

Universidade Federal do Rio de Janeiro

Escola Politécnica

Departamento de Eletrônica e de Computação

**Separação Cega de Fontes Acústicas em Ambientes com
Reverberação: Testes e Análises.**

Autor:

Gabriel de Carvalho Abi Abib

Orientadora:

Prof^a. Mariane Rembold Petraglia, Ph.D.

Co-orientador:

Prof^o. Luiz Carlos Figueira Nogueira, M. Sc.

Examinador:

Prof^o. Gelson Vieira Mendonça, Ph. D.

Examinador:

Prof^o. Julio Cesar Boscher Torres, D. Sc.

DEL

Agosto de 2013

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

Escola Politécnica – Departamento de Eletrônica e de Computação

Centro de Tecnologia, bloco H, sala H-217, Cidade Universitária

Rio de Janeiro – RJ CEP 21949-900

Este exemplar é de propriedade da Universidade Federal do Rio de Janeiro, que poderá incluí-lo em base de dados, armazenar em computador, microfilmear ou adotar qualquer forma de arquivamento.

É permitida a menção, reprodução parcial ou integral e a transmissão entre bibliotecas deste trabalho, sem modificação de seu texto, em qualquer meio que esteja ou venha a ser fixado, para pesquisa acadêmica, comentários e citações, desde que sem finalidade comercial e que seja feita a referência bibliográfica completa.

Os conceitos expressos neste trabalho são de responsabilidade do(s) autor(es) e do(s) orientador(es).

AGRADECIMENTO

À minha mãe que sempre foi um porto seguro nos momentos difíceis.

Ao meu pai, que sempre me apoiou incondicionalmente.

À minha irmã, que é um grande exemplo de dedicação e persistência.

Aos meus primos (amigos e verdadeiros irmãos), que me acompanham desde a infância.

Aos professores do Departamento de Engenharia Eletrônica e Computação, que contribuíram para a minha formação.

À minha orientadora, Mariane Rembold Petraglia, pela paciência e total disponibilidade em me ajudar.

Ao meu co-orientador, Luiz Carlos Figueira Nogueira, que além de me auxiliar com toda paciência, foi um grande companheiro nos testes desse trabalho.

RESUMO

O presente trabalho tem como objetivo avaliar o desempenho da separação cega de fontes acústicas (BSS - Blind Source Separation) em ambientes reverberantes utilizando o princípio de componentes independentes (ICA) no domínio tempo-frequência (TF).

Foram obtidos dados teóricos, a partir de simulações da acústica de salas, e dados experimentais em diferentes ambientes. O algoritmo *FastICA* no domínio TF foi utilizado nos testes para efetuar as separações. Sua eficiência foi verificada e avaliada para diferentes combinações de parâmetros.

O trabalho também apresenta a base teórica para a técnica de análise de componentes independentes que abrange a não-gaussianidade das fontes acústicas e a independência estatística entre elas.

Palavras-Chave: ICA, BSS, Reverberação, Misturas.

ABSTRACT

The aim of this study is to evaluate the performance of the Blind Source Separation in reverberant environments using the independent component principle in time-frequency domain.

Theoretical data were obtained through simulations of room acoustics, as well as experimental data in different environments. The FastICA algorithm in time-frequency domain was used for source separation. Its efficiency was verified and evaluated for different parameter combinations.

This study also presents the theoretical bases for the independent component analysis technique, which covers the non-gaussianity of acoustic sources and the statistical independence among them.

Key-words: ICA, BSS, Reverberation, Mixtures.

SIGLAS

BSS – *Blind Source Separation*

CT – Centro de Tecnologia

DFT – *Discret Fourier Transform*

EDC – *Energy Decay Curve*

ICA – *Independent Component Analysis*

LISA – Laboratório de Instrumentação e Simulação Acústica

pdf – probability density function

RIR – *Room Impulse Response*

SIR – *Source to Interferences Ratio*

STFT – *Short Time Fourier Transform*

UFRJ – Universidade Federal do Rio de Janeiro

Sumário

1	Introdução	1
	1.1 - Tema	1
	1.2 - Delimitação	1
	1.3 - Justificativa	2
	1.4 - Objetivos	3
	1.5 - Metodologia	3
	1.6 - Descrição	4
2	O problema da Separação Cega de Fontes	5
	2.1 - Sistema de Mistura	6
	2.2 - Sistema de Separação	7
	2.3 - Separação de Misturas Instantâneas e Determinadas	8
3	Análise de Componentes Independentes	10
	3.1 - Pré-Processamento para ICA	10
	3.1.1 – Centralização	10
	3.1.2 – Branqueamento	11
	3.2 - Estimativa de Componentes Independentes	11
	3.2.1 – Maximização da não-gaussianidade	12
	3.2.1.1 – Curtose	12
	3.2.1.2 – Negentropia	13
	3.2.2 – O Algoritmo <i>FastICA</i>	14
4	Misturas Convolutivas	15
	4.1 – Análise de Misturas Convolutivas no Domínio da Frequência	15

4.2 – O problema do Escalamento e da Permutação	17
5 Resposta ao Impulso de Sala Acústicas	18
5.1 – Excitação Impulsiva	18
5.2 – Resposta via FFT	18
5.3 – Tempo de Reverberação	19
6 Testes e Resultados	21
6.1 – Método de Avaliação	22
6.2 – Testes com Misturas Simuladas	22
6.2.1 – Influência do número de pontos K da STFT e o salto J entre janelas	22
6.2.2 – Influência do Tempo de Reverberação	28
6.3 – Testes com Misturas Gravadas	29
6.3.1 – Teste na Sala 1	29
6.3.2 – Teste na Sala 2	35
7 Conclusões	44
Referências Bibliográficas	45

Lista de Figuras

Figura 1.1 Cocktail Party com 2 interlocutores e 2 sensores	2
Figura 2.1 Modelo Básico BSS	5
Figura 2.2 Estrutura de sistema de separação para $N=M=2$	6
Figura 5.1 Definição do tempo de reverberação como sendo o decaimento do nível de pressão sonora em 60 dB.....	19
Figura 5.2 Estimativa de $T60=0,932s$ de uma sala pelo método de Schroeder.....	20
Figura 6.1 Sinais original e estimados (melhor e pior caso) para $T60=0,1$	23
Figura 6.2 Sinais original e estimados para $T60=0,1s$ e $T60=0,9s$ com $K=2048$ e $J=256$	26
Figura 6.3 Curvas SIR X $T60$ com os dados da Tabela 6.9	29
Figura 6.4 Estrutura da Sala 1 de testes (LISA).....	30
Figura 6.5 Fotografia do LISA mostrando os microfones utilizados nos testes.....	30
Figura 6.6 Fotografia do LISA mostrando os microfones e caixa de som utilizados nos testes .	31
Figura 6.7 Medida do $T60$ da Sala 1 (LISA) com as portas abertas	31
Figura 6.8 Medida do $T60$ da Sala 1 (LISA) com as portas fechadas.....	33
Figura 6.9 Sinais estimados e originais, com fontes nas posições 6 e 5.....	35
Figura 6.10 Estrutura da Sala 2 de testes	36
Figura 6.11 Fotografias da Sala 2 mostrando os microfones e caixa de som utilizados nos testes	36
Figura 6.12 Medida do $T60$ da Sala 2	37
Figura 6.13 SIR por faixa de frequência - Fonte de voz feminina na posição 7 e fonte de voz masculina na posição 5.....	40
Figura 6.14 SIR por tempo - Fonte de voz feminina na posição 7 e fonte de voz masculina na posição 5.....	41
Figura 6.15 SIR por faixa de frequência com $K=4096$ - Fonte de voz feminina na posição 7 e fonte de voz masculina na posição 5.....	42
Figura 6.16 SIR por tempo com $K=4096$ - Fonte de voz feminina na posição 7 e fonte de voz masculina na posição 5.....	42

Lista de Tabelas

Tabela 6.1 Separações para $T60 = 0,1s$ e $fs=8kHz$	23
Tabela 6.2 Separações para $T60 = 0,5s$ e $fs=8kHz$	24
Tabela 6.3 Separações para $T60 = 0,9s$ e $fs=8kHz$	24
Tabela 6.4 Separações para $T60 = 1,3s$ e $fs=8kHz$	25
Tabela 6.5 Separações para $T60 = 0,1s$ e $fs=16kHz$	26
Tabela 6.7 Separações para $T60 = 0,9s$ e $fs=16kHz$	27
Tabela 6.8 Separações para $T60 = 1,3s$ e $fs=16kHz$	27
Tabela 6.9 1ª Dupla Homem/Mulher – SIR x T60.....	28
Tabela 6.10 2ª Dupla Homem/Mulher – SIR x T60.....	28
Tabela 6.11 Resultados dos Testes no LISA com as Portas Abertas	32
Tabela 6.12 Resultados dos Testes no LISA com as Portas Fechadas	34
Tabela 6.13 Resultados dos Testes na SALA 2.....	377
Tabela 6.14 Separação para as posições 7 e 5 com $K=4096$ e $J=256$	41

Capítulo 1

Introdução

Ao ouvir sinais de áudios independentes misturados, sejam sinais de vozes ou músicas, nosso cérebro tem a capacidade de identificá-los, mesmo que de forma limitada. Quanto maior o número de fontes misturadas, maior a dificuldade de identificá-las. Podemos ainda ousar em afirmar que somos capazes de separar instintivamente sinais de áudio, quando focamos a atenção em um deles, em detrimento dos demais. Porém, separar sinais de áudio com qualidade, sem detrimento de nenhuma fonte passou a ser um problema computacional, pois pode demandar uma análise estatística profunda e grande capacidade de processamento.

1.1 – Tema

A separação cega de fontes consiste em um sistema de múltiplas entradas e saídas. As entradas desse sistema são os sinais de áudio misturados e as saídas são as estimativas de cada sinal separado. A separação é dita como cega, pois não é necessário nenhum conhecimento prévio das fontes a serem separadas. A única hipótese prévia necessária é que as fontes sejam estatisticamente independentes. Essa hipótese dificilmente não será atendida na prática, contribuindo para a robustez do processo de separação.

A reverberação do ambiente é um agente perturbador no processo de separação, porém é mais realista, logo deve ser modelado de forma especial. O efeito da reverberação gera versões atrasadas das fontes nas misturas, devido às reflexões dos sinais no ambiente.

1.2 – Delimitação

No fim da década de 40, as técnicas de processamento de sinais avançaram para um campo, cujo objetivo era separar sinais de fontes de ruídos e interferências. Nesse contexto, Bode e Shannon [5] avançaram com um estudo de filtragem temporal de sinais.

Com o passar das décadas, novas técnicas de filtragem e separação de sinais substituíam as antigas. A análise de componentes independentes ou ICA (*Independent Component Analyses*, do inglês) surge nos anos 80, com Héroult e Jutten, cujo trabalho

focava-se em separar um sinal neurofisiológico, que consistia numa mistura de dois sinais independentes: o deslocamento e a velocidade angular do movimento de um músculo.

A partir daí, a ICA passou a ter aplicações em diversas áreas da engenharia. Seja no monitoramento de batimentos cardíacos, cancelamento de ruído e interferências ou em sistemas de comunicação digital. É nesse último que se encontra a aplicabilidade das análises desse trabalho. Num sistema de comunicação digital, pode haver múltiplos interlocutores, cujos sinais de voz se propagam no ambiente e sofrem influências de reflexões e atenuações do meio. As técnicas de ICA são utilizadas para estimar cada um desses sinais originais de forma cega, ou seja, sem haver nenhum conhecimento prévio dos mesmos.

1.3 – Justificativa

Como já foi dito anteriormente, as técnicas de ICA podem ser utilizadas num sistema de múltiplos interlocutores, em que se deseja separar os sinais das fontes envolvidas. Esse cenário é intitulado popularmente como “*cocktail party*”, fazendo uma analogia a nossa habilidade de diferenciar sons específicos em uma barulhenta festa.

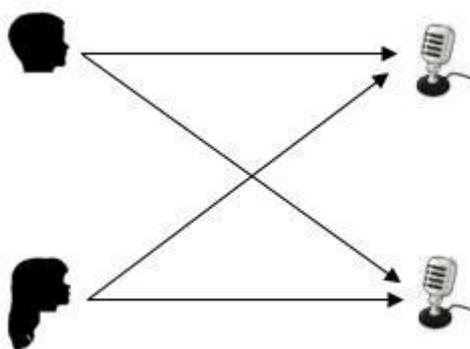


Figura 1.1 Cocktail Party com 2 interlocutores e 2 sensores

A figura 1.1 modela um cenário cujas misturas captadas pelos sensores são classificadas como instantâneas, pois não há versões atrasadas dos sinais chegando nos microfones. As técnicas de ICA [1] podem ser utilizadas, desde que o número de sensores (misturas) seja igual ou maior em relação ao número de fontes.

Sinais acústicos, como sinais de voz, por exemplo, possuem função de distribuição não gaussiana (em geral, supergaussiana). Porém ao somar variáveis aleatórias, a função de distribuição de probabilidade tende para uma curva gaussiana. A ICA trata as misturas como variáveis aleatórias independentes e para separá-las busca uma maximização da não-gaussianidade.

Porém, o foco desse trabalho será a análise de misturas convolutivas, que modelam o efeito da reverberação do ambiente. Nesse caso, métodos baseados em ICA são aplicados em raias de frequência de forma independente. Para isso, deve-se previamente fazer a transformação das misturas para o domínio da frequência, através da Transformada de Fourier por Janelas ou STFT (*Short-Time Fourier Transform*, do inglês) [4].

1.4 – Objetivos

No processo de separação de fontes, a qualidade da estimativa das fontes é extremamente sensível ao ajuste de determinados parâmetros. O objetivo desse trabalho é avaliar a resposta do processo de separação em função da variação de alguns desses parâmetros, tais como, o número de pontos K de cada janela utilizada na STFT, o tamanho J do salto entre as janelas, o tipo de voz (masculina e feminina), a taxa de amostragem dos sinais, o posicionamento das fontes no ambiente e o tipo de sala para o teste.

1.5 – Metodologia

O *software* utilizado no projeto é o Matlab, versão 2013a, desenvolvido pela Mathworks. O algoritmo para realizar as separações é o *FastICA*. As fontes são sinais de vozes masculina e feminina de aproximadamente 10 segundos cada, em versões de 8kHz e 16kHz de taxa de amostragem. As misturas foram geradas via gravação ou via simulação, dependendo do teste e da análise em questão. No primeiro caso, alto-falantes foram utilizados para reproduzir o sinal de saída de cada uma das fontes no ambiente e microfones funcionaram como sensores para a entrada das misturas. Os testes foram realizados em três ambientes reverberantes diferentes. Uma interface com o Simulink foi necessária para realizar a aquisição e reprodução dos sinais. No segundo caso, as misturas foram geradas através de um algoritmo que simula o ambiente de gravação, sendo especificados o tempo de reverberação e a posição dos sensores e das fontes na sala.

Para a avaliação da qualidade das separações, foi utilizada a relação fonte-interferência (SIR, do inglês *Source to Interferences Ratio*), que compara o sinal original com o sinal separado, sem haver necessidade do conhecimento dos dados das misturas.

1.6 – Descrição

No Capítulo 2, será introduzida a estrutura básica de um sistema de separação cega de fontes.

O Capítulo 3 disserta sobre a análise de componentes independentes, as etapas de pré-processamento e os métodos para maximização da não-gaussianidade.

O Capítulo 4 aborda uma análise de misturas convolutivas no domínio da frequência.

No Capítulo 5, são apresentados os conceitos de Resposta impulsiva de salas acústicas e tempo de reverberação.

O Capítulo 6 apresenta os resultados obtidos através dos testes realizados.

No Capítulo 7, temos a conclusão do trabalho.

Capítulo 2

O problema da Separação Cega de Fontes

O problema da separação foca-se na extração dos sinais fontes a partir de misturas obtidas nas saídas de um conjunto de sensores, cada uma recebendo uma combinação diferente dos sinais.

A separação pode ser conseguida de maneiras diferentes, de acordo com a quantidade de informação disponível. BSS (*Blind Source Separation*, do inglês) procura recuperar os sinais originais das fontes a partir de suas misturas, sem qualquer informação prévia sobre as fontes ou parâmetros das misturas. Em outras palavras, o problema de BSS pode ser interpretado como a estimativa de N fontes a partir de M medições, as quais são funções desconhecidas das fontes. O modelo de base BSS é mostrado na Figura 2.1. As componentes do vetor fonte de dados $\mathbf{s}(n)$ são combinados por uma matriz de mistura \mathbf{H} para produzir as componentes do vetor de misturas $\mathbf{x}(n)$. Algoritmos de separação, como os que se baseiam em métodos de ICA produzem uma matriz de separação \mathbf{W} , que tem a capacidade de extrair as fontes originais $\mathbf{y}_i(n)$, ou seja, réplica de $\mathbf{s}_i(n)$ a partir de misturas diferentes.

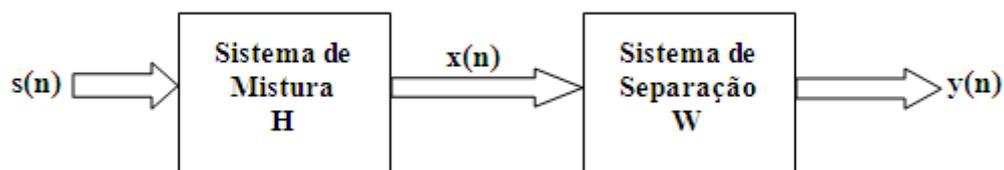


Figura 2.1 Modelo Básico BSS

Sendo N , o número de fontes e M o número de sensores, o vetor $\mathbf{s}(n)$ é composto pela concatenação das N fontes s_i , cada uma composta pelas n -ésimas amostras $s_i(n)$. O vetor $\mathbf{x}(n)$ é definido como sendo a concatenação das M misturas x_j , cada uma composta pelas n -ésimas amostras $x_j(n)$. O vetor $\mathbf{y}(n)$ dos sinais estimados é composto de forma análoga às fontes. Logo, temos que os vetores podem ser assim representados:

$$\mathbf{s}(n) = [s_1(n) \dots s_N(n)]^T \quad (2.1)$$

$$\mathbf{x}(n) = [x_1(n) \dots x_M(n)]^T \quad (2.2)$$

$$\mathbf{y}(n) = [y_1(n) \dots y_N(n)]^T \quad (2.3)$$

2.1 – Sistema de Mistura

As infinitas formas que as misturas podem assumir influenciam diretamente no processo de separação, viabilizando, ou não, a aplicação dos métodos baseados em ICA. Em um sistema de separação cega, as misturas são as únicas informações disponíveis no processo. Como cada sensor possui um posicionamento único, cada mistura contribui diferentemente para a estimativa das fontes.

Um sistema cujo número de sensores se iguala ao número de fontes ($N = M$) denomina-se como sendo determinado e possui as informações suficientes para se aplicarem os métodos baseados em ICA. O sistema é denominado sobre-determinado quando o número de sensores supera o número de fontes ($M > N$), o que implica um aumento da facilidade da separação, visto que há uma maior quantidade de informação disponível.

O presente trabalho não contemplará a análise de misturas denominadas indeterminadas, quando se tem um número de sensores menor que o número de fontes ($M < N$). Nesse caso, as técnicas de ICA não são mais aplicáveis, dando espaço para o método de Análise de Componentes Esparsos, que explora a propriedade da esparsidade das fontes para o processo de separação [10].

As técnicas de BSS se baseiam em estimar os sinais das fontes a partir de uma matriz \mathbf{W} de separação das misturas $x_j(n)$. Estas, por sua vez, são modeladas através de uma matriz \mathbf{H} de mistura das fontes $s_i(n)$. Um exemplo de estrutura de um sistema de separação para um caso de misturas determinadas, com $N = M = 2$ está esquematizado na figura 2.2.

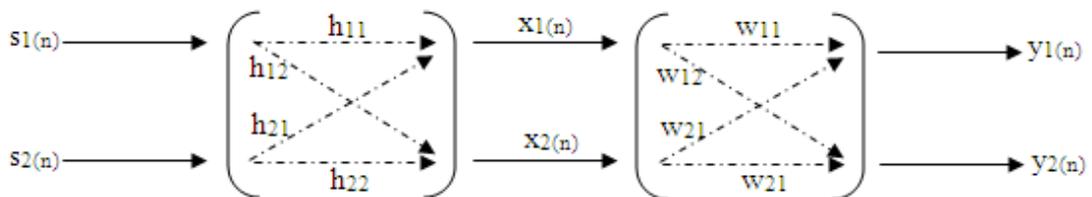


Figura 2.2 Estrutura de sistema de separação para $N=M=2$

De forma geral, \mathbf{H} é uma matriz de dimensões $M \times N$ e o vetor de misturas $\mathbf{x}(n)$ pode ser obtido pela convolução entre \mathbf{H} e o vetor de fontes $\mathbf{s}(n)$:

$$\mathbf{x}(n) = \mathbf{H} * \mathbf{s}(n) \quad (2.4)$$

Os elementos de \mathbf{H} são compostos por um conjunto de respostas ao impulso de filtros FIR de comprimento L , que simulam os múltiplos caminhos percorridos pelo sinal, no caso de haver reverberação. Logo, as amostras de cada sinal misturado podem ser calculadas como

$$x_j(n) = \sum_{i=1}^N \sum_{k=0}^{L-1} h_{ji}(k) s_i(n-k), \quad j = 1, \dots, M, \quad (2.5)$$

onde h_{ji} representa a resposta ao impulso do caminho percorrido entre a i -ésima fonte e o j -ésimo sensor. A Eq. (2.5) modela o caso de misturas convolutivas. Para $L=1$, os elementos da matriz \mathbf{H} se reduzem a constantes, modelando o caso de misturas instantâneas, isto é, sem versões atrasadas causadas pela reverberação.

2.2 – Sistema de Separação

Para o caso de misturas determinadas ou sobredeterminadas, o vetor $\mathbf{y}(n)$ dos sinais estimados é calculado da seguinte forma:

$$\mathbf{y}(n) = \mathbf{W} * \mathbf{x}(n), \quad (2.6)$$

onde \mathbf{W} é a matriz de separação de ordem $N \times M$, responsável por desfazer as misturas. A partir da Eq. (2.6), pode-se obter o cálculo de cada amostra dos sinais y_i :

$$y_i(n) = \sum_{j=1}^M \sum_{k=0}^{S-1} w_{ij}(k) x_j(n-k), \quad i = 1, \dots, N, \quad (2.7)$$

onde w_{ij} são os coeficientes do filtro FIR de separação, obtidos por um algoritmo de aprendizagem adaptativo.

Substituindo a Eq. (2.6) na Eq. (2.4), obtém-se a relação direta entre $\mathbf{y}(n)$ e $\mathbf{s}(n)$:

$$\mathbf{y}(n) = \mathbf{W} * \mathbf{H} * \mathbf{s}(n). \quad (2.8)$$

O sistema de separação é definido pela minimização ou maximização de uma determinada função custo. No caso dos métodos baseados em ICA, a otimização dessa função custo permite a estimativa de componentes independentes das misturas. Em geral, as principais funções custo são baseadas na maximização da não-gaussianidade.

2.3 – Separação de Misturas Instantâneas e Determinadas

Para o caso particular de misturas instantâneas, o comprimento dos filtros de mistura e de separação é unitário ($L = 1$) e seus elementos passam a ser meras constantes. Se assumirmos o caso determinado ($N = M$), podemos calcular as misturas pela equação matricial:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} = \begin{pmatrix} h_{11} & \dots & h_{1N} \\ \vdots & \ddots & \vdots \\ h_{N1} & \dots & h_{NN} \end{pmatrix} \begin{pmatrix} s_1 \\ \vdots \\ s_N \end{pmatrix} \quad (2.9)$$

e as saídas do sistema de separação são dadas por:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} w_{11} & \dots & w_{1N} \\ \vdots & \ddots & \vdots \\ w_{N1} & \dots & w_{NN} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}. \quad (2.10)$$

O operando da convolução na Eq. (2.8) se reduz a um simples produto matricial, ficando na seguinte forma:

$$\mathbf{y}(n) = \mathbf{W}\mathbf{H}\mathbf{s}(n). \quad (2.11)$$

Sendo a matriz \mathbf{H} não-singular, fazendo $\mathbf{W} = \mathbf{H}^{-1}$, teríamos $\mathbf{C} = \mathbf{W}\mathbf{H} = \mathbf{I}_N$, e $\mathbf{y}(n)$ seria réplica perfeita do vetor $\mathbf{s}(n)$. Porém não convém exigir das técnicas de separação que a igualdade $\mathbf{W} = \mathbf{H}^{-1}$ seja satisfeita. Uma estimativa satisfatória para a matriz \mathbf{W} seria aquela que resultasse em uma matriz \mathbf{C} aproximadamente diagonal.

No entanto, alguns problemas inerentes à BSS conhecidos, como permutação e escalamento, podem surgir quando $\mathbf{C} \neq \mathbf{I}_N$. Tais fenômenos são ilustrados no exemplo a seguir, considerando $M = N = 3$:

$$\mathbf{C} = \begin{pmatrix} 0 & 1,2 & 0 \\ 0 & 0 & -2,4 \\ 1,6 & 0 & 0 \end{pmatrix}.$$

Nesse caso, o cálculo das amostras de $\mathbf{y}(n)$ sofrerá permutação pelo fato da matriz \mathbf{C} não ser diagonal e sofrerá escalamento pelo fato dos coeficientes não nulos de \mathbf{C} serem diferentes. Tais efeitos são facilmente contornados nos casos de misturas instantâneas, mas podem prejudicar o desempenho da separação no domínio da frequência no caso de misturas convolutivas.

Capítulo 3

Análise de Componentes Independentes

Para estimar as fontes, as técnicas baseadas em ICA partem do modelo linear da Eq. (2.11) e do princípio de que os sinais originais sejam estatisticamente independentes. A restrição estatística imposta pela ICA é que as fontes possuam função de distribuição de probabilidade (pdf) não-gaussianas, ou que no máximo uma delas possua pdf gaussiana. Essa imposição é devida ao fato de que a soma de variáveis aleatórias possui uma distribuição de probabilidade conjunta gaussiana, de acordo com o Teorema do Limite Central (Papoulis, 1991), inviabilizando qualquer inferência a respeito das fontes a partir dos dados das misturas.

3.1 – Pré-Processamento para ICA

Algumas técnicas de pré-processamento devem ser empregadas nos dados das misturas antes da estimativa da matriz \mathbf{W} de separação. Para isso antes devemos definir a matriz de misturas \mathbf{X} como sendo a concatenação dos vetores $\mathbf{x}(n)$, supondo um total de k amostras, isto é,

$$\mathbf{X} = [\mathbf{x}(1) \ \mathbf{x}(2) \ \dots \ \mathbf{x}(k)]. \quad (3.1)$$

As seções 3.1.1 e 3.1.2 descrevem as duas principais etapas de pré-processamento.

3.1.1. Centralização

Essa etapa tem como objetivo tornar os vetores $\mathbf{x}(n)$ com média zero. Para isso, subtrai-se do vetor $\mathbf{x}(n)$ a sua média $E[\mathbf{x}(n)]$. Apesar da centralização não ser obrigatória, a etapa simplifica os dados para os algoritmos de separação.

3.1.2 Branqueamento

Essa outra etapa deve ser aplicada após a centralização e tem como objetivo transformar o vetor $\mathbf{x}(n)$ em um vetor $\mathbf{z}(n)$, cujos componentes sejam não correlacionados e que sua matriz de covariância seja unitária, ou seja, igual à identidade. O vetor branqueado $\mathbf{z}(n)$ é calculado através de uma matriz de branqueamento \mathbf{Q} :

$$\mathbf{z}(n) = \mathbf{Q}\mathbf{x}(n) \quad (3.2)$$

Sendo $\hat{\mathbf{C}} = E[\mathbf{x}(n)\mathbf{x}(n)^T]$ a matriz de covariância de $\mathbf{x}(n)$, um método simples para efetuar o branqueamento utiliza a decomposição em autovetores de $\hat{\mathbf{C}}$:

$$\hat{\mathbf{C}} = \mathbf{M}\mathbf{D}\mathbf{M}^T, \quad (3.3)$$

sendo \mathbf{M} e \mathbf{D} , a matriz ortogonal de autovetores e a matriz diagonal de autovalores de $\hat{\mathbf{C}}$, respectivamente. Os dados branqueados são calculados, por:

$$\mathbf{z}(n) = \mathbf{M}\mathbf{D}^{-\frac{1}{2}}\mathbf{M}\mathbf{x}(n). \quad (3.4)$$

Tal procedimento torna a matriz de covariância de $\mathbf{z}(k)$ igual à identidade, como era desejável. Podemos definir a matriz de misturas devidamente branqueada como sendo a composição dos vetores $\mathbf{z}(n)$:

$$\mathbf{Z} = [\mathbf{z}(1) \ \mathbf{z}(2) \ \dots \ \mathbf{z}(k)] \quad (3.5)$$

3.2 – Estimativa de Componentes Independentes

Em geral, a estimativa de dados independentes de um sistema de mistura é realizada a partir da minimização ou maximização de uma função custo. Para isso, o modelo para ICA depende do algoritmo de otimização, que tem como objetivo buscar a matriz de separação \mathbf{W} através de um processo iterativo.

As principais funções custo para estimar os componentes independentes se baseiam na medida e maximização da não-gaussianidade das saídas ou minimização não-gaussianidade da informação mútua.

3.2.1 Maximização da não-gaussianidade

A maximização de uma medida da não-gaussianidade é uma das principais formas de estimar as componentes independentes de uma mistura, pois como já foi dito, as variáveis aleatórias de interesse possuem distribuição de probabilidade não-gaussiana. Essa hipótese pode ser aplicada aos sinais de áudio, que têm distribuição geralmente supergaussiana.

Duas das formas de se medir a não-gaussianidade são através da curtose e da negentropia, como será explicado a seguir.

3.2.1.1 Curtose

A curtose K de uma variável aleatória $u(n)$ é dada por:

$$K(u(n)) = E[u(n)^4] - 3(E[u(n)^2])^2. \quad (3.6)$$

Supondo que a variância $E[u(n)^2]$ seja unitária, temos que

$$K(u(n)) = E[u(n)^4] - 3, \quad (3.7)$$

sendo que o termo $E[u(n)^4]$ é definido como o momento de quarta ordem. A importância do uso da curtose para o cálculo da gaussianidade advém do fato de que para uma variável com distribuição gaussiana, o momento de quarta ordem é igual a $3(E[u(n)^2])^2$, tornando nula a medida da curtose. Logo, se $K = 0$, podemos afirmar que a variável é gaussiana, caso contrário, trata-se de uma variável não gaussiana. De forma mais detalhada, temos que:

- se $K > 0$, variável sub-gaussiana;
- se $K = 0$, variável gaussiana;
- se $K < 0$, variável super gaussiana.

Com os dados da matriz de mistura devidamente branqueados, o algoritmo de otimização para a separação deve buscar uma matriz \mathbf{W} , através da minimização do valor da curtose. Para isso, o algoritmo deve inicializar \mathbf{W} com valores aleatórios e medir a direção do crescimento de K . Com os dados obtidos, recalcula-se a matriz \mathbf{W} . O processo deve ser repetido até a obtenção de todas as componentes independentes.

Medir a gaussianidade através da curtose apresenta algumas desvantagens. Além de não ser uma forma robusta para o cálculo, esse método apresenta grande complexidade computacional e é muito sensível à quantidade de amostras disponíveis, já que a estimativa é feita a partir de dados observados.

3.2.1.2 Negentropia

Seja $u(n)$ uma variável aleatória e $u_{gauss}(n)$ uma variável gaussiana de mesma variância. A negentropia J de $u(n)$ é calculada como sendo a diferença entre as entropias H de $u_{gauss}(n)$ e $u(n)$:

$$J(u(n)) = H(u_{gauss}(n)) - H(u(n)). \quad (3.8)$$

A entropia pode ser entendida como sendo a medida da informação necessária para descrever uma variável aleatória. Quanto maior a imprevisibilidade de uma variável, maior será sua entropia. Para uma variável aleatória discreta, a entropia é definida como:

$$H(u(n)) \stackrel{\text{def}}{=} -\sum p(u_i) \log_2 p(u_i), \quad (3.9)$$

sendo que $p(u_i) = \text{Prob}(u(n) = u_i)$. A importância da negentropia para o cálculo da gaussianidade parte do princípio de que variáveis gaussianas possuem a maior entropia dentre variáveis com mesma variância.

Visto que o cálculo da negentropia pode demandar muito tempo computacional, em geral são utilizadas técnicas que calculam seu valor aproximado. Um desses métodos utiliza uma função não quadrática G para esse cálculo:

$$J_G(u(n)) = a\{E[G(u(n))] - E[G(v(n))]\}^2, \quad (3.10)$$

sendo $v(n)$ uma variável gaussiana com média zero e variância unitária. Algoritmos que utilizam o cálculo aproximado da negentropia diferenciam-se pela escolha da função G . Uma escolha adequada pode resultar em um cálculo da gaussianidade com um baixo custo

computacional e garantir a robustez do processo. Um desses algoritmos é o *FastICA*, que será apresentado a seguir.

3.2.2. O Algoritmo *FastICA*

Os algoritmos da família *FastICA* [2] procuram estimar as componentes independentes através da maximização da negentropia. Isso se deve ao fato de que ao maximizar a negentropia, maximiza-se a não gaussianidade das fontes estimadas.

Considerando a matriz de misturas já branqueada \mathbf{Z} , o algoritmo busca encontrar iterativamente uma matriz de separação \mathbf{W} , maximizando a negentropia, através da seguinte aproximação:

$$\begin{aligned} \mathbf{W}_{i+1} &\leftarrow E[\mathbf{Z}G(\mathbf{W}^T\mathbf{Z})] - E[G'(\mathbf{W}^T\mathbf{Z})]\mathbf{W}^T; \\ \mathbf{W} &= \frac{\mathbf{W}_{i+1}}{\|\mathbf{W}_{i+1}\|}, \end{aligned} \tag{3.11}$$

onde G é uma função não quadrática, G' é sua derivada e $\|\mathbf{W}_{i+1}\|$ representa a norma de \mathbf{W}_{i+1} .

O algoritmo *FastICA* foi publicado primeiramente em 1999 [2]. Desde então tem sido um dos algoritmos mais utilizados na estimativa das componentes independentes, devido à sua robustez. Além disso, não demanda muito tempo computacional, visto que o método consegue ter uma boa precisão com apenas 5 a 10 iterações. O processo ainda pode ser melhorado, ajustando a função G utilizada.

Capítulo 4

Misturas Convolutivas

O caso de misturas convolutivas modela a situação mais realista, em que há reverberação no ambiente. Nesse caso os atrasos devidos às reflexões são representados por elementos de filtro FIR como descrito na Eq. (2.5). Logo, o operador da convolução da Eq. (2.8) não pode mais ser reduzido a um produto, como é feito no caso de misturas instantâneas.

Existem na literatura duas diferentes análises para misturas convolutivas. A primeira trata-se de uma análise no próprio domínio do tempo, levando em considerações características estatísticas de segunda ordem como espectro colorido e não-estacionariedade. Uma análise no domínio da frequência será discutida com detalhes na seção 4.1, visto que os testes realizados nesse trabalho contemplam essa abordagem. A separação no domínio da frequência apresenta vantagens no que diz respeito a custo computacional, porém introduz dificuldades que podem se tornar complexas, como permutação e escalamento.

4.1 – Análise de Misturas Convolutivas no Domínio da Frequência

A ideia principal desse tipo de análise é transformar uma mistura convolutiva no tempo em misturas instantâneas por raias de frequência. Assim será possível aplicar os algoritmos baseados em ICA em cada uma das raias. Para isso, é necessário realizar a transformada de Fourier dos sinais \mathbf{s} e \mathbf{x} . Para isso, definiremos $X_j(\omega)$ como sendo a transformada de $x_j(n)$ e $S_i(\omega)$, a transformada de $s_i(n)$. Logo teremos a composição dos seguintes vetores:

$$\mathbf{X}(\omega) = [X_1(\omega) \dots X_M(\omega)]^T \quad (4.1)$$

$$\mathbf{S}(\omega) = [S_1(\omega) \dots S_N(\omega)]^T. \quad (4.2)$$

Aplicando a Transformada de Fourier na Eq (2.4), temos que

$$\mathbf{X}(\omega) = \mathbf{H}(\omega)\mathbf{S}(\omega), \quad (4.3)$$

sendo que a matriz $\mathbf{H}(\omega)$ contém as resposta em frequência $H_{ji}(\omega)$ dos filtros $h_{ji}(n)$:

$$\mathbf{H}(\omega) = \begin{pmatrix} H_{11}(\omega) & \cdots & H_{1N}(\omega) \\ \vdots & \ddots & \vdots \\ H_{M1}(\omega) & \cdots & H_{MN}(\omega) \end{pmatrix}. \quad (4.4)$$

Devido à não-estacionariedade dos sinais de voz, a transformação para o domínio da frequência dos sinais misturados $x_j(n)$ é realizada aplicando a Transformada Discreta de Fourier (DFT, *Discret Fourier Transform*) em subsequências janeladas de comprimento K , resultando na Transformada curta de Fourier (STFT, *Short Time Fourier Transform*):

$$X_j(f, m) = \sum_{r=-\frac{K}{2}}^{\frac{K}{2}-1} x_j(mJ + r) \text{win}(r) e^{-j\frac{2\pi fr}{K}}, \quad 0 \leq f \leq K - 1, \quad (4.5)$$

sendo m o índice da janela de dados e J o salto entre janelas. A função $\text{win}(\cdot)$ é a função que determina o formato da janela, que deve possuir as extremidades suavizadas. A função utilizada nos testes desse trabalho é a janela *Hanning*.

Realizada a transformação as técnicas de separação baseadas em ICA podem ser aplicadas em raia de frequência f e o vetor de misturas pode ser escrito como

$$\mathbf{X}(f, t) = [X_1(f, t) \dots X_M(f, t)]^T, \quad (4.6)$$

em que $X_j(f, t)$ representa a transformada de $x_j(n)$ na raia f e janela t . A relação entre a raia f e a frequência digital ω é dada por $\omega = \frac{2\pi f}{K}$.

A matriz de separação deve ser calculada pelo algoritmo de ICA escolhido para cada raia de frequência f , ficando na seguinte forma:

$$\mathbf{W}(f) = \begin{pmatrix} w_{11}(f) & \cdots & w_{1M}(f) \\ \vdots & \ddots & \vdots \\ w_{N1}(f) & \cdots & w_{NM}(f) \end{pmatrix}. \quad (4.7)$$

A estimativa dos sinais também é realizada em janelas para cada raia de frequência:

$$\mathbf{Y}(f, t) = \mathbf{W}(f)\mathbf{X}(f, t), \quad (4.8)$$

resultando no seguinte vetor de saída:

$$\mathbf{Y}(f, t) = [Y_1(f, t) \dots Y_N(f, t)]^T. \quad (4.9)$$

Para retornar ao domínio do tempo, realiza-se a Transformada Inversa Discreta de Fourier (IDFT, *Inverse Discrete Fourier Transform*) para cada raia de frequência, obtendo-se os coeficientes dos filtros $w_{ij}(n)$.

A estimativa dos sinais no domínio da frequência introduz problemas críticos, que no caso de misturas instantâneas eram triviais de serem resolvidos. Os principais são o escalamento e a permutação.

4.2 - Os Problemas de Escalamento e da Permutação

O fato de cada raia de frequência ser tratada de forma distinta faz com que o problema do escalamento se torne crítico ao se retornarem os sinais para o domínio do tempo. Isso ocorre, porque cada frequência é amplificada por constantes de escalamento diferentes. Esse problema pode ser solucionado através do princípio da distorção mínima, aplicado na matriz $\mathbf{W}(f)$, após ser calculada pelo algoritmo de ICA:

$$\mathbf{W}(f) \leftarrow \text{diag}(\mathbf{W}^{-1}(f)) \cdot \mathbf{W}(f), \quad (4.10)$$

sendo que o operador $\text{diag}(\cdot)$ é responsável por diagonalizar a matriz, anulando todos os termos que não fazem parte da diagonal principal. Outra forma de resolver o problema do escalamento é recalculando a matriz $\mathbf{W}(f)$ para cada raia de frequência, de modo que seu determinante seja unitário:

$$\widetilde{\mathbf{W}}(f) = \mathbf{W}(f) \cdot |\mathbf{W}(f)|^{-\frac{1}{N}}, \quad (4.11)$$

onde $\widetilde{\mathbf{W}}(f)$ é a matriz de separação com determinante unitário.

O problema da permutação também se torna crítico pois as saídas estimadas podem conter raias aleatórias de qualquer uma das fontes. Os métodos mais utilizados para resolver esse problema estão relacionados com a minimização da correlação entre raias [9], ou com a estimativa do ângulo de chegada das fontes [7].

Capítulo 5

Resposta ao Impulso de Salas

A resposta impulsiva de uma sala (RIR, *Room Impulse Response*) é o resultado da excitação da mesma por um impulso sonoro semelhante à função delta de *Dirac*. A obtenção da RIR depende de pelo menos um par de emissor e receptor para a excitação e obtenção da resposta, respectivamente. Nas seções 5.1 e 5.2 serão descritos dois métodos para obtenção da RIR.

5.1 – Excitação Impulsiva

Esse método consiste em excitar o ambiente diretamente com impulso sonoro. A vantagem desse método vem do fato de não ser necessário nenhum procedimento posterior para processamento, pois a resposta coletada do receptor já é aproximação da RIR. Porém, o método possui a desvantagem atribuída à impossibilidade física de se gerar um impulso de *Dirac*.

5.2 – Resposta via FFT

Utilizando essa técnica, qualquer sinal que abranja uma faixa do espectro pode servir como excitação. Em geral usam-se varreduras de senóides ou ruídos como sinais de excitação, calculam-se as Transformadas Rápidas de Fourier (FFT, *Fast Fourier Transform*) dos sinais medidos e de referência, e dividem-se os espectros de ambos para se obter a função de transferência do sistema. A RIR é obtida através da Transformada Inversa Rápida de Fourier (IFFT, *Inverse Fast Fourier Transform*) da função de transferência.

As respostas ao impulso das salas utilizadas para os testes desse trabalho foram medidas pelo programa DSSF3E, desenvolvido pela *YMec Store* [13], através de varreduras de ruídos.

5.3 – Tempo de Reverberação

O tempo de reverberação de uma sala é o tempo no qual um som deixa de ser percebido após uma fonte em regime permanente ter cessado. A literatura costuma adotar como medida do tempo de reverberação, o parâmetro T_{60} , correspondente ao tempo para o nível de pressão sonora cair 60 dB, desde o fim da emissão do som. O tempo de reverberação pode ser estimado de forma aproximada através da fórmula de Sabine [6]:

$$T_{60} = 0,163 \frac{V}{A}, \quad (5.1)$$

onde V é o volume da sala e A é a área de absorção da sala, calculada da seguinte forma:

$$A = \sum S_i \alpha_i, \quad (5.2)$$

em que S_i representa a área da superfície ocupada pelo material “ i ” e α_i o seu coeficiente de absorção de absorção.

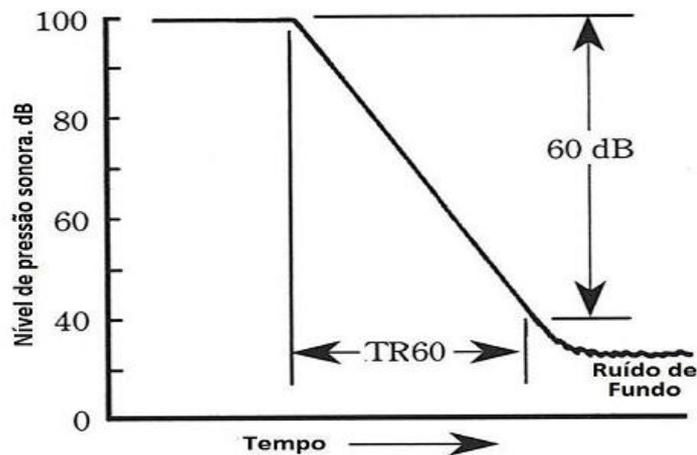


Figura 5.1 Definição do tempo de reverberação como sendo o decaimento do nível de pressão sonora em 60 dB. Referência [12]

O T_{60} pode ser medido através da estimativa da RIR do ambiente. O método utilizado nos testes desse trabalho é o proposto por Schroeder [11], a partir do qual se estima o tempo de reverberação através da curva de decaimento de energia (EDC, *energy decay curve*) da resposta ao impulso. Dada uma RIR $h(t)$, a EDC é calculada como:

$$EDC(t) = 10 \log \frac{\int_t^\infty h^2(\tau)}{\int_{-\infty}^\infty h^2(\tau)} [dB]. \quad (5.3)$$

No método Schroeder, a curva da EDC é aproximada por uma reta $s(t)$, que é posteriormente deslocada para outra reta $r(t)$ que passa pela origem. O T_{60} é estimado como sendo o tempo de decaimento de $r(t)$ em 60 dB. A figura 5.2 mostra um exemplo da estimativa de T_{60} de uma das salas utilizadas para os testes desse trabalho pelo método Schroeder.

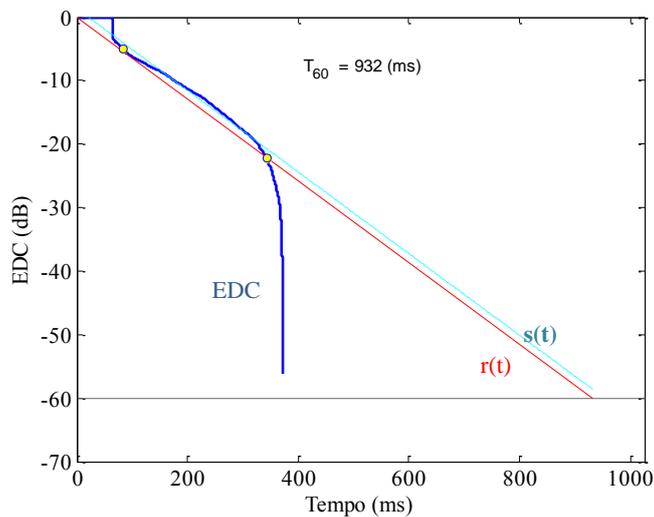


Figura 5.2 Estimativa de $T_{60}=0,932s$ de uma sala pelo método de Schroeder.

Capítulo 6

Testes e Resultados

Os testes realizados nesse trabalho tiveram como objetivo avaliar a influência de determinados parâmetros na qualidade da separação. Em geral, os experimentos avaliaram a performance da separação cega ajustando os seguintes parâmetros:

- o tamanho K da janela da STFT para realizar a transformação para o domínio da frequência como descrito na Eq.(4.7);
- o salto J entre as janelas da STFT, também evidenciado na Eq.(4.7);
- o tempo T_{60} de reverberação do ambiente;
- os tipos de vozes (masculina ou feminina) envolvidas nas misturas;
- a taxa de amostragem dos sinais de voz;
- o posicionamento das fontes no ambiente.

As misturas foram geradas de duas formas distintas. Primeiramente, os testes contemplam misturas geradas por simulação, através da criação de um banco de respostas ao impulso de acordo com as configurações de uma sala virtual e ajuste do T_{60} , de acordo com o método de Lehmann e Johansson, como descrito em [8]. A vantagem de utilizar esse tipo de abordagem (por simulação) está na liberdade do usuário de ajustar os parâmetros da sala, principalmente o tempo de reverberação. Porém, não são levadas em conta outras variáveis inerentes a uma sala física, como, por exemplo, a presença de ruído e a qualidade da instrumentação. Os testes subsequentes foram realizados a partir de misturas geradas via gravações em ambientes reais. Duas salas diferentes foram utilizadas para os testes, com respectivos tempos de reverberação medidos pelo método de Schroeder.

Todos os testes foram realizados com misturas determinadas, com $N = M = 2$ e as fontes foram sinais de voz masculina e feminina com versões em 8kHz e 16kHz.

6.1 – Método de Avaliação

A avaliação da performance das separações foi realizada utilizando a relação fonte-interferência (SIR, *Sources to Interferences Ratio*). Para a aplicação desse método, é necessário apenas o conhecimento dos dados dos sinais estimados e das fontes, sem haver necessidade dos dados das misturas.

A SIR é calculada de acordo com a seguinte equação:

$$SIR = 10 \log_{10} \frac{\|s_i^{alvo}\|^2}{\|e_i^{interf}\|^2}, \quad (6.1)$$

onde s_i^{alvo} corresponde a uma deformação aceitável da i -ésima fonte e e_i^{interf} , uma deformação causada pela interferência das outras $N - 1$ fontes. Outros métodos de avaliação podem ser utilizados como a relação fonte-artefatos (SAR, *Source Artefacts Ratio*) e a relação fonte-distorção (SDR, *Source Distortion Ratio*), que levam em consideração interferências de artefatos causados pelo algoritmo de separação e interferências causadas por ruído.

6.2 – Testes com Misturas Simuladas

A seguir serão apresentados os testes realizados com misturas geradas por simulação e seus respectivos resultados.

6.2.1 – Influência do número de pontos K da STFT e do salto J entre janelas

Nesse teste foram variados os valores de K em potências de 2 entre 256 e 4096 e os valores de J também em potências de 2, obedecendo o critério $J < K$. As separações foram repetidas para diferentes valores de T_{60} . As fontes foram sinais de voz masculina e feminina. Primeiramente a taxa de amostragem f_s foi de 8kHz e os valores da SIR para cada separação, com $T_{60}=0,1s$ estão na tabela 6.1.

Tabela 6.1 Separações para $T60 = 0,1s$ e $f_s=8kHz$

K	J	$SIR_{feminina}(dB)$	$SIR_{masculina}(dB)$
256	128	18,05362	18,84635
512	128	22,39509	23,27757
	256	22,31896	23,34312
1024	128	26,58748	26,06877
	256	26,43648	26,02283
	512	25,79088	25,93809
2048	128	26,82635	24,8321
	256	26,60922	24,5761
	512	26,13296	24,16449
	1024	24,9155	23,60085
4096	128	12,9936	15,67525
	256	19,95418	19,02578
	512	4,914828	11,14099
	1024	6,032221	11,97402
	2048	3,799656	10,10825

A separação com melhor qualidade, obtida para $K=1024$ e $J=128$ está destacada em verde na tabela. A separação com pior qualidade, obtida para $K=4096$ e $J=2048$ está destacada em vermelho. A figura 6.1 mostra o gráfico dos sinais original e estimado para a voz feminina para essas configurações.

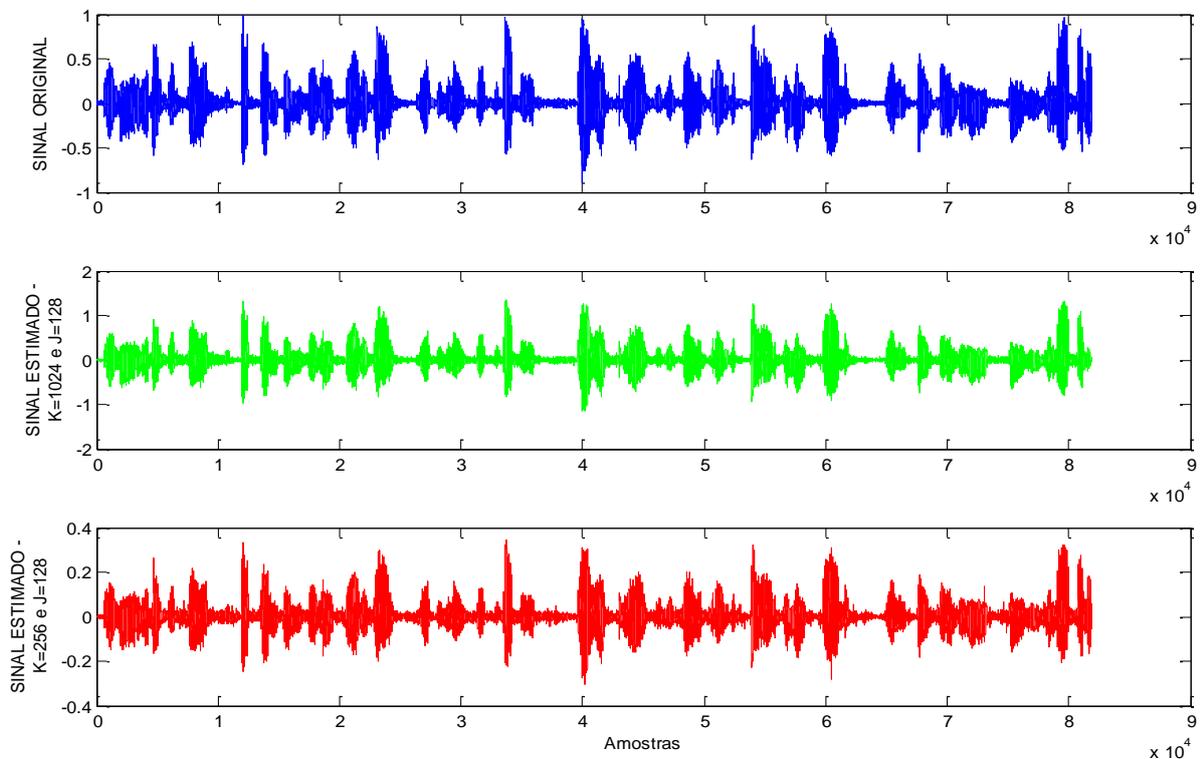


Figura 6.1 Sinais original e estimados (melhor e pior caso) para $T60=0,1s$

Pelos gráficos, ainda que sutilmente, é possível perceber que para $K=256$ e $J=128$ existe uma maior presença da interferência da outra fonte (masculina), atestando os valores obtidos para a SIR.

Os valores da SIR para outros valores de T_{60} estão nas tabelas 6.2 à 6.4. Seguindo o padrão da tabela 6.1, as estimativas de cada fonte com a melhor qualidade estão destacadas de verde e as com a pior qualidade, de vermelho.

Tabela 6.2 Separações para $T_{60} = 0,5s$ e $f_s=8kHz$

K	J	SIR _{feminine} (dB)	SIR _{masculine} (dB)
256	128	6,175433	7,554033
	256	9,886925	10,00744
512	128	9,839733	9,854594
	256	11,87195	13,33571
1024	128	11,72052	13,39712
	256	11,63485	13,18512
	512	15,92751	17,29136
2048	128	15,86934	17,09401
	256	15,6434	16,90836
	512	15,10688	16,26959
	1024	15,03256	17,57288
4096	128	15,48784	17,54596
	256	15,48171	17,17055
	512	3,847285	8,04891
	1024	6,169124	10,44119
	2048		

Tabela 6.3 Separações para $T_{60} = 0,9s$ e $f_s=8kHz$

K	J	SIR _{feminine} (dB)	SIR _{masculine} (dB)
256	128	5,506353	5,071369
	256	6,840895	7,486545
512	128	6,681114	7,369244
	256	9,03416	9,894545
1024	128	8,874779	9,845621
	256	8,725665	9,385694
	512	12,65273	12,99021
2048	128	12,62896	12,91271
	256	12,15422	12,34577
	512	11,20962	12,60226
	1024	11,19704	14,7773
4096	128	11,29794	14,63192
	256	12,88646	14,15034
	512	5,324271	9,273931
	1024	2,339603	6,395488
	2048		

Tabela 6.4 Separações para $T_{60} = 1,3s$ e $f_s=8kHz$

K	J	SIR _{feminine} (dB)	SIR _{masculine} (dB)
256	128	4,50187	3,351795
512	128	5,608008	6,262968
	256	4,183193	5,492846
1024	128	6,788062	7,539744
	256	6,944372	7,494972
	512	7,125142	7,372076
2048	128	9,996282	10,91896
	256	9,9033	10,81405
	512	9,127704	10,33023
	1024	8,993684	9,769833
4096	128	8,769515	12,91166
	256	9,614861	12,9092
	512	3,124741	7,12988
	1024	4,796971	8,205481
	2048	3,237375	6,758649

Os valores apresentados revelam que o aumento no valor de K até 2048 tende a melhorar a qualidade da separação. Porém, os testes com $K = 4096$ não acarretaram em um aumento considerável da SIR, em comparação aos testes com $K = 2048$. Para $K = 4096$, também é possível notar que a qualidade da separação se mostrou mais sensível ao incremento do salto J , acarretando na queda da qualidade das estimativas. Vale ressaltar que o incremento de K aumenta consideravelmente o custo computacional, influenciando diretamente o tempo de processamento do algoritmo de separação.

O tempo de reverberação influenciou consideravelmente a qualidade da separação. A figura 6.2 mostra uma comparação para os sinais estimados da voz feminina para $T_{60}=0,1s$ e para $T_{60}=0,9s$, ambas obtidas para $K = 2048$ e $J=256$.

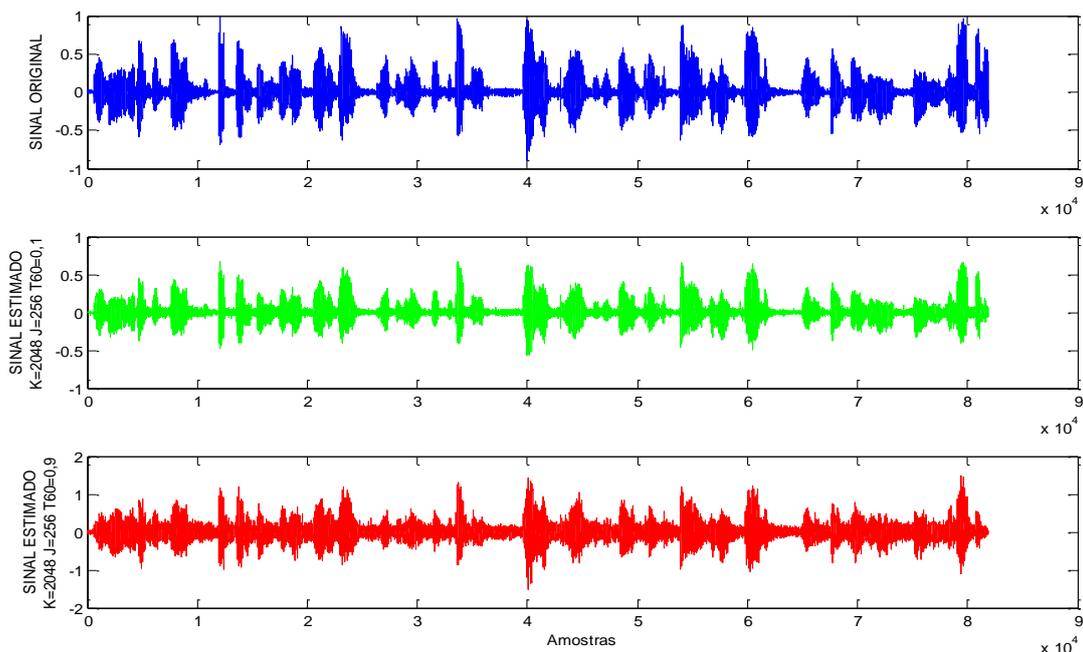


Figura 6.2 Sinais original e estimados para $T_{60}=0,1s$ e $T_{60}=0,9s$ com $K=2048$ e $J=256$

Nesta figura, é visualmente perceptível uma maior presença de interferência no sinal estimado para $T_{60}=0,9$.

Os mesmos testes foram realizados para frequência de amostragem $f_s=16kHz$. As tabelas 6.5 à 6.8 mostram os resultados.

Tabela 6.5 Separações para $T_{60} = 0,1s$ e $f_s=16kHz$

K	J	$SIR_{feminina}(dB)$	$SIR_{masculina}(dB)$
256	128	12,75108	15,13562
	256	11,54098	16,46991
512	128	13,65297	17,56919
	256	21,97456	23,28105
	512	21,77404	23,54371
1024	128	21,80233	23,15696
	256	20,60613	26,67589
	512	20,5625	26,60185
	1024	25,36714	25,68085
2048	128	20,45671	26,28929
	256	20,36394	25,54367
	512	20,32721	25,63968
	1024	20,2805	25,74967
	2048	20,40076	25,46503
4096	128	6,45117	11,54269
	256		
	512		
	1024		

Tabela 6.6 Separações para T60 = 0,5s e fs=16kHz

<i>K</i>	<i>J</i>	SIR _{feminina} (dB)	SIR _{masculina} (dB)
256	128	1,344184	5,368788
512	128	4,996369	7,803246
	256	5,595211	8,041848
1024	128	0,616301	3,813878
	256	8,86863	10,83306
	512	5,857526	10,1244
2048	128	13,31112	14,77044
	256	12,06896	14,94393
	512	12,46607	14,53405
	1024	11,66869	14,51635
4096	128	14,29247	18,70369
	256	14,22192	18,75211
	512	14,33669	18,69736
	1024	8,303919	12,60318
	2048	5,849058	9,355924

Tabela 6.7 Separações para T60 = 0,9s e fs=16kHz

<i>K</i>	<i>J</i>	SIR _{feminina} (dB)	SIR _{masculina} (dB)
256	128	0,678656	4,566619
512	128	2,329481	5,36737
	256	0,751961	3,12716
1024	128	0,862671	3,305022
	256	3,711601	6,200284
	512	5,139901	7,803404
2048	128	10,56132	11,90254
	256	9,631097	12,29447
	512	7,481155	10,94455
	1024	8,34646	11,50324
4096	128	12,0067	14,19304
	256	11,96847	14,18072
	512	11,92931	14,62839
	1024	12,86437	13,93822
	2048	4,90656	6,359339

Tabela 6.8 Separações para T60 = 1,3s e fs=16kHz

<i>K</i>	<i>J</i>	SIR _{feminina}	SIR _{masculina}
256	128	3,413244	2,941399
512	128	1,055587	3,303447
	256	1,299906	4,838998
1024	128	1,496494	3,724478
	256	3,829027	4,455348
	512	5,238067	5,604224
2048	128	9,363931	9,272941
	256	8,128297	9,157304
	512	7,545494	7,925679
	1024	5,691659	9,778978
4096	128	6,819121	10,03699
	256	10,22971	13,06565
	512	6,46379	8,995783
	1024	2,082816	4,928538
	2048	2,487557	3,814013

Para fs=16kHz, o incremento de K de 2048 para 4096 acarretou em aumentos de qualidade das separações, ao contrário do observado para fs=8kHz. No geral, a SIR foi menor em todas as separações com fs=16kHz, indicando que para separar sinais de voz é

melhor utilizar $f_s=8\text{kHz}$. Pode-se perceber também que a qualidade da separação foi melhor, na maior parte das simulações, para a voz masculina. Esse resultado é mais evidente para $f_s=16\text{kHz}$.

6.2.2 – Influência do tempo de Reverberação T_{60}

Para avaliar com mais clareza a influência do tempo de reverberação na qualidade de separação, foram variados os valores de T_{60} de 0,1s até 1,0s com incremento de 0,1s a cada teste. Foram utilizadas outras duas duplas de vozes masculina e feminina. Os valores do tamanho da janela para a STFT e o salto entre janelas foram fixados em $K=2048$ e $J=256$. A taxa de amostragem foi de $f_s=16\text{kHz}$.

Tabela 6.9 1ª Dupla Homem/Mulher – SIR x T_{60}

T_{60}	SIR feminina(dB)	SIR masculina(dB)
0,1	27,7275	25,19
0,2	20,3921	18,9994
0,3	16,0245	16,0069
0,4	15,1069	16,1088
0,5	14,1666	15,4358
0,6	13,1501	14,57
0,7	9,47283	13,3743
0,8	10,2244	12,4016
0,9	9,99043	11,6427
1	9,74697	11,0285

Tabela 6.10 - 2ª Dupla Homem/Mulher – SIR x T_{60}

T_{60}	SIR feminina(dB)	SIR masculina(dB)
0,1	26,7588	25,7124
0,2	18,0137	18,7817
0,3	14,7741	16,197
0,4	13,0479	17,1273
0,5	13,1854	16,1052
0,6	12,8711	15,3052
0,7	12,7832	13,954
0,8	12,2056	13,9975
0,9	10,9693	13,9994
1	10,7757	13,7179

Os resultados indicam que os valores da SIR decaem de forma aproximadamente exponencial. É possível perceber um decaimento abrupto da SIR quando T_{60} varia entre 0,1s e 0,3s. Porém esse decaimento é suavizado quando T_{60} atinge valores mais altos. A figura 6.3 mostra os valores da SIR em função de T_{60} , com as curvas de aproximação exponencial para a 1ª dupla Homem/Mulher:

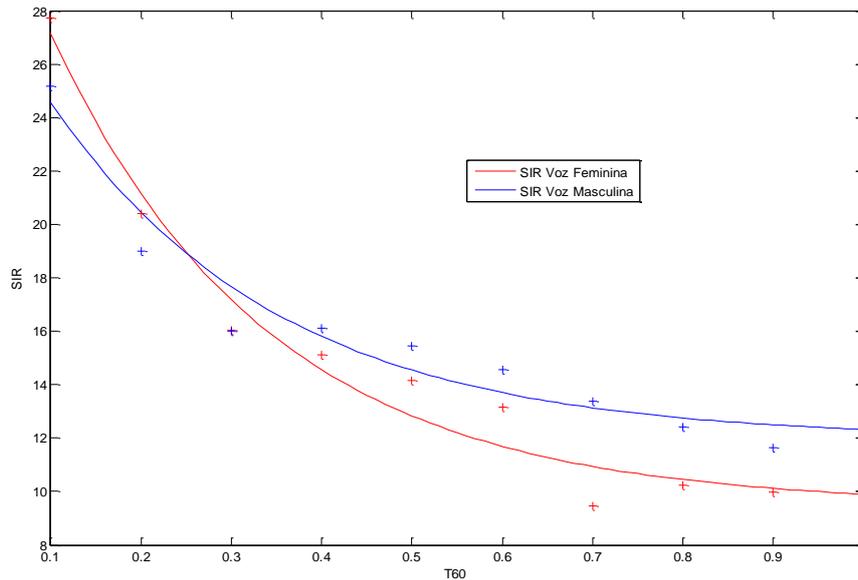


Figura 6.3 Curvas SIR X T60 com os dados da Tabela 6.9

Esse padrão de decaimento foi encontrado em outros testes similares. Pode-se perceber também que a qualidade da separação para a voz feminina foi melhor apenas para valores baixos de T_{60} .

6.3 – Testes com Misturas Gravadas

A seguir serão apresentados os testes realizados com misturas geradas por gravação em duas salas reverberantes. A resposta ao impulso de cada uma delas foi medida e os respectivos tempos de reverberação foram calculados.

6.3.1 – Teste na Sala 1

O ambiente de realização dos testes descritos nesta seção é o Laboratório de Instrumentação e Simulação Acústica (LISA), localizado na sala D110, CT/UFRJ. Foram utilizadas duas fontes de voz masculina e feminina emitidas de caixas de som, que foram posicionadas em seis pontos diferentes. Foram efetuadas as separações para todas as trinta combinações de posicionamento das duas fontes. O espaço do ambiente está representado na figura 6.4 com as suas dimensões e o posicionamento das fontes e dos microfones.

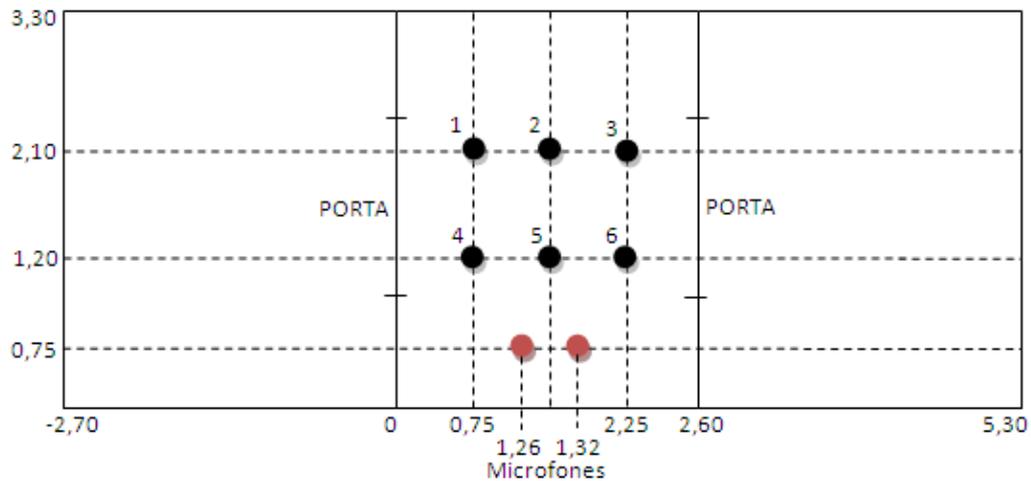


Figura 6.4 Estrutura da Sala 1 de testes (LISA)

As fontes foram colocadas nas posições 1, 2, 3, 4, 5 e 6 representadas na figura 6.4. A parte central da sala, onde foram realizadas as gravações possui duas portas de isolamento acústico. Os testes foram realizados com as portas abertas e posteriormente com as portas fechadas. O tempo de reverberação foi medido nos dois casos.

As figuras 6.5 e 6.6 apresentam as fotografias do ambiente.

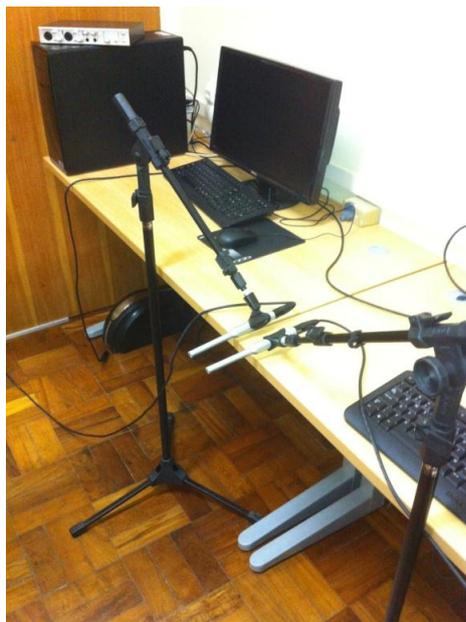


Figura 6.5 Fotografia do LISA mostrando os microfones utilizados nos testes



Figura 6.6 Fotografia do LISA mostrando os microfones e caixa de som utilizados nos testes

Tempo de Reverberação medido no LISA com Portas Abertas

Com as portas abertas, foi estimado o tempo de reverberação em $T_{60}=0,58s$, obtido da curva de decaimento do nível de pressão sonora medida a partir da resposta ao impulso da sala, como mostra a figura 6.7.

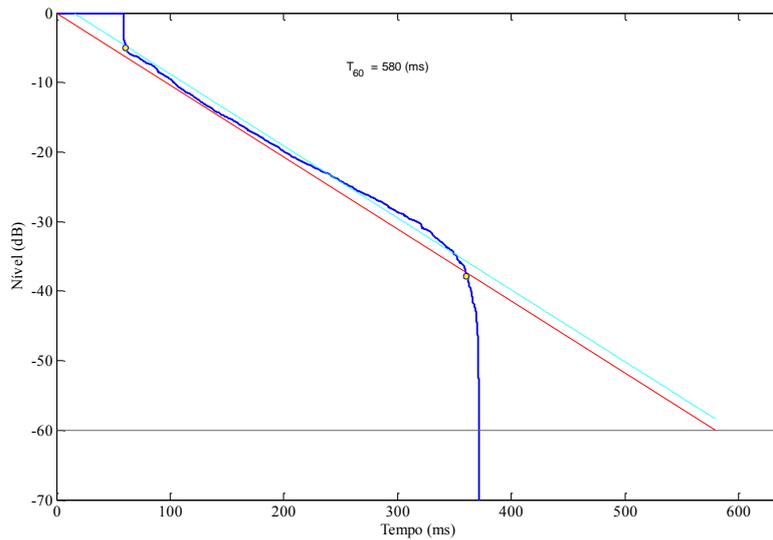


Figura 6.7 Medida do T60 da Sala 1 (LISA) com as portas abertas

A tabela a seguir mostra os valores da SIR obtidos em cada separação para $f_s=8kHz$, $K= 2048$ e $J=256$. As duas primeiras colunas indicam a posição das fontes masculina e feminina, respectivamente, com a numeração estabelecida na figura 6.4.

Tabela 6.11 Resultados dos Testes no LISA com as Portas Abertas

Pos. Masculina	Pos. Feminina	SIR _{masculina} (dB)	SIR _{feminina} (dB)
1	2	7,654005	8,433881
1	3	13,23042	9,168834
1	4	11,1475	10,73644
1	5	11,10456	8,403049
1	6	14,11383	10,40937
2	3	15,16456	10,54643
2	4	12,98835	12,26839
2	5	9,619747	6,452201
2	6	15,2341	10,16866
3	4	12,59094	12,83944
3	5	10,95353	8,680849
3	6	11,99985	9,972361
4	5	12,51982	9,077812
4	6	16,10408	11,86626
5	6	12,49929	10,30307
2	1	10,90256	11,80885
3	1	12,99394	12,13991
4	1	11,53032	7,88028
5	1	11,13408	10,12707
6	1	11,8923	12,1673
3	2	10,24571	12,22915
4	2	12,5414	10,98462
5	2	9,832791	10,98866
6	2	11,42088	14,34353
4	3	14,3937	10,29766
5	3	12,92224	9,649698
6	3	10,75795	9,997294
5	4	12,87342	11,51489
6	4	14,22378	13,51742
6	5	10,47649	10,43596

Calculando-se a média dos valores da SIR para as estimativas das duas fontes, a configuração que obteve o melhor resultado foi para a fonte de voz masculina localizada na posição 4 e a fonte de voz feminina localizada na posição 6. O segundo melhor resultado foi para a fonte de voz masculina localizada na posição 6 e a fonte de voz feminina localizada na posição 4. Provavelmente, a simetria em relação à sala e a proximidade com os microfones contribuíram para as estimativas das fontes nessas configurações. A média da SIR com as portas abertas foi de 11,37dB.

Tempo de Reverberação medido no LISA com Portas Fechadas

Os testes foram repetidos com as portas fechadas. O tempo de reverberação obtido foi de $T_{60}=0,656s$, conforme mostrado na figura 6.8.

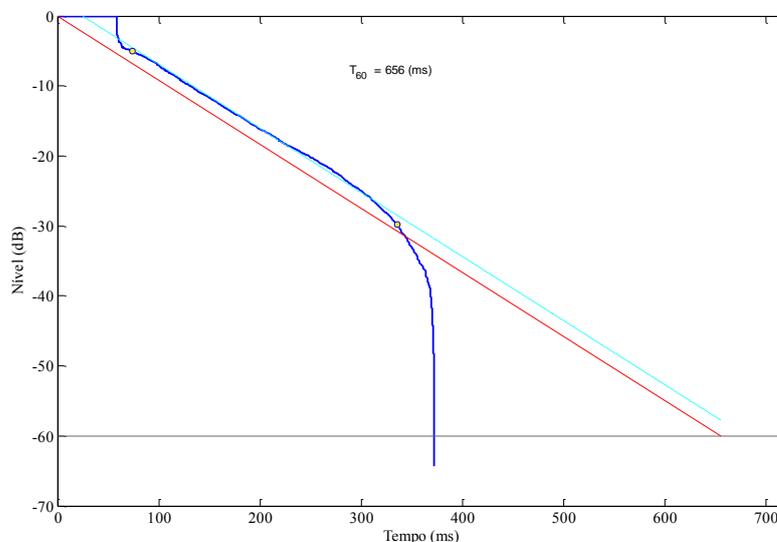


Figura 6.8 Medida do T_{60} da Sala 1 (LISA) com as portas fechadas

Logo, conclui-se que as portas contribuíram para o aumento das reflexões das ondas acústicas, aumentando a reverberação. Os resultados da separação nesse caso encontram-se na tabela 6.12.

Tabela 6.12 Resultados dos Testes no LISA com as Portas Fechadas

Pos. Masculina	Pos. Feminina	SIR _{masculina} (dB)	SIR _{feminina} (dB)
1	2	9,524724	10,20708
1	3	12,04696	11,68228
1	4	8,228132	7,578355
1	5	8,426463	9,717307
1	6	10,72567	10,87171
2	3	11,92795	11,49201
2	4	11,87639	11,3743
2	5	9,14135	9,650579
2	6	10,86452	10,80502
3	4	12,14767	11,43957
3	5	9,017045	9,753786
3	6	13,04906	12,63196
4	5	9,703815	9,579651
4	6	13,93229	13,0579
5	6	13,74831	10,99521
2	1	12,79381	11,30164
3	1	13,37553	12,30152
4	1	11,37503	9,276372
5	1	14,18248	11,95832
6	1	13,97803	11,61377
3	2	11,80172	11,05954
4	2	12,17967	11,2893
5	2	13,02126	10,72201
6	2	10,97629	10,67143
4	3	12,11723	11,04516
5	3	13,70908	12,27389
6	3	14,38769	12,31024
5	4	12,18464	10,82608
6	4	14,46485	12,36634
6	5	13,95985	11,64043

A média da SIR com as portas fechadas foi de 11,09dB. Apesar do tempo de reverberação estimado ter sido um pouco maior para o caso de portas fechadas, a qualidade dos sinais estimados não sofreu queda considerável. Em algumas posições, a separação com as portas fechadas teve melhor desempenho, como por exemplo, quando as posições 6 e 5 foram ocupadas pelas fontes de voz masculina e feminina, respectivamente. O gráfico da figura 6.8 refere-se aos sinais originais e estimados para essa configuração.

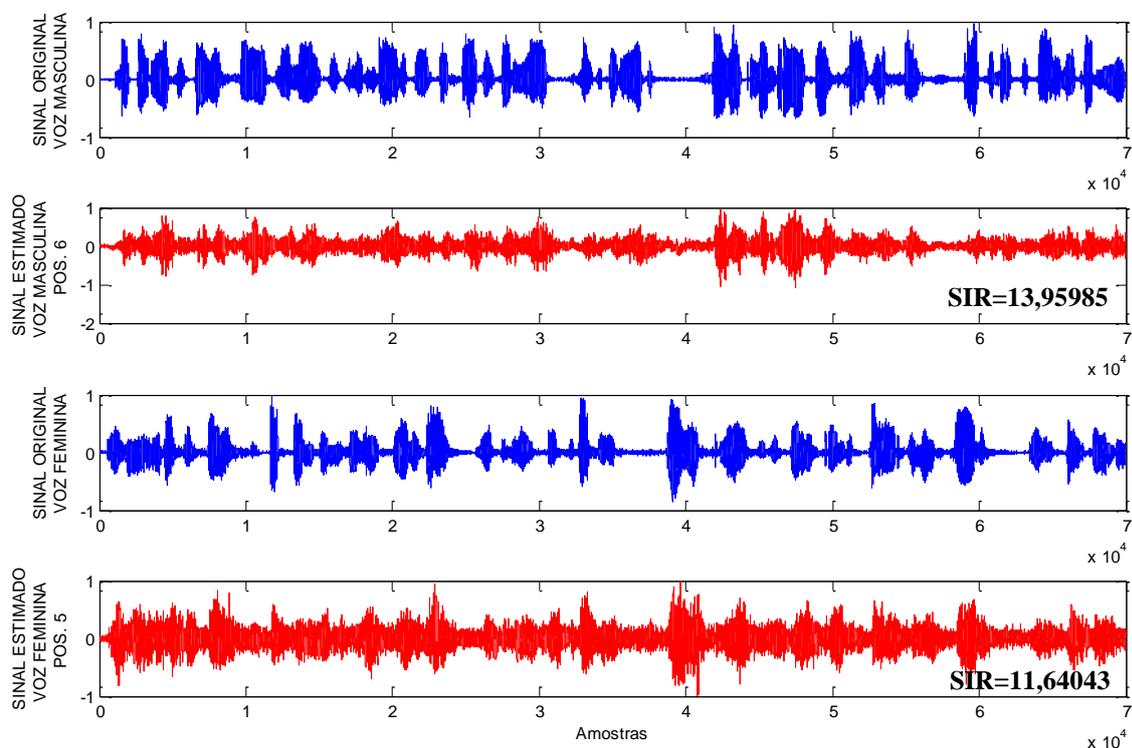


Figura 6.9 Sinais estimados e originais, com fontes nas posições 6 e 5

Percebe-se a presença de uma maior interferência da outra fonte no sinal estimado da voz feminina.

6.3.2 – Teste na Sala 2

O ambiente utilizado para os testes apresentados nessa seção, localizado na sala D105, CT/UFRJ, possui dimensões bem maiores se comparadas às do LISA. Novamente os testes foram realizados com $N = M = 2$ e as fontes foram permutadas de posição em nove diferentes pontos da sala. Foram realizadas as separações das 72 possíveis combinações de localizações das fontes. Como no teste anterior os valores $K=2048$, $J=256$ e $f_s = 8\text{kHz}$ foram mantidos. Os sinais de voz também foram os mesmos.

O espaço do ambiente está representado na figura 6.10 com as suas dimensões e o local do posicionamento das fontes e dos microfones.

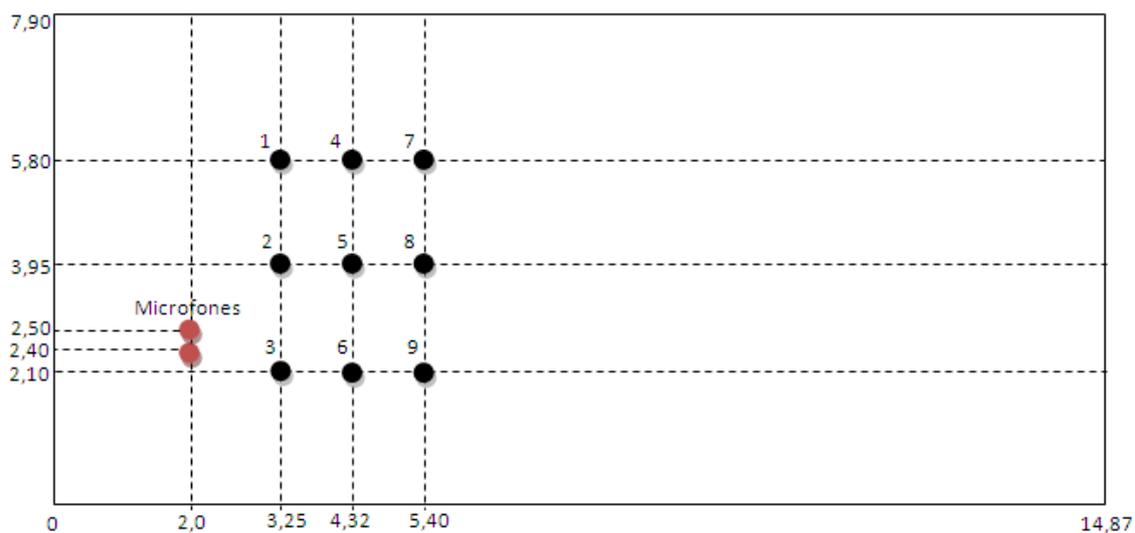


Figura 6.10 Estrutura da Sala 2 de testes

Os microfones não foram colocados em posições simétricas na sala, para ser possível avaliar com mais clareza se fatores, como proximidade com os sensores e/ou ângulos de chegada, influenciam na estimativa de uma ou outra fonte. A figura 6.11 mostra as fotos do ambiente.



Figura 6.11 Fotografias da Sala 2 mostrando os microfones e caixa de som utilizados nos testes

Tempo de Reverberação da Sala 2

O tempo de reverberação da Sala 2 foi obtido da curva de decaimento da pressão sonora medida e seu valor estimado foi de $T_{60}=0,932s$, como mostra a figura 6.12.

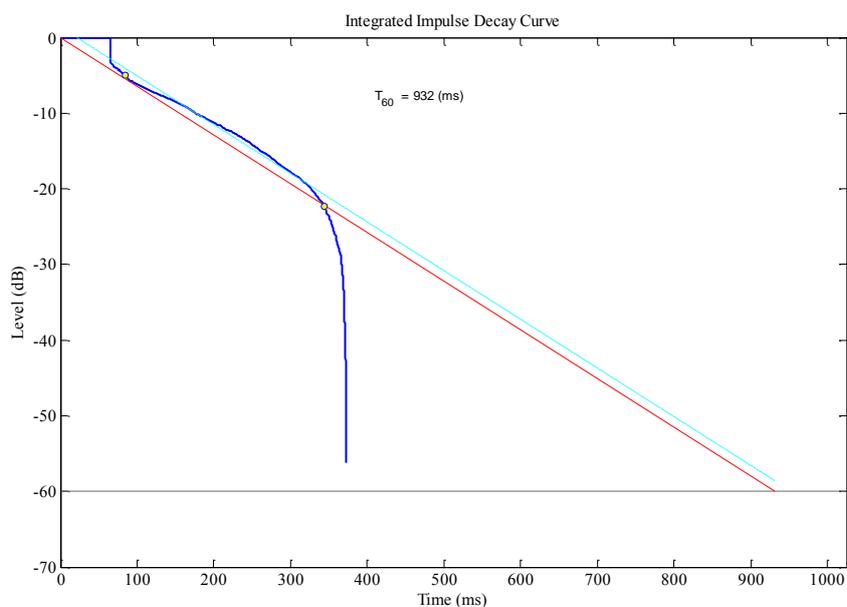


Figura 6.12 Medida do T60 da Sala 2

A tabela 6.13 mostra os valores da SIR resultantes do algoritmo de separação para todas as combinações de posições da Sala 2:

Tabela 6.13 Resultados dos Testes na SALA 2

Pos. Feminina	Pos. Masculina	SIR _{feminina}	SIR _{masculina}
1	2	4,684056	6,564868
1	3	1,070233	10,39268
1	4	1,126363	1,81308
1	5	5,51234	6,830993
1	6	1,408607	9,682777
1	7	2,587633	11,89273
1	8	4,960213	9,845672
1	9	2,774048	11,93164
2	1	1,65457	7,375502
2	3	-1,37136	10,89313
2	4	2,472439	4,840834
2	5	2,20337	7,053684

2	6	0,951586	12,5947
2	7	-4,36367	7,708723
2	8	1,966588	10,85982
2	9	-0,25481	12,10453
3	1	7,933007	4,438977
3	2	8,249368	3,879334
3	4	8,518743	2,297321
3	5	8,119787	4,710211
3	6	3,416011	6,643271
3	7	6,131273	7,252274
3	8	6,051846	6,468176
3	9	-0,53419	3,273505
4	1	3,603969	6,513593
4	2	1,639087	4,126963
4	3	-0,3773	9,984143
4	5	4,32649	6,552719
4	6	1,55054	11,35934
4	7	2,480712	11,75481
4	8	2,43784	9,244929
4	9	0,160505	9,686092
5	1	0,925383	3,916442
5	2	3,604759	5,998389
5	3	0,808106	9,874045
5	4	4,561692	4,415215
5	6	2,072383	10,31903
5	7	2,755144	11,16406
5	8	-0,54791	4,445212
5	9	0,592676	9,18663
6	1	8,312643	1,495159
6	2	8,949545	2,505677
6	3	3,65879	4,337443
6	4	8,058867	-0,19648
6	5	8,91681	2,553339

6	7	3,308122	5,90848
6	8	7,01077	4,914887
6	9	-0,64684	1,208476
7	1	8,648191	9,487612
7	2	7,086758	6,411609
7	3	2,862256	8,461781
7	4	6,662207	6,444781
7	5	7,912088	7,934416
7	6	0,388766	6,478015
7	8	6,62801	8,022916
7	9	2,608587	7,485532
8	1	8,838926	7,022687
8	2	7,236807	3,924758
8	3	4,761142	7,756364
8	4	9,194571	4,414184
8	5	1,679157	-0,07802
8	6	5,131972	6,967535
8	7	4,1964	8,390865
8	9	3,866374	7,024628
9	1	10,52438	4,828378
9	2	9,251047	4,253283
9	3	5,712086	6,576828
9	4	10,39207	3,228981
9	5	7,230421	4,356651
9	6	5,585994	5,40279
9	7	0,366886	1,867734
9	8	9,519937	7,379888

Uma análise cuidadosa dos dados da tabela 6.13, nos permite chegar a algumas conclusões. É possível perceber que houve combinações de posições das fontes que acarretaram em uma boa estimativa de uma delas, em detrimento da outra, com um valor da SIR muito alto e outro muito baixo. Algumas posições pareceram favorecer a separação das fontes que a ocuparam, por exemplo, as estimativas das fontes localizadas nas posições 3, 6 e 9 quase sempre obtiveram qualidade bem superior que a das fontes localizadas em

alguma outra posição. Porém quando as duas fontes se localizavam em duas dessas três posições, as SIR de ambas tenderam a valores mais baixos. Vale ressaltar que nessas posições, as fontes se encontravam praticamente em frente aos microfones, com ângulos de chegada próximos a 90°. Fontes na posição 7 também obtiveram bons valores da SIR, apesar de sua distância em relação aos sensores ser maior. Fontes na posição 1, 2, 4 e 5, quase sempre obtiveram valores muito baixos da SIR, provavelmente pelo fato do ângulo de chegada aos microfones em relação a essas posições serem muito pequenos.

A média da SIR para todas as separações foi de 5,39dB, deixando clara a influência da alta reverberação do ambiente. Na média, a qualidade da estimativa da voz masculina novamente foi melhor que a feminina. Foi traçada a SIR por faixa de frequência e por tempo para o caso de fonte de voz feminina na posição 7 e fonte de voz masculina na posição 5. A escolha dessa configuração para essa análise é devido ao fato de que os valores da SIR para ambas as estimativas foram muito próximos (7,912088 para a fonte de voz feminina e 7,934416 para a fonte de voz masculina) .

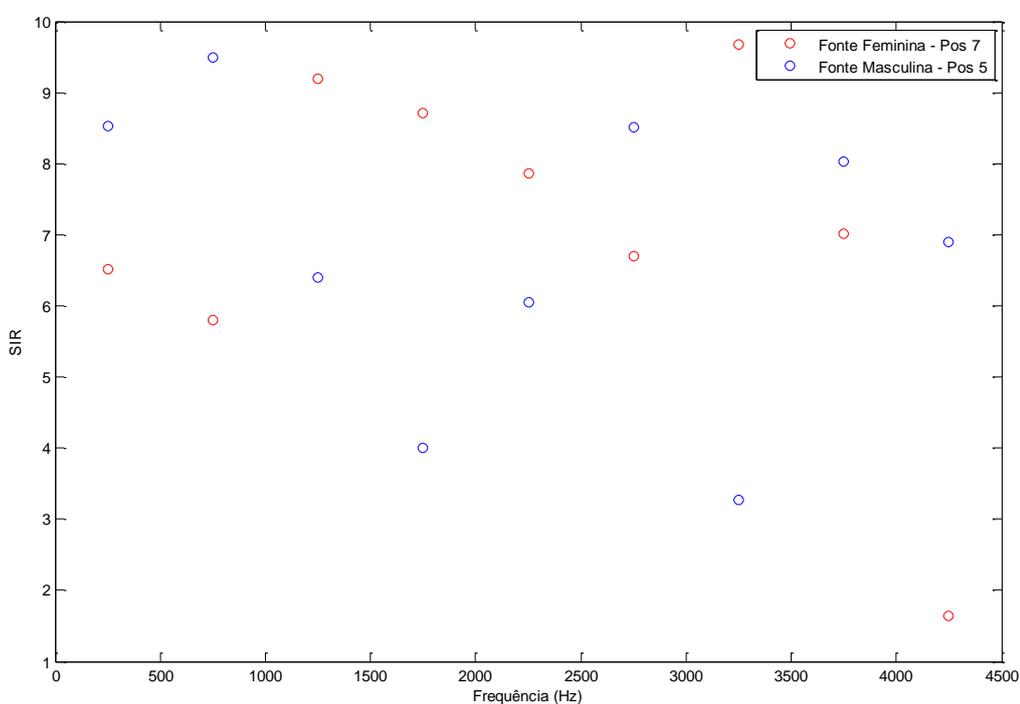


Figura 6.13 SIR por faixa de frequência - Fonte de voz feminina na posição 7 e fonte de voz masculina na posição 5

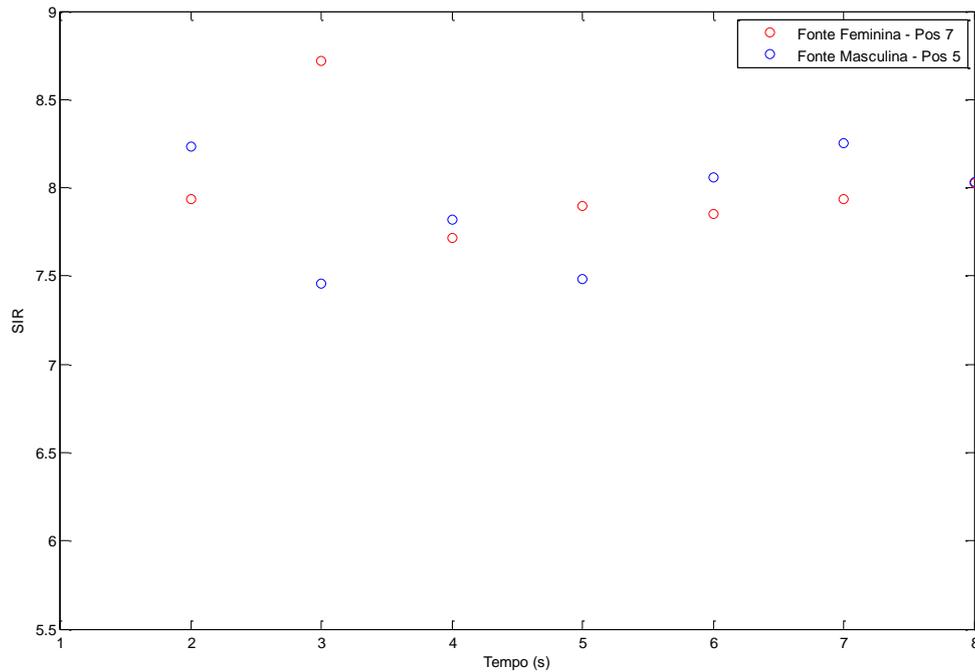


Figura 6.14 SIR por tempo - Fonte de voz feminina na posição 7 e fonte de voz masculina na posição 5

O gráfico da figura 6.13 mostra um espalhamento grande entre os valores da SIR calculados por faixa de frequência. O gráfico da figura 6.14 mostra que, para cada instante de tempo, os valores da SIR de cada uma das fontes foram bem próximos, com exceção no tempo igual a 3 segundos, em que a qualidade da estimativa da voz feminina foi bem maior em relação à da voz masculina. Para esse mesmo posicionamento das fontes foi realizada a separação com o $K=4096$ e $J=256$. Os valores da SIR obtidos estão na tabela 6.14.

Tabela 6.14 Separação para as posições 7 e 5 com $K=4096$ e $J=256$

Pos. Feminina	Pos. Masculina	SIR feminina (dB)	SIR masculina (dB)
7	5	9,698828	8,860833

Com $K=4096$, a fonte de voz feminina obteve uma melhor qualidade de estimativa. O cálculo da SIR por faixa de frequência e por tempo estão nos gráficos das figuras 6.15 e 6.16

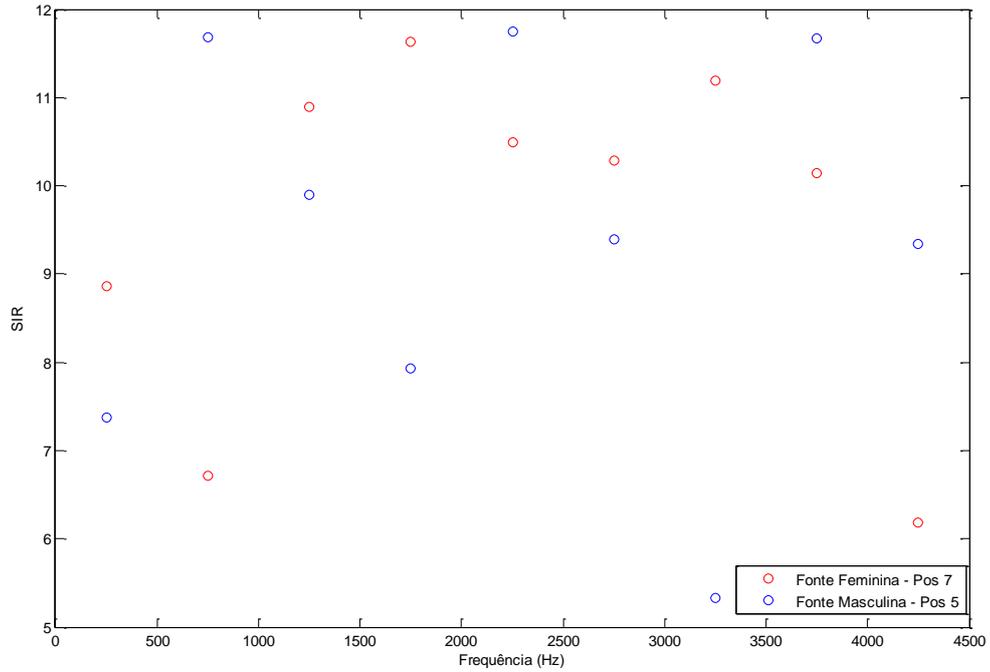


Figura 6.15 SIR por faixa de frequência com $K=4096$ - Fonte de voz feminina na posição 7 e fonte de voz masculina na posição 5

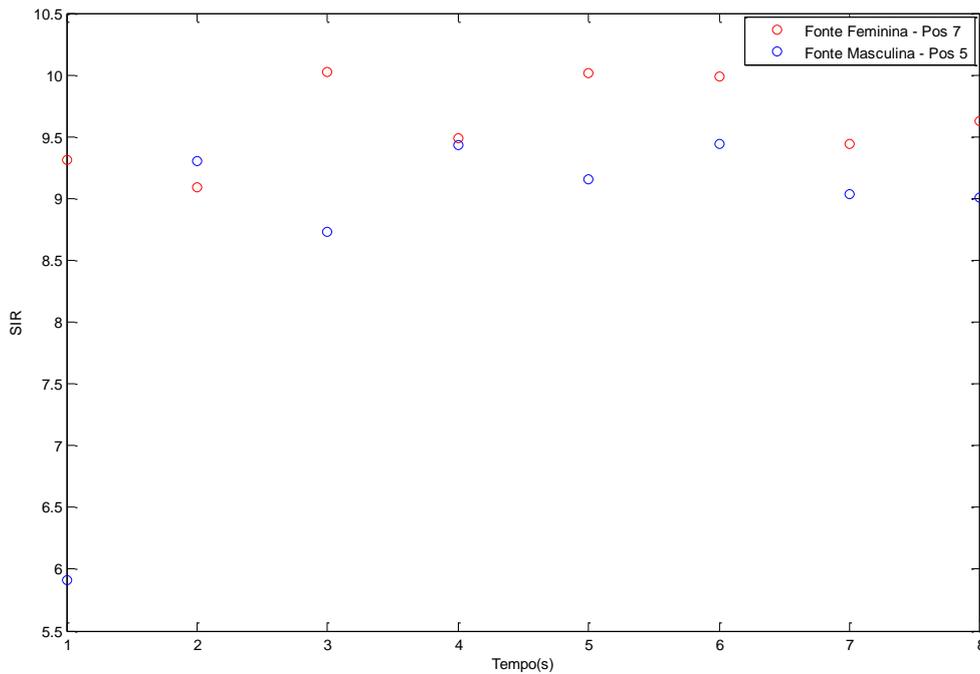


Figura 6.16 SIR por tempo com $K=4096$ - Fonte de voz feminina na posição 7 e fonte de voz masculina na posição 5

Para frequência de 250Hz, a qualidade da estimativa da fonte masculina diminuiu consideravelmente para $K=4096$, porém para outras faixas de frequência não houve

grandes variações de qualidade para ambas as fontes em relação ao teste com $K=2048$. A voz feminina obteve melhor qualidade de separação em quase todos os instantes de tempo nos testes com $K=4096$.

Capítulo 7

Conclusões

Nos Capítulos 2, 3 e 4 foi feita uma revisão teórica acerca dos tópicos relativos ao tema desse trabalho, dos quais podemos citar o modelamento matemático para os sistemas de misturas e de separação, a análise de componentes independentes e o estudo de misturas convolutivas no domínio da frequência.

No Capítulo 5, foram apresentados a definição e os métodos de medição de Resposta ao impulso de uma sala e de estimação do Tempo de reverberação.

No Capítulo 6, foi descrito o procedimento adotado na realização dos testes, com os respectivos resultados e análises.

Com base nos resultados dos experimentos realizados, podemos concluir que a separação cega de fontes de misturas convolutivas no domínio da frequência, utilizando o algoritmo *FastICA* é, em geral, um processo eficiente, mas que depende fortemente das disposições das fontes no ambiente e do ajuste correto de alguns parâmetros. A qualidade da estimativa dos sinais das fontes se mostrou extremamente sensível às variações do tamanho da janela utilizada na STFT. No geral, fixando-se o valor de K em 2048 conseguem-se boas estimativas das fontes, com um custo computacional aceitável. Valor de K igual a 4096 aumenta consideravelmente o tempo de separação sem acarretar necessariamente uma melhor qualidade dos sinais estimados. Os testes também revelaram que a taxa de amostragem igual a 8kHz é melhor para realizar as separações de sinais de voz. O posicionamento das fontes e o ângulo de chegada dos sinais aos sensores também influenciam sensivelmente a qualidade da separação.

Outros cenários para realizações de novos testes podem ser utilizados em trabalhos futuros, bem como a análise da influência da variação de outros parâmetros, tais como a distância entre os sensores e o número de fontes a serem separadas.

Referências Bibliográficas

- [1] A. Hyvärinen, J. Karhunen, and E. Oja, “*Independent Component Analysis*,” Wiley, 2001
- [2] A. Hyvärinen. Fast and Robust Fixed-Point Algorithms for Independent Component Analysis. *IEEE Transactions on Neural Networks* 10(3):626-634, 1999.
- [3] A. PAPOULIS, “*Probability, Random Variables and Stochastic Processes*”, McGraw-Hill, 1991
- [4] Avargel, Y., and Cohen I. (2010), “*Modeling and Identification of Nonlinear Systems in the Short-Time Fourier Transform Domain*,” *IEEE Transactions on Signal Processing*.
- [5] BODE H., C. SHANNON, “*A simplified derivation of linear least squares smoothing and prediction theory*” *Proc. IRE*, Vol. 38, pag 417-425, Apr. 1950
- [6] H. Kuttruff, “*Room Acoustics*”, 4th ed. Taylor & Francis, Oct. 2000.
- [7] IKRAM, M. Z., MORGAN, D. R., “*A beamforming approach to permutation alignment for multichannel frequency-domain blind speech separation*”, *Proc. ICASSP*, pp. 881–884, 2002.
- [8] LEHMANN, Eric A., JOHANSSON, Anders M., “*Prediction of energy decay in room impulse responses simulated with an image-source model*”. Crawley WA, Australia, Western Australian Telecommunications Research Institute, 2008.
- [9] MURATA, N., IKEDA, S., ZIEHE, A., “*An approach to blind source separation based on temporal structure of speech signals*”, *Neurocomputing*, v. 41, pp. 1–24, 2001.
- [10] PETRAGLIA, Mariane Rembold, BATALHEIRO, Paulo Bulkool, HADDAD, Diego Barreto, “*Análise de Componentes Esparsos e de Componentes independentes e suas respectivas aplicações em Separação cega de fontes*”. Rio de Janeiro, Universidade Federal do Rio de Janeiro, 2008.
- [11] SCHROEDER, M. R., “*New method of measuring reverberation time*”. New Jersey, EUA, Bell Telephone Laboratories, 1964.
- [12] Disponível em <http://www.gradadm.ifsc.usp.br/dados/20122/FCM02081/Acustica%20de%20Salas.pdf>
Acesso em 25 ago. 2013, 16:53.
- [13] Disponível em < <http://www.ymec.com/products/dssf3e/>>. Acesso em 28 ago. 2013, 17:52.