 2.6: Construção da wavelet de Morlet (à esquerda) e parte real e imaginária da wavelet de Morlet (à direita).

2.2.2.3 Wavelets Chapéu Mexicano

Outra wavelet típica é conhecida como chapéu mexicano (do inglês *Mexican hat wavelet*), pois sua forma é semelhante a corte transversal do chapéu mexicano. Ela é a segunda derivada de uma gaussiana, e normalizada é descrita por [34]

$$\psi_{\text{Mexh}}(t) = \frac{2}{\pi^{1/4} \sqrt{3\sigma}} \left(\frac{t^2}{\sigma^2} - 1 \right) e^{-t^2/2}, \quad (2.43)$$

sendo bem localizada no tempo e na frequência e satisfazendo a condição descrita pela Eq. (2.32).

A Fig. 2.7 ilustra $-\psi_{\text{Mexh}}(t)$ com $\sigma = 1$.

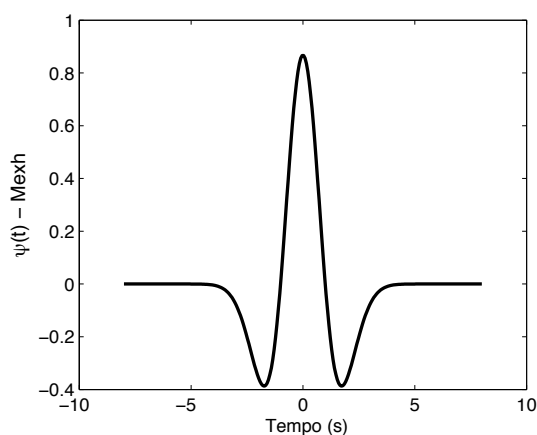


Figura 2.7: Wavelet chapéu mexicano.

2.2.2.4 Wavelets de Daubechies

As wavelets de Daubechies têm um suporte de tamanho mínimo para qualquer número de momentos de *vanishing* [34]. As wavelets de Daubechies são classificadas de acordo com o número de momentos de esvaecimento que elas possuem, e a suavidade de suas funções escala e wavelets aumenta com o aumento do número desses momentos. Para $m_v = 1$, as funções escalas e wavelet são iguais às de Haar, sendo descontínuas. Já para $m_v = 2$ as funções escala e wavelet são contínuas, mas não possuem derivadas, enquanto que $m_v = 3$ tanto função escala quanto wavelets são diferenciáveis [7].

A Fig. 2.8, à direita representa as wavelets db2, db4 e db8 e a à esquerda representa as funções escala correspondentes a essas funções.

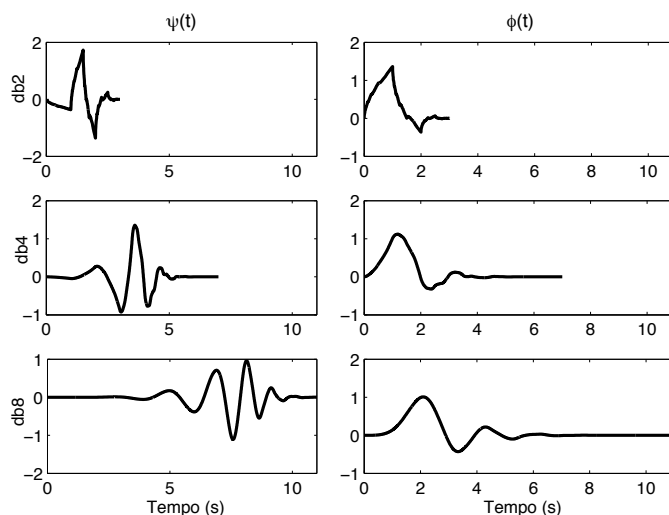


Figura 2.8: Wavelets Daubechies (à esquerda) e funções escala Daubechies (à direita).

2.2.3 Transformada contínua de wavelets

A transformada contínua de wavelets (*Continuous Wavelet Transform*), TCW, é uma ferramenta que decompõe o sinal a ser estudado através de um conjunto de funções passa-banda constante Q (ou largura de banda igual em uma escala logarítmica), e depois analisa cada componente com uma resolução correspondente à sua escala. Ela fornece uma descrição tempo-frequência, sendo descrita por [17]

$$W_{a,b}\{f(t)\} = \int_{-\infty}^{+\infty} f(t)\psi_{a,b}(t)dt. \quad (2.44)$$

Inserindo nessa equação as operações de compressão/expansão e deslocamento no atempo, Eq. (2.35), tem-se que

$$W_{a,b}\{f(t)\} = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} f(t)\psi\left(\frac{t-b}{a}\right)dt. \quad (2.45)$$

Reforçando a definição da função de expansão wavelet, $\psi(t)$ é uma wavelet mãe, a e $b \in \mathbb{R}$, são respectivamente, parâmetro de escala e de deslocamento. Assim, a Eq. (2.44) mede, de certa forma, as flutuações do sinal $f(t)$ em torno do ponto b , na escala dada por $a > 0$. E ψ é implicitamente assumida como real, mesmo que isso não seja de modo algum essencial; mas se ela não for, então o conjugado complexo têm de ser introduzido na Eq. (2.44). É importante ressaltar que a transformada wavelet ortonormal também preserva a energia entre as diferentes escalas, pois $\psi_{a,b}$ é normalizada. Já na reconstrução, a normalização se dá por $1/a^2$.

A reconstrução do sinal, $f(t)$, ou a possibilidade de decompor o sinal, ou seja, de representar o sinal como uma superposição de wavelets é obtida através da transformada inversa, conhecendo a transformada wavelet do sinal, $W_{a,b}\{f(t)\}$. Mas o processo é inversível se e somente se a condição de admissibilidade, C_ψ , Eq. (2.32), é verificada. Sendo assim, a reconstrução do sinal ocorre pela seguinte equação [?]

$$f(t) = \underbrace{\frac{1}{C_\psi} \int_{-\infty}^{+\infty} \int_0^{+\infty} \frac{1}{a^2} da db}_{\text{somatório}} \underbrace{W_{a,b}\{f(t)\}}_{\text{Coeficientes Wavelets}} \underbrace{\psi_{a,b}}_{\text{wavelets}}. \quad (2.46)$$

Existem variações possíveis nessa equação de reconstrução se restrições são impostas sobre a wavelet. Além disso, é importante lembrar que esse trabalho utiliza somente wavelets ortogonais e, portanto, a família de wavelets, $\psi_{a,b}$, para a síntese do sinal é a mesma família utilizada para a análise do sinal. No entanto, se a wavelet fosse biortogonal, haveria duas wavelets diferentes, uma para análise e outra para síntese.

Os gráficos de análises de sinal usando a transformada contínua de wavelets são

frequentemente chamados de escalogramas, em contraste com os espectrogramas que são fornecidos pela transformada breve de Fourier.

Devido a variação contínua dos valores de a e b , essa transformada é impraticável e redundante, uma vez que trata-se de uma expansão bidimensional de uma função unidimensional (transformada é sobreamostrada). Então, pode-se amostrar esses parâmetros e obter um conjunto de funções wavelets em parâmetros discretos, ou seja, obter a transformada wavelet discreta apresentada a seguir.

2.2.4 Transformada discreta de wavelets

Variando os parâmetros a e b da TCW, Eq. (2.45), de forma discreta tal que

$$a = a_0^m \quad \text{e} \quad b = nb_0 a_0^m, \quad m, n \in \mathbb{Z}, \quad (2.47)$$

e $a_0 > 1$, $b_0 > 0$ fixo, obtém-se um conjunto discreto de tempo contínuo [4]

$$\psi_{m,n}(t) = a_0^{-m/2} \psi(2^{-m}t - nb_0). \quad (2.48)$$

Valores diferentes de m correspondem a wavelets de larguras diferentes, e o parâmetro de deslocamento b torna-se dependente da escala.

A transformada discreta de wavelets, TDW, ou como também é denominada, equação de análise é [17]

$$W_{m,n}\{f(t)\} = \frac{1}{\sqrt{a_0^m}} \int_{-\infty}^{+\infty} \psi(a_0^{-m}t - nb_0) dt. \quad (2.49)$$

As wavelets estreitas (de alta frequência), comprimidas, são deslocadas por intervalos pequenos e wavelets largas (baixa frequência), expandidas, são deslocadas por intervalos maiores cobrindo todo o sinal [17, 12].

Na TDW assim como na TCW assume-se que a condição de admissibilidade seja satisfeita e portanto, a Eq. (2.31) seja válida.

Considerando um valor especial de a_0 , b_0 e uma determinada família de wavelets, ψ , em particular $a_0 = 2$ e $b_0 = 1$, haverá ψ com boas propriedades de localização tempo-frequência, tal que [17]

$$\psi_{m,n}(t) = 2^{-m/2} \psi(2^{-m}t - n). \quad (2.50)$$

constitua uma base ortonormal para $L^2(\mathbb{R})$. Essa discretização é conhecida como grade diádica.

Conhecendo os valores da TDW, se $\psi_{m,n}$ é completo em $L^2(\mathbb{Z})$, para todos os valores de $\psi(t)$, a , b , qualquer sinal de tempo contínuo $f(t) \in L^2(\mathbb{Z})$ pode ser representado

como uma superposição [36]

$$f(t) = \sum_m \sum_n d_{m,n} \psi_{m,n}(t), \quad (2.51)$$

no qual $d_{m,n}$ são os coeficientes de decomposição, Eq. 2.49.

Deve-se ressaltar que a ortonormalidade significa que o produto interno entre a wavelet mãe e suas versões deslocadas e expandidas/comprimidas sempre será zero, exceto quando as duas wavelets estiverem exatamente sob a mesma escala e deslocadas do mesmo montante no tempo, ou seja, somente quando as duas funções forem exatamente iguais. Ela permite alterar os coeficientes $d_{m,n}$ sem que se perca a capacidade de reconstrução do sinal e sem que outras partes do sinal sejam afetadas pela alteração.

Nota-se que expansão wavelet precisa de um número infinito de escalas ou resoluções m . Mas isso não é praticável, e se então, a expansão é conhecida apenas para certos valores de escalas $m < M$, é necessário um componente para completar as informações de expansão para $m > M$. Isto é feito através da introdução de uma função escala, $\phi(t)$ tal que, $\forall m \in \mathbb{R}$,

$$\phi_{j,k}(t) = 2^{-j/2} \phi(2^{-j}t - k), \quad (2.52)$$

é uma base ortonormal para o subespaço V_m de $L^2(\mathbb{R})$. Assim, ao introduzir essa componente escala o sinal $f(t)$ pode ser representado como um limite de sucessivas aproximações em diferentes resoluções [36].

2.2.5 Algoritmo de Mallat

A expansão em série wavelet para um sinal $f(t)$ pode ser descrita por [12]

$$f(t) = \sum_{k=-\infty}^{\infty} c_{j,k} 2^{j/2} \phi(2^j t - k) + \sum_{k=-\infty}^{\infty} \sum_{j=0}^{\infty} d_{j,k} 2^{j/2} \psi(2^j t - k), \quad (2.53)$$

no qual $2^{j/2}$ é um termo que matém a norma unitária das funções bases nas várias escalas utilizadas, c_k são os coeficientes de escala ou aproximação, $d_{j,k}$ são os coeficientes wavelets ou de detalhes, ϕ é a função escala e ψ é a função wavelet. As funções de escala representam um filtro passa-baixas natural e as funções wavelets são um filtro passa-banda natural [49].

Utilizando wavelets e funções escalas ortonormais, os coeficientes dessa expansão podem ser obtidos pelo produto interno entre o sinal a ser analisado e tais funções.

Sendo calculados por

$$c_{j,k} = \int f(t)\phi_{j,k}(t)dt, \quad (2.54)$$

e

$$d_{j,k} = \int f(t)\psi_{j,k}(t)dt. \quad (2.55)$$

Porém, caso a wavelet escolhida seja biortogonal, é preciso substituir as funções ϕ e ψ por suas duais, $\tilde{\phi}$ e $\tilde{\psi}$ respectivamente. Assim, as funções ϕ e ψ são a função escala e wavelet de sínteses, enquanto suas duais são a função escala e wavelet de análises. Mas nesse trabalho são consideradas apenas as wavelets ortogonais, portanto, as funções de análise e síntese são iguais, isto é, as funções wavelets e escalas utilizadas para calcular a TDW são as mesmas funções utilizadas para reconstruir o sinal.

Se as funções de escala satisfazem as assim chamadas equações de dilatação [49],

$$\phi(t) = \sum_n 2h[n]\phi(2t - n) \quad e \quad (2.56)$$

$$\tilde{\phi}(t) = \sum_n 2\tilde{h}[n]\tilde{\phi}(2t - n), \quad (2.57)$$

e as wavelets satisfazem as chamadas equações wavelets

$$\psi(t) = \sum_n 2h_1[n]\psi(2t - n) \quad e \quad (2.58)$$

$$\tilde{\psi}(t) = \sum_n 2\tilde{h}_1[n]\tilde{\psi}(2t - n), \quad (2.59)$$

utilizando algumas sequências de filtros, $h[n]$, $\tilde{h}[n]$, $h_1[n]$, $\tilde{h}_1[n]$, então os coeficientes de aproximação e de detalhes podem ser definidos por [12]

$$c_{j,k} = \sum_m h[m - 2k]c_{j+1,m} \quad e \quad d_{j,k} = \sum_m \tilde{h}_1[m - 2k]c_{j+1,m}. \quad (2.60)$$

no qual h representa a função de transferência de um filtro FIR (resposta ao impulso finita) passa-baixas e h_1 de um filtro FIR passa-altas. Assim, os coeficientes na escala 2^j podem ser calculados a partir dos coeficientes de escala, $c_{j,k}$, na escala 2^{j+1} (maior resolução) através de uma filtragem de tempo discreta seguida de uma subamostragem ou dizimação com fator 2 ($\downarrow 2$), Fig. 2.9 à esquerda. Já para o processo de síntese, Fig. 2.9 à direita, os coeficientes de aproximação na escala 2^{j+1} podem ser sintetizados

a partir dos coeficientes de aproximação e detalhes na escala 2^j (menor resolução) através de uma superamostragem ou interpolação por um fator de 2 ($\uparrow 2$) seguida de uma filtragem de tempo discreto, sendo [12]

$$c_{j+1,k} = \sum_m c_{j,m} h[k - 2m] + \sum_m d_{j,m} h_1[k - 2m]. \quad (2.61)$$

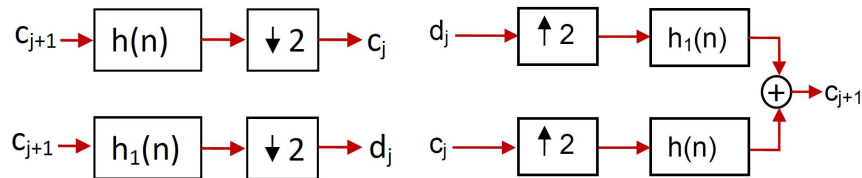


Figura 2.9: Obtenção dos coeficientes de resolução mais baixa a partir dos de alta resolução (à esquerda) e dos coeficientes de resolução mais alta a partir dos de baixa resolução (à direita).

Os processos de análise e síntese são realizados de forma recursiva produzindo uma estrutura piramidal. Logo, esse algoritmo que implementa a transformada de wavelets discreta e sua inversa é conhecido como algoritmo Mallat [49] ou como algoritmo da árvore [12].

Considerando um sinal $f(n)$ como entrada desse algoritmo, Fig. 2.10, o processo será o mesmo descrito acima. O sinal é decomposto em duas faixas de frequências diferentes em cada nível de decomposição, após a filtragem, realiza-se a operação de dizimação ($\downarrow 2$) eliminando todos os coeficientes de índice par da sequência resultante. A sub-amostragem garante que o número total de pontos permaneça constante após a TDW. Essa sequência pode ser repetida por mais níveis, sendo que após o primeiro nível, o sinal de entrada para a filtragem são os coeficientes de aproximação. Para reconstruir o sinal original, Fig. 2.11, tem-se a interpolação ($\uparrow 2$) e em seguida a filtragem.

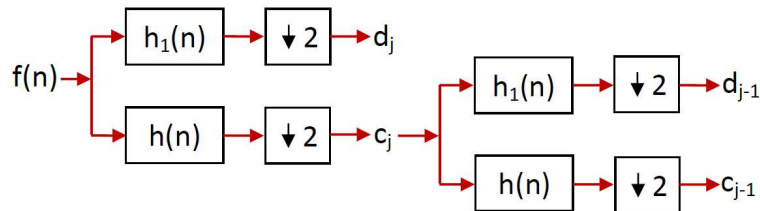


Figura 2.10: Implementação da transformada de wavelets discreta através de um banco de filtros digitais.

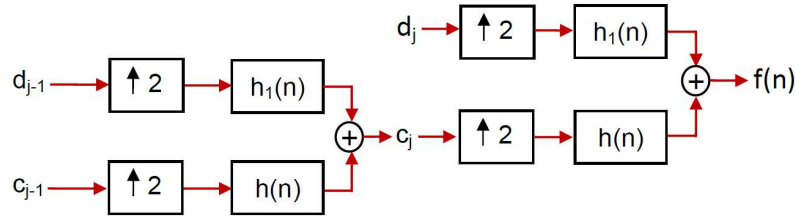


Figura 2.11: Reconstrução do sinal original através do banco de filtros.

2.3 Wavelets na remoção de ruído

A remoção de ruído é realizada aplicando conjuntamente a transformada de wavelets discreta, utilizando o algoritmo da árvore, com a escolha adequada de um limiar (*thresholding*), λ , e o tipo de regra associado a esse limiar. Sendo resumida nas seguintes etapas [21]

1. Analisar o sinal realizando a transformada de wavelets discreta: $W_{m,n}(x(n)) = X$;
2. Determinar qual será o tipo de regra associado ao *thresholding* e o valor do limiar apropriado para cada nível de decomposição e aplicá-los em cada nível: $\hat{X} = T(X, \lambda)$;
3. Realizar a síntese do sinal, aplicando a transformada de wavelets discreta inversa: $x(n) = W^{-1}\hat{X}$.

O *thresholding* inadequado poderá resultar numa significativa perda do sinal original, e os tipos de regras associados a ele são “hard”, T_h , ou “soft”, T_s , [21]. Tais regras são descritas por

$$T_h(X, \lambda) = \begin{cases} X, & |X| \geq \lambda \\ 0, & |X| < \lambda \end{cases} \quad (2.62)$$

e

$$T_s(X, \lambda) = \begin{cases} \text{sgn}[X][|X - \lambda|], & |X| \geq \lambda \\ 0, & |X| < \lambda \end{cases} \quad (2.63)$$

em que sgn representa a função sinal.

Geralmente a regra de *thresholding hard* é referida como *wavelet thresholding*, enquanto que a regra de *thresholding soft* é referida como *wavelet shrinkage*, uma vez que os coeficientes com alta amplitude vão diminuindo em direção a zero.

A obtenção do limiar λ não é simples e o limiar considerado universal é [28]

$$\lambda = \sigma\sqrt{2\ln N}, \quad (2.64)$$

no qual σ representa uma estimativa do desvio padrão do ruído (ou perfil do ruído), e N é o número de amostras do sinal no qual foi realizado a TDW. Outro limiar que também pode ser utilizado é uma percentagem do coeficiente de maior módulo em cada nível.

Esse procedimento pode ser aplicado a diversos sinais acústicos a fim de remover o ruído presente nesses sinais. Por exemplo, sinais medidos por microfones em um tubo de impedância, ou sinais medidos por microfones em um ambiente ruidoso.

2.4 Metodologia

Na simulação descrita a seguir, considerou-se o modelo do tubo de impedância mostrado na Fig. 2.1 modelado por um programa de simulação numérica que utiliza os métodos dos elementos finitos sendo implementado em Matlab pelo professor Alexander Mattioli Pasqual, que gentilmente cedeu o código para a realização deste trabalho. São fornecidos ao programa o sinal de excitação do alto-falante e a impedância característica, $v(t)$ e $Z_s(\omega)$, retornando os sinais que simulam as medições dos microfones 1 e 2, denotados respectivamente por $y_1(t)$ e $y_2(t)$.

Como sinal de excitação $v(t)$, utilizou-se uma varredura exponencial, Fig. 2.12, descrita na Seção 2.1.1, com frequências inicial e final de 30 Hz e 6 kHz, duração de 8 s, amplitude de 0,02 V e frequência de amostragem de 120 kHz, dez vezes superior à taxa de Nyquist (12 kHz), conforme discutirei a seguir. O início e o término da varredura foram suavizados adicionando um crescimento exponencial de amplitude durante 0,1 s ao início do sinal, bem como um decaimento exponencial de duração 0,01 s ao término. Além disso, foi adicionado um intervalo de silêncio de 4 s ao fim da varredura. Isso é necessário, visto que a deconvolução para estimar $H(\omega)$ foi realizada através da transformada de Fourier discreta (DFT), o que geraria erros devido ao *aliasing* no domínio do tempo caso o tamanho dos sinais não fosse aumentado para contemplar a duração da resposta ao impulso do sistema. Porém, antes de estimar $H(\omega)$, os sinais de cada microfone foram filtrados por um filtro passa-baixas e subamostrados por um fator de 6. Esse procedimento é realizado porque o programa de simulação do tubo de impedância realiza uma integração numérica no tempo, requerendo uma frequência de amostragem muito superior à taxa de Nyquist (12 kHz), logo, os sinais y_1 e y_2 são gerados por uma frequência de amostragem igual a 120 kHz, possuindo um número muito elevado de amostras. Assim, realiza-se a dizimação para que os demais

processamentos sejam feitos com um número menor de amostras correspondendo a uma amostragem a 20 kHz.

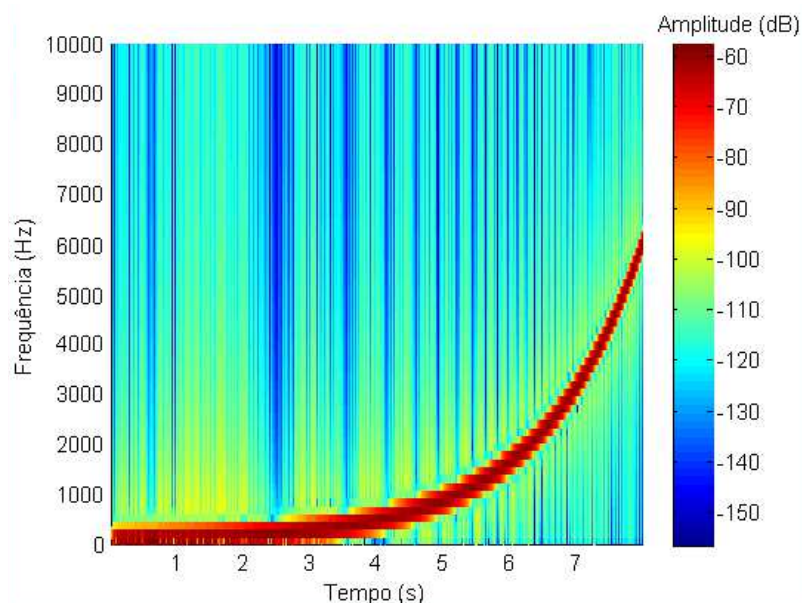


Figura 2.12: Espectrograma representativo da varredura logarítmica com frequências iniciais e finais de 30 Hz e 6 kHz.

Quanto à amostra de teste, os parâmetros utilizados foram:

- $m_s = m_l = 0,1$ g;
- $c_s = c_l = 0,27$ Nm/s;
- $k_s = k_l = 2320,2$ N/m;
- $S = 0,0012$ m² (correspondendo a um diâmetro do tubo de 0,039 m);
- $c = 340$ m/s, $\rho = 1,21$ kg/m³, $Bl_e = 3,17$ Tm;
- $R_e = 6,3$ Ω .

Eles simulam um alto-falante absorvedor, conforme descrito na ISO 10534-2 [1]. Utilizou-se um tubo circular de diâmetro 0,039 m e comprimento $L = 0,32$ m, e adotou-se $l = 0,15$ m e $s = 0,02$ m, Fig. 2.2. Com esses dados, as Eqs. (2.18) e (2.8) permitem calcular $Z_s(\omega)$ e $R(\omega)$, a partir da qual obtém-se a curva de absorção “exata” da amostra, que foi usada para avaliar a qualidade das curvas estimadas via deconvolução.

Os parâmetros referentes aos osciladores correspondem ao alto-falante Aurasound NSW2-326-8A montado em uma caixa de volume $V_b = 10^{-3}$ m³, que é um falante de duas polegadas. Esses dados se encontram no datasheet do alto-falante.

A cada sinal de pressão sonora simulado ($y_1(t)$ e $y_2(t)$) foi acrescentado um ruído branco gaussiano com desvio padrão σ , simulando o ruído de medição. Duas situações foram avaliadas: $\sigma = 0,002$ e $\sigma = 0,006$. A partir desses sinais contaminados, foi estimada a função de resposta em frequência, FRF, entre os microfones 1 e 2 utilizando DFTs, considerando que $y_1(t)$ é o sinal de entrada e $y_2(t)$ é o sinal de saída. Através de $H(\omega)$, calculou-se $\alpha(\omega)$.

A fim de reduzir o ruído adicionado a $y_1(t)$ e $y_2(t)$, aplica-se a TDW para que, em seguida, $H(\omega)$ possa ser estimado e, finalmente o coeficiente de absorção calculado. Assim, a transformada foi aplicada considerando o "algoritmo da árvore", e esse mesmo princípio foi utilizado para reconstruir o sinal. Especificou-se um limiar para comparação dos coeficientes de detalhe, zerando todos os coeficientes com valor inferior a esse limiar. Em seguida, utilizando os coeficientes de aproximação do primeiro nível e todos os coeficientes de detalhe, reconstruiu-se o sinal analisado. Esse procedimento permite reduzir o ruído incorporado nos sinais simulados, tornando as curvas de absorção estimadas mais próximas da "exata". O limiar considerado foi uma porcentagem da maior amplitude dos coeficientes de detalhes de cada nível. Além disso, utilizou-se para cálculo da TDW as famílias Daubechies, coiflet e symlet.

2.5 Resultados

Conforme descrito anteriormente, uma varredura senoidal logarítmica foi fornecida a um programa que simula o tubo de impedância, e os sinais de saída $y_1(t)$ e $y_2(t)$ foram gerados. Tais sinais são apresentados na Fig. 2.13.

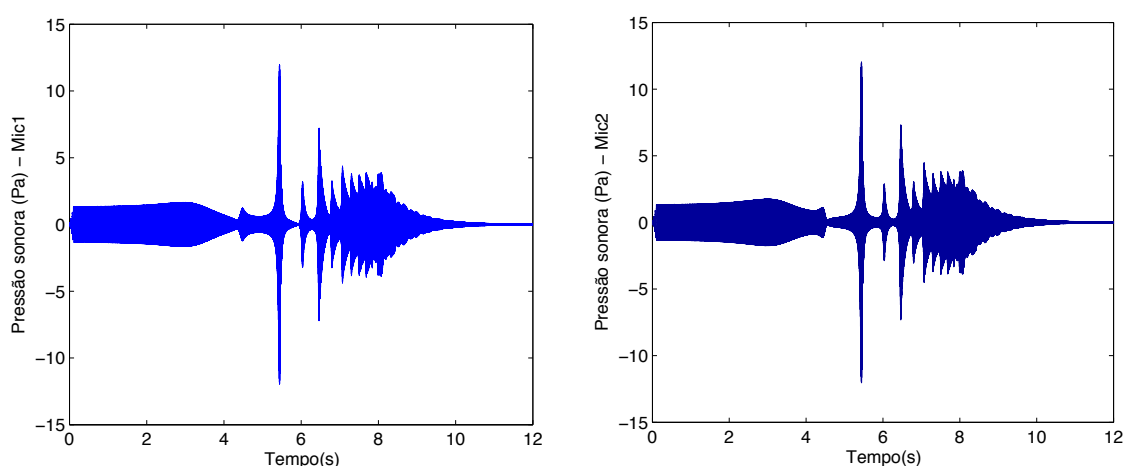


Figura 2.13: Sinais dos microfones 1 (à esquerda) e 2 (à direita).

Apesar dos sinais $y_1(t)$ e $y_2(t)$ parecerem semelhantes, há uma diferença em suas amplitudes conforme mostra a Fig. 2.14.

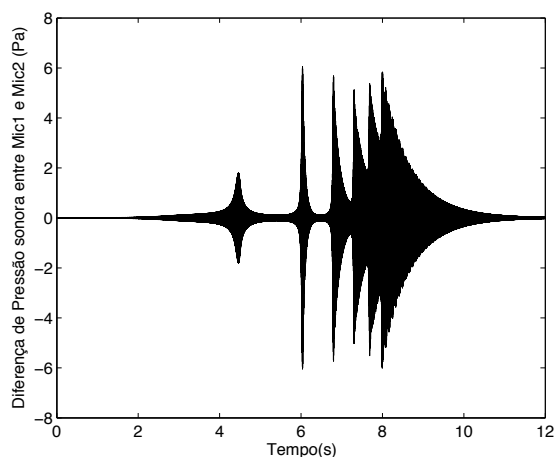


Figura 2.14: Diferença da pressão sonora medida entre os microfones 1 e 2.

Na Fig. 2.15 à esquerda tem-se a curva do sinal ruidoso do microfone 2 em azul, e em vermelho esse mesmo sinal filtrado utilizando a wavelet db45 com 11 níveis e um limiar de 0,0005% do valor máximo dos coeficientes de detalhes em cada nível. Nessa figura não é possível notar diferença entre os sinais original e filtrado, porém, ao aplicar um "zoom" nos instantes de 4,5 a 4,7 segundos tem-se a Fig. 2.15 à direita, na qual tal diferença pode ser observada.

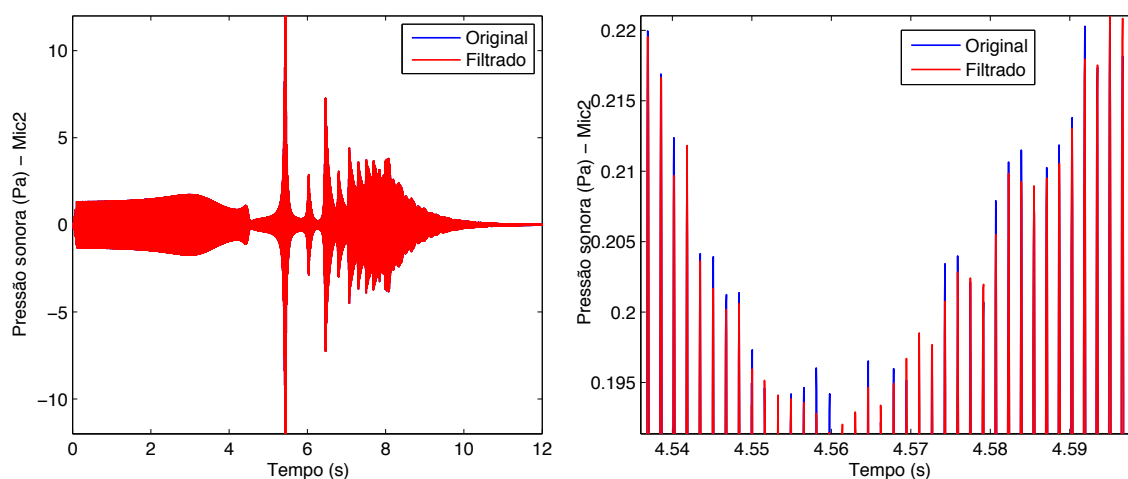


Figura 2.15: Sinais original e filtrado (db45 com 11 níveis, $\sigma = 0,006$, limiar = 0,0005%) do microfone 2 (à esquerda) e um zoom nesses sinais (à direita).

A Fig. 2.16 mostra a curva de absorção exata. E as curvas de absorção sonora estimadas apresentadas a partir desse momento, sempre serão acompanhadas de outra figura que mostra a diferença entre o sinal do microfone 2 com ruído e esse mesmo sinal filtrado pelo processo de *denoising* via wavelets. Essa diferença não é apresentada para

o microfone 1 porque sua forma de onda é semelhante ao do microfone 2, conforme visto na Fig. 2.13.

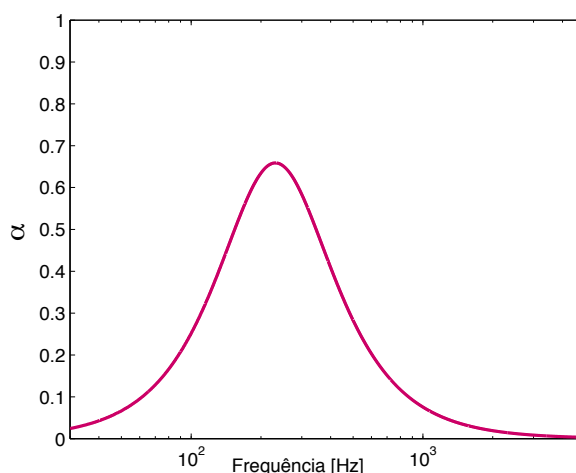


Figura 2.16: Coeficiente de absorção exato.

As Figs. 2.21, 2.22, 2.23, 2.24 e 2.25 foram obtidas considerando um ruído gaussiano branco com desvio padrão (σ) igual a 0,006. A Fig. 2.21 apresenta os resultados obtidos utilizando a wavelet Daubechies 45 (db45), Fig. 2.17, com 11 níveis e um limiar de 0,0005% do valor máximo do coeficiente de detalhe em cada nível. À esquerda, apresentam-se as curvas de absorção estimadas com os sinais originais (sem remoção de ruído) e com os sinais filtrados. Observa-se que a remoção do ruído foi significativa para frequências abaixo e em torno de 100 Hz. À direita, apresenta-se a diferença entre os sinais com e sem filtragem, onde nota-se que esta é próxima ao valor do desvio padrão do ruído. Já a Fig. 2.22 utiliza essa mesma wavelet (db45 com 11 níveis), porém com um limiar de 0,001%. Quanto à Fig. 2.23, a wavelet aplicada foi a db20 com 13 níveis, Fig. 2.18, e um limiar de 0,0005%. O resultado foi satisfatório em baixas frequências. Porém, em altas frequências, a filtragem via wavelets não conduziu a uma melhora na estimativa das curvas de absorção.

A Fig. 2.24 refere-se à Symlet 20 (sym20), Fig. 2.20, com 10 níveis e um limiar de 0,0005%. O resultado foi satisfatório em baixas frequências. E a Fig. 2.25 refere-se à Coiflet 5 (coif5), Fig. 2.19, com 9 níveis e um limiar de 0,0005%, que apresentou uma pequena melhora em frequências em torno de 100 Hz.

Por fim, considerando um ruído com menor desvio padrão ($\sigma = 0,002$), tem-se a Fig. 2.26 com os resultados obtidos utilizando a db45 com 11 níveis e um limiar de 0,0005%. O resultado final foi satisfatório, mostrando-se altamente eficiente abaixo de 100 Hz.

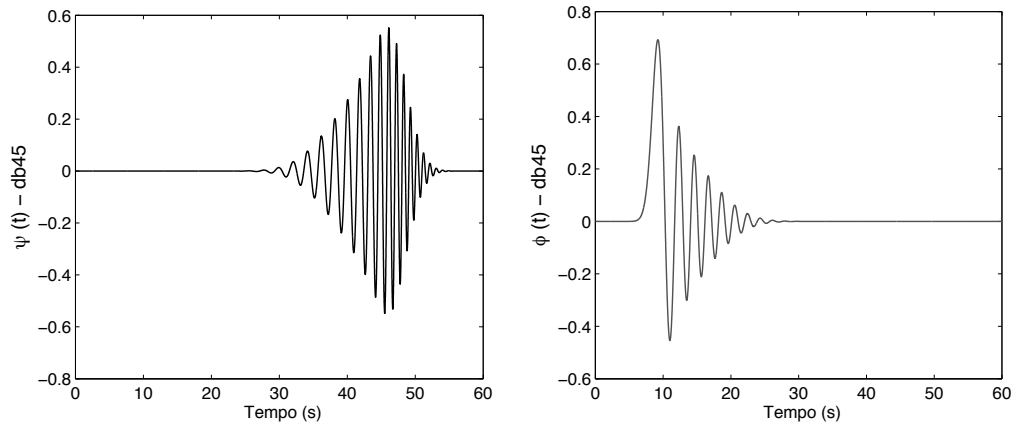


Figura 2.17: Função wavelet db45 (à esquerda) e função escala db45 (à direita).

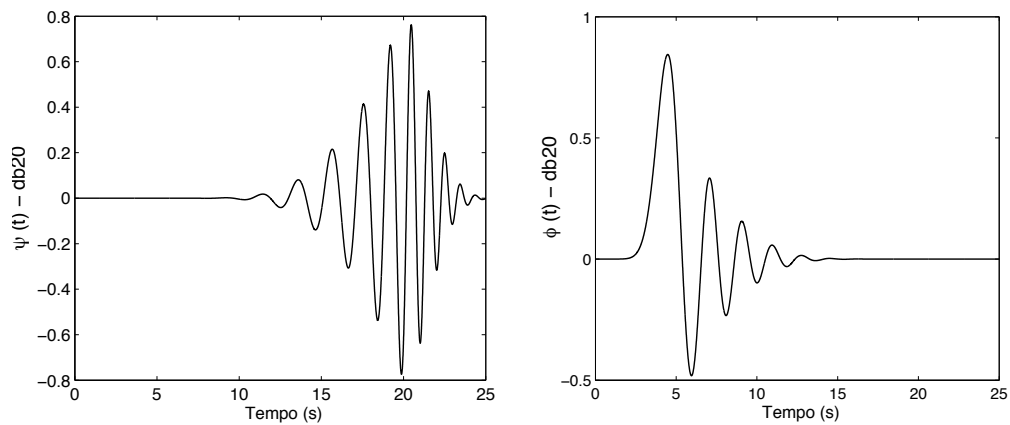


Figura 2.18: Função wavelet db20(à esquerda) e função escala db20 (à direita).

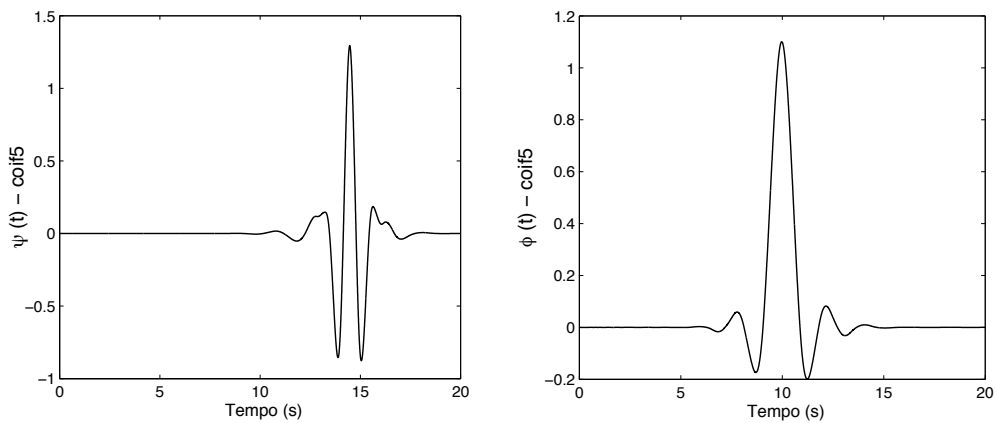


Figura 2.19: Função wavelet coif5 (à esquerda) e função escala coif5 (à direita).

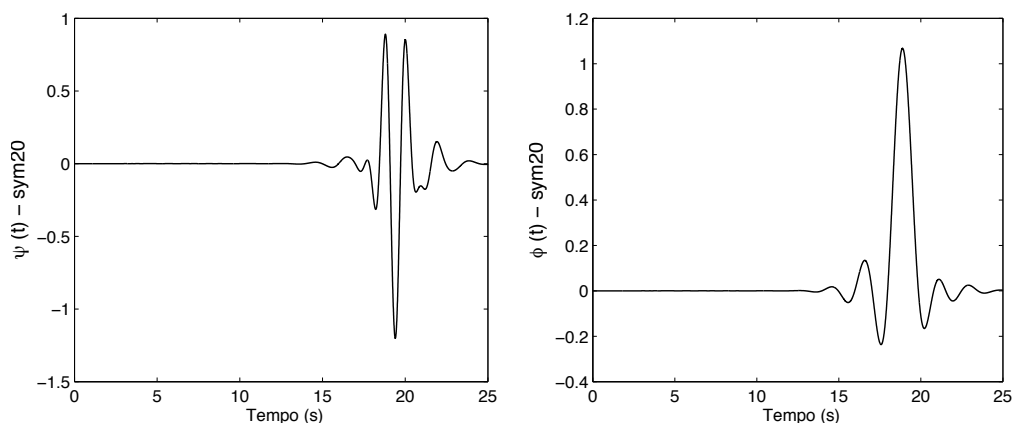


Figura 2.20: Função wavelet sym20 (à esquerda) e função escala sym20 (à direita).

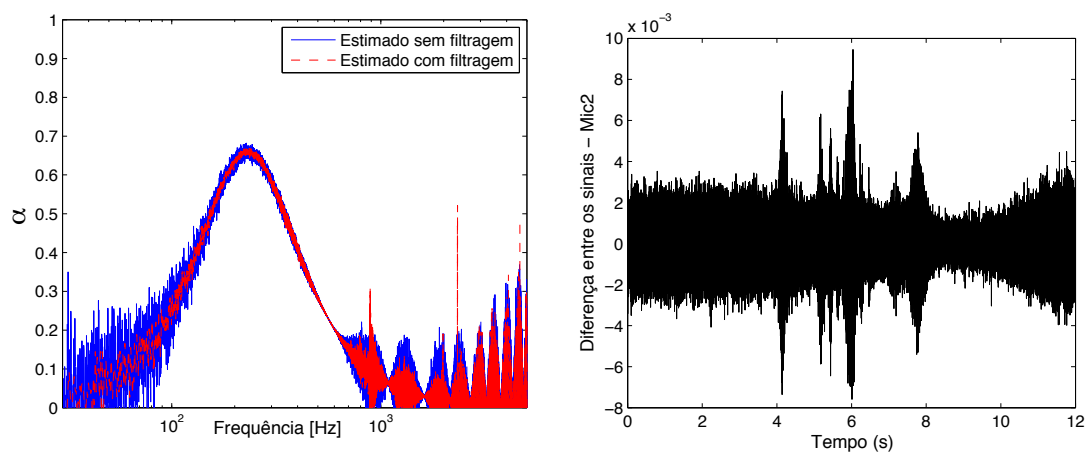


Figura 2.21: Curvas de absorção sonora estimadas (à esquerda) e diferença entre os sinais original e filtrado no microfone 2 (à direita) para db45 com 11 níveis, $\sigma = 0,006$, limiar = 0,0005%.

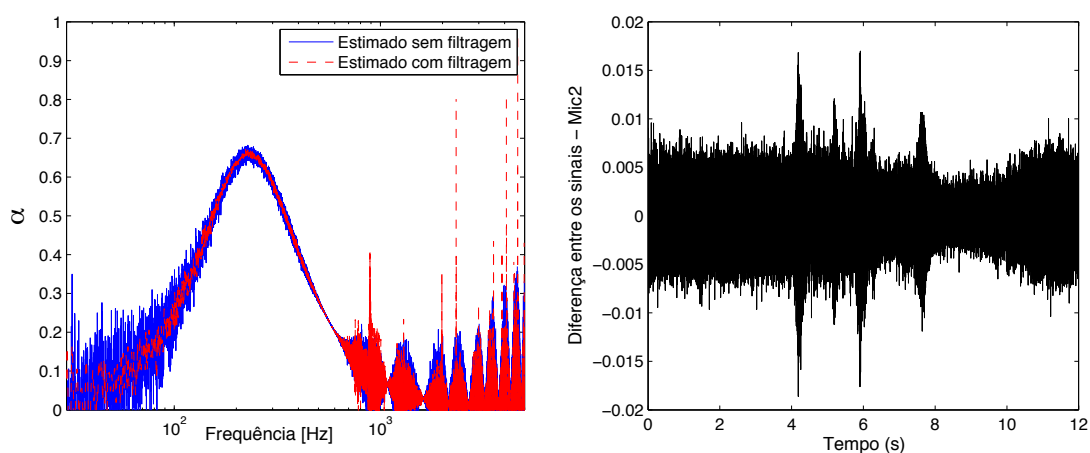


Figura 2.22: Curvas de absorção sonora estimadas (à esquerda) e diferença entre os sinais original e filtrado no microfone 2 (à direita) para db45 com 11 níveis, $\sigma = 0,006$, limiar = 0,001%.

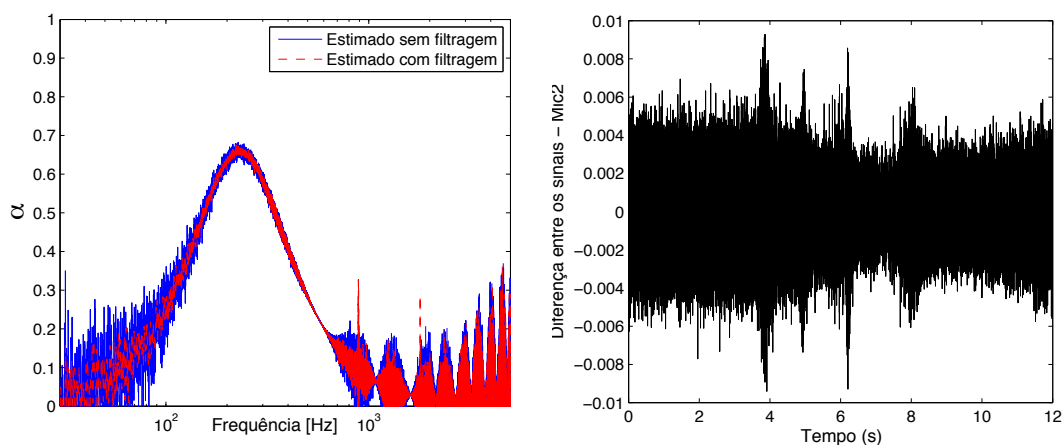


Figura 2.23: Curvas de absorção sonora estimadas (à esquerda) e diferença entre os sinais original e filtrado no microfone 2 (à direita) para db20 com 13 níveis, $\sigma = 0,006$, limiar = 0,0005%.

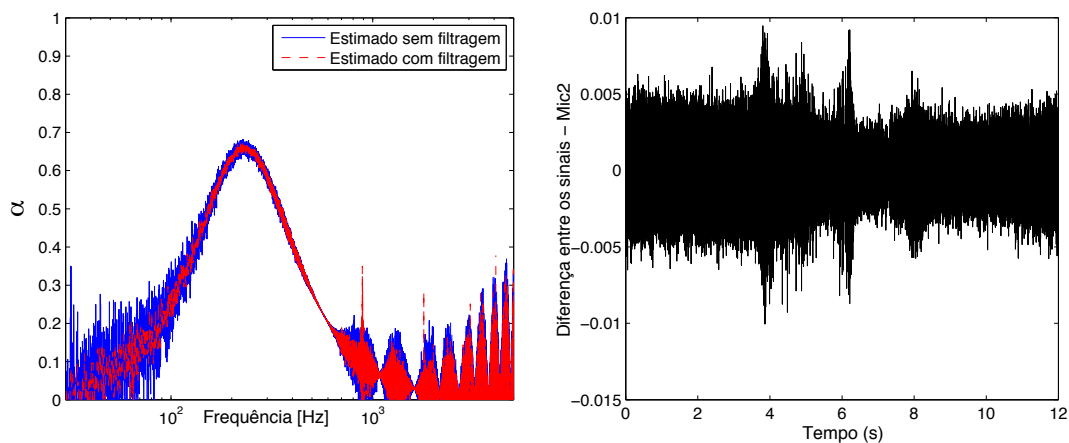


Figura 2.24: Curvas de absorção sonora estimadas (à esquerda) e diferença entre os sinais original e filtrado no microfone 2 (à direita) para sym20 com 10 níveis, $\sigma = 0,006$, limiar = 0,0005%.

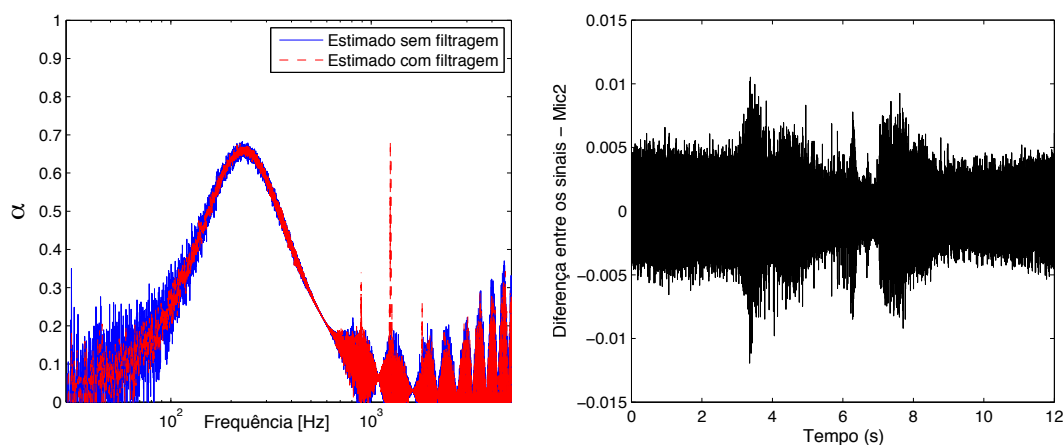


Figura 2.25: Curvas de absorção sonora estimadas (à esquerda) e diferença entre os sinais original e filtrado no microfone 2 (à direita) para coif5 com 9 níveis, $\sigma = 0,006$, limiar = 0,0005%.

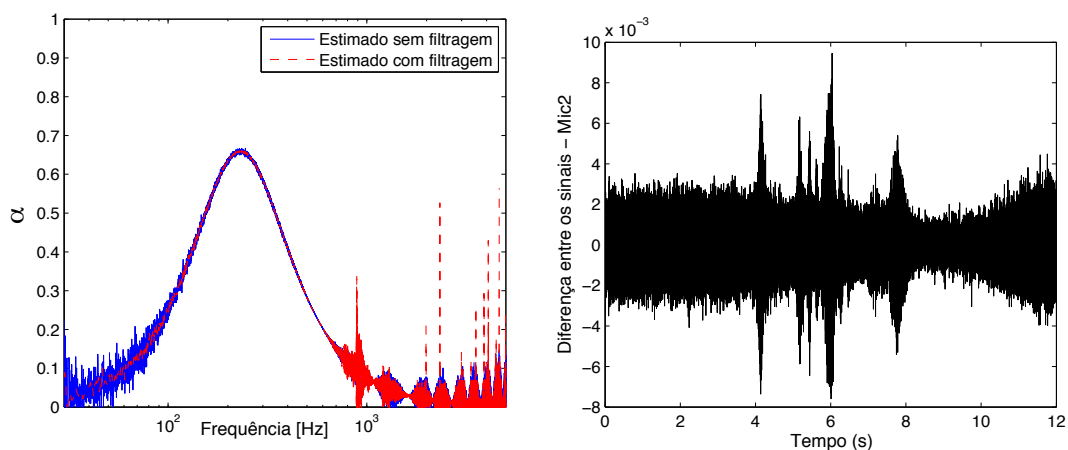


Figura 2.26: Curvas de absorção estimadas (à esquerda) e diferença entre os sinais original e filtrado no microfone 2 (à direita) para db45 com 11 níveis, $\sigma = 0,002$, limiar = 0,0005%.

2.6 Conclusões

O tubo de impedância é um equipamento utilizado para avaliar as propriedades de absorção acústica de materiais. Porém a presença de ruído de medição degrada a qualidade das análises obtidas através deste. Assim, este trabalho utilizou métodos baseados em wavelets para reduzir o ruído em sinais de dois microfones obtidos de uma simulação de tubo de impedância, sendo que tal simulação foi realizada como um sistema caixa preta. Foram fornecidos a esse sistema a varredura senoidal como sinal de excitação do alto-falante e a impedância acústica, e obteve-se os sinais correspondentes aos dois microfones presentes no tubo. Nesses sinais foram adicionados ruídos gaussianos brancos para que posteriormente pudesse avaliar a eficiência do processo de *denoising*. Algumas famílias de wavelets foram testadas, sendo as melhores a db, coif e sym. O resultado foi mais satisfatório nas baixas frequências, ou seja, na faixa que o alto-falante utilizado produz uma baixa relação sinal ruído. Além disso, a escolha do *thresholding* mostrou-se extremamente importante para um bom desempenho do método empregado. Neste caso, utilizou-se uma porcentagem do valor máximo dos coeficientes de detalhes em cada nível, sendo testados aleatoriamente alguns valores de porcentagem, e o melhor encontrado foi de 0,0005%. Por fim, constatou-se que a diferença entre os sinais original e filtrado dos microfones foi próxima ao desvio padrão do ruído, e as wavelets se mostraram capazes de melhorar curvas de absorção estimadas em tubos de impedâncias.

Capítulo 3

Localização de fontes sonoras

Duas técnicas principais para a localização de fontes acústicas são: 1) *Beamforming* que se caracteriza por uma estimativa da distribuição de potência do espectro espacial dos sinais e, a partir daí, da direção de chegada como aquela em que há a maior potência espectral; 2) Estimação de atraso temporal (*Time Delay Estimation* - TDE) que refere-se à estimativa do tempo de chegada de uma frente de onda em um sensor receptor medida utilizando um arranjo de sensores, partindo do pressuposto de que a velocidade de propagação do som é conhecida ou pode ser estimada [11]. Assim, este capítulo irá descrever um método de estimativa de TDE e como utilizá-lo para localizar a posição da fonte sonora. Irá descrever também a metodologia utilizada para simular dois espaços bidimensional e tridimensional para analisar o desempenho dos algoritmos de localização da fonte sonora, resultados e conclusões obtidas.

3.1 Fundamentos de localização de fontes sonoras

O TDE pode ser dividido em duas categorias dependendo da natureza de sua aplicação, sendo estas nomeadas TDOA (*Time Difference Of Arrival*), diferença entre os tempo de chegada de uma frente onda em duas posições do sensor, e TOA (*Time Of Arrival*), tempo que a frente de onda leva para viajar da fonte até a posição do receptor [15]. O TOA tem como objetivo medir o tempo de atraso entre a transmissão de pulso e a recepção do seu eco, sendo primordial para um sistema ativo, como o radar e o sonar ativo. Embora exista uma relação intrínseca entre a estimativa TOA e TDOA, no primeiro caso, o sinal de referência "limpo", ou seja, o sinal transmitido, é conhecido, de modo que a estimativa de atraso de tempo pode ser obtida com base em um único sensor [15].

Os algoritmos dessas duas técnicas são, geralmente, divididos em três grandes grupos, baseados em [10]

1. Estimativa do atraso temporal principalmente por meio de correlação cruzada e correlação cruzada generalizada, sendo o meio de estimação mais simples. E estimação pela diferença de tempo (*Interaural Time Difference*, ITD), que visa a simular de forma simples a audição humana;
2. Encontrar pontos de alta energia, normalmente, denominados de respostas de potência direcionada;
3. Algoritmos baseados em autovalores, tais como o *Multiple Signal Classification*, MUSIC, sendo complexos devido à decomposição em autovalores. Trata-se de um autovetor, correspondente ao autovalor mínimo da matriz de covariância dos sinais dos microfones, que contém a resposta ao impulso entre a fonte e os sinais dos microfones, e, portanto, todas as informações necessárias para a estimativa do tempo de atraso.

3.1.1 Modelo de um ambiente reverberante

Os caminhos acústicos entre a fonte sonora e os microfones podem ser modelados como sistemas lineares, visto que a propagação da onda sonora em um ambiente é modelada por uma equação linear. Tais modelos lineares são válidos para condições reais, como por exemplo, salas de aula, salas de conferência e outros. Na presença de superfícies refletoras de som, as ondas sonoras produzidas por uma única fonte se propagam ao longo de múltiplos caminhos acústicos, dando origem a reverberação. Caso a reverberação do local seja suave, isto é, ocorram pequenas reflexões do sinal sonoro, as pessoas presentes no local podem não perceber esse fenômeno. Mas isso não significa que a reverberação não seja prejudicial à inteligibilidade da fala e a outros fatos como a determinação da posição de uma fonte sonora. Portanto, a propagação multi-caminho deve ser incorporada ao modelo do ambiente para o processamento de sinais, conforme mostra a Fig. 3.1 .

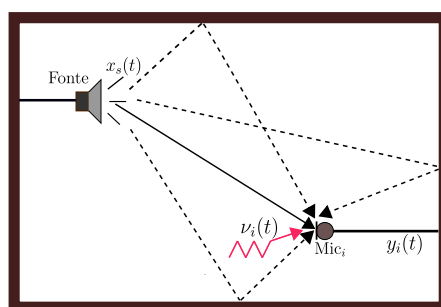


Figura 3.1: Representação de um ambiente reverberante com uma fonte sonora e um microfone.

No modelo da Fig. 3.1 há um microfone, Mic_i e uma fonte sonora capaz de emitir uma onda, $x_s(t)$, que irá se propagar no interior do ambiente, sendo ilustrada pela linha contínua quando percorre o caminho direto da fonte até o microfone e pela linha tracejada quando incide sobre a superfície do ambiente e é refletida para o microfone, e por fim, um ruído que será captado por tal microfone, $\nu_i(t)$. Seja, então, $\boldsymbol{\theta}_f$ e $\boldsymbol{\theta}_{mi}$ os vetores contendo as coordenadas espaciais da posição da fonte sonora e do microfone Mic_i , respectivamente, pode se representar o sinal medido pelo i -ésimo microfone como [45]

$$y_i(t; \boldsymbol{\theta}_f, \boldsymbol{\theta}_{mi}) = g_{mi}(t; \boldsymbol{\theta}_f, \boldsymbol{\theta}_{mi}) * x_s(t) + \nu_i(t), \quad (3.1)$$

em que $*$ denota convolução de duas funções, g_{mi} é a resposta ao impulso acústico entre a fonte e o i -ésimo microfone e $\nu_i(t)$ é o ruído presente no ambiente juntamente com o ruído de medição. Porém, geralmente, o ruído presente no ambiente é mais significativo do que o ruído do canal de microfones e tende a ser dominante no termo $\nu_i(t)$. Além disso, assume-se que o ruído possui distribuição independente e normal para cada microfone.

A resposta ao impulso da sala contém as propriedades do ambiente o qual ele foi gerado e pode ser dividida em três partes: a parte inicial que contém o som direto, a intermediária onde estão presentes as primeiras reflexões e o final do sinal, no qual tem-se as reflexões consideradas tardias. As primeiras reflexões devem ser identificadas até um instante de tempo determinado, e para determinar tal instante é necessário considerar o conhecimento prévio e segundo trabalhos presentes na literatura esse valor varia de 1 a 4 ms [45]. Como estimar esse valor é discutido nas Refs. [19, 20, 26, 29]. Além disso, em um ambiente reverberante a resposta ao impulso será a resposta ao sinal emitido pelo alto-falante que atinge o microfone por um caminho direto, g_{mi}^d , mais a resposta ao sinal que é refletido, g_{mi}^k , pela superfície do ambiente, em que k representa o índice de cada reflexão. Assim, tem-se que [45]

$$g_{mi} = g_{mi}^d + \sum_{k=1}^k g_{mi}^k(t) + \nu_i'(t), \quad (3.2)$$

No entanto, quando a onda sonora incide na superfície parte da sua energia é absorvida pelo meio material e o próprio meio no qual a onda se propaga também consome parte da energia do sinal resultando em uma onda refletida semelhante à onda incidente, mas com magnitude menor. Essa diferença de magnitude entre o som direto e o refletido é quantificada pelo fator de atenuação, β_i^k . Além disso irão ocorrer atrasos na resposta das reflexões, já que o sinal emitido pelo alto-falante ao ser refletido irá percorrer uma distância maior para alcançar o microfone em relação ao sinal recebido

pelo microfone emitido diretamente pela fonte sonora. Além disso, uma onda refletida pode sofrer mais de uma reflexão antes de alcançar o observador/sensor, ou seja, a onda emitida pela fonte incide em uma superfície e reflete para outra superfície sendo refletida novamente na direção do observador. Nesse caso, a reflexão é de segunda ordem. Mas se a onda emitida incide sobre a superfície e é refletida em seguida na direção do observador ela será de primeira ordem. Desse modo, as reflexões podem ser de diferentes ordens, e quanto maior a ordem, maior é a atenuação da onda.

Para determinar os atrasos das reflexões é preciso determinar a distância percorrida pelo sinal refletido, e isso pode ser realizado utilizando um método conhecido na literatura como o *método da fonte imagem*, Fig. 3.2. Esse método consiste em construir uma imagem espelhada da fonte sonora, isto é, uma fonte imagem, F_{img} , é construída com as características direcionais simétricas as da fonte real (F_{re}), porém do outro lado da parede, e então, a superfície da parede pode ser desconsiderada, e seu efeito é substituído pela fonte imagem. Assume-se que a imagem emite exatamente o mesmo sinal sonoro que a fonte original, e então, a distância percorrida pela reflexão será igual à distância da fonte imagem ao ponto de medição (linha vermelha pontilhada na Fig. 3.2). Para uma parede de extensão finita, as direções de emissão de F_{img} são restringidas em conformidade, ou de forma diferente, para certos pontos de observação O a fonte imagem pode se tornar invisível. Isto ocorre, caso a linha que conecta O com a fonte imagem não cruzar a parede real [32].

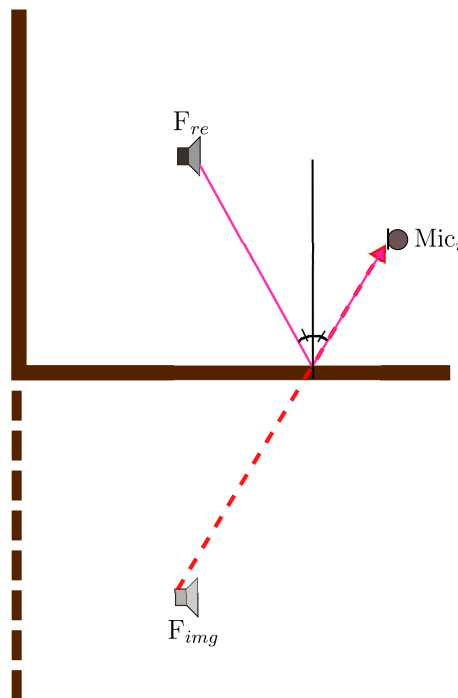


Figura 3.2: Representação do método da fonte imagem.

Normalmente, nem toda a energia que atinge uma parede é refletida a partir dela,

sendo que parte dessa energia também pode ser absorvida pela parede ou até mesmo ser transmitida por ela para o outro lado da parede. Logo a resposta ao impulso das reflexões, Eq. (3.2), terá a mesma forma do sinal emitido pela fonte sonora mas atrasado por um intervalo de tempo t_i^k e com magnitude atenuada por um valor β_i^k sendo descrita como

$$g_{mi}^k(t) = \beta_i^k x(t - t_i^k(d_r)) + \nu''_i(t), \quad (3.3)$$

no qual d_r representa a distância do caminho percorrido por uma reflexão da sua localização até o microfone, sendo que sua localização refere-se a localização da fonte imagem. O fator de atenuação depende de alguns fatores como:

- As propriedades de absorção da superfície e do ar;
- Distância entre a fonte sonora e os microfones;
- Diretividade da fonte e dos microfones.

Por fim, na resposta da sala a média temporal da densidade das reflexões, ρ_e , isto é, o número médio de reflexões por segundo, é aproximadamente proporcional ao quadrado de tempo t [32],

$$\frac{d\rho_e}{dt} = 4\pi c^3 \frac{t^2}{V} \quad (3.4)$$

sendo c a velocidade do som em m/s, V o volume da sala em m³. Esta expressão pode ser facilmente demonstrada para salas retangulares e generalizada para salas de qualquer geometria [29].

3.2 Algoritmos de TDOA

Existem vários algoritmos desenvolvidos na literatura para determinar o valor de TDE, mas dois métodos bastante utilizados são os métodos de correlação cruzada e correlação cruzada generalizada que serão descritos a seguir. Embora existam outros métodos comumente usados baseados na função da diferença média, estes não serão abordados nesse trabalho, mas são estudados por Chen et al. [14]. Além disso, a acurácia na estimação desses valores depende da frequência de amostragem.

Vale ressaltar que uma estimação precisa e robusta de TDOA é fundamental para a eficácia dos algoritmos de localização da fonte sonora. Porém, há dois fatores que complicam essa estimativa, sendo eles o ruído de fundo e as reverberações da sala.

Para um par de microfones, $n = 1, 2$, o TDOA associado a ele, τ_{12} , é definido como [45]

$$\tau_{12} \equiv \tau_1 - \tau_2, \quad (3.5)$$

sendo τ_1 e τ_2 o tempo necessário para a onda sonora alcançar por um caminho direto os microfones 1 e 2.

3.2.1 Método da correlação cruzada

Seja $x_s(t)$ um sinal emitido por uma fonte sonora, e um conjunto de pelo menos dois microfones. Os sinais recebidos por tais microfones, conforme as equações descritas na Seção 3.1.1, serão

$$\begin{cases} y_1(t), & = g_{m1}(t - \tau_1) * x_s(t) + \nu_1(t), \\ y_2(t), & = g_{m2}(t - \tau_1 - |\tau_{12}|) * x_s(t) + \nu_2(t), \end{cases} \quad (3.6)$$

em que: $*$ representa a operação de convolução; $y_1(t)$ e $y_2(t)$ são, respectivamente, os sinais medidos pelos microfones 1 e 2; $g_{m1}(t)$ é a resposta ao impulso entre a fonte e o microfone 1, enquanto que $g_{m2}(t)$ é a resposta ao impulso entre a fonte e o microfone 2; $\nu_1(t)$ e $\nu_2(t)$ são, respectivamente, ruídos presentes no ambiente e na medição em relação ao microfone 1 e 2; os parâmetros τ_1 e τ_2 representam o tempo necessário para a onda sonora alcançar por um caminho direto, respectivamente, os microfones 1 e 2; e por fim, τ_{12} representa a diferença de tempo de chegada, TDOA, entre os microfones 1 e 2.

As respostas ao impulso em relação a cada microfone possuem formato semelhantes. Portanto, a Eq. (3.6) mostra que o sinal do microfone 2 é uma versão deslocada no tempo do sinal medido pelo microfone 1. Logo, ao realizar a correlação cruzada entre os dois sinais, o resultado será um sinal com um pico no instante de tempo no qual as versões do microfone 1 e a deslocada se alinham, correspondendo ao valor de TDOA, τ_{12} . Portanto, calcular função de correlação cruzada entre os sinais dos microfones é uma alternativa para estimar os valores de TDOA. Essa função, então, compara o quão semelhantes os sinais são e fornece um pico de máximo quando a semelhança for máxima. Assim sendo, τ_{12} é estimado como o argumento que maximiza a função de correlação cruzada dos sinais medidos pelos microfones [31], isto é,

$$\widehat{\tau}_{12}^{cc} = \left(\arg \max_{\tau \in d} R_{y_1 y_2}^{cc}(\tau) \right) \frac{1}{f_s}, \quad (3.7)$$

em que [3]

$$R_{y_1 y_2}^{cc}(\xi, t) = E[y_1(t)y_2(t + \xi)] \quad (3.8)$$

é a função de correlação cruzada (FCC) entre os sinais y_1 e y_2 , $E[\cdot]$ é a esperança matemática, f_s é a frequência de amostragem e $1/f_s$ é o intervalo de amostragem.

Como a correlação cruzada é sempre realizada entre os sinais medidos por dois microfones, deve-se adotar como microfone de referência aquele no qual o som direto chega primeiro, e assim, a FCC é calculada entre o sinal do microfone de referência, y_r , com os demais microfones. Além disso, para os casos, como o deste trabalho, em que o sinal emitido pela fonte sonora é conhecido, pode-se substituir o sinal do microfone de referência pelo sinal da fonte sem atrasos e, então, obter uma melhor estimativa de TDOA.

Ainda sobre a FCC, para um sistema ergódico a média temporal é igual à média instantânea, e assume-se a estacionariedade do sistema para eliminar a dependência da função de covariância cruzada com o tempo, portanto a esperança matemática descrita na Eq. (3.8) pode ser substituída pela média temporal. Assim a função de correlação cruzada torna-se [3]

$$R_{y_r y_i}^{cc}(\xi) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T y_r(t) y_i(t + \xi) dt. \quad (3.9)$$

A fim de ilustrar a aplicação da FCC no cálculo de TDOAs, considere um espaço bidimensional simulado com uma frequência de amostragem de 40 kHz contendo 4 microfones em diferentes posições desse espaço no qual o sinal emitido pela fonte sonora é um pulso senoidal modulado por uma gaussiana, Fig. 3.3 à esquerda. Os sinais captados por esses microfones contém algumas reflexões devido as superfícies do ambiente e ruído. Além disso, nota-se que o som direto alcança primeiro o microfone 1, portanto ele é considerado o microfone de referência. Porém, como o sinal emitido pela fonte é conhecido, o sinal adotado como referência é o sinal da fonte. Assim, ao calcular a correlação cruzada entre os sinais de entrada e do microfone 2, obtém-se o sinal da Fig. 3.3 à direita. Em tal sinal, há um pico de maior amplitude, em -139 atrasos, que corresponde ao τ_{12} (τ_{xy_2}) em -139 atrasos, e outros três picos com amplitudes menores devido as reflexões presentes no sinal, ou seja, caso não houvesse reflexões no ambiente, ao calcular a correlação cruzada, só haveria um pico. Vale ressaltar que o número negativo dos atrasos significa que o sinal do microfone 2 está atrasado em relação ao 1, isto é, o som direto alcança primeiro o microfone 1. Por fim, para calcular o TDOA entre os microfones basta substituir esses valores na Eq. 3.7, sendo

$$\hat{\tau}_{xy_2}^{cc} = -(-139)/f_s = 139/40000 = 3,5ms. \quad (3.10)$$

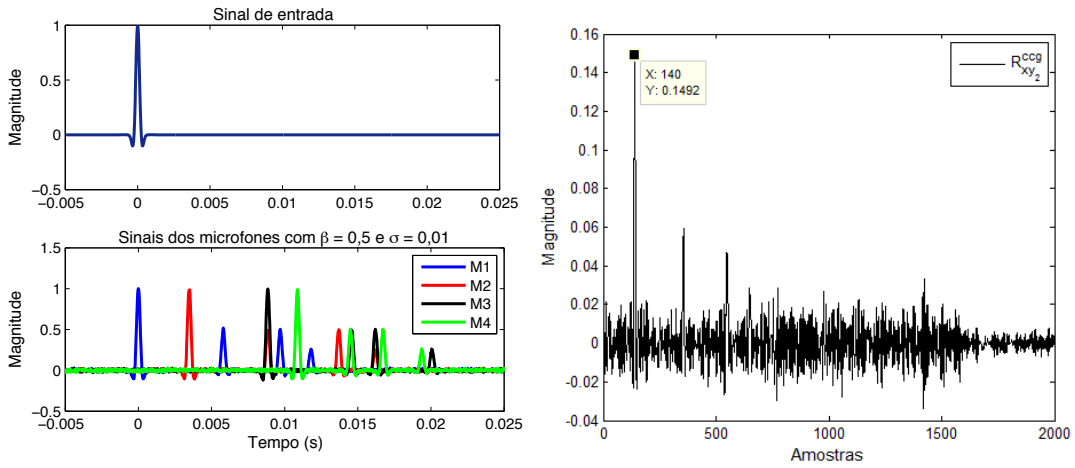


Figura 3.3: Simulação de um espaço bidimensional com quatro microfones considerando três reflexões e ruído ambiente (à esquerda) e correlação cruzada entre o sinal de entrada e o microfone 2 (à direita).

3.2.2 Método da correlação cruzada generalizada

O método da correlação cruzada generalizada (CCG), *Generalized Cross-Correlation*, foi publicado por Knapp e Carter em 1976. Ele pode ser considerado uma versão melhorada do método da correlação cruzada, isso porque nele é incorporado o conhecimento prévio dos sinais, melhorando seu desempenho de estimação do TDE [15].

Sejam os sinais medidos pelos microfones, $y_r = r$ e $y_i = i$, a função de CCG (FCCG), $R_{ri}^{ccg}(\xi)$, é definida como a correlação cruzada de duas versões filtradas dos sinais $r(t)$ e $i(t)$, com as transformadas de Fourier desses filtros denotadas respectivamente, por $H_r(\omega)$ e $H_i(\omega)$. A CCG pode ser expressa em termos da transformada de Fourier dos sinais dos microfones, $Y_r(\omega)$ e $Y_i(\omega)$ [9],

$$R_{ri}^{ccg}(\xi) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} (H_r(\omega)Y_r(\omega))(H_i(\omega)Y_i(\omega))^* e^{i\omega\xi} d\omega, \quad (3.11)$$

em que o símbolo $*$ indica que deve-se utilizar o conjugado complexo da expressão entre os parênteses.

Reorganizando a ordem dos sinais e filtros e definindo uma função de ponderação dependente da frequência,

$$\Gamma_{ri} \equiv H_r(\omega)H_i(\omega)^*. \quad (3.12)$$

a função de CCG pode ser expressa como [9]

$$R_{ri}^{ccg}(\xi) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \Gamma_{ri}(\omega)Y_r(\omega)Y_i(\omega)^* e^{i\omega\xi} d\omega. \quad (3.13)$$

Idealmente, $R_{ri}^{ccg}(\xi)$ exibirá um máximo global explícito no ponto correspondente

ao tempo de atraso entre os microfones. Assim, o TDOA é estimado como o valor que maximiza a função de correlação cruzada generalizada [9]

$$\hat{\tau}_{ri}^{ccg} = (\arg \max_{\tau \in d} R_{y_r y_i}^{ccg}(\xi)) \frac{1}{f_s}, \quad (3.14)$$

sendo que a faixa de potenciais valores de TDOA é restrita a um intervalo finito, d , que é determinado pela distância física entre os microfones.

Geralmente, nas situações em que há reverberação nos ambientes, tanto no caso da FCC quanto na FCCG, poderá ocorrer vários máximos locais que podem obscurecer o verdadeiro pico de TDOA produzindo uma estimativa incorreta. As amplitudes e os atrasos de tempo correspondentes a esses máximos errôneos dependem de um número de fatores já discutidos anteriormente. Para tentar resolver esse problema, a função de correlação cruzada generalizada conta com a função de ponderação, definida na Eq. (3.12). O objetivo dessa função é enfatizar a FCCG no ponto correspondente ao verdadeiro valor de TDOA sobre os extremos locais indesejados. Há várias tipos dessas funções, como por exemplo o processador Roth [39], o Hassab- Boucher transform [25], o filtro Echart, a transformação de fase (do inglês *phase transform*, PHAT), o processador ML [31] e outras. Pode-se utilizar também, a função de ponderação unitária, e assim a FCCG torna-se a função de correlação cruzada.

Neste trabalho, a função de ponderação utilizada é a transformação de fase decrita por [9],

$$\Gamma_{ri} \equiv \frac{1}{|Y_r(\omega)Y_i(\omega)^*|}. \quad (3.15)$$

Ao enfatizar cada componente de fase dos sinais para realizar a comparação, o pico resultante na função CCG-PHAT corresponde ao atraso dominante no sinal reverberado. Embora isso seja eficaz na redução de algumas das degradações devidas aos multi-caminhos, a transformação de fase acentua os componentes do espectro que possuem baixa relação sinal ruído (SNR) e têm o potencial de fornecer resultados ruins, particularmente sob baixa reverberação e condições de alto ruído [9].

A aplicação desse método pode ser ilustrada utilizando a simulação do espaço bi-dimensional na Fig. 3.3 à esquerda. Utilizando os sinais emitidos pela fonte e do microfone 2 para calcular a correlação cruzada generalizada com a transformação de fase como função de ponderação, tem-se como resultado a Fig. 3.4, no qual um pico de maior amplitude ocorre em 140 amostras. O valor de TDOA é calculado substituindo esse valor na Eq. 3.14,

$$\hat{\tau}_{xy_2}^{ccg} = 140/f_s = 140/40000 = 3,5ms. \quad (3.16)$$

Esse procedimento é repetido para os demais microfones para calcular todos os valores

de TDOAs.

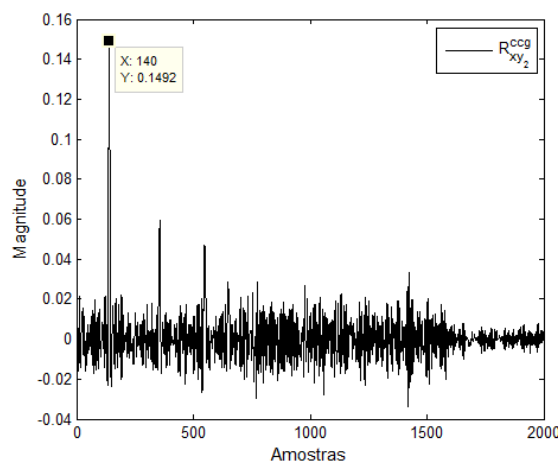


Figura 3.4: Resultado da correlação cruzada generalizada utilizando transformação de fase (CCG-PHAT) entre o sinal emitido pela fonte sonora e o microfone 2.

3.2.3 Transformada wavelets na estimação de TDOA

A localização de fontes sonoras requer alta taxa de amostragem para uma estimação precisa. Porém, essa taxa elevada resulta em um alto volume de dados, exige memória elevada para armazená-los e uma grande largura de banda para transmiti-los, além de gerar um alto consumo de energia. Isso dificulta a utilização de uma rede de sensores acústicos sem fio para a localização de fontes sonoras, visto que é necessário um hardware mais complexo, o consumo de energia da bateria será alto e os custos financeiros ficam elevados. Por isso, há um interesse de desenvolver técnicas de baixa frequência na utilização de sensores sem fio.

Na Ref. [8], os autores propõem, para resolver esse problema, utilizar baixa taxa de amostragem combinando informações no domínio do tempo e da frequência resultando em um análise tempo-frequência. Esse domínio fornece informações sobre como o conteúdo do sinal muda com o tempo, proporcionando assim uma técnica ideal para processar e interpretar os sinais recebidos a baixas taxas de amostragem. Esses autores sugerem utilizar baixas taxas de amostragem e aplicar um envelope sobre o sinal e em seguida transformada wavelets antes de estimar os valores de TDOA.

A utilização do envelope do sinal em vez das amplitudes no processo de estimação de TDE é um dos métodos preferidos para melhorar a precisão. Ele é capaz de minimizar a ambiguidade que existe em torno do início das amplitudes do sinal e dos índices de máximos se for utilizada a correlação cruzada convencional [8]. Existem vários métodos de extrair os envelopes dos sinais medidos, e qualquer um deles pode ser usado nesse

trabalho, desde que ele não retarde a execução da operação. Porém o método escolhido é descrito por [8].

$$\tilde{y}_i(t) = \sqrt{(y_i(t))^2 + (\text{Im}(\hat{y}_i(t)))^2}, \quad (3.17)$$

no qual $(y_i(t))^2$ é a energia do sinal medido pelo microfone i , $\hat{y}_i(t)$ é a transformada Hilbert do sinal $y_i(t)$, Im representa a parte imaginária do sinal complexo e, por fim \tilde{y}_i é o envelope obtido para esse sinal $y_i(t)$. Essa equação, Eq. (3.17), pode ser obtida, simplesmente, pelo módulo da transformada Hilbert do sinal em questão,

$$\tilde{y}_i(t) = |\hat{y}_i(t)|. \quad (3.18)$$

A Fig. 3.5 ilustra um pulso senoidal com modulação gaussiana com frequência central igual a 500 Hz e banda de passagem de 0,5, e seu envelope gerado a partir da Eq. (3.17).

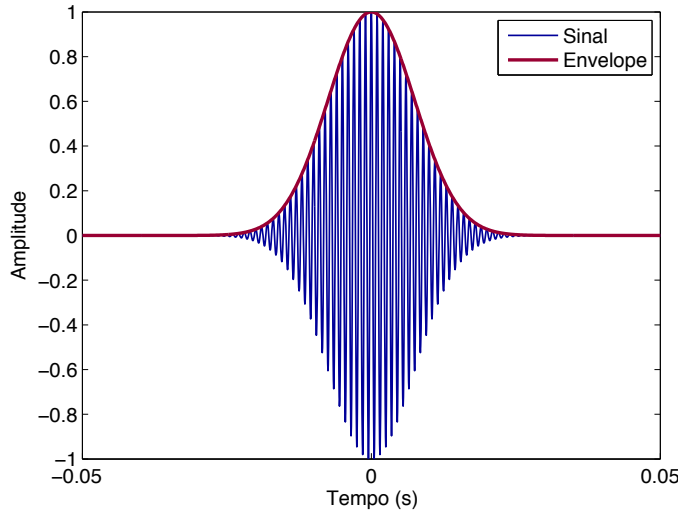


Figura 3.5: Pulso senoidal com modulação gaussiana e seu envelope.

Após extrair os envelopes dos sinais dos microfones é preciso calcular a transformada contínua de wavelets, TCW, para neutralizar o impacto da utilização da baixa taxa de amostragem. E conforme descrito na Seção 2.2.3, a TCW dos envelopes extraídos de cada microfone será dada por

$$W_{a,b}\{\tilde{y}_i(t)\} = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} \tilde{y}_i(t) \psi^* \left(\frac{t-b}{a} \right) dt, \quad (3.19)$$

no qual $\psi(t)$ é a wavelet mãe, a é o parâmetro de escala, b é o parâmetro de deslocamento e $*$ denota o conjugado complexo caso a wavelet escolhida para decomposição seja complexa. Tal decomposição permite a análise do sinal no domínio tempo-frequência.

A escolha da wavelet mãe a ser utilizada é dependente das características do sinal a ser processado, então, para um bom resultado deve-se buscar uma família de wavelets semelhante a tal sinal [8].

Todo esse procedimento pode ser resumido nas seguintes etapas:

1. Extrair os envelopes dos sinais medidos pelos microfones;
2. Escolher a wavelet para decomposição, $\psi(t)$, e determinar quais escalas serão utilizadas para expandir a wavelet mãe e calcular a TCW, sendo que a escala será um vetor contendo de 1 a N variações, $a = 1, 2, \dots, N$. Por exemplo, $N = 5$, então os valores de escala serão $a = [1 \ 2 \ 3 \ 4 \ 5]$;
3. Calcular a TCW dos envelopes para cada valor de escala escolhido. Assim, para sinal de envelope, ao calcular a TCW é criada uma matriz de tamanho $(N \times L_y)$, sendo L_y o número de amostras do sinal do envelope (= ao número de amostras dos sinais dos microfones).
4. Aplicar algum método de estimação de TDOAs, por exemplo a CC ou CCG em cada escala individualmente;
5. Calcular as médias de todos os valores de TDOAs calculados em cada escala;
6. A partir dos valores de TDOAs estimar a posição da fonte sonora.

Por fim, é preciso lembrar que em situações no qual o sinal da fonte sonora é conhecido, a estimação de TDOA é realizada sempre entre o sinal do microfone em relação ao sinal da fonte que é tido como sinal de referência, obtendo um melhor desempenho. Assim, para esse procedimento utilizando wavelets, as três primeiras etapas listadas anteriormente devem ser executadas sobre o sinal da fonte sonora.

3.3 Localização de fontes sonoras

A precisão da localização de fontes depende de alguns fatores como a quantidade e qualidade dos microfones utilizados, a posição relativa dos microfones entre si e em relação à fonte, os níveis de ruído ambiente e a reverberação, e o número de fontes ativas e seu conteúdo espectral. Geralmente o desempenho na localização da fonte melhora com o aumento do número de microfones.

Há vários algoritmos para determinar a posição de uma fonte sonora no espaço, mas este trabalho utiliza o método de estimação por mínimos quadrados sem restrição, LS (*least-square*) [44].

Considere um sistema, Fig. 3.6, formado por uma única fonte sonora e pelos microfones, Mic_r , Mic_i , com o índice r representando o microfone de referência e $i = 1, 2, 3 \dots N_m$ os demais, sendo N_m o número total de microfones. A posição desses elementos pertence ao plano de coordenadas cartesianas, x e $y \in \mathbb{R}^2$. O vetor contendo as posições da fonte sonora nesse sistema é denotado por θ_f , enquanto o vetor posição de Mic_r e Mic_i é, respectivamente, θ_{mr} e θ_{mi} . Todos esses vetores são vetores colunas, isto é, cada eixo da coordenada cartesiana ocupa uma linha do vetor.

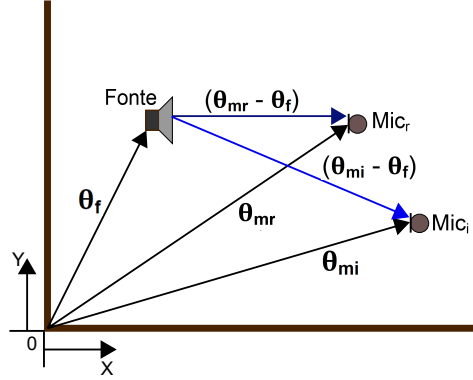


Figura 3.6: Representação de um ambiente de duas dimensões com uma fonte sonora e dois microfones.

As distâncias dos microfones Mic_r e Mic_i em relação a fonte sonora, respectivamente, d_{rf} e d_{if} , são descritas por

$$\begin{cases} d_{rf} = \|\theta_{mr} - \theta_f\|, \\ d_{if} = \|\theta_{mi} - \theta_f\|, \end{cases} \quad (3.20)$$

no qual $\|\cdot\|$ representa a norma Euclidiana do vetor. Relacionando as Equações (3.20) tem-se o valor de quanto o microfone i está mais distante da fonte em relação ao microfone de referência, d_{ri}^f [44],

$$d_{ri}^f = \|\theta_{mr} - \theta_f\| - \|\theta_{mi} - \theta_f\| \quad (3.21)$$

Reorganizando a Eq. (3.21), e elevando ao quadrado os dois lados da equação,

$$\underbrace{(\|\theta_{mi} - \theta_f\|)^2}_A = \underbrace{(d_{ri}^f + \|\theta_{mr} - \theta_f\|)^2}_B. \quad (3.22)$$

A fim de facilitar a compreensão do desdobramento da Eq. (3.22), ela será resolvida

por partes, primeiro a parte A e em seguida a parte B [44],

$$\begin{aligned}
A &= (\boldsymbol{\theta}_{mi} - \boldsymbol{\theta}_f)^T (\boldsymbol{\theta}_{mi} - \boldsymbol{\theta}_f) \\
&= (\boldsymbol{\theta}_{mi}^T - \boldsymbol{\theta}_f^T) (\boldsymbol{\theta}_{mi} - \boldsymbol{\theta}_f) \\
&= \boldsymbol{\theta}_{mi}^T \boldsymbol{\theta}_{mi} - \boldsymbol{\theta}_{mi}^T \boldsymbol{\theta}_f - \boldsymbol{\theta}_f^T \boldsymbol{\theta}_{mi} + \boldsymbol{\theta}_f^T \boldsymbol{\theta}_f \\
&= \|\boldsymbol{\theta}_{mi}\|^2 - \boldsymbol{\theta}_{mi}^T \boldsymbol{\theta}_f - \boldsymbol{\theta}_{mi}^T \boldsymbol{\theta}_f + \|\boldsymbol{\theta}_f\|^2 \\
&= \|\boldsymbol{\theta}_{mi}\|^2 - 2\boldsymbol{\theta}_{mi}^T \boldsymbol{\theta}_f + \|\boldsymbol{\theta}_f\|^2.
\end{aligned}$$

$$B = (d_{ri}^f)^2 + 2d_{ri}^f \|\boldsymbol{\theta}_{mr} - \boldsymbol{\theta}_f\| + \|\boldsymbol{\theta}_{mr} - \boldsymbol{\theta}_f\|^2. \quad (3.23)$$

Retornando com a resolução de A e B (Eq. (3.23)) na Eq. (3.22),

$$\begin{aligned}
\|\boldsymbol{\theta}_{mi}\|^2 - 2\boldsymbol{\theta}_{mi}^T \boldsymbol{\theta}_f + \|\boldsymbol{\theta}_f\|^2 &= (d_{ri}^f)^2 + 2d_{ri}^f \|\boldsymbol{\theta}_{mr} - \boldsymbol{\theta}_f\| + \|\boldsymbol{\theta}_{mr} - \boldsymbol{\theta}_f\|^2, \\
d_{ri}^f \|\boldsymbol{\theta}_{mr} - \boldsymbol{\theta}_f\| + (\boldsymbol{\theta}_{mi}^T - \boldsymbol{\theta}_{mr}^T) \boldsymbol{\theta}_f &= \frac{\|\boldsymbol{\theta}_{mi}\|^2 - (d_{ri}^f)^2 - \|\boldsymbol{\theta}_{mr}\|^2}{2}.
\end{aligned} \quad (3.24)$$

Mas considerando que há mais de dois microfones, tal equação pode ser expandida em matrizes, sendo elas [44]

$$\boldsymbol{\theta} = \begin{bmatrix} \|\boldsymbol{\theta}_{mr} - \boldsymbol{\theta}_f\| \\ \boldsymbol{\theta}_f \end{bmatrix}, \quad (3.25)$$

$$\mathbf{D} = \begin{bmatrix} d_{r1}^f & (\boldsymbol{\theta}_{m1}^T - \boldsymbol{\theta}_{mr}^T) \\ d_{r2}^f & (\boldsymbol{\theta}_{m2}^T - \boldsymbol{\theta}_{mr}^T) \\ \vdots & \vdots \\ d_{ri}^f & (\boldsymbol{\theta}_{mi}^T - \boldsymbol{\theta}_{mr}^T) \end{bmatrix}, \quad (3.26)$$

$$\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_i \end{bmatrix}, \quad d_i = \frac{\|\boldsymbol{\theta}_{mi}\|^2 - (d_{ri}^f)^2 - \|\boldsymbol{\theta}_{mr}\|^2}{2}. \quad (3.27)$$

As dimensões das matrizes dependem do número de microfones e se o espaço utilizado é bi-dimensional ou tri-dimensional. Logo, considerando que a posição da fonte é um vetor coluna, a matriz $\boldsymbol{\theta}$ tem o número de linhas igual à dimensão do espaço mais um, e sempre uma única coluna. Já a matriz \mathbf{D} possui o número de linhas igual ao número de microfones menos um, e o número de colunas igual à dimensão do espaço mais um. A matriz \mathbf{d} tem o número de linhas igual ao número de microfones menos

um, e uma coluna [44]. Para o exemplo Fig. 3.6, as dimensões das matrizes são:

$$\begin{cases} \boldsymbol{\theta} &= (3 \times, 1), \\ \mathbf{D} &= (N_m - 1) \times 3 \\ \mathbf{d} &= (N_m - 1) \times 1 \end{cases} \quad (3.28)$$

Assim, a Eq. (3.24), torna-se

$$\mathbf{D}\boldsymbol{\theta} = \mathbf{d}. \quad (3.29)$$

Considerando as condições reais, nos quais a posição da fonte é desconhecida, a distância entre um determinado microfone e o microfone de referência em relação à posição da fonte sonora, Eq. (3.21), pode ser estimada a partir dos valores que foram estimados de TDOAs, sendo

$$\tilde{\mathbf{d}}_{ri}^f = c\tilde{\tau}_{ri}, \quad (3.30)$$

em que c é a velocidade do som em m/s e $\tilde{\tau}_{ri}$ representa a diferença do tempo de chegada, TDOA, entre o microfone de referência, Mic_r , e o Mic_i , com $i = 1, 2, \dots, N_m$. Logo, a Eq. (3.29), é reescrita como

$$\tilde{\mathbf{D}}\tilde{\boldsymbol{\theta}} \approx \tilde{\mathbf{d}}. \quad (3.31)$$

A solução dessa equação, Eq. (3.31), neste trabalho, é obtida utilizando o método de mínimos quadrados sem restrição, LS [44],

$$\tilde{\boldsymbol{\theta}} = (\tilde{\mathbf{D}}^T \tilde{\mathbf{D}})^{-1} \tilde{\mathbf{D}}^T \tilde{\mathbf{d}}, \quad (3.32)$$

sendo que

$$\tilde{\boldsymbol{\theta}} = \arg \min \|\tilde{\mathbf{D}}\tilde{\boldsymbol{\theta}} - \tilde{\mathbf{d}}\|. \quad (3.33)$$

A matriz $\tilde{\mathbf{D}}$ possui posto completo se o arranjo de microfones for pelo menos igual ao número de dimensões do espaço mais dois, caso um microfone seja tido como referência. Isso porque a matriz $\boldsymbol{\theta}$ possui o número de linhas igual à dimensão do espaço mais um, ou seja, esse é o número de incógnitas estimadas e para cada uma delas é preciso ter um microfone. Além disso, é preciso ter mais um microfone para servir como referência para o cálculo de TDOAs. Então, para o exemplo discutido nessa seção, o número mínimo de microfones utilizados é 4, pois a matriz $\boldsymbol{\theta}$ possui 3 linhas e assim há estimação de 3 incógnitas, necessitando de 3 microfones mais 1 usado como referência, dando um total de 4 microfones. Porém, neste trabalho, o sinal emitido pela fonte é conhecido,

sendo utilizado como sinal referência em vez do sinal do microfone. Essa substituição é realizada para obter um melhor desempenho na estimação dos TDOAs, pois o sinal emitido pela fonte é um sinal "limpo", isto é, não contém ruído natural do ambiente e reflexões devido aos limites das superfícies e objetos presentes no local.

3.4 Metodologia

A fim de estudar alguns dos métodos existentes na literatura para determinar a posição de uma fonte sonora, foi simulado um espaço bidimensional (\mathbb{R}^2), Fig. 3.7, e tridimensional, \mathbb{R}^3 com 6 superfícies, ou seja, uma sala. Nesses espaços há uma fonte sonora que emite um sinal $x_s(t)$ e microfones denominados de Mic_i , sendo $i = 1, \dots, 4$ para o espaço \mathbb{R}^2 e $i = 1, \dots, 5$ para \mathbb{R}^3 . A posição desses objetos é descrita em relação ao sistema de coordenadas cartesianas x e y para \mathbb{R}^2 , e x , y e z para \mathbb{R}^3 . Para tais simulações considerou-se que os ambientes são reverberantes, e para determinar os atrasos das ondas refletidas utilizou-se o *método das fontes imagens* descrito na Seção 3.1.1.

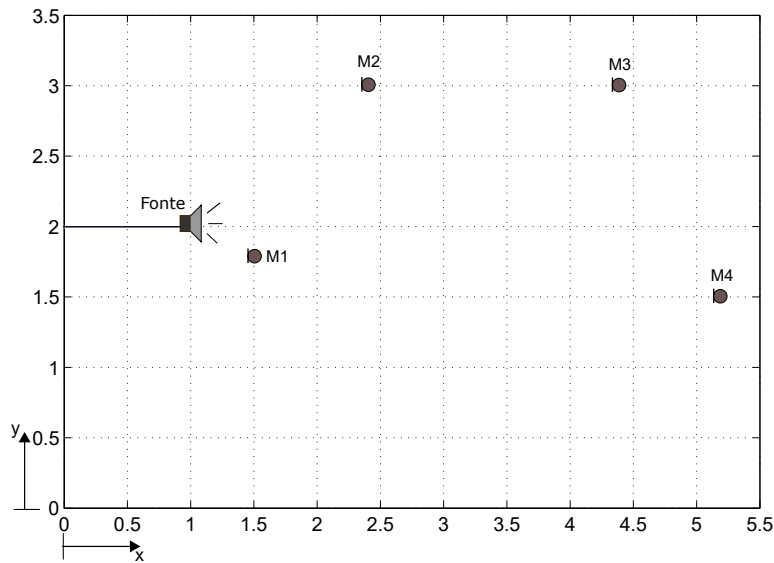


Figura 3.7: Posições da fonte sonora e dos microfones em \mathbb{R}^2 .

As posições dos objetos utilizadas na simulação do espaço bidimensional estão representadas na Fig. 3.7. Assumiu-se que ocorrem apenas três reflexões, sendo duas de primeira ordem e uma de segunda ordem. Assim, tem-se a Fig. 3.8, na qual F_{re} representa a fonte real, F_1 , F_2 e F_3 representam as três fontes imagens e Mic_i o microfone. Os vetores θ contém as coordenadas das posições de todos os objetos, sendo que os subscritos desses vetores são para diferenciá-los. Assim, o θ_f corresponde à posição da fonte real, θ_{mi} à posição do microfone e os demais θ s correspondem às posições das

fontes imagens. As linhas tracejadas em vermelho, saindo da fonte imagem até o microfone referem-se as reflexões de primeira ordem. Já a linha tracejada em verde, que sai da terceira fonte imagem em direção ao microfone, refere-se à reflexão de segunda ordem.

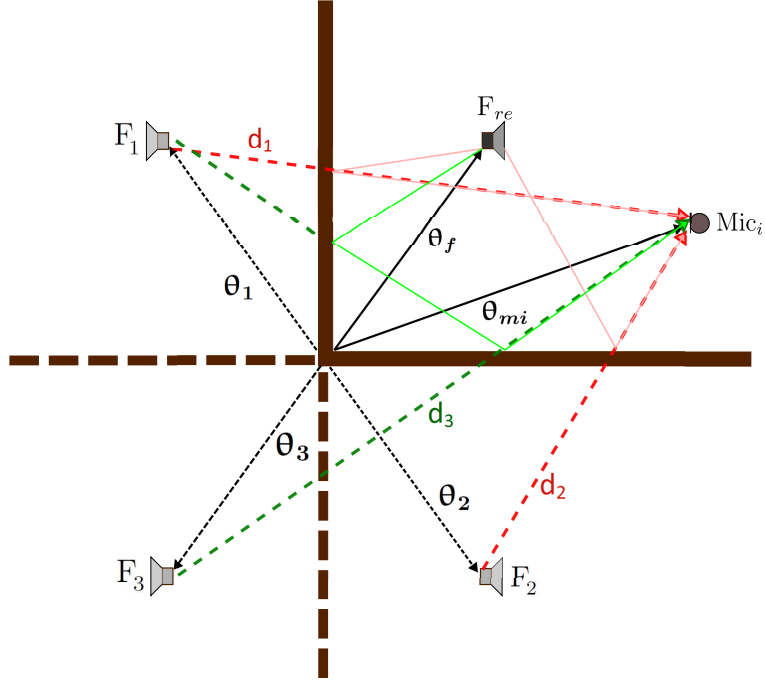


Figura 3.8: Método das fontes imagens em \mathbb{R}^2 considerando duas reflexões de primeira ordem e uma de segunda ordem.

As distâncias entre as fontes imagens e o microfone, Fig. 3.8, são descritas por

$$\begin{aligned} d_1 &= \|\theta_{mi} - \theta_1\|, \\ d_2 &= \|\theta_{mi} - \theta_2\|, \\ d_3 &= \|\theta_{mi} - \theta_3\|, \end{aligned}$$

em que as posições das fontes, real e imaginárias, e microfones são descritas na Tab. 3.1.

Objeto	Eixo x	Eixo y
F_{re}	1,0	2,0
F_2	1,0	-2,0
F_1	-1,0	2,0
F_3	-1,0	-2,0
Mic_1	1,5	1,8
Mic_2	2,4	3,0
Mic_3	4,4	3,0
Mic_4	5,2	1,5

Tabela 3.1: Posições das fontes sonoras, real e imaginárias, e dos microfones em \mathbb{R}^2 .

Todo esse procedimento é realizado para calcular os atrasos das reflexões em relação a cada microfone, e o processo é bem semelhante para o espaço tridimensional. Porém, para esse novo espaço, consideram-se apenas as reflexões de primeira ordem, sendo 6 reflexões e a posição dos objetos é dada na Tab. 3.2. As dimensões desse espaço são: 7,10 m de comprimento, 7,30 m de largura e 3,5 m de altura, simulando uma sala de aula da Escola de Engenharia da Universidade Federal de Minas Gerais.

Objeto	Eixo x	Eixo y	Eixo z
F_{re}	1,5	3,6	2,0
F_1	-1,5	3,6	2,0
F_2	1,5	-3,6	2,0
F_3	(2xComprimento - 1,5)	3,6	2,0
F_4	1,5	(2xLargura - 3,6)	2,0
F_5	1,5	3,6	-2,0
F_6	1,5	3,6	(2xAltura - 2,0)
Mic_1	3.0	2.5	1.0
Mic_2	4.2	4.0	1.5
Mic_3	5.6	3.4	1.10
Mic_4	3.8	5.2	1.8
Mic_5	6.0	6.5	0.8

Tabela 3.2: Posição das fontes sonoras, real e imaginária, e dos microfones em \mathbb{R}^3 .

Os sinais emitidos pela fonte sonora nas simulações foram: um pulso senoidal modulado por uma gaussiana, um sinal aleatório, um sinal de decaimento exponencial e uma varredura senoidal. Para cada um desses sinais foram simulados os sinais que seriam medidos pelos microfones. E para cada tipo de sinal emitido pela fonte, os seguintes algoritmos foram testados para calcular os TDOAs:

1. Simulação do ambiente bidimensional e tridimensional gerando os sinais dos microfones;
2. Adição do ruído gaussiano branco em cada sinal dos microfones;
3. Remoção do ruído utilizando a transformada discreta de wavelets e o algoritmo de Mallat;
4. Extração dos envelopes dos sinais de cada microfone e do sinal que é emitido pela fonte;
5. Decomposição dos sinais utilizando a transformada contínua de wavelets;
6. Cálculo dos valores de TDOAs utilizando a correlação cruzada entre os sinais da fonte e de cada microfone;

7. Cálculo dos valores de TDOAs utilizando a correlação cruzada generalizada e transformação de fase entre os sinais da fonte e de cada microfone;

Após o cálculo dos valores de TDOAs a posição da fonte foi estimada utilizando mínimos quadrados conforme descrito na Seção 3.3. Como a posição exata da fonte é conhecida, calculou-se o erro percentual da posição estimada, $\tilde{\theta}_f$, em relação à posição exata, θ_f , sendo,

$$\text{Erro} = 100 \left(\frac{\|\tilde{\theta}_f - \theta_f\|}{\|\theta_f\|} \right). \quad (3.34)$$

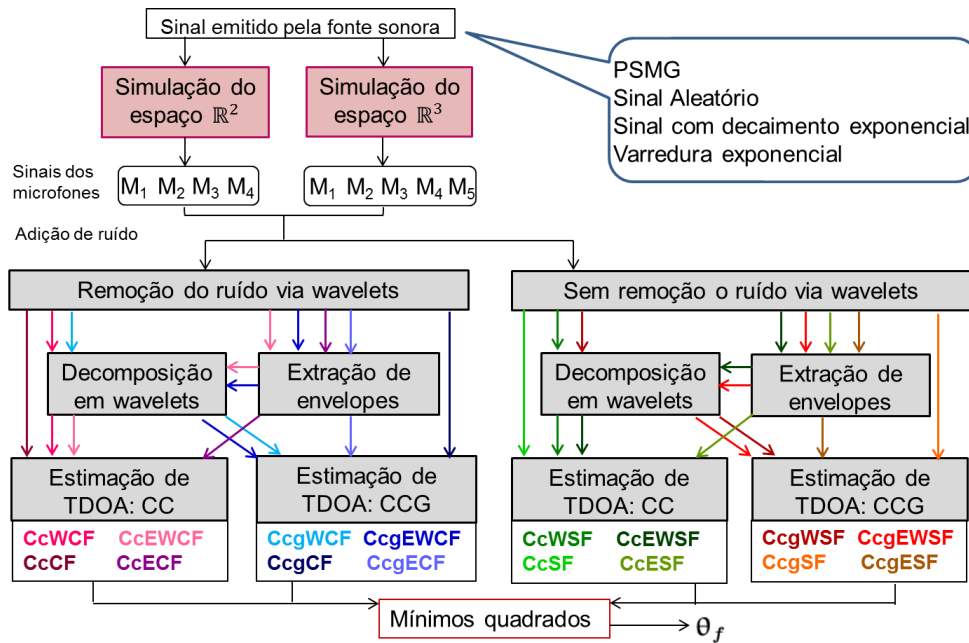


Figura 3.9: Diagrama das simulações realizadas para estimação da posição da fonte sonora no espaço.

Pode-se resumir as simulações realizadas conforme o diagrama da Fig. 3.9. Nesse diagrama há 2 processos principais: estimação de TDOA por correlação cruzada e estimação de TDOA por correlação cruzada generalizada, e 3 processos secundários: remoção de ruído, decomposição em wavelets e extração dos envelopes. Assim, os algoritmos foram implementados com os processos principais separadamente e combinando-os com os demais processos, e em seguida foram comparados ao mesmo procedimento, porém, utilizando os sinais originais, isto é, sinais sem filtrar o ruído. Desse modo, foram implementados 16 algoritmos diferentes, resultando em 16 tipos de TDOAs que foram, em seguida, utilizados para estimação da posição da fonte sonora por mínimos quadrados. Cada tipo de TDOA está representando por uma cor e seu nome é dado pelas iniciais dos algoritmos, sendo:

1. Cc = correlação cruzada;
2. Ccg = correlação cruzada generalizada;
3. W = decomposição em wavelets;
4. E = extração de envelopes;
5. CF = sinal filtrado utilizando wavelets (remoção de ruído);
6. SF = sinal sem filtro (sem remoção de ruído);

Além disso, para as simulações em que o ruído foi adicionado aos sinais dos microfones, as simulações foram executadas por 30 vezes e em seguida o erro da localização da fonte foi calculado.

3.5 Resultados

Os resultados apresentados foram obtidos nas simulações de localização de fontes sonoras descritas na metodologia na Seção 3.4. Inicialmente, tem-se os resultados obtidos no espaço bidimensional, no qual para cada tipo de sinal de excitação sonora há uma subseção. Em seguida tem-se a subseção contendo os resultados do espaço tridimensional. Além disso, para simplificar, os algoritmos testados são identificados apenas por suas iniciais, sendo:

- 1 = CcWSF; 2 = CcWCF; 3 = CcgWSF; 4 = CcgWCF;
- 5 = CcEWSF; 6 = CcEWCF; 7 = CcgEWSF; 8 = CcgEWCF;
- 9 = CcSF; 10 = CcCF; 11 = CcgSF; 12 = CcgCF;
- 13 = CcESF; 14 = CcECF; 15 = CcgESF; 16 = CcgECF,

no qual as letras significam:

- Cc = função de correlação cruzada;
- Ccg = função de correlação cruzada generalizada utilizando a transformação de fase;
- W = decomposição em wavelets contínuas;
- SF = sinais dos microfones utilizados para cálculo dos TDOAs apresentam ruído e não são filtrados através das wavelets;

- CF = sinais dos microfones são filtrados através das wavelets para minimização do ruído incorporado a eles;
- E = envelope extraídos dos sinais dos microfones.

Todos os resultados apresentados nesse capítulo de localização de fontes sonoras foram obtidos utilizando uma frequência de amostragem de $f_s = 40$ kHz.

3.5.1 Pulso senoidal com modulação gaussiana

A Fig. 3.10 representa um pulso senoidal modulado por uma gaussiana, PSMG ($f_c = 1000$ Hz e $bw = 2$), e os sinais de resposta a esse sinal nos 4 microfones presentes no espaço bidimensional, sem considerar as reflexões, à esquerda, e considerando as reflexões presentes no local com um coeficiente de atenuação de 0,6, à direita.

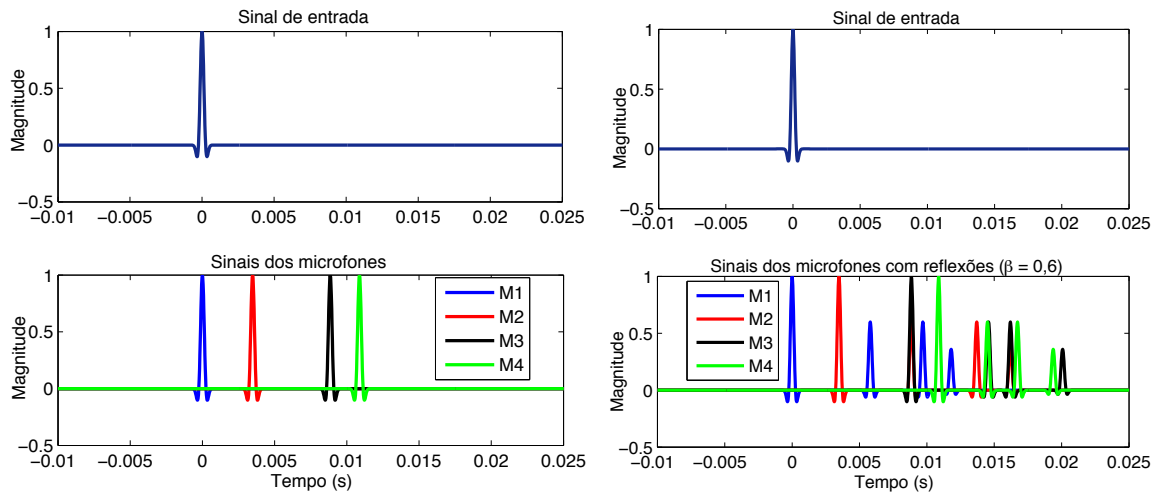


Figura 3.10: Sinal emitido pela fonte sonora, PSMG, e sinais de resposta simulados para os 4 microfones em \mathbb{R}^2 sem considerar as reflexões à esquerda, e considerando as reflexões à direita.

Utilizando o sinal de entrada da Fig. 3.10 à esquerda, aplicou-se a wavelet de Meyer, dmey, para decompor os sinais em escalas de frequência discretas de 1 a 5. Tal wavelet foi escolhida devido a semelhança da sua curva com a curva do pulso senoidal modulado pela gaussiana. E, alterando o coeficiente de atenuação, β (Eq. 3.3), isto é, variando a amplitude das reflexões, os erros relativos da estimação da posição da fonte são apresentadas na Tab. 3.3. Nas duas linhas iniciais da última coluna dessa tabela, NC significa que não foi possível estimar a posição da fonte sonora para o algoritmo CcgESF, pois a matriz D na estimação por mínimos quadrados é singular. O mesmo ocorre para o algoritmo CcgEWSF em todos os valores de β , por isso ele não foi apresentado nessa tabela. Observa-se que para valores de atenuação abaixo de 0,8 os

algoritmos CcWSF, CcSF, CcgSF e CcESF apresentam erros percentuais menores do que 1% na determinação da posição da fonte sonora. Mas para $\beta = 0,98$ e $\beta = 0,999$ apenas o método da correlação cruzada tradicional foi capaz de identificar a posição da fonte com um erro de 0,35%, enquanto os demais algoritmos apresentaram erro superior a 90%.

β	CcWSF	CcgWSF	CcEWSF	CcSF	CcgSF	CcESF	CcgESF
0	0,35	132,25	92,66	0,35	0,35	0,35	NC
0,4	0,72	132,25	1249,84	0,35	0,35	0,35	NC
0,6	0,72	132,25	88,62	0,35	0,35	0,35	111,65
0,8	0,72	132,25	174,55	0,35	0,35	4,02	88,13
0,98	59,99	109,68	50,32	0,35	243,29	216,19	6160,52
0,999	93,95	107,94	152,32	0,35	255,99	97,85	6160,52

Tabela 3.3: Erros (%) relativos da estimação da posição da fonte sonora quando o sinal da fonte sonora é o PSMG considerando alguns valores do coeficiente de atenuação (β).

Adiciona-se um ruído gaussiano branco com desvio padrão de 0,2 (SNR = 0,3920 db) aos sinais dos microfones da Fig. 3.10 à esquerda (sem reflexões) e obtêm-se os sinais da Fig. 3.11.

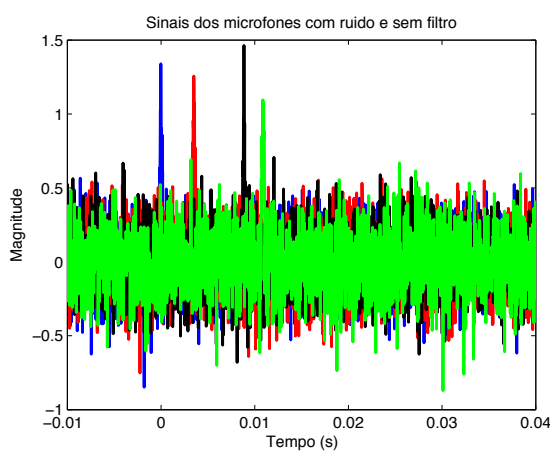


Figura 3.11: Sinais dos microfones sem considerar as reflexões e com um ruído de $\sigma = 0,2$.

Para minimizar o ruído utilizou-se a wavelet dmey com *thresholding hard*, um limiar de 0,18 e 5 níveis de decomposição. Ao repetir esse procedimento por 30 vezes, o erro relativo da fonte é apresentado na Fig. 3.12 à esquerda, e os melhores resultados estão nessa figura à direita. Já a Fig. 3.13 segue esse mesmo processo, porém, os sinais dos microfones contêm reflexões com o coeficiente de atenuação de 0,4, então, a relação sinal ruído média é de 0,58 db. Essas imagens com os valores de erros relativos são *boxplot* geradas através do *software* Matlab[®]. Assim, os retângulos verticais azuis são o terceiro quartil (parte superior) e primeiro quartil (parte inferior), as barras horizontais

em vermelho representam a mediana do erro (ou segundo quartil) para aquele tipo de algoritmo testado, a linha contínua azul mostra o valor inferior e superior e os sinais "+" vermelhos representam os valores discrepantes (*outliers*), isto é, os valores que estão fora do limite superior e inferior. Em ambas as figuras todos os algoritmos que aplicam a decomposição dos sinais através da transformada contínua de wavelets tiveram erros acima de 100%.

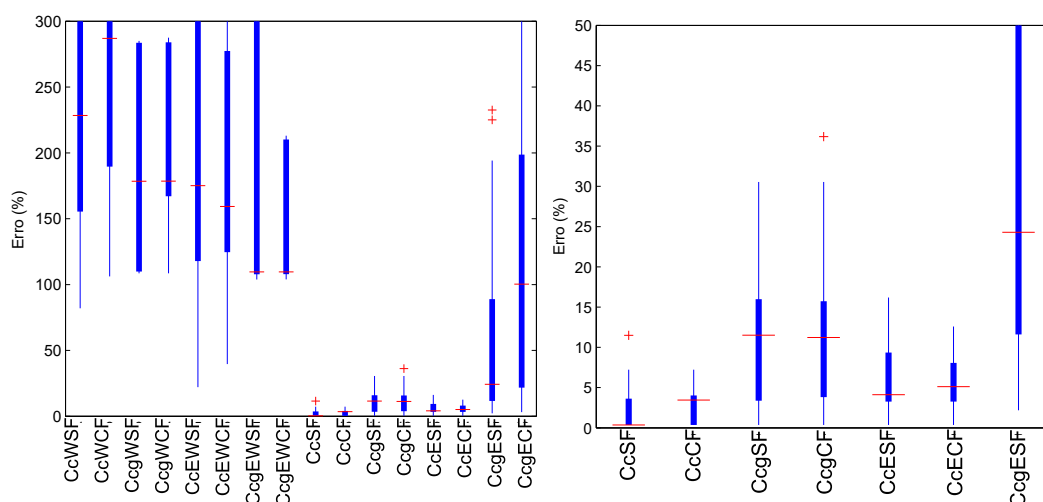


Figura 3.12: Erros relativos de todos os algoritmos testados para os sinais dos microfones com ruído de $\sigma = 0,2$ e sem a presença de reflexões quando o sinal de entrada é o PSMG à esquerda, e os melhores resultados à direita.

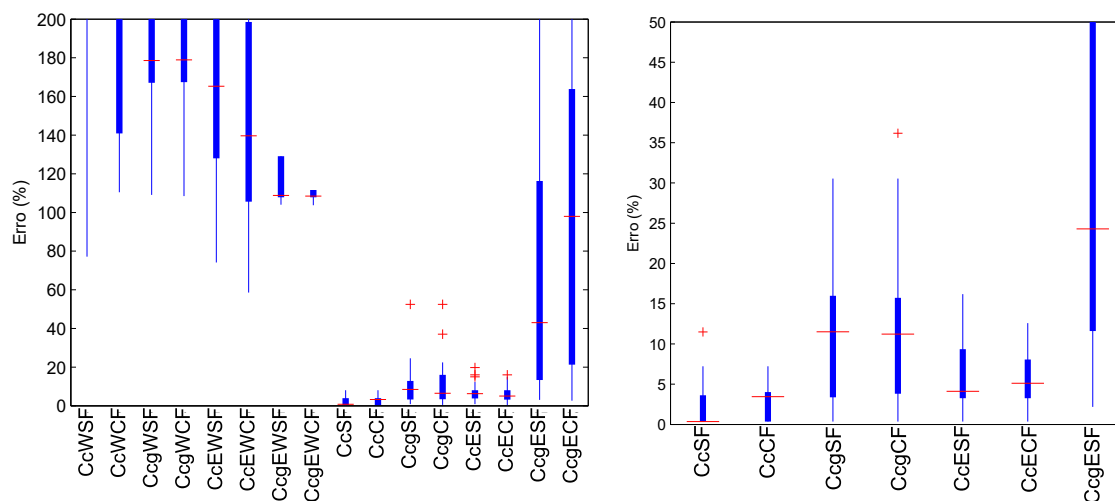


Figura 3.13: Erros de todos os algoritmos testados para sinais dos microfones com ruído de $\sigma = 0,2$ e $\beta = 0,4$ quando o sinal de entrada é o PSMG à esquerda, e os melhores resultados à direita.

3.5.2 Sinal Aleatório

A Fig. 3.14 representa um sinal aleatório, SA, com $\sigma = 1$, e os sinais de resposta a esse sinal nos 4 microfones presentes no espaço bidimensional, sem considerar as reflexões presentes no local.

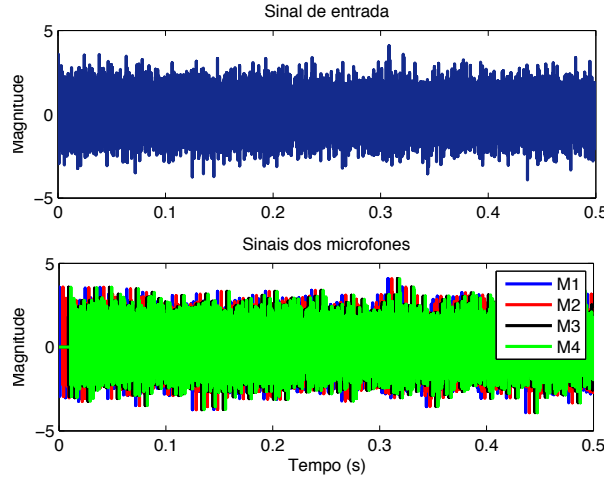


Figura 3.14: Sinal emitido pela fonte sonora, sinal aleatório, e sinais de resposta simulados para os 4 microfones em \mathbb{R}^2 sem considerar as reflexões locais.

β	CcWSF	CcgWSF	CcEWSF	CcgEWSF	CcSF	CcgSF	CcgESF
0	0,35	132,25	0,35	109,23	0,35	0,35	0,35
0,4	0,35	132,25	0,35	109,61	0,35	0,35	0,35
0,6	0,35	389,57	0,35	103,92	0,35	0,35	0,35
0,8	0,35	222,62	0,35	107,85	0,35	0,35	0,35
0,98	1028,64	107,11	40,76	111,65	0,35	0,35	107,67
0,999	813,36	107,67	121,54	108,05	235,30	0,35	147,88

Tabela 3.4: Erros (%) relativos da estimação da posição da fonte sonora quando o sinal da fonte sonora é aleatório e os sinais dos microfones não possuem ruídos.

Utilizou-se a wavelet de Meyer, *dmey*, para decompor os sinais da Fig. 3.14 em escalas de frequência discretas de 1 a 5. Adicionando reflexões com diferentes valores do coeficiente de atenuação (β), os erros relativos da estimação da posição da fonte são apresentadas na Tab. 3.4, sendo que os erros para o algoritmo CcESF são iguais ao CcgSF, por isso eles não constam na tabela. Para os valores de $\beta = 0, 0,4, 0,6$ e $0,8$ os algoritmos CcWSF, CcEWSF, CcSF, CcgSF e CcgESF tiveram erro de 0,35%, enquanto que os demais algoritmos da tabela mostram erros superiores a 100%. Já para $\beta = 0,98$ somente os erros dos algoritmos CcSF e CcgSF é inferior a 10%, sendo 0,35%, e por fim, quando $\beta = 0,999$ o erro de CcgSF é 0,35% e os demais são maiores do que 107%. Esses resultados indicam que quando o sinal da fonte sonora é aleatório e

os sinais medidos pelos microfones não contém ruído, mas possuem reflexões, o melhor algoritmo para determinar a posição da fonte sonora é a correlação cruzada generalizada com transformação de fase.

Adicionando aos sinais dos microfones da Fig. 3.14 um ruído gaussiano branco de $\sigma = 0,2$ e reflexões com atenuação de 0,4, $\text{SNR} = 15,36$ db, tem-se os sinais resultantes na Fig. 3.15. A fim de minimizar o ruído desses sinais escolheu-se a wavelet dmey, com *thresholding hard* e limiar de 0,18, e as diferenças entre os sinais com ruídos e os sinais filtrados utilizando essa wavelet estão na Fig. 3.16. Nota-se que tais diferenças são da mesma ordem de grandeza do desvio padrão do ruídos, portanto, a wavelet é capaz de minimizar os ruídos nos sinais. Apesar disso, o desempenho da localização da fonte sonora não é significativamente alterado, visto que para o sinal aleatório, os algoritmos de correlação cruzada (CC) são capazes de determinar a posição da fonte mesmo com a presença do ruído nos sinais. Isso pode ser comprovando pelo fato de que ao calcular por 30 vezes todos os algoritmos testados nesse trabalho, o valor de estimação da posição da fonte foi o mesmo em cada algoritmo. Para os métodos CcWSF, CcWCF, CcEWSF, CcEWCF, CcSF, CcCF, CcgSF, CcgCF, CcESF, CcECF, CcgESF e CcgECF o erro foi igual a 0,3502%, já para os métodos CcgWSF e CcgWCF foi 132,25% e para os métodos 7 e 8 foram em média superior a 1000%. Os erros elevados são justificados porque o cálculo de estimação de TDOA a partir do envelope extraído dos sinais medidos e a decomposição em wavelets contínuas são aplicados para baixas frequências de amostragem, o que não se aplica nesse caso.

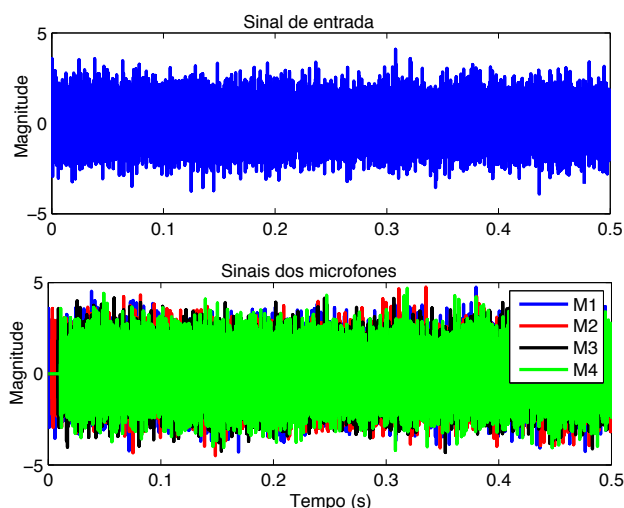


Figura 3.15: Sinal aleatório da fonte sonora e sinais dos microfones com $\sigma = 0,2$ e $\beta = 0,4$.

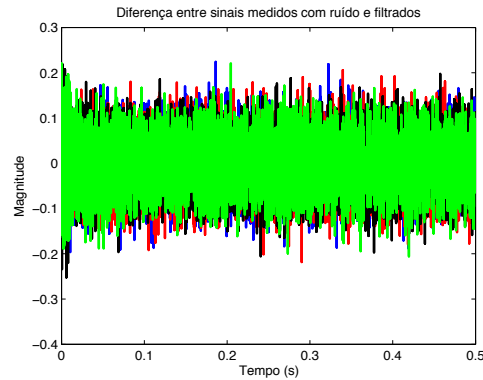


Figura 3.16: Diferença entre sinais dos microfones com ruídos, $\sigma = 0,2$ e $\beta = 0,4$, e os filtrados pela wavelet.

3.5.3 Sinal com decaimento exponencial

As figuras a seguir foram geradas utilizando a $f_s = 40$ kHz e considerando um sinal de decaimento exponencial, SD, de 100 Hz de frequência e um fator de decaimento de 0,9991 como sinal emitido pela fonte sonora. A Fig. 3.17 representa esse sinal e suas repostas em cada microfone sem considerar as reflexões presentes no local.

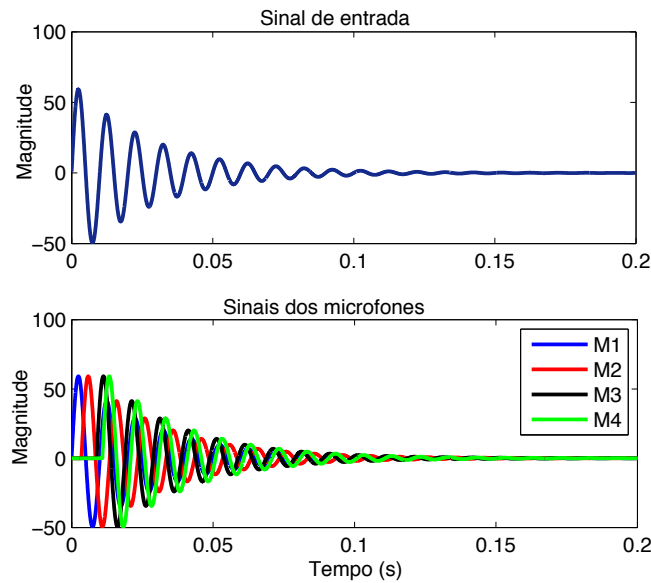


Figura 3.17: Sinal emitido pela fonte sonora, decaimento exponencial, e sinais de resposta simulados para os 4 microfones em \mathbb{R}^2 sem considerar as reflexões locais.

Considerando os sinais da Fig. 3.17 e utilizando a wavelet dmey para decompor o sinal em escalas de frequência discretas de 1 a 5, ao adicionar reflexões com diferentes valores de atenuação, os erros relativos da estimação da posição da fonte são apresentadas na Tab. 3.5. Porém, os algoritmos CcEWSF, CcgEWSF e CcgESF não foram capazes de localizar a fonte sonora independente do valor do coeficiente de atenua-

ção, por isso eles não são apresentados nessa tabela. Observando essa tabela pode-se concluir que utilizando o sinal de decaimento exponencial como sinal de teste e considerando que o ambiente gera reflexões com amplitudes atenuadas em relação ao som direto, o melhor algoritmo para determinar a posição da fonte sonora é a correlação cruzada generalizada com transformação de fase (CcgSF).

β	CcWSF	CcgWSF	CcSF	CcgSF	CcESF
0	0,35	132,96	0,35	0,35	19,30
0,4	24,97	132,96	86,83	0,35	316,09
0,6	35,38	132,96	89,82	0,35	190,30
0,8	114,92	132,96	44,54	0,35	185,49
0,98	100,66	125,57	53,55	0,35	151,81
0,999	100,68	125,57	54,66	156,73	152,14

Tabela 3.5: Erros (%) relativos da estimação da posição da fonte sonora quando o sinal da fonte é o sinal de decaimento exponencial com frequência de 100 Hz e não há ruído nos sinais dos microfones.

Adicionando ruído de $\sigma = 0,2$ e reflexões com $\beta = 0,4$ aos sinais da Fig. 3.17, tem-se a Fig. 3.18 com uma relação sinal ruído média de 30,54 db. A fim de minimizar tal ruído escolheu-se a wavelet *dmey*, com *thresholding hard* e limiar de 0,18. Nesse caso, os melhores resultados obtidos para os erros relativos na estimação da posição da fonte sonora estão na Fig. 3.19. Nesses resultados os algoritmos CcgSF e CcgCF resultaram em erros de 0,3502% nas 22 vezes que foram executados e um erro máximo de 7,2322%.

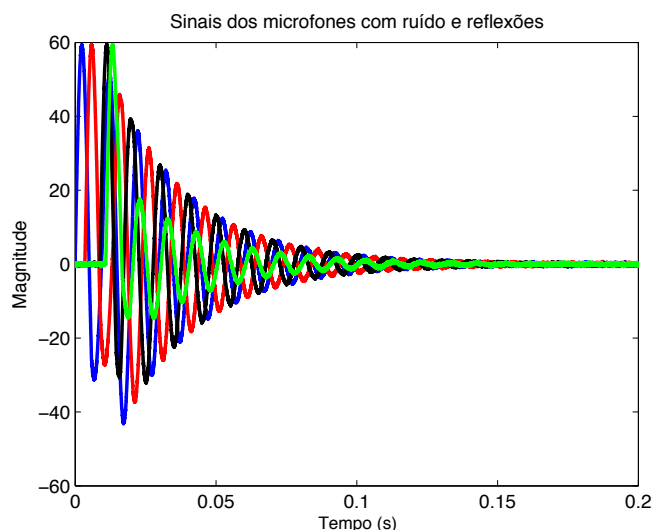


Figura 3.18: Sinais dos microfones com ruído de $\sigma = 0,2$ e $\beta = 0,4$ quando o sinal da fonte sonora é um decaimento exponencial.

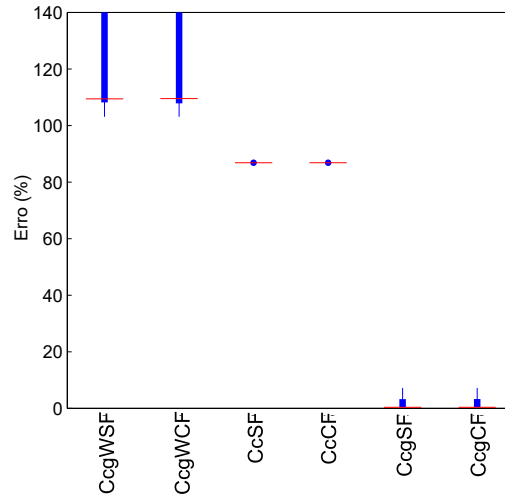


Figura 3.19: Melhores resultados de erros na estimação da posição da fonte sonora quando os sinais dos microfones possuem $\sigma = 0,2$ e $\beta = 0,4$ e o sinal de teste é o SD.

3.5.4 Varredura exponencial

Uma varredura exponencial com $\omega_1 = 2\pi i100$ e $\omega_2 = 2\pi i1000$ [rad/s] e amplitude igual a 0,9, VE, foi gerada e utilizada como sinal emitido pela fonte sonora para gerar as demais figuras dessa seção.

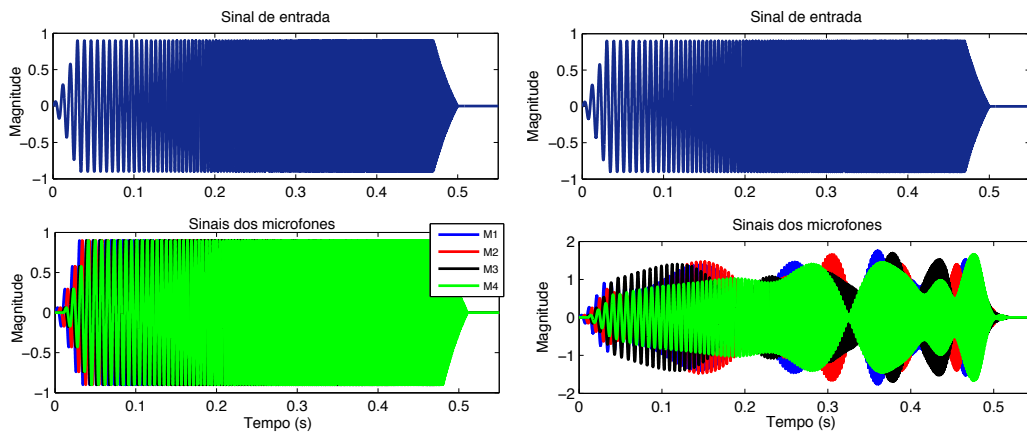


Figura 3.20: Sinal emitido pela fonte sonora, varredura exponencial, e sinais de resposta simulados para os 4 microfones com $\beta = 0$ (à esquerda) e $\beta = 0,4$ (à direita).

A Fig. 3.20 à esquerda mostra o sinal emitido pela fonte sonora, varredura exponencial, e os sinais medidos pelos microfones sem considerar as reflexões presentes no ambiente. Já a Fig. 3.20 à direita apresenta esse mesmo sinal da fonte, porém, aos sinais dos microfones são acrescentadas reflexões com um coeficiente de atenuação de 0,4. Assim, utilizando os sinais da Fig. 3.20 à esquerda, e a wavelet dmey para decompor o sinal em escalas de frequência discretas de 1 a 5, ao adicionar reflexões com diferentes valores de atenuação (β), os erros relativos da estimação da posição da fonte são apre-

sentadas na Tab. 3.6. Nessa tabela, não são apresentados os algoritmos CcgEWSF e CcgESF porque eles não foram capazes de localizar a posição da fonte sonora. Para $\beta = 0, 0,4, 0,6$ e $0,8$ os melhores algoritmos foram CcWSF, CcSF e CcgSF com erros de 0,35%, e para $\beta = 0,98$ apenas o algoritmo CcgSF é capaz de estimar a posição da fonte sonora com erro de 0,35%, enquanto os demais algoritmos apresentam erro superior a 108%.

β	CcWSF	CcgWSF	CcgEWSF	CcSF	CcgSF	CcESF
0	0,35	132,25	99,75	0,35	0,35	0,35
0,4	0,35	132,25	108,15	0,35	0,35	122,04
0,6	0,35	132,25	108,36	0,35	0,35	144,44
0,8	0,35	132,25	108,34	0,35	0,35	162,09
0,98	114,26	132,25	108,41	107,89	0,35	183,73
0,999	1107,40	95,36	215,89	107,89	156,73	220,08

Tabela 3.6: Erros (%) relativos da estimação da posição da fonte sonora para a varredura exponencial como sinal de teste e sinais dos microfones sem ruído.

Adicionando um ruído gaussiano branco com $\sigma = 0,2$ aos sinais dos microfones da Fig. 3.20 à direita ($\beta = 0,4$), tem-se uma relação sinal ruído média de 10,46 db. Os erros relativos a posição da fonte sonora foram calculados e apresentados na Fig. 3.21 à esquerda. Os melhores resultados são apresentados novamente na Fig. 3.21 à direita com uma escala menor. Nota-se que nesse caso, ou seja, ao utilizar a varredura exponencial como sinal de teste, desde que haja ruído nos sinais medidos, o melhor algoritmo é a correlação cruzada tradicional.

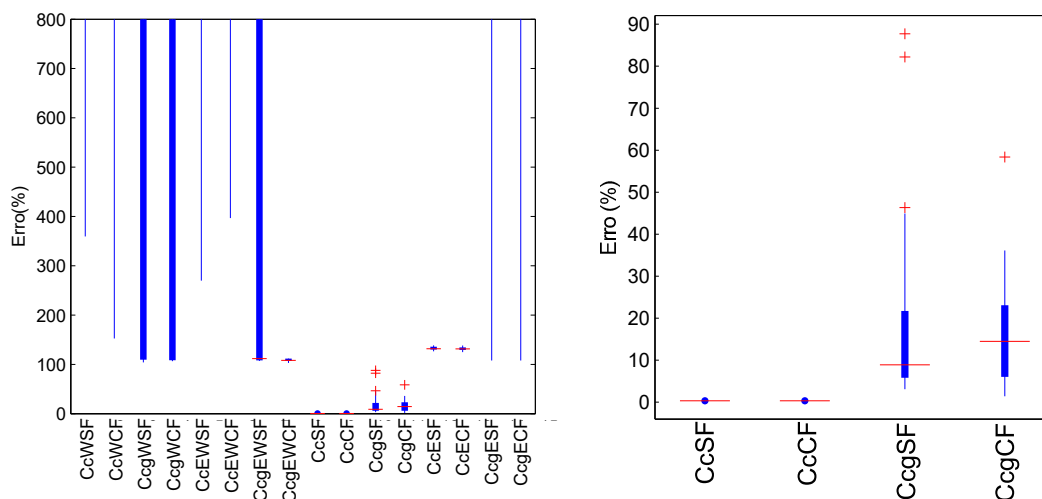


Figura 3.21: Resultados para todos os algoritmos testados quando o sinal da fonte é a varredura exponencial e os sinais dos microfones possuem $\sigma = 0,2$ e $\beta = 0,4$ (à esquerda) e melhores resultados (à direita)

3.5.5 Simulações no espaço tridimensional

Os resultados apresentados a seguir referem-se ao espaço tridimensional, \mathbb{R}^3 , descrito nesse capítulo na seção 3.4, e os algoritmos testados foram os mesmos do espaço bidimensional.

Considerando que não existam reflexões e ruído natural no recinto, o erro na estimação da posição da fonte sonora é apresentado na Tab. 3.7 para todos os sinais de teste, pulso senoidal modulado por uma gaussiana (PSMG), sinal aleatório (SA), sinal de decaimento (SD) e varredura exponencial (VE). Para obter esses resultados, utilizou-se a wavelet meyer, e uma escala de frequências discretas de 1 a 5, para decompor os sinais e seus envelopes através desta wavelet. Nessa tabela não são apresentados os valores obtidos para o algoritmo CcgSF, pois esses valores são iguais aos do CcSF. Nota-se que o algoritmo CcgWSF apresentou valores de erros elevados em todos os sinais de teste, enquanto que os algoritmos CcWSF, CcSF e CcgSF apresentaram menos de 0,1% de erro nos sinais PSMG, SD e SA. Além disso, ao utilizar para teste o sinal aleatório ou a varredura exponencial, quase todos os algoritmos conseguiram estimar a posição da fonte com erro inferior a 0,5%, com exceção dos algoritmos CcgWSF e CcEWSF.

Sinal	CcWSF	CcgWSF	CcEWSF	CcgEWSF	CcSF	CcESF	CcgESF
PSMG	0,09	48,51	0,09	NC	0,09	0,09	NC
SD	0,09	48,51	NC	NC	0,09	5,50	NC
SA	0,09	48,51	0,09	48,51	0,09	0,09	0,09
VE	0,35	222,62	0,35	107,85	0,35	0,35	0,35

Tabela 3.7: Erros (%) relativos da estimação da posição da fonte sonora para todos os sinais de teste no espaço tridimensional.

A Fig. 3.22 representa a varredura exponencial emitida pela fonte e os sinais simulados para os cinco microfones quando o coeficiente de atenuação do ambiente é $\beta = 0,4$. O erro da localização da fonte quando esse sinal e os demais sinais de teste são utilizados são apresentados na Tab. 3.8 e 3.9 para os valores de $\beta = 0,4$ e $\beta = 0,7$. Nota-se que utilizando o pulso senoidal modulado por uma gaussiana, PSMG, e considerando reflexões com $\beta = 0,4$, os melhores resultados são dos algoritmos CcWSF, CcSF e CcgSF com erros de 0,09%. Porém, para esse mesmo sinal quando $\beta = 0,7$ não é possível determinar a posição da fonte sonora com erro inferior a 89,4%. Ao utilizar o sinal com decaimento exponencial o erro foi de 0,09% para o algoritmo CcgSF quando $\beta = 0,4$. Já para o sinal aleatório, os erros foram de 0,09% ao executar os algoritmos CcWSF, CcEWSF, CcSF e CcgSF quando os sinais dos microfones consideravam as reflexões com $\beta = 0,4$. Mas, para esse mesmo sinal, quando $\beta = 0,7$ os erros foram maiores que 102%. Utilizando a varredura senoidal, os menores erros ocorreram quando

$\beta = 0,4$, nos algoritmos CcWSF (erro = 2,68%), CcSF (erro = 1,1%) e CcgSF (erro = 0,09%).

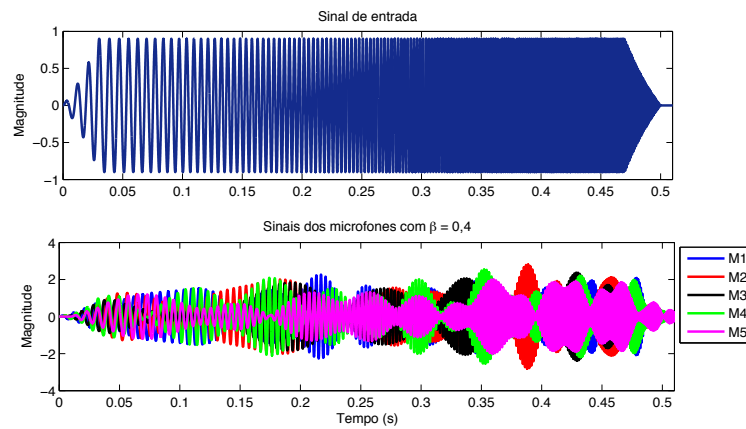


Figura 3.22: Sinal da fonte sonora, varredura exponencial, e sinais dos microfones com $\beta = 0,4$ e sem ruído em \mathbb{R}^3 .

Tipo de sinal	β	CcWSF	CcgWSF	CcEWSF	CcgEWSF
PSMG	0,4	0,09	48,51	126,21	NC
PSMG	0,7	99,81	129,37	89,4	1 NC
SD	0,4	198,42	48,51	NC	NC
SD	0,7	119,02	129,37	NC	NC
SA	0,4	0,09	48,51	0,09	48,51
SA	0,7	102,67	129,37	102,67	129,37
VExp	0,4	2,68	48,51	NC	NC
VExp	0,7	124,41	129,37	NC	NC

Tabela 3.8: Erros (%) relativos da estimação da posição da fonte sonora para os algoritmos que utilizam decomposição em wavelets no espaço tridimensional.

Tipo de sinal	β	CcSF	CcgSF	CcESF	CcgESF
PSMG	0,4	0,09	0,09	1,27	282,75
PSMG	0,7	102,67	102,67	102,86	282,75
SD	0,4	192,92	0,09	219,57	NC
SD	0,7	137,10	102,67	198,03	NC
SA	0,4	0,09	0,09	0,09	0,09
SA	0,7	102,67	102,67	102,67	102,67
VExp	0,4	1,11	0,09	1850,54	NC
VExp	0,7	140,91	102,67	150,28	NC

Tabela 3.9: Erros (%) relativos da estimação da posição da fonte sonora para todos os sinais de teste no espaço tridimensional considerando as reflexões do ambiente.

Adicionando, 30 vezes, um ruído gaussiano branco com $\sigma = 0,2$ aos sinais dos microfones obtidos na simulação quando o sinal de teste foi a varredura exponencial

sem considerar as reflexões do local e com relação sinal ruído média de 9,37 db, todos os algoritmos foram capazes de estimar a posição da fonte sonora. Porém alguns deles apresentaram erro superior a 100%, e assim seus resultados não serão apresentados. Para os demais, o valor do erro obtido encontra-se na Fig. 3.23 à esquerda. Nesse caso, utilizou-se a wavelet db4 e um limiar de 0,18 para remover o ruído, e para decompor os sinais utilizou-se a wavelet meyer e as frequências das escalas de 1 a 5. Observa-se que o erro obtido utilizando o TDOA calculado por correlação cruzada foi um único valor nas 30 simulações, independente da presença do ruído, sendo igual a 0,087%. Utilizando as mesmas wavelets descritas na obtenção da Fig. 3.23 à esquerda (db4 e meyer), porém com um ruído de $\sigma = 0,4$ e um limiar igual a 0,39, tem-se os resultados apresentados na Fig. 3.23 à direita. Nessa figura são apresentados apenas os algoritmos que foram capazes de localizar a fonte sonora com um erro inferior a 50%.

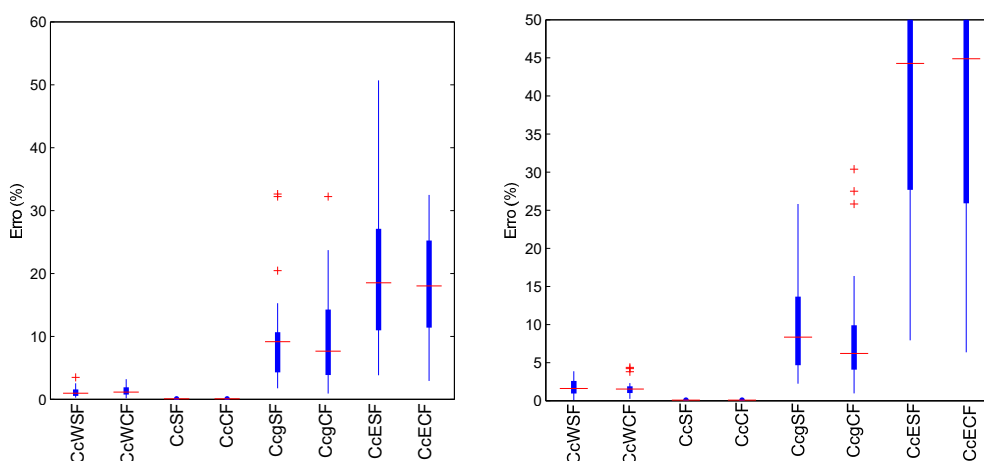


Figura 3.23: Resultados de alguns dos algoritmos testados para os sinais dos microfones com $\sigma = 0,2$ à esquerda e $\sigma = 0,4$ à direita.

A Fig. 3.24 apresenta o sinal de decaimento exponencial e os sinais medidos pelos microfones quando a fonte sonora emite tal sinal na simulação do espaço tridimensional sem considerar as reflexões presentes no ambiente. Aos sinais dos microfones são adicionados ruídos gaussianos brancos com $\sigma = 0,2$, $\text{SNR} = 31,63$ db, e utilizou-se a wavelet de dmey com um limiar de 0,18 para remover esse ruído e a wavelet meyer para decompor os sinais em uma escala de frequências discretizadas de 1 a 5. Realizando esse procedimento por 30 vezes, o valor do erro relativo é apresentado na Fig. 3.25, sendo que os algoritmos CcEWSF, CcEWCF, CcgEWSF, CcgEWCF, CcgESF e CcgECF não conseguiram determinar a posição da fonte sonora porque as matrizes utilizadas no método dos mínimos quadrados são singulares. Já os algoritmos CcgWSF e CcgWCF tiveram a média de erro superior a 50%, por isso, esses algoritmos não são apresentados nessa figura. Os melhores resultados foram obtidos para os algoritmos CcSF e CcCf considerando que não há dispersão dos resultados, enquanto que

os algoritmos CcgSF e CcgCF tiveram um bom resultado considerando suas medianas, mas, maior dispersão dos erros calculados.

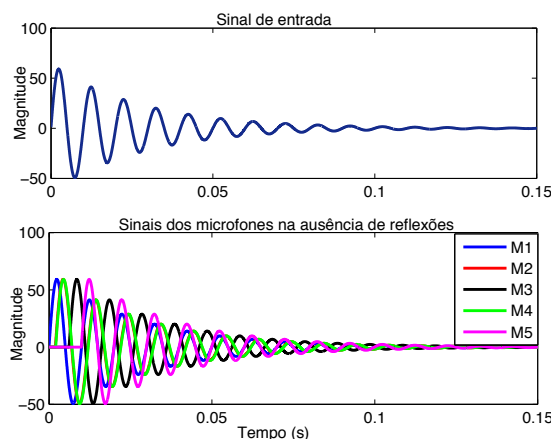


Figura 3.24: Sinal emitido pela fonte, SD, e sinais medidos pelos microfones no espaço tridimensional.

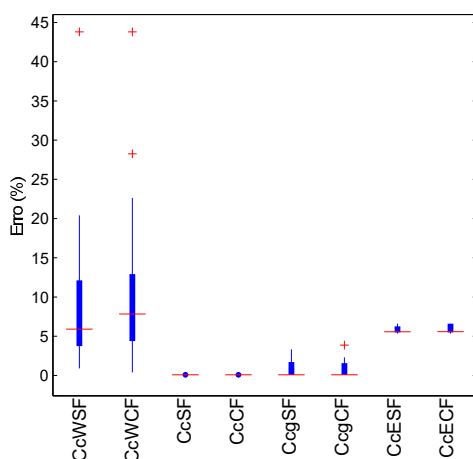


Figura 3.25: Resultados para alguns dos algoritmos testados quando os sinais dos microfones possuem ruído com $\sigma = 0,2$, para o sinal de decaimento exponencial como sinal de teste.

3.6 Conclusões

A localização de fontes sonoras é a estimativa da posição da fonte sonora em relação a um ponto no espaço. Ela pode ser realizada através do cálculo dos TDOAs (diferença de tempo de chegada entre os microfones), sendo a correlação cruzada e a correlação cruzada generalizada os métodos mais clássicos para estimação destes. A fim de estudar esses métodos foi simulado um espaço bidimensional e tridimensional (sala). Além disso, avaliou o método de extração de envelopes dos sinais que simulam os sinais medidos pelos microfones e posterior decomposição desses sinais em wavelets contínuas

para finalmente estimar a diferença de tempo de chegada, TDOA, entre os microfones e determinar a posição da fonte. Foram testados vários sinais de testes e em algumas simulações adicionou-se um ruído gaussiano branco aos sinais dos microfones para avaliar o processo de remoção de ruído via wavelets.

Os resultados obtidos foram satisfatórios para os métodos tradicionais de correlação cruzada e em alguns casos para a correlação cruzada generalizada. Quanto ao ruído, a utilização de wavelet possibilitou que ele fosse minimizado, mas isso não melhorou os resultados, visto que apesar do ruído, a correlação cruzada foi capaz de localizar a fonte com um erro inferior a 1% na maioria dos casos. Já para os métodos utilizando CCG-PHAT, talvez o resultado não tenha sido satisfatório porque as reflexões presentes no sinal não eram em quantidades elevadas quando comparadas com um ambiente reverberante real. Para o método de decomposição em wavelets contínua, na literatura, o método é proposto para baixas frequências e considerando uma expansão da wavelet mãe por um fator de deslocamento no tempo igual ao período de amostragem, T_a , $b = T_a$, e por um escalonamento variável. Porém, nesse trabalho, o deslocamento no tempo da wavelet mãe foi igual a 1, $b = 1$, e a frequência de amostragem escolhida foi elevada.

Por fim, para o espaço bidimensional sem considerar as reflexões geradas devido as superfícies desse espaço, os melhores resultados obtidos encontram-se na Tab. 3.10. Adicionando ruído gaussiano branco aos sinais, utilizando o PSMG, SA e VE como sinais de teste, os menores erros foram obtidos utilizando o método da correlação cruzada tradicional sem remover o ruído dos sinais (CcSF) e removendo o ruído através das wavelets (CcCF). Mas ao utilizar o sinal de decaimento, SD, como sinal emitido pela fonte, o algoritmo com menor erro foi a correlação cruzada generalizada com transformação de fase.

Nas simulações do espaço tridimensional, desconsiderando as reflexões presentes no ambiente, os menores erros foram obtidos nos algoritmos CcWSF, CcSF e CcgWSF. Nas presença de ruído, quando o sinal de teste foi a VE os melhores foram CcWSF, CcWCF, CcSF e CcCF, e quando foi o SD, os menores erros obtidos ocorreram para os algoritmos CcSF, CcCF, CcgSF e CcgCF.

β		PSMG	SA	SD	VE
0	Algoritmo	CcWSF CcSF CcgSF CcESF	CcWSF CcEWSF CcSF CcESF CcgSF CcgESF	CcWSF CcSF CcgSF	CcWSF CcSF CcgSF CcESF
	Erro(%)	0,35	0,35	0,35	0,35
0,4	Algoritmo	CcSF CcgSF CcESF	CcWSF CcEWSF CcSF CcESF CcgSF CcgESF	CcgSF	CcWSF CcSF CcgSF
	Erro(%)	0,35	0,35	0,35	0,35
0,6	Algoritmo	CcSF CcgSF CcESF	CcWSF CcEWSF CcSF CcESF CcgSF CcgESF	CcgSF	CcWSF CcSF CcgSF
	Erro(%)	0,35	0,35	0,35	0,35
0,8	Algoritmo	CcSF CcgSF	CcWSF CcEWSF CcSF CcESF CcgSF CcgESF	CcgSF	CcWSF CcSF CcgSF
	Erro(%)	0,35	0,35	0,35	0,35
0,98	Algoritmo	CcSF	CcSF CcgSF	CcgSF	CcgSF
	Erro(%)	0,35	0,35	0,35	0,35
0,98	Algoritmo	CcSF	CcgSF	CcSF	CcgWSF
	Erro(%)	0,35	0,35	54,66	95,36

Tabela 3.10: Menores erros (%) da estimação da posição da fonte e seus respectivos algoritmos para simulação do espaço bidimensional.

Capítulo 4

Considerações finais

4.1 Conclusões gerais

Os objetivos desse trabalho foram remover o ruído presente em sinais que simulam sinais medidos em tubo de impedância para posterior cálculo do coeficiente de absorção sonora, e comparar os métodos mais clássicos, existentes na literatura, de estimação da posição da fonte sonora. Assim, este trabalho utilizou métodos baseados na transformada discreta de wavelets para remover o ruído de sinais gerados por um programa de simulação numérica do tubo de impedância. Algumas famílias de wavelets foram testadas, sendo as melhores a db, coif e sym. O resultado foi mais satisfatório nas baixas frequências, ou seja, na faixa que o alto-falante utilizado produz uma baixa relação sinal ruído. Constatou-se que a diferença entre os sinais original e filtrado dos microfones foi próxima ao desvio padrão do ruído, e as wavelets se mostraram capazes de melhorar curvas de absorção estimadas em tubos de impedâncias. Além disso, este trabalho simulou um espaço bidimensional e tridimensional para estudar os métodos de correlação cruzada e correlação cruzada generalizada com transformação de fase na estimação da posição da fonte sonora. Nesses métodos foram incorporados os processos de extração de envelopes dos sinais e decomposição em wavelets, porém, tais processos não apresentaram vantagens significativas para esse tipo de estimação, visto que os métodos tradicionais obtiveram bons resultados na maioria dos testes realizados. Analisou-se também a possibilidade de remover ruído dos sinais medidos por microfones presentes nos dois espaços (\mathbb{R}^2 e \mathbb{R}^3) através das wavelets, entretanto, apesar das wavelets serem capazes de remove o ruído, não houve melhoria relevante.

4.2 Sugestões de trabalhos futuros

1. Estudar outros tipos de regras de *thresholding* para utilizar as transformadas wavelets na remoção de ruído;
2. Realizar experimentos para validar os resultados obtidos na estimação da localização da fonte sonora;
3. Estudar métodos de remoção de reflexões acústicas.

4.3 Publicações

1. LARA, L. T.; BOAVENTURA, W. C. ; PASQUAL, A. M. . Aplicação de wavelets para reduzir ruído em curvas de absorção acústica obtidas em tubos de impedância (*Improving the estimated acoustic absorption curves in impedance tubes by using wavelet-based denoising methods*). In: X Congresso Iberoamericano de Acústica, 2016, Buenos Aires. Proceedings of the 22nd International Congress on Acoustics.. Buenos Aires: Asociación de Acústicos Argentinos, 2016. p. 1-10.
2. LARA, L. T.; PASQUAL, A. M. ; VECCHI, M. A. M. . The effects of the excitation source directivity on some room acoustic descriptors obtained from impulse response measurements. In: 22nd International Congress on Acoustics, 2016, Buenos Aires. Proceedings of the 22nd International Congress on Acoustics. Buenos Aires: Asociación de Acústicos Argentinos, 2016. p. 1-10.
3. PASQUAL, ALEXANDER MATTIOLI ; LARA, LUANA TORQUETE . Time-domain simulation of acoustic impedance tubes. Journal of the Brazilian Society of Mechanical Sciences and Engineering (Impresso), v. 39, p. 67-79, 2017.

Referências Bibliográficas

- [1] (1998). *Acoustics – Determination of sound absorption coefficient and impedance in impedance tubes — Part 2: transfer — function method*. ISO 10534-2, Switzerland.
- [2] (2012). *Standard test method for impedance and absorption of acoustical materials using a tube, two microphones and a digital frequency analysis system*. ASTM E1050.
- [3] Aguirre, L. A. (2015). *Introdução à identificação de sistemas: Técnicas lineares e não lineares aplicadas a sistemas: Teoria e Aplicação*. Editora UFMG.
- [4] Akansu, A. N., Tazebay, M. V., Medley, M. J., and Das, P. K. (1997). Wavelet and subband transforms: fundamentals and communication applications. *IEEE Communications Magazine*, 35(12):104–115.
- [5] Astapov, S., Berdnikova, J., and Preden, J. (2013). Optimized acoustic localization with srp-phat for monitoring in distributed sensor network. *International Journal of Electronics and Telecommunications*, 59(4):383–290.
- [6] Benesty, J. (2000). Adaptive eigenvalue decomposition algorithm for passive acoustic source localization. *Journal of the Acoustical Society of America*, 107(1):384–391.
- [7] Boggess, A. and Narcowich, F. J. (2009). *A First Course in Wavelets with Fourier Analysis*. Wiley.
- [8] Bouzid, O. M., Tian, G. Y., Neasham, J., and Sharif, B. (2012). Envelope and wavelet transform for sound localisation at low sampling rates in wireless sensor networks. *Journal of Sensors*, 2012:1–9.
- [9] Brandstein, M. and Ward, D. (2001). *Microphone Arrays: Signal Processing Techniques and Applications*. Springer.
- [10] Broeck, B. V. D., Bertrand, A., Karsmkers, P., Vanrumste, B., Hamme, H. V., and Moonen, M. (2012). Time-domain generalized cross correlation phase transform sound source localization for small microphone arrays. In *Education and Research Conference (EDERC), 2012 5th European DSP*, pages 76–80. IEEE.

- [11] Brutti, A., Omologo, M., and Svaizer, P. (2008). Comparison between sound source localization techniques based on a real data collection. In *Proceedings of Hands Free Speech Communication and Microphone Arrays*, Trento, Itália.
- [12] Burrus, C. S., Gopinath, R. A., and Guo, H. (1988). *Introduction to Wavelets and Wavelet Transforms: A Primer*. Prentice Hall.
- [13] Castilho, J. E., Domingues, M. O., Pagamisse, A., and Mendes Jr., O. (2012). *Notas*. SBMAC.
- [14] Chen, J., Benesty, J., and Huang, Y. (2005). Performance of gcc and amdf-based time-delay estimation in practical reverberant environments. *EURASIP Journal on Applied Signal Processing*, 2005(1):25–36.
- [15] Chen, J., Benesty, J., and Huang, Y. (2006). Time delay estimation in room acoustic environments: an overview. *EURASIP Journal on Applied Signal Processing*, 2006:170–170.
- [16] Chen, W. and Holm, S. (2003). Modified szabo’s wave equation models for lossy media obeying frequency power law. *Journal Acoustical Society of America*, 114(5):2570–2574.
- [17] Daubechies, I. (1992). *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics.
- [18] Daubechies, I. (1996). Where do wavelets come from? — a personal point of view. In *Proceedings of the IEEE*, volume 84, pages 510–513.
- [19] Defrance, G., Daudet, L., and Polack, J. D. (2009). Using matching pursuit for estimating mixing time within room impulse responses. *Acta Acustica united with Acustica*, 95:1082–1092.
- [20] Defrance, G. and Polack, J. D. (2010). Estimating the mixing time of concert halls using the extensible fourier transform. *Applied Acoustics*, 71:777–792.
- [21] Donoho, D. L. (1995). De-noising by soft-thresholding. *IEEE Transactions on Information Theory*, 41(3):613–627.
- [22] Duarte, M. A. Q., Filho, J. V., and Villarreal, F. (2004). Um novo método de redução de ruído e sinais de voz baseado em wavelets. In *XXI Simpósio Brasileiro de Telecomunicações*, Belém, Brasil.

- [23] Farina, A. (2000). Simultaneous measurement of impulse response and distortion with a swept-sine technique. In *Proceedings of the 108th Audio Engineering Society Convention*, number 5093, pages 1–23, Paris.
- [24] Goswami, J. C. and Chan, A. K. (2011). *Fundamentals of Wavelets: Theory, Algorithms, and Applications*. Wiley.
- [25] Hassab, J. C. and Boucher, R. E. (1981). Performance of the generalized cross correlator in the presence of a strong spectral peak in the signals. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 29(3):549–555.
- [26] Hidaka, T., Yamada, Y., and Nakagawa, T. (2007). A new definition of boundary point between early reflections and late reverberation in room impulse responses. *The Journal of the Acoustical Society of America*, 122:326–332.
- [27] Huether, B. M., Gustafson, S., and Broussard, R. P. (2001). Wavelet preprocessing for high range resolution radar classification. *IEEE Transactions on Aerospace and Electronic Systems*, 37(4):5598–5605.
- [28] Jansen, M. (2001). *Noise Reduction by Wavelet Thresholding. Lecture notes in statistics*. Springer.
- [29] Jot, J. M., Cerveau, L., and Warusfel, O. (1997). Analysis and synthesis of room reverberation based on a statistical time-frequency model. In *Proceedings of 103th Audio Engineering Society Convention*, number 4629, Trento, Itália.
- [30] Kinsler, L. E., Frey, A. R., Coppens, A. B., and Sanders, J. V. (1999). *Fundamentals of Acoustics*. Wiley, 4th edition edition.
- [31] Knapp, C. and Carter, G. (1976). The generalized correlation method for estimation of time delay. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 24:320–327.
- [32] Kuttruff, H. (2009). *Room acoustics*. CRC Press, NY, USA.
- [33] Lang, M., Guo, H., Odegard, J. E., Burrus, C. S., and Wells, R. O. (1996). Noise reduction using an undecimated discrete wavelet transform. *IEEE Signal Processing Letters*, 3(1):10–12.
- [34] Mallat, S. (2009). *A wavelet tour of signal processing: the sparse way*. Academic Press.
- [35] Müller, S. and Massarani, P. (2001). Transfer-function measurement with sweeps. *Journal of the Audio Engineering Society*, 49(6):443—471.

- [36] Pham, T. V. (2007). *Wavelet Analysis For Robust Speech Processing and Applications*. PhD thesis, Technischen Universität Graz.
- [37] Reed, F., Feintuch, P., and Bershad (1981). Time delay estimation using the lms adaptive filter — static behavior. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 29(3):561–571.
- [38] Ribeiro, P. F. (1994). Wavelet transform: an advanced tool for analyzing non-stationary harmonic distortions in power systems. In *Proceedings of the IEEE ICHPS VI: International Conference on Harmonics in Power Systems*, Bologna, Itália.
- [39] Roth, P. R. (1971). Effective measurements using digital signal analysis. *IEEE Spectrum*, 8:62–70.
- [40] Santoso, S., Powers, E. J., Grady, W. M., and Hofmann, P. (1996). Power quality assessment via wavelet transform analysis. *IEEE Transactions on Power Delivery*, 11(2):924–930.
- [41] Sardy, S., Tseng, P., and Bruce, A. (2001). Robust wavelet denoising. *IEEE Transactions on Signal Processing*, 49(6):1146–1152.
- [42] Sharma, R. and Pyara, V. P. (2013). A robust denoising algorithm for sounds of musical instruments using wavelet packet transform. *Circuits and Systems*, 4:459–465.
- [43] Small, R. H. (1972). Direct-radiator loudspeaker system analysis. *Journal Audio Eng Society*, 20(5):383–395.
- [44] Stoica, P. and Li, J. (2006). Source localization from range-difference measurements. *IEEE Signal Processing Magazine*, 23:63–69.
- [45] Tervo, S., Pätynen, J., and Lokki, T. (2012). Acoustic reflection localization from room impulse responses. *Acta Acustica United with Acustica*, 98(3):418–440.
- [46] Tian, J. and Wells, R. O. (1996). Image compression by reduction of indices of wavelet transform coefficients. In *Proceedings of the Data Compression Conference - DCC*.
- [47] Treeby, B. and Cox, B. (2010). Modeling power law absorption and dispersion for acoustic propagation using the fractional laplacian. *Journal Acoustical Society of America*, 127(5):25741–2748.
- [48] Vetterli, M. and kovačević (2007). *Wavelets and Subband Coding*. Prentice Hall PTR.

-
- [49] Wei, D. (1998). *Type Wavelets: Theory, Design And Applications*. PhD thesis, Universidade do Texas em Austin.
- [50] Wei, D., Odegard, J. E., Guo, H., Lang, M., and Burrus, C. S. (1995). Simultaneous noise reduction and sar image data compression using best wavelet packet basis. In *Proceedings of the IEEE International Conference on Image Processing*, volume 3, pages 200–203, Washington, USA.
- [51] Wickerhauser, M. V. (1992). *Acoustic Signal compression with wavelet packets. Wavelets: A Tutorial in Theory and Applications. Vol. 2 na série: Wavelet Analysis and its Applications*.
- [52] Wikinson, W. A. and Cox, M. D. (1996). Discrete wavelet aalysis of power system transients. *IEEE Transactions on Power Systems*, 11(4):2038–2044.