

LISELENE DE ABREU BORGES

**EXTRAÇÃO DE PARÂMETROS CARACTERÍSTICOS PARA  
DETECÇÃO ACÚSTICA DE VAZAMENTO DE ÁGUA.**

SÃO PAULO  
FEVEREIRO/2011

LISELENE DE ABREU BORGES

**EXTRAÇÃO DE PARÂMETROS CARACTERÍSTICOS PARA  
DETECÇÃO ACÚSTICA DE VAZAMENTO DE ÁGUA.**

Tese apresentada à Escola Politécnica da  
Universidade de São Paulo para a obtenção  
do título de Doutor em Engenharia.

Área de Concentração: Sistemas Eletrô-  
nicos.

Orientador: Prof. Dr. Miguel Arjona Ramírez

SÃO PAULO  
FEVEREIRO/2011

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

São Paulo, 31 de maio de 2011.

Assinatura do autor

Assinatura do orientador

## FICHA CATALOGRÁFICA

Borges, Liselene de Abreu

Extração de parâmetros característicos para detecção acústica de vazamento de água / L.A Borges. – ed. rev. – São Paulo, 2011.

60 p.

Tese (Doutorado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Sistemas Eletrônicos.

1. Processamento Digital de Sinais 2. Vazamentos de Água  
3. Predição Linear Perceptual 4. Máquina de Vetores de Suporte  
I. Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Sistemas Eletrônicos II.t.

# Agradecimentos

Aos meus pais, Liamar e Jair e as minhas irmãs, Sílvia e Lisiane.

Ao meu orientador prof. Dr. Miguel e aos meus colegas de trabalho do projeto DVAP.

Aos meus amigos Bruna, Guilherme, Letícia, Claudete e José Eduardo.

Ao Instituto Nokia de Tecnologia e ao Venturus Centro de Inovação Tecnológica que incentivaram através de seus programas de bolsa de estudos.

E a todas aquelas pessoas que me auxiliaram, de forma direta ou indireta, no desenvolvimento desta tese.

Dedico à minha amiga  
Silvana Dacol (in memoriam).

# Resumo

Este trabalho apresenta a pesquisa sobre a extração de parâmetros característicos de sinais acústicos para fins de detecção automática de vazamento de água em tubulações enterradas. Os sinais acústicos foram adquiridos com o auxílio de um geofone eletrônico e também catalogados por técnicos especialistas em detecção acústica. De todos os sinais foram extraídos os modelos de predição linear perceptual de várias ordens, determinando-se como melhor a ordem 2. A partir de um conjunto de modelos de referência de sinais de vazamento, a distância média de Itakura dos outros modelos em relação a estas referências foram calculadas. Em conjunto com estas distâncias, quatro características espectrais são também extraídas do sinal a fim de compor o vetor de parâmetros característicos do sinal. Parte destes vetores de parâmetros característicos são utilizados para treinar o classificador de máquina de vetores de suporte. O restante dos dados são, então, submetidos a este classificador que obteve a taxa de acerto de classificação em torno de 93%. Experimentos anteriores, utilizando modelos de predição linear, de ordem 10, obtiveram uma taxa de acerto em torno de 82%. Isso demonstra que estes novos parâmetros característicos propostos alcançam os objetivos deste trabalho, que são algoritmos com melhor taxa de acerto na detecção de vazamentos.

**Palavras-chave:** Vazamento de água, Processamento de sinais, Predição linear, Máquina de vetores de suporte.

# Abstract

This work presents a research about feature extraction of acoustic signals for detection of water leak in buried pipes. Acoustic signals were acquired by means of an electronic geophone and also labeled by technicians specialized in acoustic water leak detection. For every signals, its linear predictive model was estimated for a range of prediction orders, concluding for the best order 2. Out of this group of models, some leaky ones are used as reference for calculating the Itakura mean distance with respect to the other models. Completing this measure, four spectral features are extracted to compose the signal feature vector. Some of these vectors were used to train a support vector machine to be used as a classifier. The remaining ones were used to evaluate the classification. The resulting accuracy rate achieved is around 93%. Earlier experiments, which use linear prediction of order 10 had an accuracy rate around 82%. This shows that this novel proposal of feature vector achieves the main goal of this research, which is the increase in the leak detection accuracy rate.

**Keywords:** Water Leak, Signal processing, Linear prediction, Support vector machine.

# Sumário

<b>Lista de Abreviaturas</b>	<b>vii</b>
<b>Lista de Figuras</b>	<b>viii</b>
<b>Lista de Tabelas</b>	<b>xi</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Motivação . . . . .	1
1.2 Objetivos e Contribuições . . . . .	1
1.3 Organização do Trabalho . . . . .	2
<b>2 Detecção acústica de vazamento de água</b>	<b>3</b>
2.1 Os tipos de vazamento . . . . .	3
2.2 O procedimento de detecção do vazamento . . . . .	4
2.3 O sinal sonoro de vazamento . . . . .	7
<b>3 Sistema de detecção dos padrões acústicos de vazamento</b>	<b>13</b>
3.1 Resumo histórico . . . . .	13
3.2 Reconhecimento de padrão acústico de vazamento . . . . .	14
3.2.1 Máquina de vetores de suporte . . . . .	17
3.3 Métodos de avaliação dos parâmetros característicos . . . . .	19
3.3.1 Avaliação individual dos parâmetros . . . . .	20
3.3.2 Avaliação dos parâmetros em conjunto . . . . .	21

<b>4</b>	<b>Extração de parâmetros característicos</b>	<b>23</b>
4.1	Introdução . . . . .	23
4.2	Parâmetros característicos individuais . . . . .	24
4.2.1	Energia de Curto Prazo . . . . .	24
4.2.2	Taxa de cruzamento por zero . . . . .	25
4.2.3	Centróide espectral . . . . .	25
4.2.4	Largura do espectro . . . . .	25
4.2.5	Fluxo espectral . . . . .	27
4.2.6	Ponto de <i>Roll-off</i> . . . . .	27
4.2.7	Taxa de espalhamento . . . . .	27
4.3	Parâmetros característicos estatísticos . . . . .	28
4.3.1	Modelo de predição linear . . . . .	28
4.3.2	Modelo de predição linear perceptual . . . . .	29
<b>5</b>	<b>Desenvolvimento da pesquisa</b>	<b>35</b>
5.1	Construção da Base de Dados . . . . .	35
5.1.1	Definição das Bases para Treinamento e Teste . . . . .	36
5.2	Desenvolvimento prático . . . . .	36
5.3	Avaliação dos parâmetros individualmente . . . . .	39
5.4	Avaliação em grupo . . . . .	47
5.4.1	Classificação de acordo com a ordem do modelo de predição linear . . . . .	48
5.4.2	Classificação para grupos de diversas combinações de parâmetros . . . . .	51
<b>6</b>	<b>Conclusões e Propostas de Trabalhos</b>	<b>56</b>
6.1	Conclusões . . . . .	56
6.2	Propostas de Trabalhos Futuros . . . . .	57
	<b>Referências Bibliográficas</b>	<b>58</b>

# Lista de Abreviaturas

ABENDI	Associação Brasileira de Ensaios Não Destrutivos e Inspeção
AR	Modelo Auto-regressivo
dB	decibel
DEP	Densidade Espectral de Potência
DFT	do inglês <i>Discrete Fourier Transform</i>
LP	do inglês <i>Linear Prediction</i>
PCM	do inglês <i>Pulse Code Modulation</i>
PEAD	Polietileno de Alta Densidade
SPL	do inglês <i>Sound Pressure Level</i>
SVM	do inglês <i>Support Vector Machine</i>
TA	Taxa de Acerto Total
TANV	Taxa de Acerto de Não Vazamento
TAV	Taxa de Acerto de Vazamento

# Lista de Figuras

2.1	Tipos de vazamentos possíveis.[SAB05]	3
2.2	Haste de escuta metálica[PN03].	4
2.3	Manuseio da haste de escuta metálica pelo operador[SR 10].	5
2.4	Correlacionador de ruído[PN03].	5
2.5	Foto Correlacionador	6
2.6	Geofone eletrônico[PN03].	6
2.7	Manuseio do geofone eletrônico[SR 10].	7
2.8	Fontes de ruído de vazamento.	7
2.9	Amostra de sinal sem ruído de vazamento	8
2.10	Trecho de 20 ms da amostra do sinal sem ruído de vazamento	9
2.11	DEP de sinal sem ruído de vazamento	9
2.12	Amostra de sinal com ruído de vazamento	10
2.13	Trecho de 20 ms da amostra do sinal com ruído de vazamento	10
2.14	DEP de sinal com ruído de vazamento	11
2.15	DEP de sinais em tubulação de PEAD	11
2.16	DEP de sinais em tubulação de Ferro dúctil	12
3.1	Sistema detecção de padrão acústico de vazamento	14
3.2	Diagrama em blocos das técnicas de pré-processamento dos sinais acústicos.	14
3.3	Janelas de segmentação do tipo retangular e de Hamming no tempo.	16
3.4	Espectro de frequência das janelas retangular e de Hamming.	16
3.5	Processo de segmentação do sinal, adaptado de[Wee10].	17
3.6	Exemplo de dados linearmente separáveis para dimensão dois.	18

3.7	Espaço com dados não separáveis linearmente . . . . .	19
3.8	Representação gráfica da estatística-M. . . . .	20
4.1	Diagrama de blocos da extração de parâmetros característicos. . . . .	24
4.2	Ilustração do centróide espectral de um sinal. . . . .	26
4.3	Ilustração da largura espectral de um sinal. . . . .	26
4.4	Ilustração do ponto de <i>roll-off</i> de um sinal. . . . .	27
4.5	Diagrama em blocos da análise por predição linear perceptual. . . . .	29
4.6	Anatomia do ouvido humano. . . . .	30
4.7	Diagrama em blocos do processamento psicoacústico. . . . .	30
4.8	Curva da relação entre as escalas de frequência hertz e Bark. . . . .	31
4.9	Filtro de banda crítica de frequência central de aproximadamente 4 Barks. . . . .	32
4.10	Limiar de percepção mínimo. . . . .	32
5.1	DEP do sinal sem vazamento original e pré-processado. . . . .	37
5.2	DEP do sinal com vazamento original e pré-processado. . . . .	38
5.3	Histograma : energia de curto prazo. . . . .	40
5.4	Histograma : taxa de cruzamento por zero. . . . .	41
5.5	Histograma : centróide espectral. . . . .	41
5.6	Histograma : largura de espectro . . . . .	42
5.7	Histograma : fluxo espectral. . . . .	42
5.8	Histograma : ponto de <i>roll-off</i> . . . . .	43
5.9	Histograma : taxa de espalhamento . . . . .	43
5.10	Histograma : Máxima distância de Itakura - AR . . . . .	43
5.11	Histograma : Distância Média de Itakura - AR . . . . .	44
5.12	Histograma : Máxima diferença espectral - AR . . . . .	44
5.13	Histograma : Diferença média espectral - AR . . . . .	45
5.14	Histograma : Distância máxima de Itakura - PLP . . . . .	45
5.15	Histograma : Distância média de Itakura - PLP . . . . .	46
5.16	Histograma : Diferença máxima espectral - PLP . . . . .	46

5.17	Histograma : Diferença média espectral - PLP . . . . .	47
5.18	Taxa de acerto da distância máxima de Itakura por ordem de predição. . .	49
5.19	Taxa de acerto da distância média de Itakura por ordem de predição. . . .	49
5.20	Taxa de acerto da distância máxima espectral por ordem de predição. . . .	50
5.21	Taxa de acerto da diferença média espectral por ordem de predição. . . .	50

# Lista de Tabelas

3.1	Tipos de função Kernel. . . . .	19
5.1	Subgrupos de dados . . . . .	35
5.2	Divisão do grupo de dados em teste e treino . . . . .	36
5.3	Taxa de acerto para distância de Itakura para $p = 10$ . . . . .	38
5.4	Medida de separabilidade . . . . .	39
5.5	Melhores resultados da taxa de acerto por parâmetro característico . . . . .	48
5.6	Taxa de acerto dos parâmetros característicos mais relevantes . . . . .	51
5.7	Classificação com sinal original AR e PLP . . . . .	52
5.8	Taxa de acerto com pré-ênfase de primeira ordem para AR e PLP . . . . .	52
5.9	Classificação com pré-ênfase de segunda ordem AR e PLP . . . . .	52
5.10	Resultado do classificador para o sinal original . . . . .	53
5.11	Resultado do classificador para o sinal com pré-ênfase de primeira ordem . . . . .	54
5.12	Resultado do classificador para o sinal com pré-ênfase de segunda ordem . . . . .	54

# Capítulo 1

## Introdução

### 1.1 Motivação

Um sistema de distribuição de água potável, que é constituído de tubulações, submersas ou não, está sempre vulnerável a apresentar perdas em algum ponto da sua rede. Estas perdas podem ocorrer tanto em função de ligações clandestinas quanto em função de vazamentos. Dentro das perdas decorrentes de vazamentos, os motivos mais comuns são problemas de instalação da tubulação, desgaste natural do material ou mesmo rupturas em função de intervenções humanas.

Além disso, estudos da concessionária de tratamento e distribuição de água de São Paulo, a SABESP[SAB10], mostram que o volume de água tratada desperdiçado por dia está em torno de 26%, sendo que 40% deste é relativo às perdas em vazamentos. Isto equivale ao abastecimento diário de uma cidade de 600 mil habitantes. Portanto, mais do que nunca, métodos de detecção e prevenção de desperdícios de água são imprescindíveis hoje em dia.

### 1.2 Objetivos e Contribuições

Esta pesquisa se propôs a pesquisar métodos de extração de parâmetros característicos para serem utilizados em um sistema automático de detecção acústica de vazamento. Atualmente, a detecção acústica ainda é executada de forma manual, na qual um técnico especializado ouve o sinal sonoro com um geofone, e, a partir do seu conhecimento, identifica a presença ou não de ruído de vazamento.

Na literatura atual já existem algumas propostas de automação deste processo através da aplicação de técnicas de processamento digital de sinais. Porém, nenhuma ainda demonstrou resultados satisfatórios em se tratando de detecção para sinais coletados em campo.

Um dos principais objetivos desta pesquisa foi o de sugerir novos parâmetros caracte-

rísticos a serem extraídos dos sinais desta natureza que garantam a melhor taxa de acerto de detecção de vazamentos possível. Partiu-se então da avaliação das propostas atuais encontradas na literatura. Como referência, escolheu-se a proposta apresentada por Yuriko e Akira [TM08]. Nesta proposta, eles utilizam como parâmetros característicos a medida máxima e média de Itakura dos modelos de predição em relação ao modelo de predição de referência (sinal com vazamento) e também o maior autovalor do sinal; e submetem estes três parâmetros a um classificador de máquina de vetores de suporte. Em um estudo anterior [BR10] esta proposta foi avaliada utilizando dados reais coletados em campo e a taxa de acerto de detecção foi de 82%. Foi então que se começou a analisar outras alternativas de parâmetros a serem extraídos dos sinais.

Acompanhando a maneira com a qual os técnicos fazem a identificação dos sinais com ruído de vazamento e levando em consideração que o ouvido humano tem uma forma em particular de perceber os sinais acústicos, propôs-se a utilização de modelos de predição linear perceptual, ao invés do modelo normal. Com este modelo perceptual as particularidades do sistema auditivo humano poderiam ser incorporadas ao modelo de predição linear. Adicionalmente, em substituição ao parâmetro maior autovalor, sugeriu-se a utilização de outras características obtidas a partir do espectro do sinal. Com esta combinação de parâmetros, consegue-se obter melhores resultados, que são apresentados no capítulo de resultados desta tese.

Uma contribuição que agregou um diferencial a este trabalho foi a utilização de filtros de pré-ênfase de primeira e de segunda ordem. Com a utilização dos filtros, a taxa de acerto utilizando somente os parâmetros espectrais é superior à obtida na avaliação anterior. Assim, os resultados obtidos contribuem para a melhora na qualidade do monitoramento de falhas na rede de distribuição de água.

### 1.3 Organização do Trabalho

Esta tese está organizada da seguinte forma: no Capítulo 2 são apresentados os métodos manuais hoje utilizados na detecção acústica de vazamento, além dos sinais acústicos que contêm ou não ruído de vazamento. Em seguida, no Capítulo 3, é descrito um sistema de detecção automática de vazamento incluindo as fases de pré-processamento sugeridas. Também serão descritos os métodos com os quais os parâmetros característicos foram avaliados. No Capítulo 4 são descritos os métodos de extração dos parâmetros característicos que foram avaliados. Já no Capítulo 5 é descrito o processo de criação da base de dados é descrito e são apresentados os resultados obtidos através dos experimentos realizados. Por fim, no Capítulo 6, são apresentadas as conclusões desta pesquisa e as propostas de trabalhos futuros.

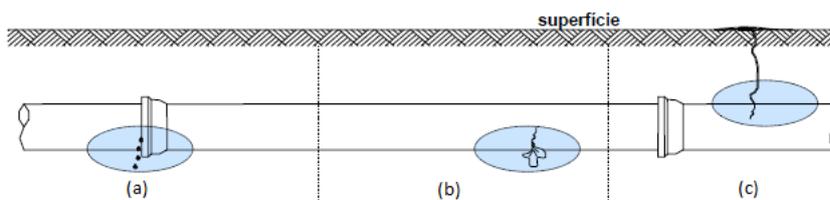
## Capítulo 2

# Detecção acústica de vazamento de água

Neste capítulo são apresentados, primeiramente, os tipos de vazamentos possíveis encontrados na rede de distribuição de água. Logo após são descritos os métodos manuais utilizados hoje em dia na detecção acústica de vazamento. Por fim, são apresentados exemplos de sinais acústicos que contêm ou não ruído de vazamento.

### 2.1 Os tipos de vazamento

Os vazamentos de água em uma rede de distribuição de água potável em geral são classificados em função da forma como podem ou não ser detectados. São basicamente três os tipos de vazamentos existentes. Os vazamentos visíveis são aqueles tipos de vazamento que podem ser identificados visualmente, uma vez que a água aflora à superfície onde a tubulação está enterrada, Figura 2.1(c). Já os vazamentos inerentes, são vazamentos que não são visíveis e tampouco podem ser detectados, uma vez que são vazamentos de uma granularidade muito pequena, Figura 2.1(a). Por fim, os vazamentos não visíveis são os tipos de vazamento que, apesar de não poderem ser vistos, são passíveis de detecção através de métodos acústicos, Figura 2.1(b).



**Figura 2.1:** *Tipos de vazamentos possíveis.*[SAB05]

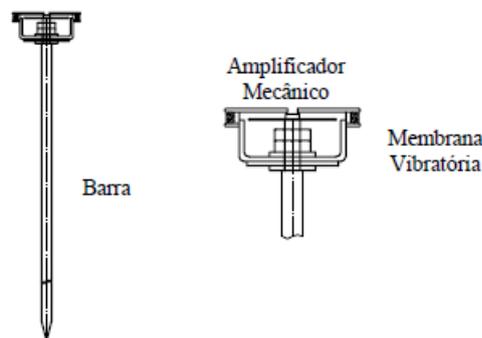
## 2.2 O procedimento de detecção do vazamento

O procedimento de detecção de vazamento em um sistema de distribuição de água é composto por diversas fases. Cada fase deste procedimento procura reduzir a região onde é feita busca pelo local exato do vazamento.

Normalmente, uma rede de distribuição de água é monitorada por diversos tipos de sensores, como sensores acústicos, sensores de pressão e sensores de vazão, que estão instalados na rede de tubulação. Uma vez identificado através de algum destes sensores que existe algum comportamento estranho na rede é então iniciado um processo pela busca do local exato do defeito na rede. As perdas na rede distribuição, além das causadas por vazamentos, também podem ser decorrentes de ligações clandestinas.

Uma vez identificada a região de distribuição onde o defeito foi encontrado, entram em cena os técnicos especializados na detecção acústica de vazamento. Este tipo de técnico é treinado na escuta de ruídos característicos de vazamento de água neste tipo de cenário e, para isto, segue os procedimentos definidos pela norma PR-051 da Abendi[ABE04].

Esta norma define que o primeiro equipamento a ser utilizado é a haste de escuta, mostrada na Figura 2.2. Com este equipamento, o técnico verifica, no ponto de entrada de água do consumidor, os ruídos existentes na rede. O ponto de entrada de água do consumidor é o chamado "pé-de-cavalete" da rede. É neste local que se encontra o medidor de consumo. Veja exemplo na Figura 2.3. No momento em que o técnico encontrar algum ruído suspeito é solicitado que o registro de entrada de água do consumidor seja fechado. Uma vez que o registro esteja fechado, o técnico faz novamente uma sessão de escuta. Com este procedimento, o técnico procura verificar se o vazamento é dentro da área de responsabilidade do consumidor ou se é na tubulação externa, de responsabilidade da concessionária de distribuição.



**Figura 2.2:** Haste de escuta metálica[PN03].

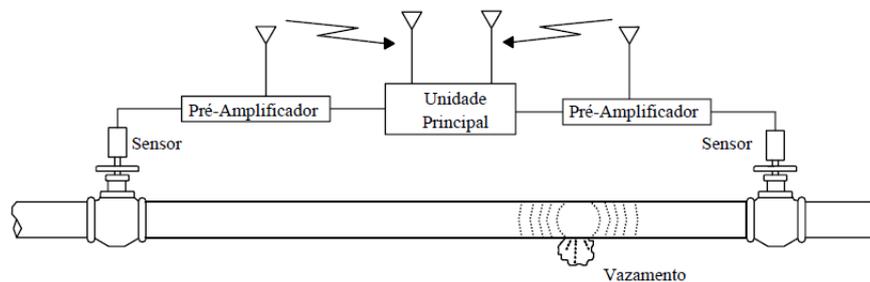
Uma vez confirmado que o ruído é na área de responsabilidade da concessionária, inicia-se, uma segunda fase da escuta. Esta segunda fase pode ser feita por um equipamento chamado de geofone ou por outro chamado de correlacionador de ruído. De acordo também com a norma PR-051, a pressão mínima recomendada da rede de distribuição para se



**Figura 2.3:** *Manuseio da haste de escuta metálica pelo operador[SR 10].*

fazer a detecção acústica de vazamento é de  $1,5 \text{ kgf/cm}^2$  (o equivalente a 15 metros de coluna de água). Em redes onde a pressão está abaixo deste valor, a detecção acústica é comprometida uma vez que em tubulações de baixa pressão geralmente não se pode garantir um mesmo padrão auditivo do ruído.

Quando o técnico utiliza o correlacionador de ruído, Figura 2.4, o equipamento aponta dentre dois pontos de medição uma aproximação da distância entre o ponto medido e o local do vazamento. Este equipamento ouve o ruído em um dos sensores e, ao escutar o mesmo ruído no segundo sensor, calcula o tempo que este ruído teve para se propagar de um sensor ao outro. Tendo medido este tempo e a distância entre os sensores, a localização do vazamento é feita através do processo de correlação entre estas duas informações. Este tipo de equipamento não é muito utilizado, visto que ele exige que os sensores estejam em contato direto com a tubulação e que também a tubulação não tenha derivações. Veja exemplo de utilização do correlacionador na Figura 2.5.



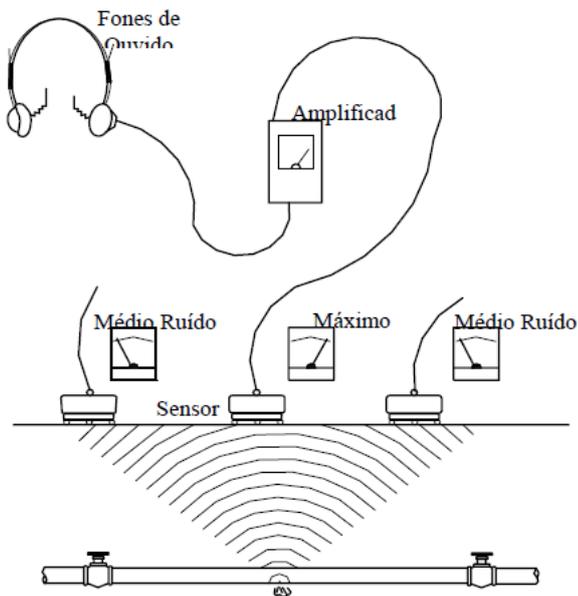
**Figura 2.4:** *Correlacionador de ruído[PN03].*

A outra forma de se localizar o vazamento é através do uso do geofone. O geofone, Figura 2.6, é um equipamento que escuta o ruído de vazamento através do contato do sensor com o solo. Este equipamento pode ser do tipo mecânico ou eletrônico. A diferença entre os tipos é basicamente devido à existência, no geofone eletrônico, de circuitos eletrônicos



**Figura 2.5:** Foto de um técnico manuseando um correlacionador de ruído[SR 10].

que adicionam funcionalidades extra ao equipamento. Porém, ambos os equipamentos requerem que a sua utilização seja feita por um técnico especializado na escuta deste tipo de sinal, veja exemplo na Figura 2.7.



**Figura 2.6:** Geofone eletrônico[PN03].

Um geofone eletrônico é basicamente constituído por um cristal piezoelétrico, um amplificador de sinal e alguns filtros. O cristal piezoelétrico serve como sensor, pois, quando é submetido a uma pressão, gera um campo elétrico (em um eixo transversal àquele onde foi aplicada a pressão) que pode ser coletado na forma de tensão. Como a tensão que o sensor gera é de baixa amplitude, é necessário a utilização de um amplificador. Os filtros, por sua vez, são úteis na remoção de ruídos indesejados que podem prejudicar o processo de escuta do ruído de vazamento.

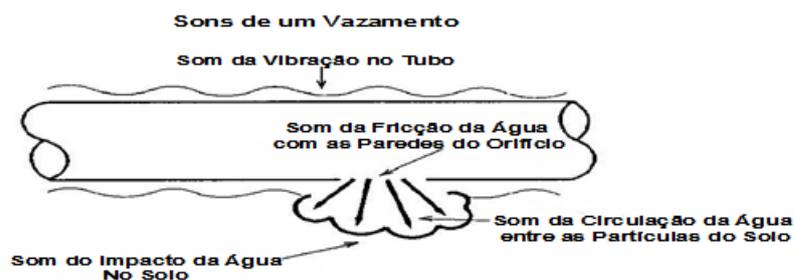


**Figura 2.7:** *Manuseio do geofone eletrônico[SR 10].*

No presente trabalho, os sinais de interesse serão adquiridos através de um geofone eletrônico, sem a utilização de nenhum tipo de filtro a ele associado. Com isto, espera-se obter sinais com e sem a presença do ruído de vazamento o mais fielmente possível para que eles possam ser utilizados no desenvolvimento desta pesquisa.

## 2.3 O sinal sonoro de vazamento

Um sinal sonoro de vazamento é uma combinação de diversos sinais sonoros emitidos por fontes distintas, como mostra a Figura 2.8 [Zan09]. São basicamente quatro diferentes tipos de fontes. Estas fontes são oriundas do impacto da água com o solo onde a tubulação está enterrada, da vibração da tubulação em função da movimentação da água para fora da tubulação, da circulação da água entre as partículas do solo e da fricção da água com as paredes da tubulação. Quando a escuta é feita no local exato do vazamento, todas estas fontes podem ser percebidas. Já na escuta em locais próximos ao do vazamento, somente algumas fontes podem ser percebidas e, mesmo assim, na forma de propagação do sinal através da tubulação.



**Figura 2.8:** *Diferentes tipos de fontes sonoras que compõem um sinal acústico de vazamento em rede subterrânea.*

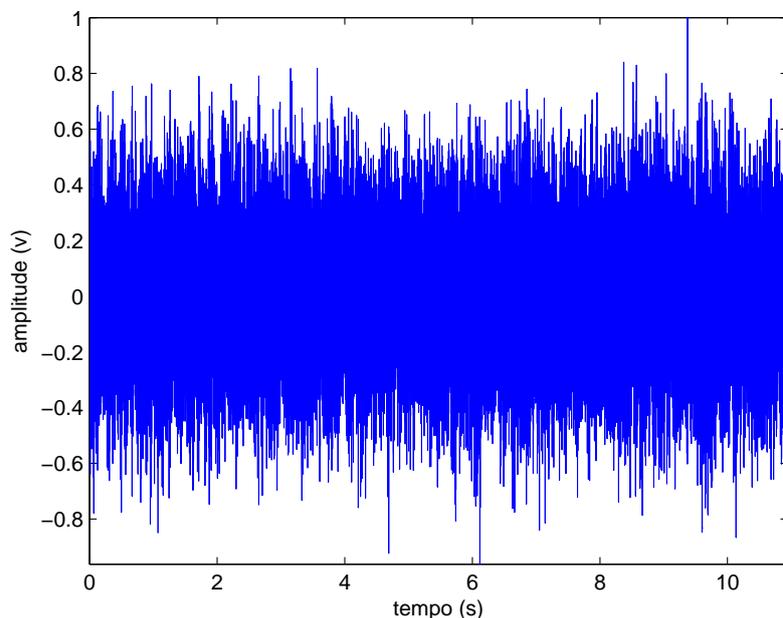
O formato do ruído de vazamento também pode ser influenciado pelo tipo do material de que é feita a tubulação, pelo tipo de solo onde está enterrada, pelo diâmetro e pelo comprimento da tubulação e também pela profundidade na qual está enterrada.

Segundo [Hun00], as tubulações feitas de material metálico possuem mais componentes de alta frequência e também propagam o sinal gerado com mais facilidade. Por outro lado, tubulações feitas de material plástico produzem mais componentes de baixa frequência e, ao contrário das metálicas, não propagam o sinal gerado com tanta facilidade.

Os sinais de vazamento são mais audíveis em tubulações enterradas em solos arenosos do que em solos argilosos. Também os sinais são mais audíveis onde a tubulação está sob asfalto ou concreto do que sob gramado.

Nesta tese, restringimos a utilização de somente sinais acústicos de vazamento e não vazamento em tubulações de PEAD e de Ferro dúctil. Esta escolha deve-se ao fato de que estes são os tipos mais comuns de materiais utilizados nas redes de distribuição no Brasil.

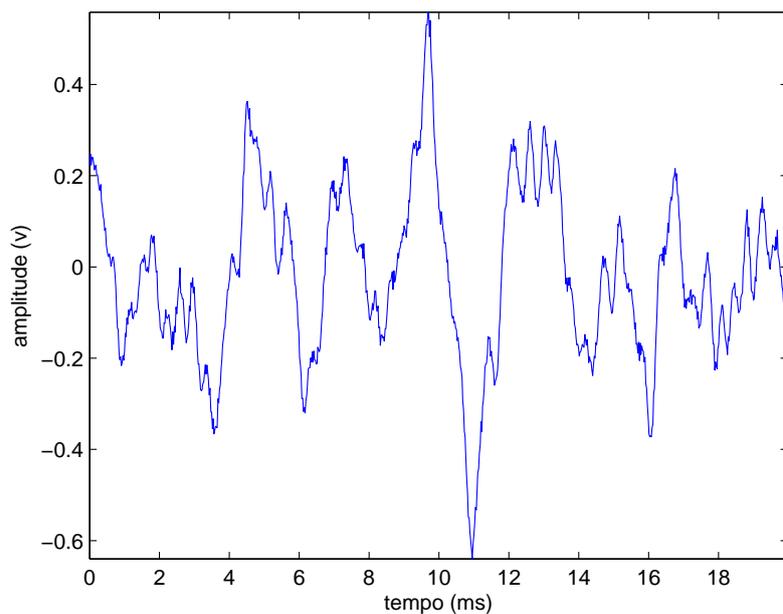
Na Figura 2.9 é apresentada uma amostra de sinal sonoro de não vazamento e, na Figura 2.12, outra amostra, desta vez de sinal de vazamento. Avaliando ambos os sinais visualmente não se pode dizer muito a respeito, uma vez que ambos têm formato muito parecido.



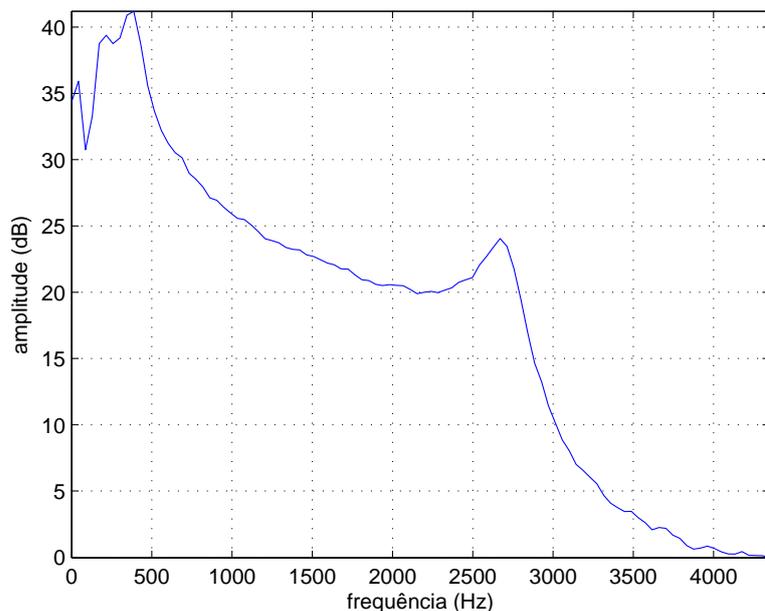
**Figura 2.9:** *Amostra de sinal sem ruído de vazamento*

Para melhor avaliar os sinais foi extraído de cada sinal um segmento de duração de 20ms e estes são apresentados nas Figuras 2.10 e na Figura 2.13, sendo a primeira figura a de um sinal de não vazamento e a segunda a de um sinal de vazamento.

A partir destes segmentos menores do sinal percebe-se nitidamente que o sinal de não vazamento é bem mais ruidoso em relação ao sinal de vazamento. Uma das explicações

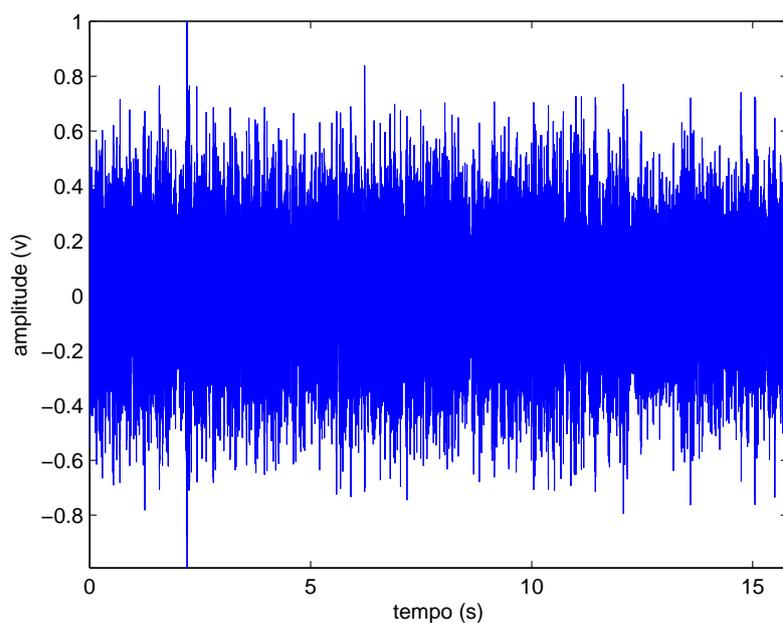


**Figura 2.10:** Trecho de 20 ms da amostra do sinal sem ruído de vazamento

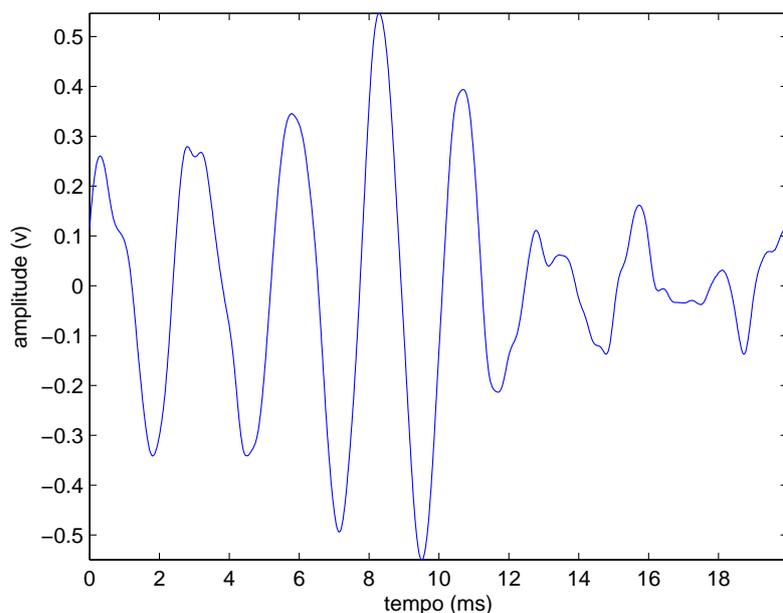


**Figura 2.11:** Densidade espectral de potência média entre janelas do sinal sem ruído de vazamento

para este fenômeno é de que como o sensor é muito sensível e capta sinais de amplitude muito pequena, quando não há sinal de vazamento o sinal captado é basicamente o ruído de ambiente, que tende a uma característica de ruído branco. Ambos os sinais aparecem normalizados para uma melhor comparação, uma vez que os sinais sem ruído de vazamento têm amplitude da ordem de milésimos de volts, ao contrário dos sinais com ruído de vazamento, que têm amplitude na ordem de décimos de volts

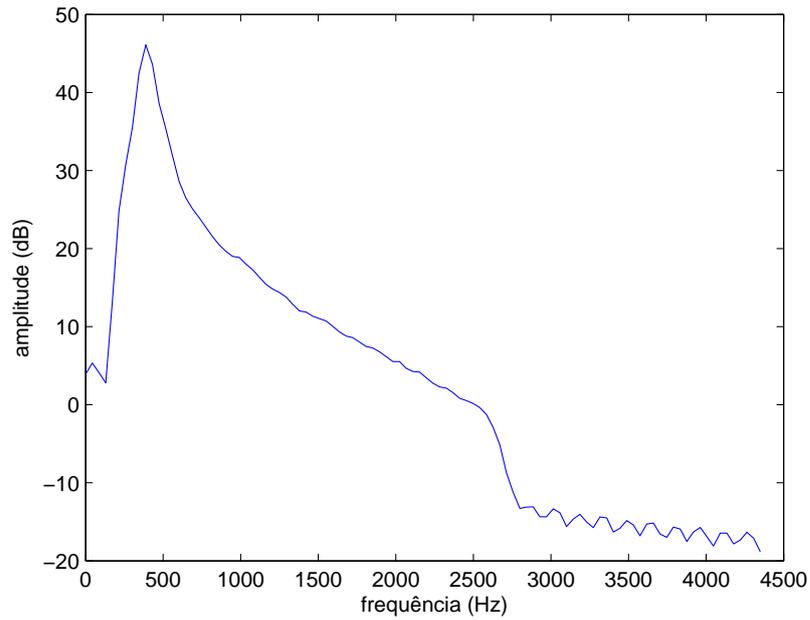


**Figura 2.12:** *Amostra de sinal com ruído de vazamento*

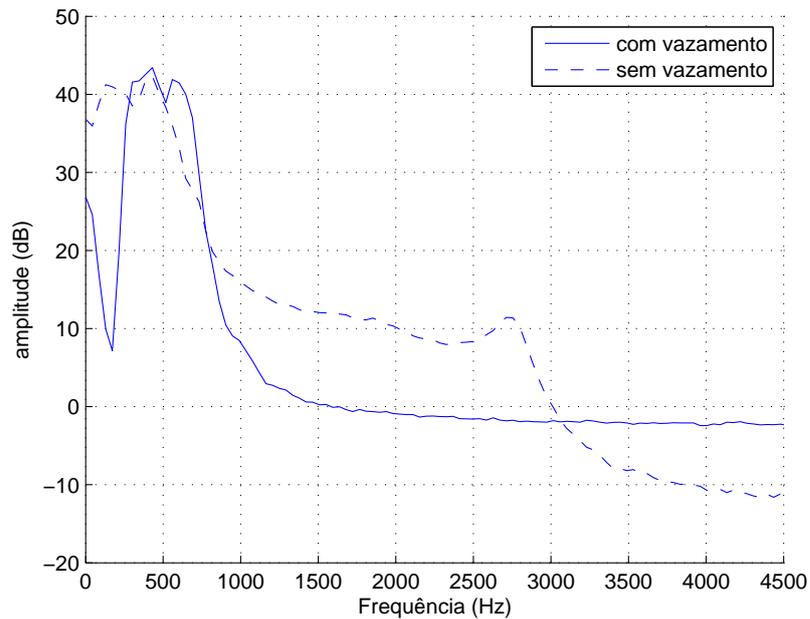


**Figura 2.13:** *Trecho de 20 ms da amostra do sinal com ruído de vazamento*

E, por fim, a densidade espectral potência (DEP) média entre frames do sinal, de ambos os sinais é apresentada. Na Figura 2.11 é apresentada a DEP do sinal de não vazamento e na Figura 2.14, a DEP do sinal de vazamento. Para o cálculo da DEP, os sinais foram segmentados, utilizando janelas do tipo Hamming, em segmentos de duração de 20ms, e transformados para o domínio espectral através da transformada discreta de fourier (DFT) [OSB99].

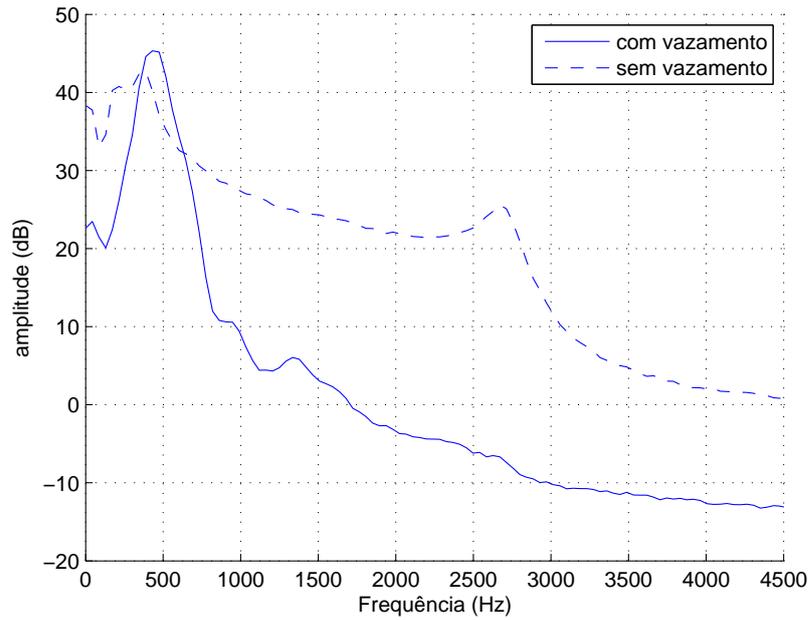


**Figura 2.14:** *Densidade espectral de potência média entre janelas do sinal com ruído de vazamento*



**Figura 2.15:** *Densidade espectral de potência média entre janelas dos sinais em tubulação de PEAD*

A partir de uma análise espectral dos sinais pode-se notar que ambos os espectros estão ceifados em 3k Hz. Esta é uma característica do próprio sensor piezoelétrico. Também é possível notar que o sinal de não vazamento tem altos valores de energia praticamente em toda a banda de passagem do sensor, com dois picos, sendo um de lóbulo, mais largo, entre 200 e 400 Hz e o outro, mais estreito em torno de 2.6k Hz. Já o sinal de vazamento



**Figura 2.16:** *Densidade espectral de potência média entre janelas do sinais em tubulação de Ferro dúctil*

tem somente um pico, bem marcado, em torno de 400 Hz.

A partir do gráfico do espectro dos sinais, verifica-se que a banda de frequência dos sinais adquiridos com ruído de vazamento é bem mais estreita, em torno de 1,5kHz enquanto que a banda dos sinais que não contém vazamento ocupa boa parte do espectro disponível, que é de aproximadamente 4kHz.

Buscando verificar diferenças entre sinais obtidos em diferentes tipos de tubulação, são apresentadas, nas Figuras 2.15 e 2.16, exemplos de DEP de sinais obtidos em dutos de PEAD e de Ferro dúctil, respectivamente. Verifica-se que para ambos os espectros dos sinais de não vazamento o seu formato é muito semelhante. Com relação aos sinais de vazamento, os de tubulação de Ferro dúctil possuem um pico mais marcado que o de tubulação de PEAD.

# Capítulo 3

## Sistema de detecção dos padrões acústicos de vazamento

Neste capítulo descrevemos um sistema básico de detecção de padrões acústicos de vazamento. Primeiramente é apresentado um breve resumo histórico da pesquisa na área de detecção dos padrões de vazamento de água por métodos acústicos. Em seguida, apresentamos as diversas implementações deste tipo de sistema encontradas na literatura. Por fim, abordamos o método de classificação que foi escolhido para ser utilizado na seleção dos parâmetros característicos mais relevantes.

### 3.1 Resumo histórico

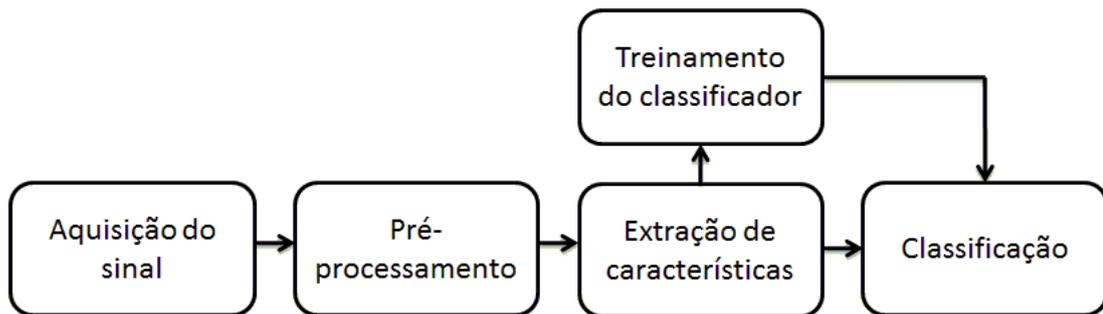
Um dos primeiros sistemas de detecção dos padrões de vazamento através do processamento dos sinais acústicos por ele produzidos foi apresentado em 1994 por Seaford Harry, apud J. Yang, Y. Wen e P. Li, 2008 [YWL08]. Este sistema é baseado na extração de padrões do espectro do sinal em alta frequência. Já em 1995, Savic e Michel [Sav95] propuseram um novo sistema utilizando os coeficiente de predição linear cepstrais. Em 2006, Ai et al [ASZ<sup>+</sup>06] descreveram um sistema que utilizava os coeficientes cepstrais de predição linear e modelos ocultos de Markov.

Entre 2007 e 2008, Toshitaka, Akira e Sato [SM07, TM08] propuseram sistemas utilizando máquina de vetores de suporte como classificador e algumas medidas com relação aos modelos auto-regressivos das amostras dos sinais como parâmetros característicos.

Também em 2008, Jin et al [YWL08] propuseram um sistema que utilizava a entropia aproximada como parâmetro de um sistema de decisão baseado em rede neural.

## 3.2 Reconhecimento de padrão acústico de vazamento

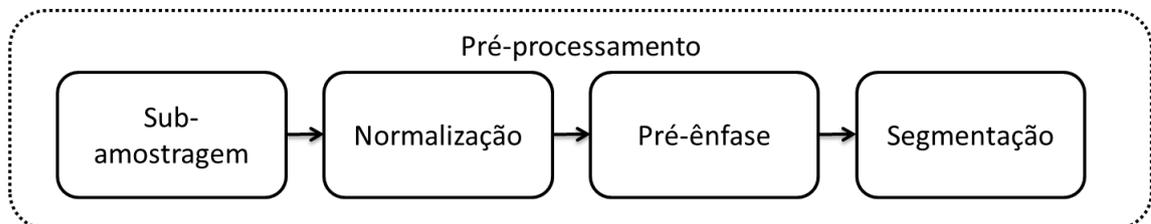
Um sistema de detecção é uma especificação de um sistema de reconhecimento de padrões cujo objetivo é classificar padrões a partir de informações previamente conhecidas. Um sistema de reconhecimento de padrões é composto basicamente de quatro etapas distintas. Estas quatro etapas são: a aquisição do sinal, o pré-processamento, um mecanismo de extração de parâmetros característicos e um sistema de classificação de padrões baseado nas parâmetros característicos extraídos, conforme ilustra a figura 3.1.



**Figura 3.1:** Sistema detecção de padrão acústico de vazamento

A primeira etapa, a de aquisição do sinal, é a mais importante para o sistema como um todo. Esta importância se deve ao fato de que tudo que é feito nas fases seguintes depende diretamente da qualidade das amostras de sinal adquiridas. As amostras de sinal adquiridas devem corresponder fielmente ao formato do sinal original.

Após a aquisição do sinal é usual, em processamento de sinais, a utilização de técnicas de pré-processamento. Estas técnicas têm o intuito de melhorar a qualidade do sinal assim como também ressaltar as informações mais relevantes. As técnicas escolhidas de pré-processamento são definidas de acordo com as características do sistema. Na Figura 3.2 são apresentadas, em forma de diagrama de blocos, as técnicas de pré-processamento utilizadas nesta tese.



**Figura 3.2:** Diagrama em blocos das técnicas de pré-processamento dos sinais acústicos.

O primeiro pré-processamento é baseado na análise do espectro de potência do sinal, onde se verifica a possibilidade de subamostrar o sinal. Com a subamostragem é possível diminuir o número de amostras do sinal a ser processado e, com isto, eliminar informações redundantes contidas no sinal.

O segundo pré-processamento é o de normalização do sinal. A normalização é útil uma vez que ela busca eliminar as diferenças de amplitude entre sinais capturados com diferentes equipamentos em diferentes situações. O cálculo da normalização é descrito na Equação 3.1, onde  $s(n)$  é a amostra do sinal  $s$  no tempo  $n$ .  $N$  é o número total de amostras do sinal a ser analisado e  $\mu_s$  e  $\sigma_s$  são a média e o desvio padrão do sinal  $s$ , calculados sobre todo o sinal, conforme ilustra a equação abaixo:

$$\tilde{s}(n) = \frac{s(n) - \mu_s}{\sigma_s}. \quad (3.1)$$

A terceira técnica de pré-processamento utilizada, que visa ressaltar as componentes de alta frequência com relação às demais, é a pré-ênfase. Um filtro de pré-ênfase de primeira ordem, no domínio temporal, é definido por

$$\hat{s}(n) = \tilde{s}(n) - \alpha_1 \tilde{s}(n - 1), \quad (3.2)$$

onde  $\alpha_1$  é derivado dos coeficientes de autocorrelação do sinal [RS78]. O valor do coeficiente  $\alpha_1$  deve estar entre  $0.9 \leq \alpha_1 \leq 1.0$ .

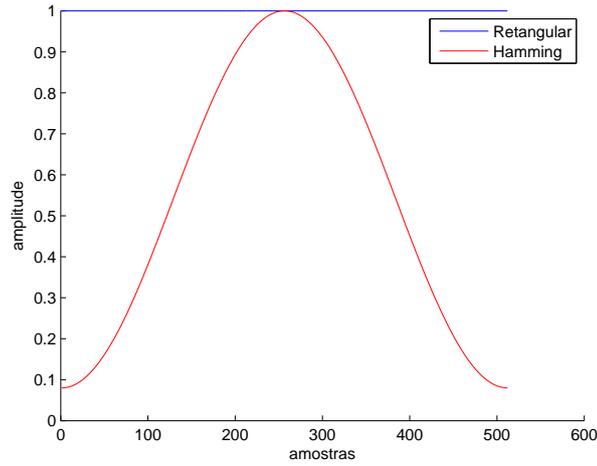
O filtro de pré-ênfase nada mais é que um sinal passa-altas com ganho  $1 + \alpha_1$  na frequência de Nyquist. Como procuramos evidenciar as pequenas nuances entre os sinais com e sem ruído de vazamento, propomos então a aplicação de uma segunda pré-ênfase, mas, neste caso, sobre o sinal já modificado pela primeira pré-ênfase. Com a utilização deste filtro de segunda ordem espera-se um realce ainda maior das componentes de maior frequência com relação às de menor frequência.

A última etapa do pré-processamento é a segmentação. A segmentação é a divisão do sinal original em segmentos de curta duração, em geral algo em torno de 10 ms a 40 ms. A segmentação é necessária para que possamos fazer a análise de curto prazo, pois, dentro deste intervalo de tempo restrito, pode-se considerar o sinal como sendo estacionário. Uma vez que o sinal é estacionário é possível aplicar técnicas de processamento de sinais, como a transformada de Fourier discreta, sem o risco de os resultados não serem coerentes, uma vez que estas técnicas se baseiam em sinais de comportamento estacionário.

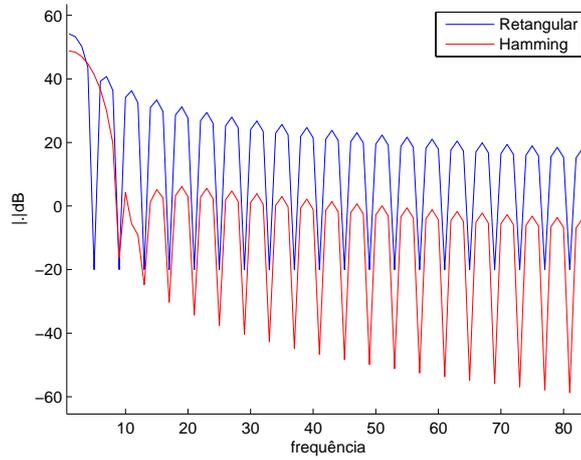
Para fazer a segmentação do sinal é necessária a utilização de janelas de segmentação. Existem diversos tipos de janelas de segmentação. Dois exemplos são as janelas retangular e a de Hamming, como ilustra a Figura 3.3. A janela retangular é descrita por:

$$w(n) = \begin{cases} 1 & \text{se } 0 \leq n \leq N - 1; \\ 0 & \text{para demais valores de } n \end{cases} \quad (3.3)$$

e a janela de Hamming pela Equação 3.4, onde  $N$  é o número total de pontos da janela.



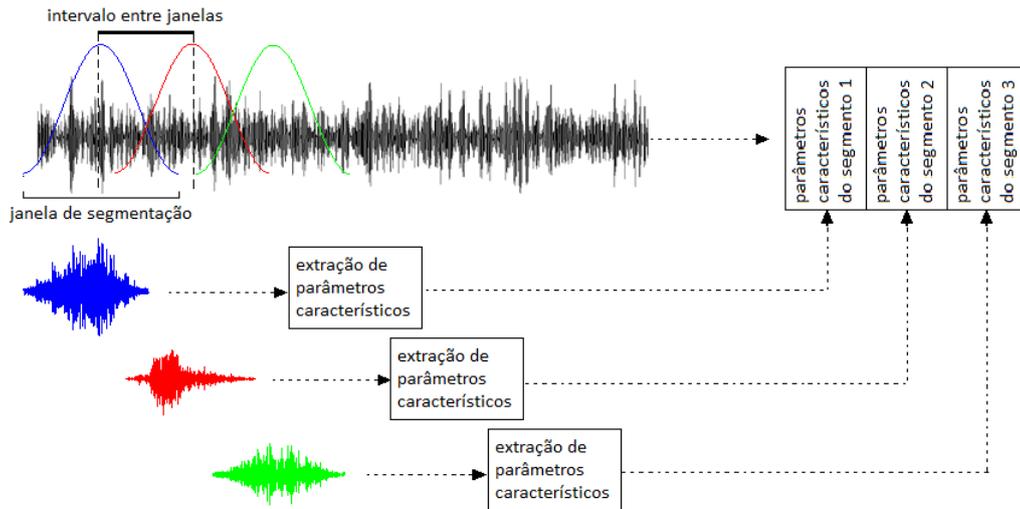
**Figura 3.3:** *Janelas de segmentação do tipo retangular e de Hamming no tempo.*



**Figura 3.4:** *Espectro de frequência das janelas retangular e de Hamming.*

$$w(n) = \begin{cases} 0.54 - 0.46 \cos\left(\frac{2\pi n}{N}\right) & \text{se } 0 \leq n \leq N - 1; \\ 0 & \text{para demais valores de } n \end{cases} \quad (3.4)$$

Normalmente quando o processamento é feito no domínio do tempo utiliza-se janela do tipo retangular. Já quando o processamento é no domínio espectral, a utilização deste tipo de janela não é recomendada. Apesar de esta janela identificar uma dada frequência com maior precisão, a sua definição de amplitude é pior, uma vez que o espectro da janela retangular possui lóbulos secundários com grande amplitude [OSB99]. Para resolver este problema é utilizada a janela de Hamming, que possui lóbulos secundários de menor amplitude, ver Figura 3.4. Contudo, como este tipo de janela possui no tempo um decaimento nas bordas, se faz necessária a sobreposição entre janelas consecutivas durante a segmentação, como mostra a Figura 3.5. Esta solução garante que nenhuma informação será perdida durante o processo de segmentação.



**Figura 3.5:** *Processo de segmentação do sinal, adaptado de [Wee10].*

Após a etapa de pré-processamento, vem a etapa de extração de parâmetros característicos. Esta etapa consiste em extrair do sinal pré-processado informações que sejam relevantes à caracterização do sinal ou mesmo à caracterização de medidas estocásticas do mesmo. Cada tipo de sinal possui um conjunto de parâmetros característicos que podem ser extraídos. A escolha de quais parâmetros característicos serão extraídos está associada diretamente à natureza do sinal. Para que se consiga extrair os parâmetros característicos mais importantes, uma análise detalhada do sinal de interesse deve ser feita a fim de buscar quais são as informações mais relevantes de cada tipo de sinal. A extração de parâmetros característicos é o foco principal desta tese e por este motivo será descrita com mais detalhes em um capítulo à parte.

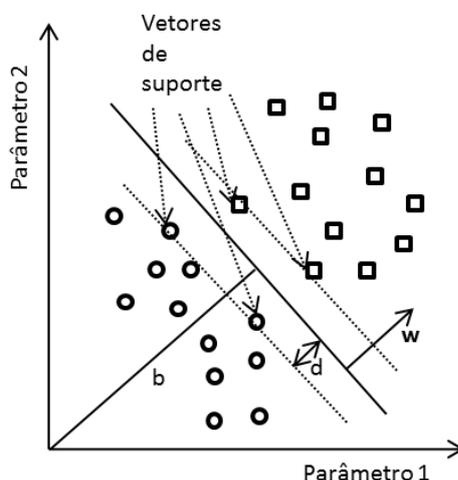
Por fim, temos a etapa de classificação. Nesta etapa é feita efetivamente a classificação dos parâmetros característicos de entrada de acordo com as classes definidas no sistema. Para que seja possível a classificação é necessário que seja feito um treinamento da máquina de classificação. Este treinamento pode ser feito a partir de informações previamente disponíveis no sistema, que é o chamado aprendizado supervisionado, ou então a partir de relações estabelecidas através da análise estatísticas dos padrões, no momento da classificação, o chamado aprendizado não-supervisionado. Nesta tese, o classificador a ser utilizado é o de máquina de vetores de suporte, que será descrito no capítulo seguinte.

### 3.2.1 Máquina de vetores de suporte

A Máquina de vetores de suporte, mais conhecida pela sigla SVM (do inglês *Support Vector Machine*) [CV95, Bur98] é um sistema de aprendizado de máquina muito utilizado para fins de classificação de padrões. Este sistema busca separar um conjunto de dados em

dois subconjuntos distintos. Os subconjuntos são separados a partir da definição de um hiperplano que divide os dados no espaço dos parâmetros característicos. Este hiperplano de separação é construído a partir de vetores pertencentes aos dados de treinamento e estes então são chamados de vetores de suporte.

Seja um conjunto de dados de treinamento  $\{\mathbf{x}_i, y_i\}$ , sendo  $i = 1, 2, \dots, T$ , onde  $x_i$  é o vetor dos parâmetros característicos da amostra  $i$  de treinamento e  $y_i$  é o rótulo da sua classe, onde  $y_i \in \{-1; 1\}$  sendo  $-1$  o rótulo para a classe *nvaz* e  $1$  o rótulo para a classe *vaz*. A máquina de vetores de suporte busca definir um hiperplano ótimo de separação entre os dados das classes que maximize a distância entre as amostras mais próximas de ambas as classes. No caso de os dados já serem linearmente separáveis no espaço ao qual pertencem, este hiperplano é definido por  $\mathbf{x} \cdot \mathbf{w} + b = 0$ , onde  $\mathbf{w}$  é a normal ao hiperplano e  $|b| / \|\mathbf{w}\|$  é a distância perpendicular ao hiperplano de origem, conforme ilustra a Figura 3.6.



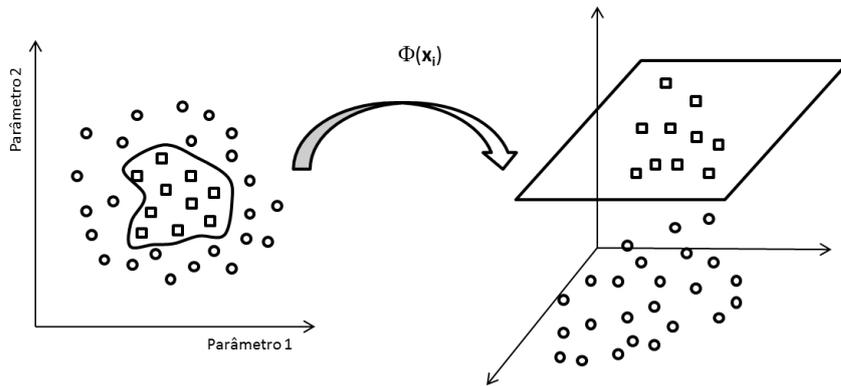
**Figura 3.6:** Exemplo de dados linearmente separáveis para dimensão dois.

Já para os casos onde os dados das classes não podem ser separados no seu espaço de origem, ver Figura 3.7, o SVM propõe que estes dados sejam projetos em uma dimensão maior de tal forma que, neste novo espaço, os dados das classes sejam separáveis. Dado que o mapeamento dos dados das classes no seu espaço de origem para o espaço de alta dimensão seja  $\Phi(\mathbf{x}_i)$ , a função Kernel é definida como o produto interno dos vetores  $\mathbf{x}_i$  e  $\mathbf{s}_j$  mapeados no espaço de alta dimensão.

O hiperplano de separação no espaço de maior dimensão é então definido por

$$f(\mathbf{x}) = \alpha_i y_i K(\mathbf{x}_i, \mathbf{s}_j) + b, \quad (3.5)$$

onde  $K(.,.)$  é a função kernel,  $\mathbf{s}_i$  são os vetores de suporte e  $\alpha_i$  as ponderações dos vetores de suporte. Exemplos de função Kernel são apresentados na Tabela 3.1. Uma função Kernel deve representar um mapeamento onde seja possível o cálculo de produtos internos que seguem as condições estabelecidas pelo teorema de Mercer [Mer09]. Todos



**Figura 3.7:** Espaço com dados não separáveis linearmente e uma projeção em um espaço maior.

os parâmetros da função Kernel são encontrados durante a fase de treinamento [CV95].

Kernel	produto interno do kernel
Linear	$K(\mathbf{x}, \mathbf{s}_j) = (\mathbf{x} \cdot \mathbf{s}_j)$
Gaussiano	$K(\mathbf{x}, \mathbf{s}_j) = e^{\frac{-\ \mathbf{x}-\mathbf{s}_j\ ^2}{\sigma^2}}$
Polinomial	$K(\mathbf{x}, \mathbf{s}_j) = (\mathbf{x} \cdot \mathbf{s}_j)^q$

**Tabela 3.1:** Tipos de função Kernel.

Para se classificar os novos vetores de parâmetros característicos dos dados  $\mathbf{x}$  como pertencentes as classes *nvaz* ou *vaz*, ou -1 ou 1, a decisão é descrita por

$$\text{sgn}(f(\mathbf{x})) = \begin{cases} \mathbf{x} \in \textit{nvaz} & \text{se } f(\mathbf{x}) < 0 \\ \mathbf{x} \in \textit{vaz} & \text{se } f(\mathbf{x}) > 0. \end{cases} \quad (3.6)$$

Ou seja, se a variável de decisão for maior que zero o vetor de teste é classificado como pertencente a classe *nvaz*, caso contrário, o vetor é classificado como pertence à classe *vaz*.

### 3.3 Métodos de avaliação dos parâmetros característicos

Como um dos objetivos desta tese é identificar parâmetros característicos que melhor representem os sinais de vazamento ou de não vazamento, é necessário que sejam definidos métodos de avaliação para os mesmos. A avaliação dos parâmetros característicos pode ser feita de forma individual ou em conjunto. A avaliação dos parâmetros característicos de forma individual busca identificar quais parâmetros possuem os maiores fatores de discriminação em relação às classes a que os parâmetros possam pertencer. Já a avaliação em conjunto busca determinar quais as combinações possíveis dos parâmetros que alcançam os melhores resultados de classificação. Segundo [Guy03], nem sempre os parâmetros que

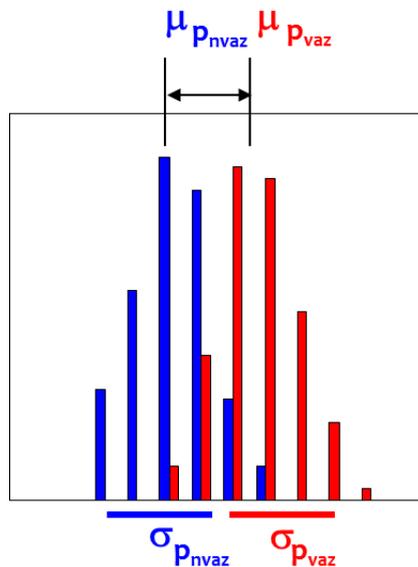
se destacam individualmente, quando avaliados em conjunto, apresentarão os melhores resultados de classificação. Este efeito deve-se provavelmente ao fato de que os melhores parâmetros carregam informações redundantes do sinal, entre si, não podendo contribuir para melhorar a discriminação das classes. Nas próximas sessões vamos detalhar as duas formas de avaliação a serem aplicadas nesta pesquisa.

### 3.3.1 Avaliação individual dos parâmetros

Uma das formas mais simples de se avaliar isoladamente um parâmetro característico é verificando a separabilidade entre as distribuições dos valores dos parâmetros de acordo com as classes às quais pertencem. Existem diversas medidas de se calcular a separabilidade entre classes. Para o caso de duas classes distintas onde a distribuição dos valores dos seus parâmetros característicos são do tipo Gaussiana, um método muito utilizado é o da estatística $_M$ , também conhecida na estatística por relação sinal-ruído. A estatística-M foi definida em [GST<sup>+</sup>99] e é descrita por

$$\text{estatística}_M = \frac{|\mu_{\mathbf{p}_{\mathit{vaz}}} - \mu_{\mathbf{p}_{\mathit{nvaz}}}|}{\sigma_{\mathbf{p}_{\mathit{vaz}}} + \sigma_{\mathbf{p}_{\mathit{nvaz}}}}. \quad (3.7)$$

Na equação acima,  $\mathit{vaz}$  representa a classe das amostras de sinais de vazamento e  $\mathit{nvaz}$  representa a classe das amostras de sinais de não vazamento.  $\mu_{\mathbf{p}}$  e  $\sigma_{\mathbf{p}}$  são respectivamente a média e variância do parâmetros característicos  $\mathbf{p}$  de cada classe, como ilustra a Figura 3.8.



**Figura 3.8:** Representação gráfica da estatística-M.

Através desta medida, podemos inferir quais os parâmetros que apresentam maior relevância dentre os diversos existentes com relação a separabilidade entre as classes. Valores de estatística $_M$  menores que 1 indicam que as classes apresentam sobreposição e

que a possibilidade de separação das classes é baixa. Por outro lado, valores maiores que 1 indicam que as classes estão menos sobrepostas e, portanto, a possibilidade de separação entre elas é alta.

Este tipo de avaliação serve para inferir quais parâmetros característicos poderão ser descartados, uma vez que parâmetros com baixo poder discriminativo entre suas classes tendem a deteriorar o desempenho do classificador.

### 3.3.2 Avaliação dos parâmetros em conjunto

Os parâmetros característicos também devem ser avaliados em conjunto e esta avaliação deve ser feita com a utilização de um classificador. Esta avaliação em conjunto se dá devido ao fato de que dificilmente somente um parâmetro característico, de forma isolada, obteria o melhor desempenho do sistema de classificação. Para se obter a melhor resposta em termos de taxa de acerto de um classificador o ideal é testar o maior número possível de combinações de parâmetros característicos disponíveis.

A partir da medida de avaliação individual pode-se ter um indicativo de quais parâmetros devem ser avaliados em conjunto. Porém, como já foi apresentado anteriormente, isto não garante a melhora no desempenho do classificador. Portanto, combinações com parâmetros de baixa relevância também devem ser avaliadas.

Para avaliar os parâmetros em grupo, optou-se pela utilização do classificador chamado máquina de vetores de suporte, conforme descrito no sessão 3.2.1. Esta escolha se deu devido ao fato de que este classificador apresenta bons resultados em termos de classificação binária e também apresenta alta eficiência quando a dimensão dos dados for alta.

O classificador SVM define, a partir dos dados de treinamento, um hiperplano de separação entre duas classes existentes. A partir da definição do hiperplano, novos dados são avaliados a partir de sua projeção neste hiperplano. Variáveis de decisão com valores negativos indicam que o dado pertence a um lado do hiperplano, ou seja, à classe *nvaz*. Já variáveis de decisão com valores positivos indicam que o dado pertence ao outro lado do hiperplano, ou seja, à classe *vaz*.

A partir do resultado de classificação dos dados de teste, calcula-se a taxa de acerto do classificador. A taxa de acerto mede a relação entre o número total de classificações corretas (NC) pelo número total das classificações feitas (NT). Sendo NT o número total amostras de sinais submetidas ao classificador e NV o número total de amostras de sinal classificadas verdadeiramente, a taxa de acerto do classificador (TAC) é definida por

$$\text{TAC} = \frac{\text{NC}}{\text{NT}}. \quad (3.8)$$

Também costuma-se calcular as taxas de acerto para cada uma das classes do classificador, que seriam as taxas de acerto de para sinais de vazamento e para sinais de não

vazamento. A taxa de acerto de vazamento (TAV) é a relação entre o número de amostras de vazamento que foram classificadas corretamente ( $NC_{vaz}$ ) e o número total de amostras de vazamento ( $NT_{vaz}$ ). Por sua vez, a taxa de acerto de não vazamento (TANV) é a relação entre o número de amostras de não vazamento que foram classificadas corretamente ( $NC_{nvaz}$ ) e o número total de amostras de não vazamento ( $NT_{nvaz}$ ). Estas duas medidas só são válidas para o caso de classificação binária. A TAV é descrita por

$$TAV = \frac{NC_{vaz}}{NT_{vaz}} \quad (3.9)$$

e a TANV por

$$TANV = \frac{NC_{nvaz}}{NT_{nvaz}}. \quad (3.10)$$

# Capítulo 4

## Extração de parâmetros característicos

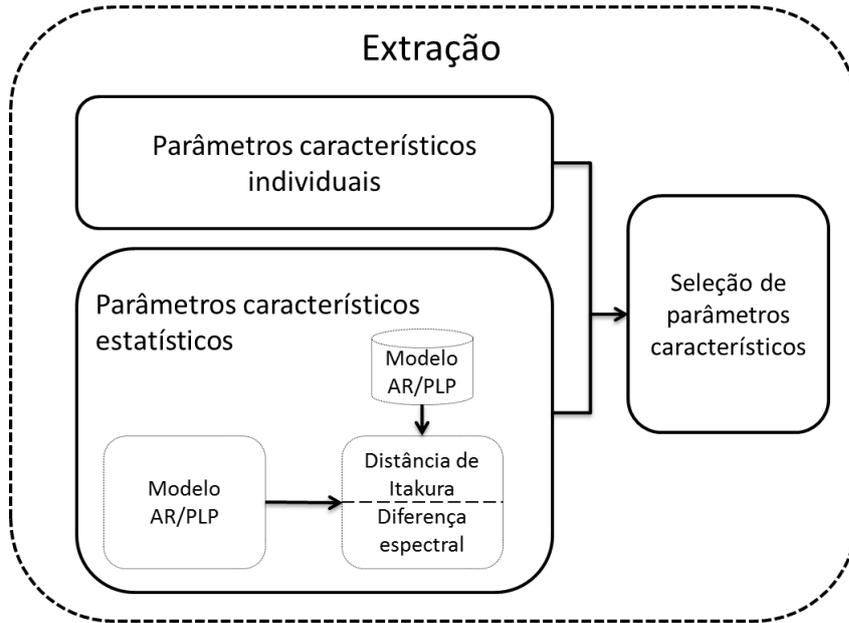
Neste capítulo são apresentados os métodos de extração de parâmetros característicos avaliados. Primeiramente são apresentados os parâmetros característicos individuais extraídos do sinal. Em seguida são apresentados os parâmetros característicos estatísticos estimados a partir de medidas estatísticas relativas aos modelos de predição dos sinais de vazamento.

### 4.1 Introdução

Conforme apresentado no capítulo anterior, a etapa central do processamento do sinal é a extração de parâmetros característicos. Estes parâmetros, nesta tese, serão divididos em duas categorias que são a dos parâmetros característicos individuais e dos estatísticos, como mostra a Figura 4.1. Os parâmetros individuais são estimados de tal forma que representem as informações relevantes para a identificação da origem primordial do sinal. Já os parâmetros característicos estatísticos são estimados de tal forma que representem o comportamento estatístico dos modelos de predição relacionado aos modelos de sinais de vazamento. Os parâmetros característicos individuais constituem uma representação do sinal no domínio do tempo ou no domínio da frequência e os parâmetros característicos estatísticos são uma medida de pertinência do sinal na classe de sinais de vazamento.

Nas próximas sessões serão descritas as estimativas de alguns parâmetros individuais representativos do domínio do tempo, que são a Energia de curto prazo e a Taxa de cruzamentos por zero. Também serão apresentadas estimativas de alguns parâmetros representativos no domínio da frequência, que são o centróide espectral, a largura de banda, o fluxo espectral, o ponto de roll-off e a taxa de espalhamento.

Por fim, alguns parâmetros estatísticos relacionados aos modelos de predição linear (AR) e nos modelos de predição linear perceptual (PLP) terão suas medidas apresentadas. Para descrever os parâmetros característicos estatísticos primeiramente é necessário descrever a estimação de modelos de predição linear normal e perceptual. Este sequenciamento é necessário porque os parâmetros estatísticos são obtidos a partir de medidas de



**Figura 4.1:** Diagrama de blocos da extração de parâmetros característicos.

relação dos modelos de predição dos sinais de entrada e os modelos de predição dos sinais de vazamento definidos como referência.

Ao final será descrita a estimação do modelo de predição linear, ao que se segue a estimação do modelo de predição linear perceptual. Por fim, são descritos os métodos de cálculo da distância de Itakura e da diferença espectral, que são as medidas que definem os parâmetros característicos estatísticos.

## 4.2 Parâmetros característicos individuais

### 4.2.1 Energia de Curto Prazo

Através de uma análise visual dos sinais com ruído de vazamento e sem ruído de vazamento, notou-se que os sinais de vazamento, em geral, possuem amplitude média maior do que as dos sinais de não vazamento. Um dos motivos para este fato é que os sinais de não vazamento são basicamente sinais do ruído de fundo do ambiente e do fluxo de água pelos dutos da tubulação. Já os sinais de vazamento têm maior amplitude, uma vez que a energia que é perdida em função da água que é expelida, em alta pressão, e se choca contra os dutos das paredes da tubulação e contra o solo é alta.

A energia de curto prazo [RS78] de um segmento de sinal  $x_k$  é descrita por

$$E_k = \sum_{n=0}^{N-1} [x_k(n)w(n)]^2 \quad (4.1)$$

onde  $w(\cdot)$  é a janela retangular, Equação(3.3) e  $N$  é o tamanho da janela.

## 4.2.2 Taxa de cruzamento por zero

Como também analisado visualmente, os sinais demonstram ter uma diferenciação com relação à sua variação em torno de zero e em função disto optou-se por também extrair a taxa de cruzamento por zero [RS78] do segmento de sinal. Este parâmetro característico representa a quantidade de vezes que o sinal cruza o eixo do tempo, ou seja, toda a vez que uma amostra muda o seu sinal de positivo para negativo ou vice versa, com relação à amostra anterior, como ilustra a equação

$$Z_k = \sum_{n=0}^{N-1} |sgn[x_k(n)] - sgn[x_k(n-1)]| w(n) \quad (4.2)$$

onde a função  $sgn(\cdot)$  é definida por

$$sgn[x(m)] = \begin{cases} 1 & \text{se } x(n) > 0; \\ 0 & \text{se } x(n) = 0; \\ -1 & \text{se } x(n) < 0. \end{cases} \quad (4.3)$$

e a janela  $w(\cdot)$  é a janela retangular, definida na Equação 3.3, ponderada por  $\frac{1}{2N}$ .

## 4.2.3 Centróide espectral

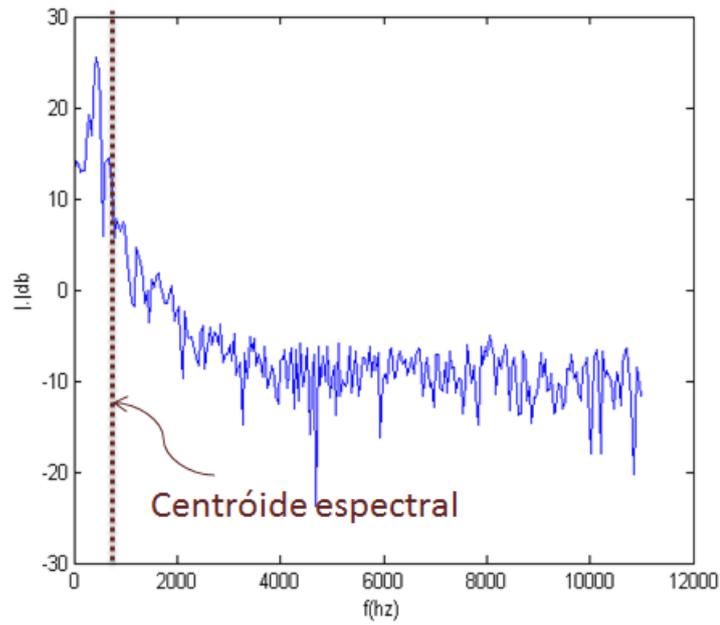
O centróide espectral [TC02] é um parâmetro característico que representa a frequência média ponderada do espectro de um sinal. Seria o equivalente na física ao "centro de gravidade", da distribuição espectral da energia do sinal. O centróide espectral é calculado através da equação

$$c_k = \frac{\sum_{m=0}^{M-1} f(m) |X_k(m)|^2}{\sum_{m=0}^{M-1} |X_k(m)|^2}, \quad (4.4)$$

onde  $X_k(m)$  é a magnitude do espectro de frequência resultante do cálculo da transformada de Fourier do quadro  $k$  no tempo, do sinal  $x$  e  $f(m)$  é a frequência espectral respectiva. A constante  $M$  é definida como sendo o número de pontos da metade superior da amostras da densidade espectral de potência do sinal, uma vez que a outra metade inferior é a reflexão da metade superior. Altos valores de centróide espectral representam que a componente espectral principal do sinal é de alta frequência.

## 4.2.4 Largura do espectro

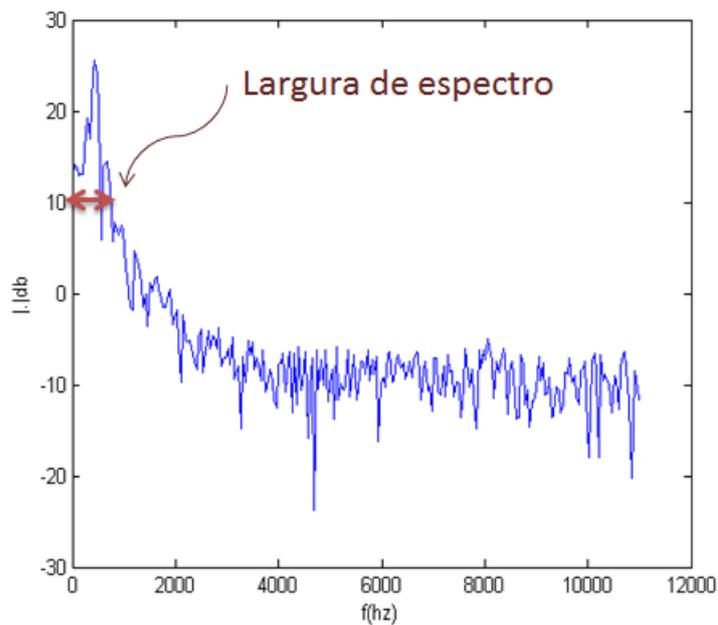
O parâmetro característico largura de espectro [SS97] descreve como é a distribuição do espectro do sinal com relação ao seu centróide espectral. Valores baixos deste parâmetro indicam que o espectro do sinal está mais concentrado próximo ao centróide espectral. Por outro lado, altos valores deste atributo indicam que o espectro do sinal está mais



**Figura 4.2:** Ilustração do centróide espectral de um sinal.

espalhado ao longo do eixo espectral. A Equação 4.5 apresenta o cálculo da largura de espectro.

$$le_k = \sqrt{\frac{\sum_{m=0}^{M-1} (c_k - m)^2 |X_k(m)|^2}{\sum_{m=1}^M |X_k(m)|^2}}. \quad (4.5)$$



**Figura 4.3:** Ilustração da largura espectral de um sinal.

## 4.2.5 Fluxo espectral

O fluxo espectral [SS97, TC02] é definido como o somatório das diferenças quadráticas entre as magnitudes de sucessivas componentes espectrais. Com este parâmetro característico busca-se caracterizar as mudanças no formato do espectro ao longo de todo o espectro do sinal. O fluxo espectral é definido por:

$$fe_k = \sum_{m=0}^{M-1} (|X_k(m)| - |X_{k-1}(m)|)^2. \quad (4.6)$$

## 4.2.6 Ponto de *Roll-off*

O parâmetro característico ponto de *roll-off* [TC02] mede a frequência abaixo da qual se encontra 85% da energia do sinal. Assim como no caso do parâmetro característico centróide espectral, este parâmetro também busca indicar o formato geral do espectro do sinal. O ponto de *Roll-off* é o menor valor de  $R$  que satisfaz a equação:

$$\sum_{m=1}^R |X_k(m)| = 0.85 \sum_{m=1}^{N/2} |X_k(m)|, \quad (4.7)$$

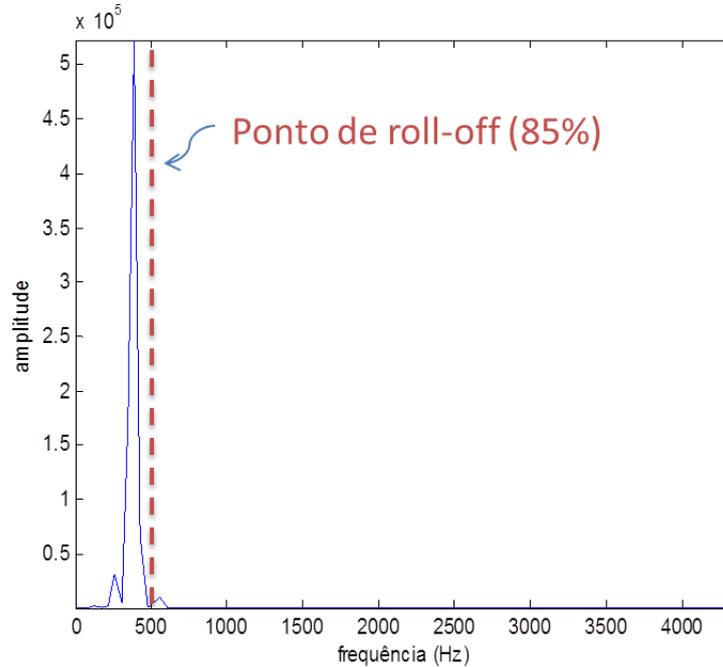


Figura 4.4: Ilustração do ponto de roll-off de um sinal.

## 4.2.7 Taxa de espalhamento

O parâmetro característico taxa de espalhamento [SS97] é dado pelo desvio padrão do espectro de magnitudes de cada quadro. Sinais constituídos de ruído branco tendem

a ter um valor de espalhamento maior do que aqueles onde há a concentração de energia somente em certas bandas. A taxa de espalhamento é dada por:

$$esp_k = \sqrt{\frac{1}{M-1} \sum_{m=0}^{M-1} (X_k(m) - \mu_{X_k})^2}, \quad (4.8)$$

onde  $\mu_{X_k}$  é a média do espectro do sinal  $X_k$ .

## 4.3 Parâmetros característicos estatísticos

### 4.3.1 Modelo de predição linear

Em linhas gerais, a análise por predição linear [PM06] é um método linear de se modelar um sinal  $x$  a partir de suas amostras em tempos passados. Dado um segmento do sinal  $x(n)$ , com  $n = (0, 1, \dots, N-1)$  onde  $N$  é o número de amostras, a análise por predição linear de ordem  $p$  busca estimar os coeficientes da amostra do sinal  $x$  no tempo  $n$ ,  $x(n)$ , através da combinação linear das  $p$  amostras em tempos anteriores,  $x(n-1), x(n-2), \dots, x(n-p)$ , constituindo o sinal predito no instante  $n$  como:

$$\tilde{x}(n) = \sum_{k=1}^p a_k x(n-k). \quad (4.9)$$

O sinal de erro residual,  $e(n)$ , entre a amostra do sinal original  $x(n)$  e a amostra do sinal predito  $\tilde{x}(n)$  é dado por

$$e(n) = x(n) - \tilde{x}(n). \quad (4.10)$$

O vetor  $\mathbf{a} = [a_1, a_2, \dots, a_p]$  contém os  $p$  coeficientes estimados do modelo AR. Os coeficientes do modelo AR são estimados através da solução das equações de Yule-Walker abaixo apresentadas:

$$\begin{bmatrix} r_{xx}(0) & r_{xx}(1) & \cdots & r_{xx}(p-1) \\ r_{xx}(1) & r_{xx}(0) & \cdots & r_{xx}(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ r_{xx}(p-1) & r_{xx}(p-2) & \cdots & r_{xx}(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} r_{xx}(1) \\ r_{xx}(2) \\ \vdots \\ r_{xx}(p) \end{bmatrix}, \quad (4.11)$$

onde  $r_{xx}(\tau)$  é a função de autocorrelação do sinal  $x(n)$ , que é definida pela sequência

$$\begin{aligned} r_{xx}(\tau) &= \langle x(n)x(n+\tau) \rangle \\ &= \frac{1}{N} \sum_{n=-\infty}^{+\infty} x(n)x(n+\tau), \quad \tau = 0, 1 \dots p, \end{aligned} \quad (4.12)$$

onde  $\tau$  é a medida de deslocamento. A autocorrelação mede o nível de correlação entre a amostra do sinal  $x$  no tempo  $n$  e sua amostra deslocada no tempo por  $\tau$  amostras. A partir da medida da autocorrelação é possível avaliar a periodicidade do sinal  $x(n)$ .

### 4.3.2 Modelo de predição linear perceptual

Uma vez que a detecção de vazamento é uma técnica que se baseia no conhecimento prévio dos sinais sonoros percebidos pelos ouvidos de técnicos especializados e como sabe-se que o sistema auditivo humano possui algumas particularidades no processamento do sinal, o chamado processamento psicoacústico, propôs-se então a inclusão deste processamento no sinal original antes de ser feita a análise por predição linear.

Segundo [Rob99] as componentes de frequência que são melhores percebidas pelo sistema auditivo humano estão entre a faixa de 20 Hz a 20k Hz. Componentes de frequências fora desta faixa não conseguem ser captados pelos sensores do sistema auditivo humano típico. Também segundo apresentou [Rob99], o sistema auditivo não capta as componentes de frequência de uma forma linear. Esta relação não linear pode ser aproximada por duas diferentes escalas de percepção que são a escala Mel e a escala Bark. Nesta tese vamos utilizar a escala Bark.

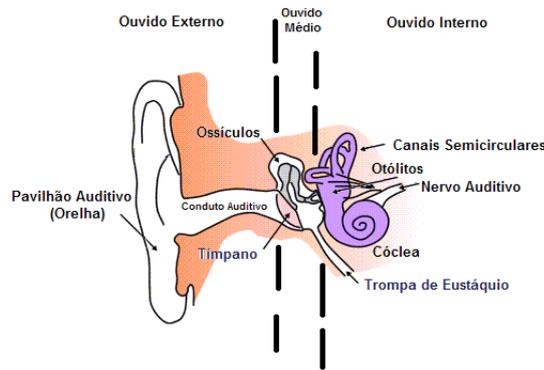
A análise de predição linear perceptual [Her90], então, busca melhorar a estimativa do modelo de predição linear considerando as características psicoacústicas do sistema auditivo humano, como ilustra a Figura 4.5. E isto é feito pré-processando o sinal original antes de ele ser utilizado para estimar o modelo de predição.



**Figura 4.5:** Diagrama em blocos da análise por predição linear perceptual.

Primeiramente vamos apresentar alguns conceitos sobre a constituição e sobre o funcionamento do sistema auditivo humano. O sistema auditivo humano é constituído basicamente de três partes: o ouvido externo, ouvido médio e ouvido interno, como ilustra a Figura 4.6. O ouvido externo tem como função principal coletar os sons e conduzi-los, através do conduto auditivo, até o tímpano, que é parte integrante do ouvido médio. Outra função do ouvido externo é de fazer uma espécie de filtragem do som com o intuito de auxiliar a localização da origem das fontes sonoras. Também é função do ouvido externo servir de amplificador dos sinais da faixa de frequência da fala humana (entre 2k Hz e 5k Hz).

Já o ouvido médio tem a função de transformar as variações de pressão causadas pelas ondas sonoras recebidas pelo ouvido externo em energia mecânica que é então transmitida ao ouvido interno. É também função do ouvido médio regular a diferença de

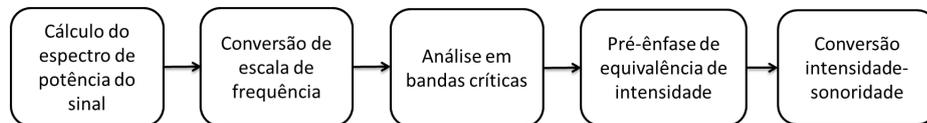


**Figura 4.6:** Anatomia do ouvido humano. Fonte: <http://pt.wikipedia.org/wiki/Orelha>

pressão entre os ouvidos externo e interno.

Por fim, o ouvido interno tem a função de transformar as ondas mecânicas do som recebido pelo ouvido médio em impulsos neurais codificados que são enviados ao cérebro para serem interpretados.

O processamento psicoacústico do sistema auditivo pode ser aproximado pelo diagrama em blocos apresentado na Figura 4.7. O diagrama divide-se basicamente em quatro etapas distintas que são: a conversão de escala de frequência, a análise em bandas críticas, a pré-ênfase de equivalência de volume e a conversão intensidade-sonoridade.

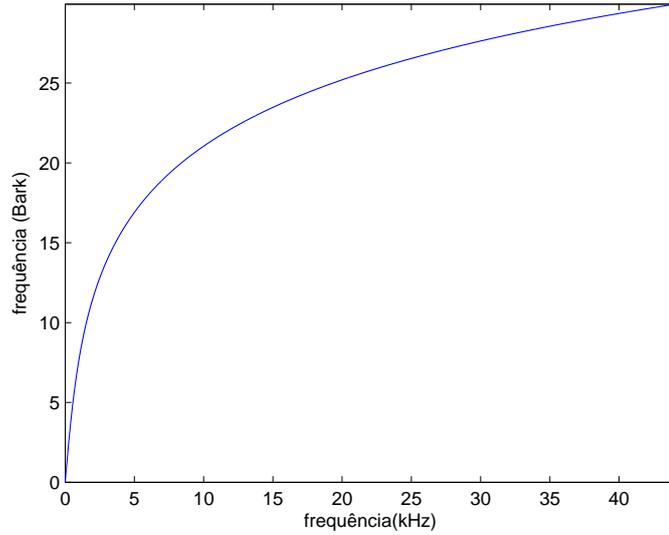


**Figura 4.7:** Diagrama em blocos do processamento psicoacústico.

Uma das primeiras constatações feitas com relação à percepção de frequência do sistema auditivo humano [Rob99] é a de que ele não percebe as componentes de frequência de forma linear e por este motivo se faz necessário a conversão do espectro do sinal da escala Hz para a escala Bark. A relação entre a escala Hz e a escala Bark é definida na Equação 4.13. Na escala Bark, as componentes de baixa frequência têm uma relação de escala de frequência linear e as componentes de alta frequência têm uma relação de escala de frequência logarítmica, como ilustra a Figura 4.8.

$$\nu = 6 \arcsin \left( \frac{f_{Hz}}{600} \right) \quad (4.13)$$

Além de as componentes de frequência não serem percebidas de forma linear, elas também não são percebidas de forma isolada. Os sensores que captam as diferentes frequências funcionam como uma espécie de filtro de faixas de frequência. Cada sensor capta uma faixa de frequências diferentes em torno de uma frequência central que é chamada de banda crítica. Para uma dada frequência central, componentes vizinhas a ela são interpretadas como pertencentes à banda crítica centrada nela. Um filtro de banda crítica é definido



**Figura 4.8:** Curva da relação entre as escalas de frequência hertz e Bark.

através de

$$H_c(\nu) = \begin{cases} 0 & \text{para } -2.5 \leq \nu - \nu_c \\ 10^{(\nu - \nu_c + 0.5)} & \text{para } -2.5 < \nu - \nu_c < -0.5 \\ 1 & \text{para } -0.5 \leq \nu - \nu_c \leq 0.5 \\ 10^{-2.5(\nu - \nu_c - 0.5)} & \text{para } 0.5 < \nu - \nu_c < 1.3 \\ 0 & \text{para } \nu - \nu_c \geq 1.3 \end{cases}, \quad (4.14)$$

onde  $\nu$  são as frequências na escala Bark e  $\nu_c$  a frequência central,  $c$ , do filtro, também em escala Bark. A Figura 4.9 ilustra um filtro de banda crítica para a frequência central de aproximadamente 4 Bark. A resposta em frequência do espectro de potência do sinal sonoro após a análise em bandas críticas é dado por:

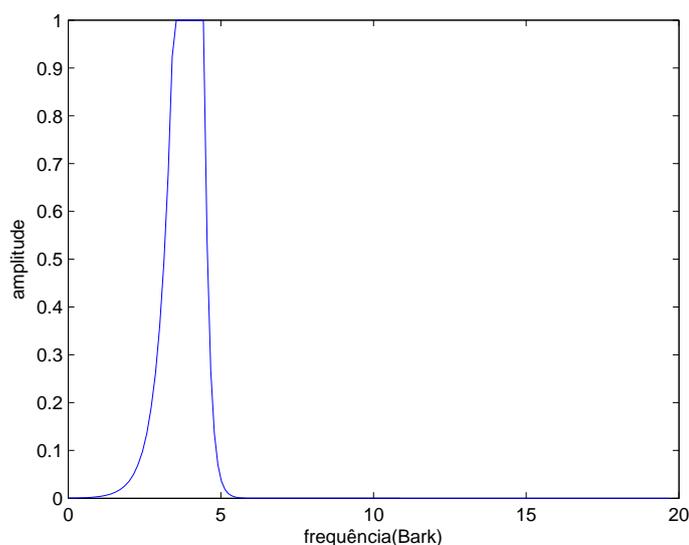
$$I_k(\nu) = \sum_{\nu \in \{\nu_c\}} H(\nu) X_k(\nu). \quad (4.15)$$

Também em função da anatomia do sistema auditivo humano, o nível mínimo de percepção de uma componente de frequência varia. Este comportamento do nível mínimo de percepção é estimado pela curva descrita na Equação 4.16 e é ilustrado pela Figura 4.10. Para frequências mais baixas o limiar de percepção em db de nível de pressão de sonoro está em torno de 9 db. Para as frequências médias o limiar está em torno de 0 db e para as frequências altas o limiar é de acima de 20 db.

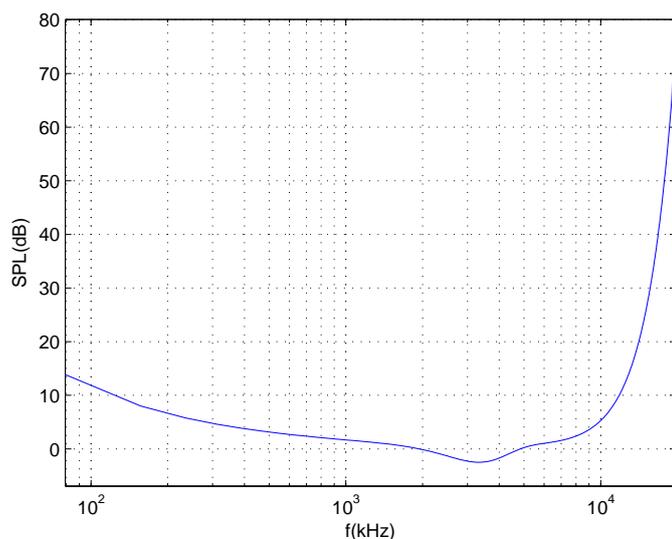
$$E(\nu) = \frac{(\nu^2 + 56.8 \cdot 10^{+6})\nu^4}{(\nu^2 + 6.3 \cdot 10^{+2})^2(\nu^2 + 0.38 \cdot 10^{+9})} \quad (4.16)$$

e ilustrado na Figura 4.10. O espectro de potência do sinal após a pré-ênfase é dado por

$$EI_k(\nu) = I_k(\nu)E(\nu). \quad (4.17)$$



**Figura 4.9:** *Filtro de banda crítica de frequência central de aproximadamente 4 Barks.*



**Figura 4.10:** *Nível mínimo de percepção do aparelho auditivo humano de acordo com a frequência do sinal.*

Para finalizar o processamento psicoacústico temos a etapa de conversão da intensidade-sonoridade e isto é feito a partir da extração da raiz cúbica do espectro sinal, ou seja:

$$L(\nu) = \sqrt[3]{EI_k(\nu)}. \quad (4.18)$$

Por sua vez, para finalizar a análise de predição linear perceptual, é necessário fazer a transformação do espectro do sinal processado para o domínio temporal, através da utilização da transformada discreta de Fourier inversa. Tendo então novamente o sinal no domínio do tempo, os coeficientes de predição linear são então estimados, conforme já fora descrito na Seção 4.3.1.

## Diferença espectral

A medida de diferença espectral [Ram06] calcula a diferença entre as densidades espectrais de potência de dois sinais reconstruídos a partir de seu modelo de predição linear. Dados o modelo de predição  $\mathbf{a}$  de um sinal  $x_1(n)$  e  $\mathbf{b}$  de um sinal  $x_2(n)$ , a medida de diferença espectral é calculada como:

$$d_{esp} = \sqrt{\frac{1}{N} \sum_{f=0}^{N-1} (10 \log_{10} P_a(f) - 10 \log_{10} P_b(f))^2}, \quad (4.19)$$

onde  $P_a(f)$  e  $P_b(f)$  são as densidades espectrais de frequência dos modelos autoregressivos  $\mathbf{a}$  e  $\mathbf{b}$ , respectivamente.

## Distância de Itakura

Para definirmos a distância de Itakura, primeiro é necessário ser definida a distorção de Itakura [Ita75]. A distorção de Itakura mede a diferença entre o modelo de predição linear  $\mathbf{b}$  de um sinal  $x_2(n)$ , com relação ao modelo de predição  $\mathbf{a}$  de um sinal  $x_1(n)$ . A distorção é definida como sendo o logaritmo da relação entre o erro quadrático médio do modelo  $\mathbf{a}$  e o erro quadrático médio do modelo  $\mathbf{b}$ ,

$$d(\mathbf{a}, \mathbf{b}) = \log \left( \frac{E_{\mathbf{b}}}{E_{\mathbf{a}}} \right). \quad (4.20)$$

A distância de Itakura, por sua vez, é a média entre a distorção do modelo  $\mathbf{a}$  com relação ao modelo  $\mathbf{b}$  e a distorção entre o modelo  $\mathbf{b}$  com relação ao modelo  $\mathbf{a}$ , conforme descreve a equação abaixo:

$$d_{it}(\mathbf{a}, \mathbf{b}) = \frac{d(\mathbf{a}, \mathbf{b}) + d(\mathbf{b}, \mathbf{a})}{2}, \quad (4.21)$$

uma vez que a distorção do modelo  $\mathbf{a}$  com relação ao modelo  $\mathbf{b}$  não é necessariamente igual à distorção do modelo  $\mathbf{b}$  com relação ao modelo  $\mathbf{a}$ , ou seja,  $d(\mathbf{a}, \mathbf{b}) \neq d(\mathbf{b}, \mathbf{a})$ .

Seja  $E_a$  o erro quadrático médio de predição do sinal  $x(n)$ , definido por

$$E = \frac{1}{N} \sum_{n=0}^{N-1} e(n)^2 \quad (4.22)$$

e  $e(n)$  o erro residual do sinal  $x(n)$ , definido pela Equação 4.10, então temos que

$$E_{\mathbf{a}} = \sum_{n=0}^{N-1} \left( x(n) - \sum_{k=1}^p a_k x(n-k) \right)^2$$

$$\begin{aligned}
&= \sum_{n=0}^{N-1} \left( x(n)^2 - 2 \sum_{k=1}^p a_k x(n-k)x(n) + \sum_{k=1}^p \sum_{j=1}^p a_k x(n-k)a_j x(n-j) \right) \\
&= \sum_{n=0}^{N-1} x(n)^2 - 2 \sum_{k=1}^p a_k \sum_{n=0}^{N-1} x(n-k)x(n) + \sum_{k=1}^p \sum_{j=1}^p a_k a_j \sum_{n=0}^{N-1} x(n-k)x(n-j) \\
&= r_{xx}(0) - 2 \sum_{k=1}^p a_k r_{xx}(k) \sum_{k=1}^p \sum_{j=1}^p a_k a_j r_{xx}(k-j). \tag{4.23}
\end{aligned}$$

Reescrevendo a Equação 4.23 no formato matricial temos:

$$\begin{aligned}
E\mathbf{a} &= \begin{bmatrix} -1 & a_1 & \dots & a_p \end{bmatrix} \begin{bmatrix} r_{xx}(0) & r_{xx}(1) & r_{xx}(2) & \dots & r_{xx}(p) \\ r_{xx}(1) & r_{xx}(0) & r_{xx}(1) & \dots & r_{xx}(p-1) \\ r_{xx}(2) & r_{xx}(1) & r_{xx}(0) & \dots & r_{xx}(p-2) \\ \dots & \dots & \dots & \dots & \dots \\ r_{xx}(p) & r_{xx}(p-1) & r_{xx}(p-2) & \dots & r_{xx}(0) \end{bmatrix} \begin{bmatrix} -1 \\ a_1 \\ \dots \\ a_p \end{bmatrix} \\
&= \mathbf{a}^T R_{xx} \mathbf{a}, \tag{4.24}
\end{aligned}$$

onde a dimensão da matriz  $R_{xx}$  é de  $((p+1) \times (p+1))$ . Por sua vez, o erro quadrático médio do modelo de predição  $\mathbf{b}$ , pode ser escrito como sendo:

$$E_b = \mathbf{b}^T R_{xx} \mathbf{b}. \tag{4.25}$$

Substituindo as Equações 4.25 e 4.24 na Equação 4.21, tem-se que a distância de Itakura é dada por:

$$d(\mathbf{a}, \mathbf{b}) = \log \left( \frac{\mathbf{b}^T R_{xx} \mathbf{b}}{\mathbf{a}^T R_{xx} \mathbf{a}} \right). \tag{4.26}$$

# Capítulo 5

## Desenvolvimento da pesquisa

Primeiramente serão apresentados os resultados de classificação (TA) para os parâmetros sugeridos em [SM07] e também o resultado da aplicação dos filtros de pré-ênfase. Em seguida serão apresentados os histogramas de todos os parâmetros característicos onde será feita uma análise inicial da estatística-M e da TA. Por fim serão apresentados os resultados de TA para diversos conjuntos distintos de dados.

### 5.1 Construção da Base de Dados

Uma vez que não existe nenhuma base de dados, de uso comercial ou acadêmico, de sinais acústicos de vazamento de água, fez-se necessário a criação de uma base de dados específica de áudio desta natureza. A base de dados criada foi adquirida em parceria com uma empresa brasileira especializada em detecção acústica de vazamentos de água.

Foram adquiridas 194 amostras de áudio com e sem a presença de sinal de vazamento, em condições reais. Do total de amostras adquiridas 123 delas foram de sinal de vazamento e 71 foram de sinais de não vazamento, como mostra a Tabela 5.1. Todas as amostras de sinais têm duração de aproximadamente 10 s e foram gravadas em codificação PCM, com taxa de amostragem de 44.1k Hz e com 24 bits por amostra.

Vazamento	Não Vazamento	Total
71	123	194

**Tabela 5.1:** *Subgrupos de dados*

As aquisições foram feitas pelos técnicos especializados na detecção acústica de vazamento. Para a aquisição dos sinais foi utilizado um gravador digital de alta resolução. Este gravador foi ligado diretamente à saída de áudio do geofone eletrônico. Também tomou-se o cuidado de desabilitar todas as filtragens disponíveis no geofone, para garantir que os sinais adquiridos sejam sinais puros.

Após a coleta das amostras, todas elas foram rotuladas como vazamento ou não vazamento. Além disto, todas as amostras foram validadas manualmente para garantir sua qualidade. Sinais com interferências do ambiente não foram utilizados.

Uma vez avaliado que o sinal de interesse, o sinal de vazamento, possui uma banda de frequência que vai de aproximadamente 20 Hz até 4k Hz, foi possível, então, fazer uma subamostragem com fator 5 no sinal. Esta subamostragem modificou a frequência de amostragem original de 44.1kHz para 8.820kHz.

Todos os experimentos e simulações foram desenvolvidos na ferramenta Matlab [Mat11]. A rotina de cálculo da predição linear perceptual foi baseada na rotina desenvolvida por [Ell05]. Também foram utilizadas algumas rotinas de aquisição de áudio e de cálculo da distância de Itakura desenvolvidas por [Bro11].

### 5.1.1 Definição das Bases para Treinamento e Teste

Dado o grupo total de amostras, que é de 194, separou-se primeiramente os sinais de vazamento dos sinais de não vazamento. Foram selecionados 5 sinais de vazamento para serem utilizados como base de referência para o cálculo da distância de Itakura e da diferença espectral. O restante dos sinais foi dividido em dois diferentes subgrupos. O primeiro subgrupo, de aproximadamente 40% do total das amostras, como a base de dados de treinamento. O segundo subgrupo, de aproximadamente 60% do total das amostras, como a base de dados de teste, conforme apresenta a Tabela 5.2. Para fins de robustez dos resultados obtidos, foi feito um rodízio entre os dados dos conjuntos de treinamento e de teste. Com isto, cada resultado final é uma média de 10 repetições do experimento.

Subgrupo	Vazamento	Não Vazamento
Referência		
de vazamento	5	0
Treinamento	28	49
Teste	43	74
total	71	123

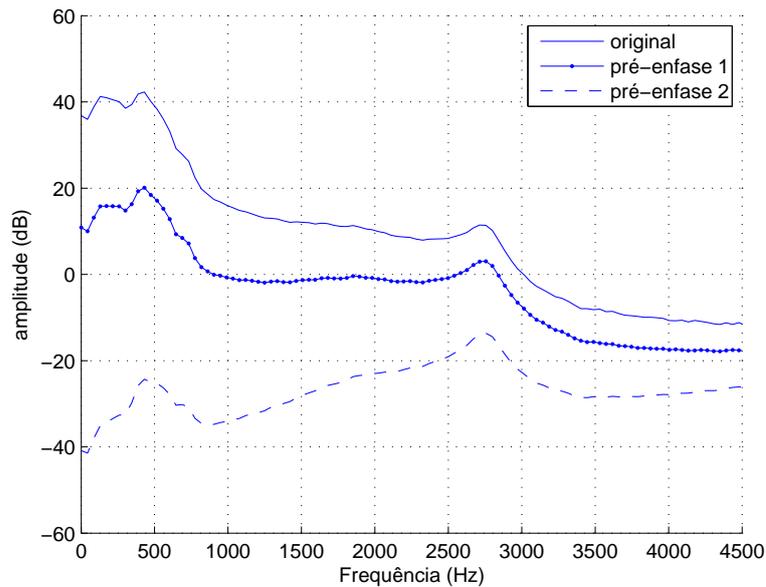
**Tabela 5.2:** *Divisão do grupo de dados para treinamento e teste dos parâmetros característicos*

## 5.2 Desenvolvimento prático

Conforme foi apresentado no primeiro capítulo, esta tese se propôs a estudar e avaliar a tecnologia de detecção de vazamento de água a partir de métodos acústicos e com este estudo propor novas formas de extração de parâmetros característicos a serem utilizados num sistema de detecção de vazamento. A primeira etapa desta pesquisa, e talvez a mais trabalhosa, foi a validação de todos os sinais adquiridos para formar a base de dados de

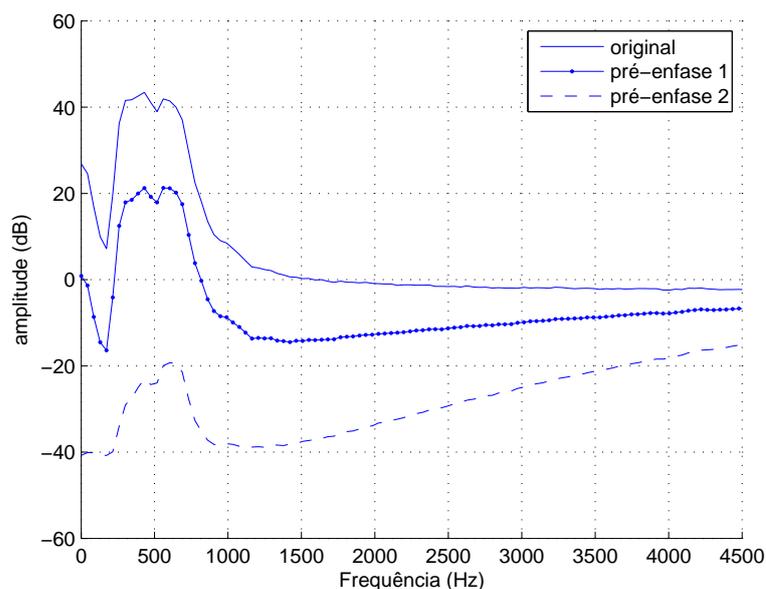
ruído de vazamento. Todas as amostras foram coletadas e rotuladas por técnicos especializados, porém para que as amostras pudessem ser realmente utilizadas nesta pesquisa foi necessário ouvir todas as amostras, manualmente, buscando verificar se a amostra capturada não apresentava nenhum tipo de perturbação indesejada que pudesse vir a prejudicar os resultados dos experimentos.

De posse da base de dados, o primeiro teste realizado foi o de reproduzir os experimentos descritos em (Yuriko e Akira, 2008) que propuseram a utilização de parâmetros característicos estatísticos obtidos a partir da distância máxima e média de Itakura com relação aos modelos de predição linear de sinais de vazamento definidos como referência. Estes experimentos já haviam sido avaliados em uma base de dados menor em [BR10], onde a TA ficou em torno de 80%. Para esta nova avaliação, somente as medidas de distância foram consideradas, uma vez que o terceiro parâmetro característico, o maior autovalor, não havia apresentado bons resultados anteriormente. Para esta nova base de dados a TA obtida ficou em torno de 65%.



**Figura 5.1:** DEP do sinal sem vazamento original e após as pré-ênfase de primeira e segunda ordem.

Um segundo experimento foi de extrair os mesmos parâmetros característicos, porém desta vez considerando a utilização da pré-ênfase. Nas Figuras 5.1 e 5.2 são apresentados os espectros de potência dos sinais de não vazamento e de vazamento, respectivamente. Para os sinais de não vazamento, com a utilização do filtro de pré-ênfase, algumas componentes de alta frequência que não eram percebidas no sinal original puderam agora ser consideradas. Com o filtro de pré-ênfase de segunda ordem o pico de frequência em torno de 3k Hz destacou-se mais que o pico de 500 Hz presente no sinal original. Já para o sinal de vazamento, ficou evidente que este tipo de sinal não possui componentes de alta



**Figura 5.2:** DEP do sinal com vazamento original e após as pré-ênfase de primeira e segunda ordem.

frequência, ao contrário do que apresentou o sinal de não vazamento.

Parâmetros característicos	Sinal original	Sinal com pré-ênfase de ordem=1	Sinal com pré-ênfase de ordem=2
Modelo AR de ordem $p = 10$			
Distância máxima e média de Itakura	65.60	60.80	79.86
Diferença máxima e média espectral	59.73	70.93	71.06
Modelo PLP de ordem $p = 10$			
Distância máxima e média de Itakura	62.93	66.93	69.60
Diferença máxima e média espectral	66.00	75.06	84.13

**Tabela 5.3:** TA para os parâmetros distância máxima e média de Itakura para os modelo AR e PLP de ordem de predição  $p = 10$ .

Após finalizado este experimento, foram então testados os parâmetros característicos agora incluindo o processamento psicoacústico, que é a análise de predição linear perceptual. Infelizmente, os resultados obtidos, para o caso do sinal original, foram piores que os sem processamento psicoacústico, modelos AR. Porém, com a utilização dos filtros de pré-ênfase foi obtida uma melhora significativa para os parâmetros de diferença espectral dos modelos PLP com relação aos modelos AR, de 71% para 84%.

Uma vez que os resultados obtidos, utilizando modelos de predição linear perceptual, não obtiveram a melhora esperada com relação ao modelo linear, outros parâmetros característicos, usualmente utilizados em processamento de sinais de áudio musical, foram propostos. Na próxima sessão faremos uma avaliação individual dos diversos parâmetros característicos propostos de tal forma que possamos determinar combinações de parâme-

tros que possam alcançar o melhor desempenho possível.

### 5.3 Avaliação dos parâmetros individualmente

A primeira análise feita sobre os parâmetros característicos foi a verificação do potencial de separabilidade das classes para cada parâmetro característico. Para fazer esta avaliação utilizou-se a medida estatística-M, conforme fora apresentada no Capítulo 3. A seguir são ilustrados os histogramas de cada um dos parâmetros característicos, para as classes do sistema, que são *vaz* e *nvaz*.

Na Tabela 5.4 são apresentados os valores da estatística-M para cada um dos parâmetros característicos avaliados. Na segunda coluna desta tabela estão dispostos os valores obtidos para a extração feita sobre o sinal original. Na terceira coluna estão os valores obtidos a partir do sinal resultante da pré-ênfase de primeira ordem. E, por fim, na quarta coluna, os valores obtidos a partir do sinal resultante da pré-ênfase de segunda ordem.

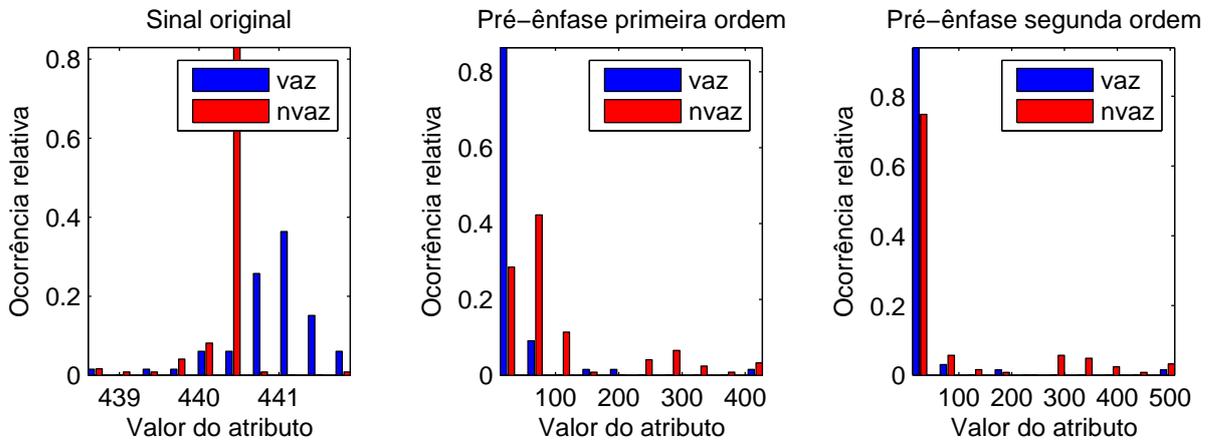
Parâmetro	Sinal original	Sinal com pré-ênfase de primeira ordem	Sinal com pré-ênfase de segunda ordem
Energia de curto prazo	0.46	0.22	0.66
Taxa de cruzamento por zero	0.43	1.54	1.66
Centróide espectral	1.21	1.63	1.53
Largura de espectro	1.66	1.38	0.67
Fluxo espectral	1.81	1.15	0.94
Ponto de roll-off	1.52	1.55	1.34
Taxa de espalhamento	0.80	0.36	0.70
Modelo AR ordem=2			
Distância máxima de Itakura	0.19	0.48	0.69
Distância média de Itakura	0.15	0.48	0.74
Diferença máxima espectral	0.05	0.85	0.74
Diferença média espectral	0.23	0.88	0.97
Modelo PLP ordem=2			
Distância máxima de Itakura	0.49	0.21	0.27
Distância média de Itakura	0.37	0.32	0.17
Diferença máxima espectral	0.47	0.40	0.40
Diferença média espectral	0.44	0.42	0.14

**Tabela 5.4:** Medida de separabilidade entre as classes *vaz* e *nvaz* em parâmetros extraídos a partir do sinal original e com pré-ênfase.

A partir da Tabela 5.4 pode-se dizer que, para o sinal original, os parâmetros característicos que melhor distinguem as duas classes são: o centróide espectral, a largura de espectro, o fluxo espectral e o ponto de *roll-off*. Todos os parâmetros que apresentaram separabilidade maior são parâmetros relacionados basicamente a distribuição das frequências no espectro. Isto já era esperado uma vez que a DEP dos sinais com e sem ruído de

vazamento apresentam formatos distintos com relação ao ponto central do espectro e a banda na qual ele ocupa.

Já para o caso dos sinais com pré-ênfase de primeira ordem não trouxe aumento significativo da separabilidade entre as classes para todos os parâmetros, exceto para o caso do Centróide espectral, taxa de cruzamento por zero e todas as distâncias e as distorções para o modelo AR. Já com a utilização da pré-ênfase de segunda ordem, pode-se observar um aumento na estatística-M para os parâmetros característicos energia de curto prazo, taxa de cruzamento por zero, as distâncias máxima e média de Itakura AR e diferença média espectral AR.

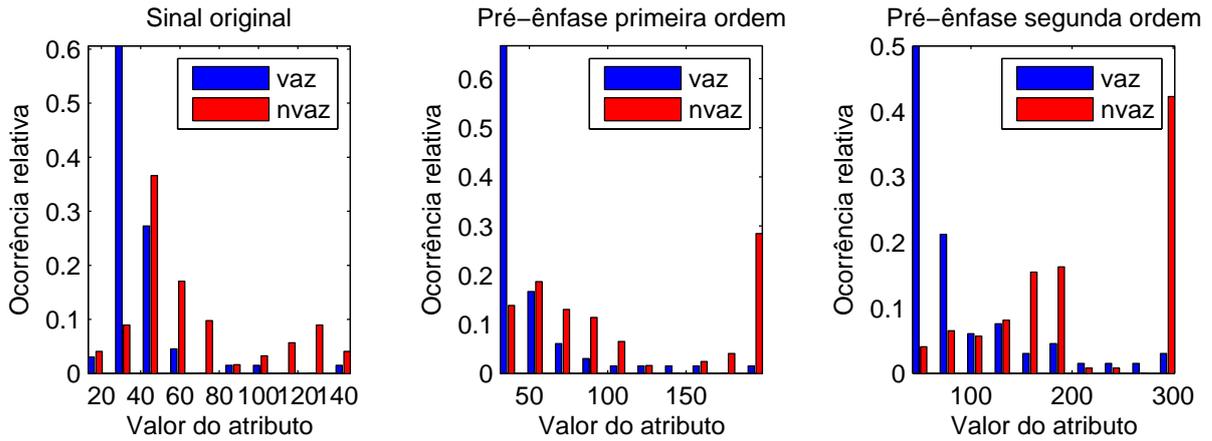


**Figura 5.3:** Histograma : energia de curto prazo.

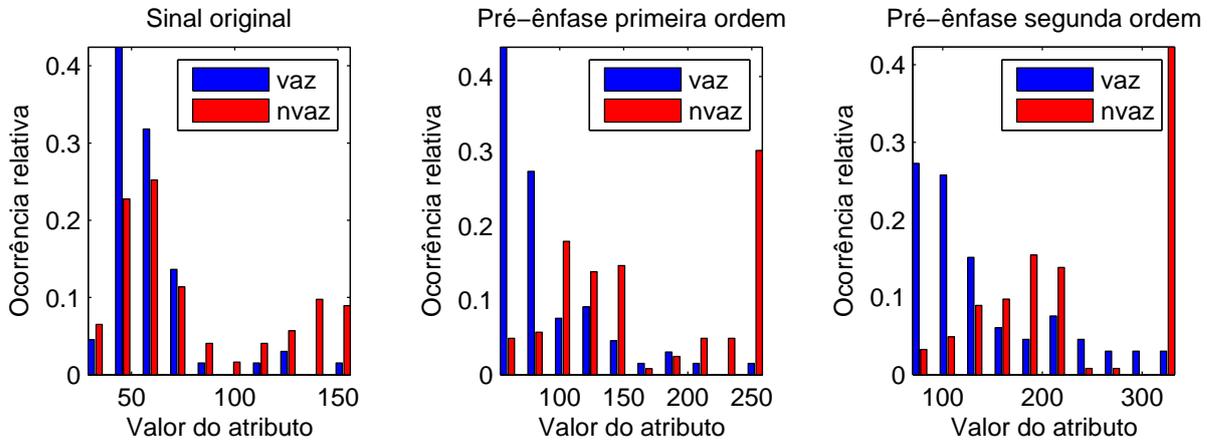
Conforme ilustra a Figura 5.3 o parâmetro característico energia de curto prazo não demonstra ser uma informação relevante quando à separabilidade entre as classes. Para o sinal original, a distribuição dos valores do parâmetro para ambas as classes está muito sobreposta. Tanto para o sinal original quanto para os sinais com pré-ênfase, a distribuição dos valores do parâmetro para ambas as classes está muito sobreposta. Esta sobreposição pode ser verificada através do valor da estatística-M que é de 0.44 para o sinal original, 0.22 para o sinal com pré-ênfase de primeira ordem e de 0.66 para o sinal com pré-ênfase de segunda ordem. Apesar dos valores da estatística-M serem diferentes para cada sinal, em termos de taxa de acerto de reconhecimento não houve variação, em todos os casos ela se manteve em torno de 64%.

O parâmetro taxa de cruzamento por zero, obtido a partir do sinal original apresenta um valor de estatística-M de 0.42 que vai de encontro com o resultado obtido na TA que foi de 64%. Por outro lado para os sinais com pré-ênfase o valor da estatística-M foi maior, de 1.54 para o de primeira ordem e de 1.60 para o de segunda ordem. Em termos de TA os resultados para os sinais com pré-ênfase obtiveram 81.70% e 84.04%, respectivamente para primeira e segunda ordem, obtendo uma melhora em termos absolutos de aproximadamente 20% maior se comparado com o sinal original.

O parâmetro característico centróide espectral do sinal original apresenta um bom



**Figura 5.4:** Histograma : taxa de cruzamento por zero.



**Figura 5.5:** Histograma : centróide espectral.

valor de estatística-M, que foi de 1.20. Para os sinais com pré-ênfase, a valor aumentou para 1.63, para a primeira ordem, e de 1.53 para a segunda ordem. Em termos de taxa de acerto, o melhor resultado foi obtido para o sinal com pré-ênfase de primeira ordem, que foi de 83.08%. Para os sinais com pré-ênfase, a TA do de primeira ordem foi 77.87% e para o de segunda ordem 81.91%.

A informação da banda do sinal, representada pelo parâmetro característico, largura de espectro, foi uma dos parâmetros que se mostrou mais discriminativo de forma isolada. Para o sinal original, a estatística-M obtida foi de 1.66 e sua respectiva taxa de acerto de classificação de 85.31%. Quando da utilização da pré-ênfase de primeira ordem, tanto a estatística-M quando a sua taxa de acerto tiveram uma queda, obtendo uma estatística-M de 1.38 e a taxa de acerto de 83.29%. Por outro lado, utilizando a pré-ênfase de segunda ordem a estatística-M foi de 0.66 e a taxa de acerto foi de 87.55%, o que demonstra que para este caso, o valor da estatística-M não reflete o desempenho do classificador, uma vez que o esperado era de reduzir a taxa de acerto. Uma explicação para este fato é de que aparentemente a distribuição dos valores para este parâmetro não segue um padrão

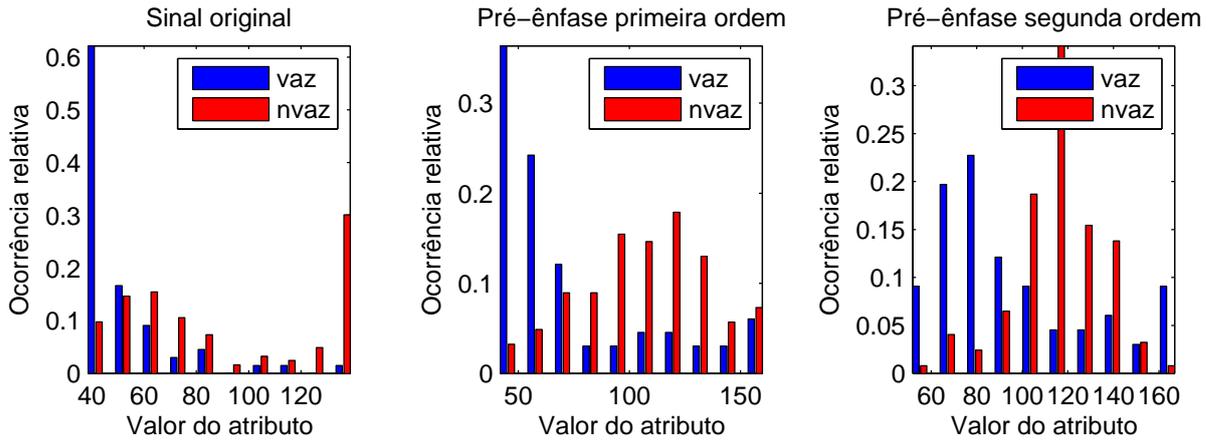


Figura 5.6: Histograma : largura de espectro

de distribuição normal.

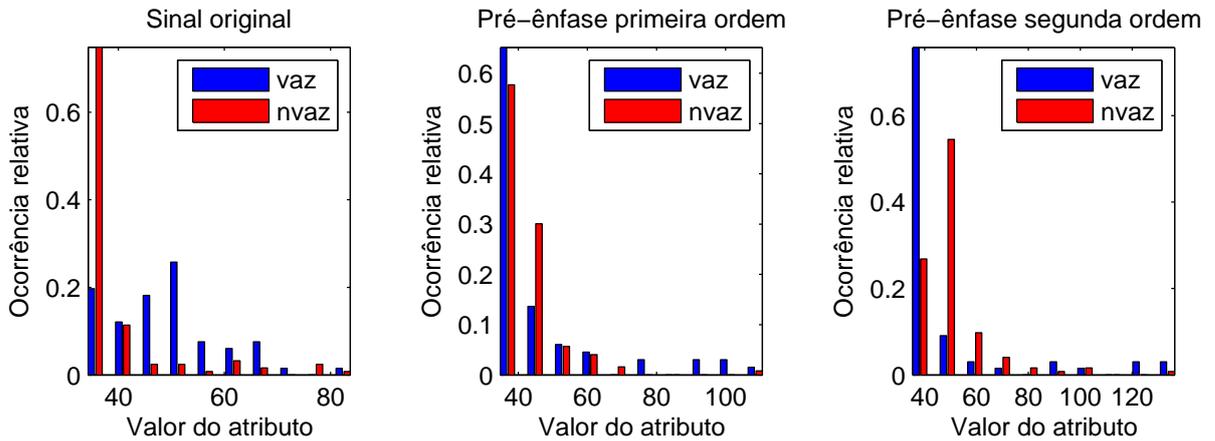


Figura 5.7: Histograma : fluxo espectral.

A partir da análise visual dos histogramas do parâmetro fluxo espectral pode-se dizer que o efeito da pré-ênfase sobre a separação das classes não teve destaque. Para o sinal original o valor da estatística-M que foi de 1.81 e sua TA foi de 83.61%. Para o sinal com pré-ênfase de primeira ordem a estatística-M foi de 1.15 porém a TA não sofreu alteração. Já para o sinal com pré-ênfase de segunda ordem o valor da estatística-M foi de 0.94 e sua respectiva TA de 76.70%.

Para o parâmetro característico ponto de *roll-off* podemos dizer que não houve mudança significativa entre os valores da estatística-M para o sinal original e para os sinais com pré-ênfase, que no caso foi de 1.52, 1.55 e 1.34 respectivamente. pra pré-ênfase de primeira ordem para a segunda observou-se uma leve queda e esta queda também pode ser percebida na TA, que passou de 83.29% para 79.89%.

Para o caso do parâmetro taxa de espalhamento, os valores da estatística-M para o sinal original e para o sinal com pré-ênfase de segunda ordem, indicaram uma bom grau de separação, que foi de 0.80 e 0.70, respectivamente. Porém em termo de TA, não

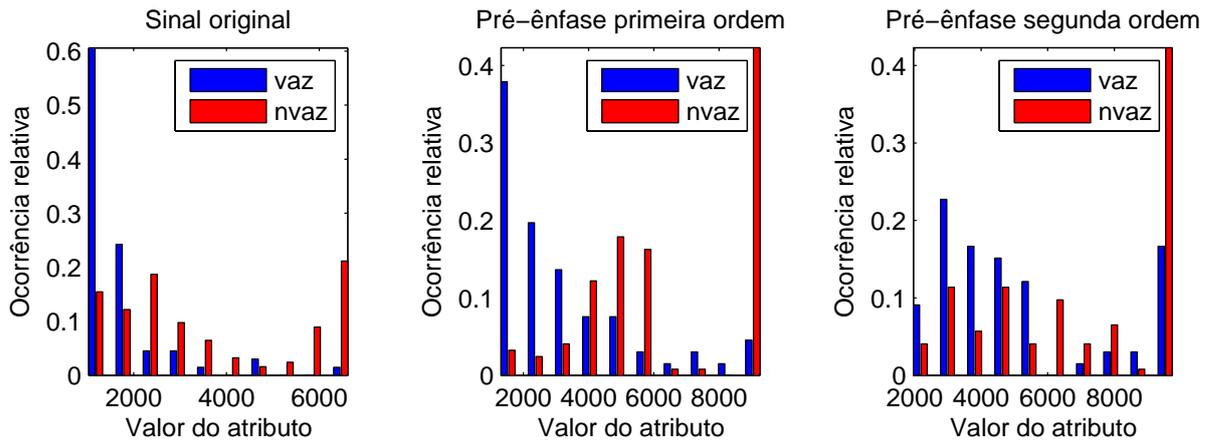


Figura 5.8: Histograma : ponto de roll-off.

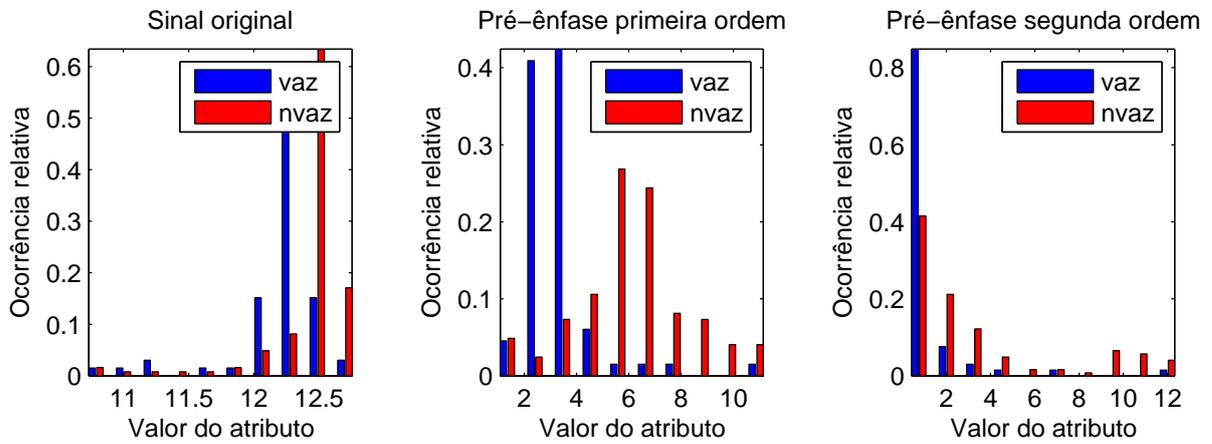


Figura 5.9: Histograma : taxa de espalhamento

houve nenhuma mudança pois em ambos os casos a TA foi em média 66%. E apesar de a estatística-M para o sinal com pré-ênfase de primeira ordem ter sido menor que as anteriores, a TA também permaneceu inalterada.

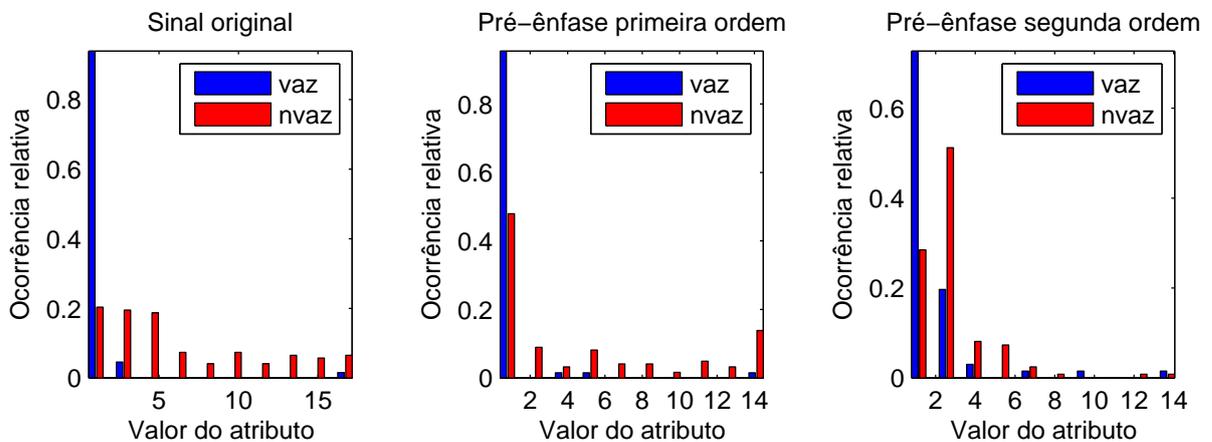
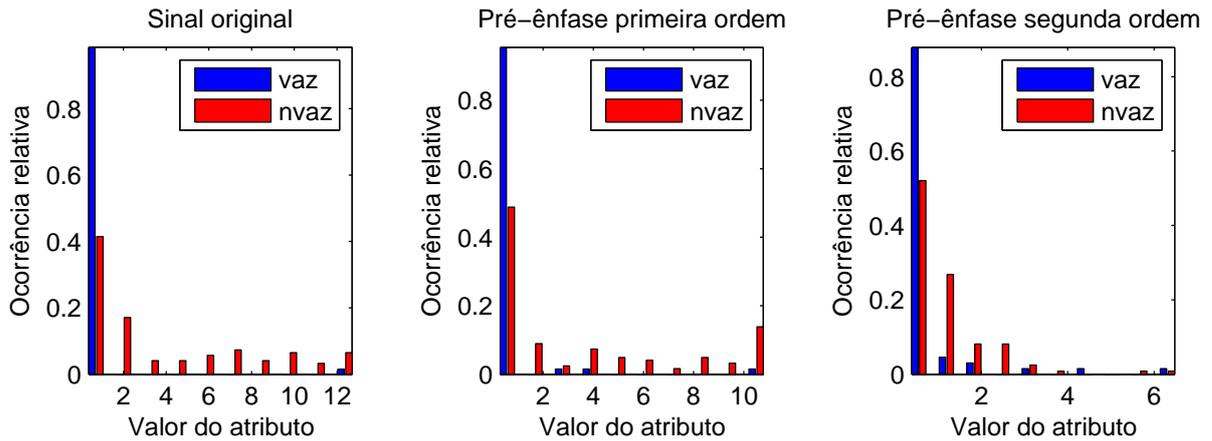


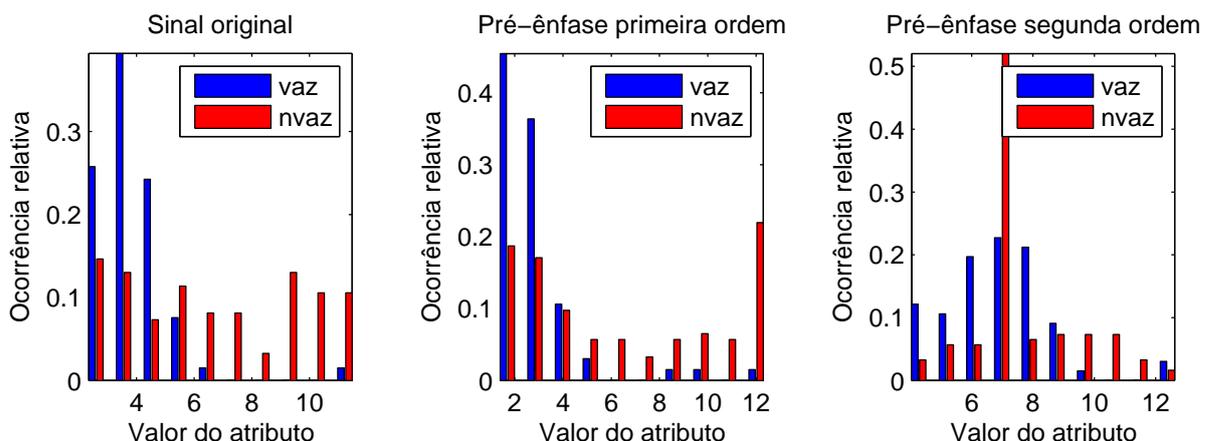
Figura 5.10: Histograma : Máxima distância de Itakura - AR

O parâmetro característico Máxima distância de Itakura, para o modelo AR, apresentou um valor de de 0.19 para o sinal original. Com a utilização das pré-ênfases, os valores da estatística-M subiu para 0.48 no caso do sinal com pré-ênfase de primeira ordem e de 0.70 para o de segunda ordem. Com relação à TA o melhor resultado, de 75.33%, foi obtido para o sinal com pré-ênfase de segunda ordem. Já para o sinal original a taxa foi de 66.13% e para o sinal com pré-ênfase de segunda ordem a taxa caiu para 61.46%.



**Figura 5.11:** Histograma : Distância Média de Itakura - AR

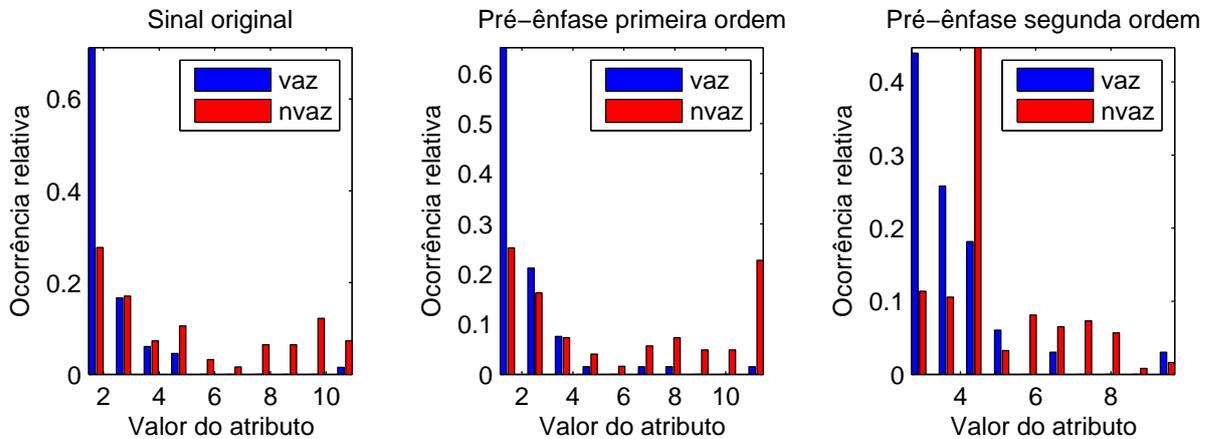
A variação da estatística-M para o parâmetro distância média de Itakura, reflete o comportamento obtido no classificador. Os valores de estatística-M foram de 0.15 para o sinal original, 0.48 para o sinal com pré-ênfase de primeira ordem e de 0.74 para o de segunda. Os resultados do classificador, foram de 59.73%, 70.93% e 71.06%, respectivamente para o sinal original, sinal com pré-ênfase de primeira e de segunda ordem.



**Figura 5.12:** Histograma : Máxima diferença espectral - AR

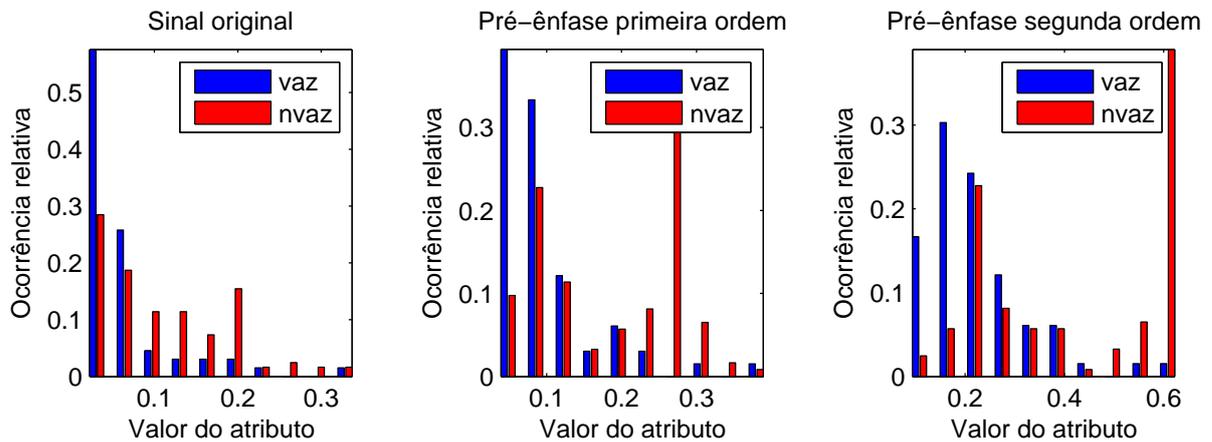
Para o parâmetro característico máxima diferença espectral do sinal original a estatística-M foi de 0.05, para o sinal de pré-ênfase de primeira ordem o valor da estatística-M aumentou para 0.85 e para de segunda ordem o valor diminuiu para 0.74. Porém em termos

de TA, o ganho não foi tão significativo, sendo de 62.93% para o sinal original, de 66.90% para o sinal de pré-ênfase de primeira ordem e de 69.6% para o de segunda.



**Figura 5.13:** Histograma : Diferença média espectral - AR

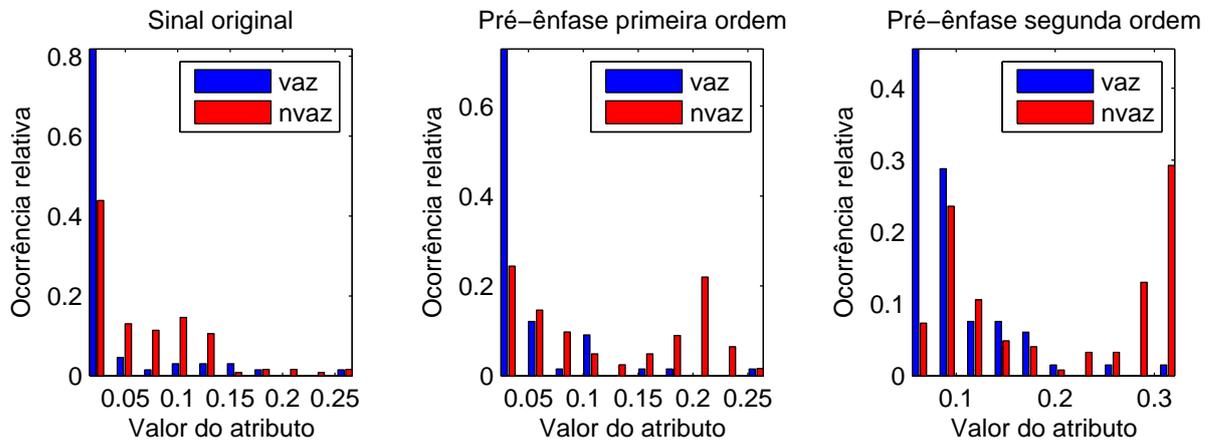
O parâmetro característico distância espectral média para o modelo AR obteve melhora em seus valores de estatística-M com a utilização das pré-ênfases. Para o sinal original o valor foi de 0.23 e já para os demais foi de 0.87 e 0.96, respectivamente. Em termos da TA do classificador foi obtida uma melhora significativa com relação a do sinal original, que foi de 66.00%, para os sinais com pré-ênfase, que obtiveram 75.07% para a pré-ênfase de primeira ordem e de 84.13% para o de segunda ordem.



**Figura 5.14:** Histograma : Distância máxima de Itakura - PLP

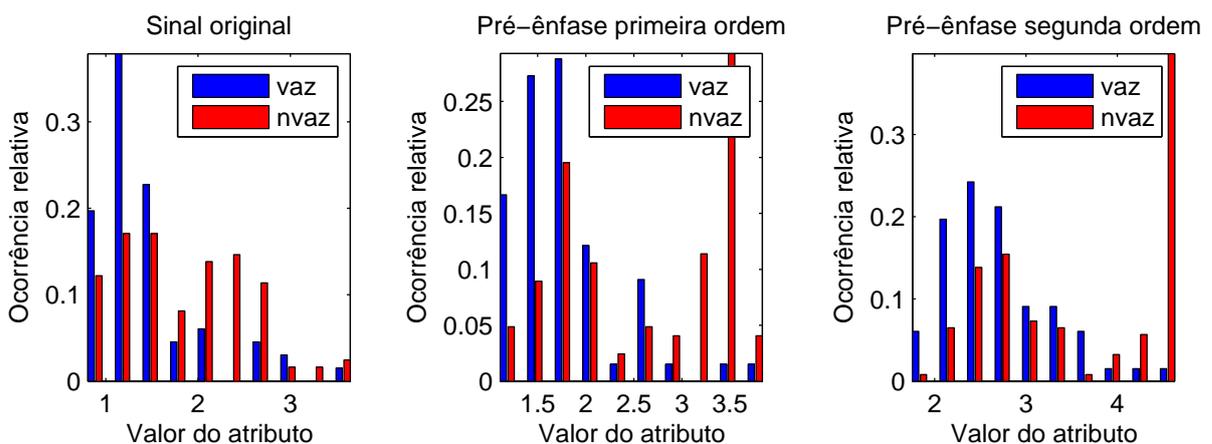
A partir da avaliação dos valores da estatística-M para o parâmetro característico estatístico distância máxima de Itakura, para o modelo de predição perceptual, não se espera obter bons resultados em termos de TA, pois os mesmos indicam baixo poder de separação entre as classes, que foi de 0.50 para o sinal original, 0.21 para o sinal com pré-ênfase de primeira ordem e de 0.27 para o de segunda ordem. Já com relação às taxas de acerto, os resultados foram diferentes do esperado. Onde se esperava que a TA

fosse menor, para o sinal com pré-ênfase de primeira ordem, obteve-se o mesmo resultado que para o sinal original, que deveria ser de 66.27%. Já para o sinal com pré-ênfase de segunda ordem, a TA foi a mais elevada, chegando a 78.13%. Uma justificativa para este comportamento se dá em função do formato da distribuição dos dados.



**Figura 5.15:** Histograma : Distância média de Itakura - PLP

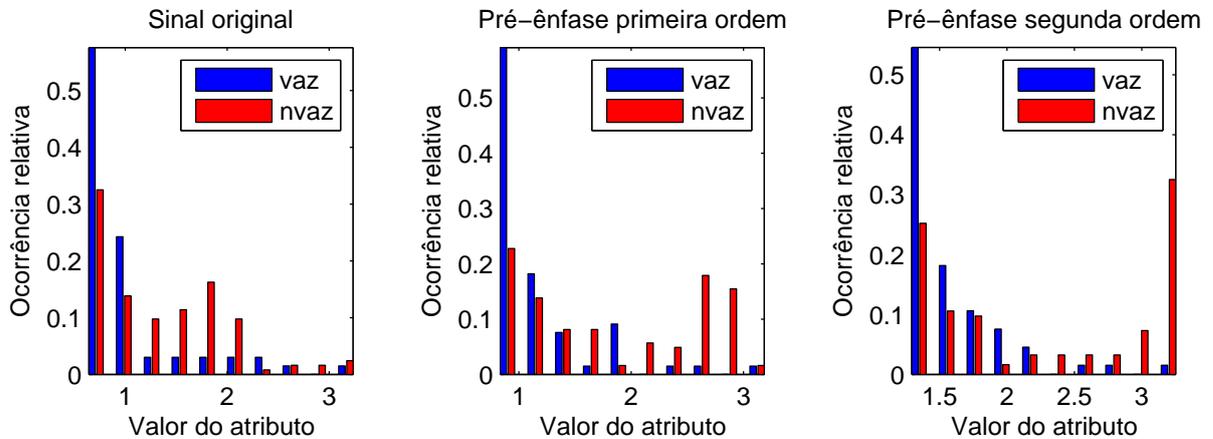
A distribuição dos valores do parâmetro característico distância média de Itakura para o modelo PLP apresenta o mesmo comportamento do parâmetro distância máxima de Itakura, uma vez que os resultados da taxa de acerto do classificador não correspondem ao valor de estatística-M calculados. A TA para o sinal original foi de 66.80% e para os sinais com pré-ênfase de primeira e segunda ordem foi de 70.93%. Os valores de estatística-M calculados para cada uma dos sinais foi de 0.37, 0.32 e 0.17, respectivamente para o sinal original e para os sinais com pré-ênfase de primeira e segunda ordem.



**Figura 5.16:** Histograma : Diferença máxima espectral - PLP

O histograma do parâmetro característico diferença máxima espectral para o modelo PLP é ilustrado a Figura 5.16. Os valores de estatística-M obtidos para os diferentes tipos de sinais de entrada foram próximos, em torno de 0.40. Em termos de TA o melhor

resultado foi obtido pelo sinal com pré-ênfase de segunda ordem, onde obteve-se 72.27%. Nos demais casos, o resultado da TA ficou abaixo, em torno de 65.00%.



**Figura 5.17:** Histograma : Diferença média espectral - PLP

Por fim, na Figura 5.17 temos o histograma das distribuições dos dados do parâmetro característico distância média espectral para o modelo PLP. O valor da estatística-M para o sinal original e de pré-ênfase de primeira ordem foi de 0.40; e suas taxas de acerto algo em torno de 66.00%. Já para o sinal com pré-ênfase de segunda ordem a taxa de acerto foi de 76.53% apesar de a estatística-M para este sinal de entrada ter sido de 0.15.

Na Tabela 5.5 é apresentado um resumo das TAs de cada parâmetro característico, para os três diferentes sinais de entrada: sinal original, sinal com pré-ênfase de primeira e de segunda ordem.

A partir destes resultados pode ser verificado que alguns parâmetros característicos já apresentam TA relativamente alta, como é o caso do parâmetro largura de espectro, que obteve uma TA em torno de 87% de acerto. O pior desempenho foi apresentado pelo parâmetro energia de curto prazo, que independente da utilização ou não do filtro de pré-ênfase, sua TA é muito inferior aos dos outros parâmetros característicos. Pode-se dizer que com a utilização de dois filtros de pré-ênfase a TA aumentou para todos os parâmetros característicos. Esta melhora possivelmente se deveu ao fato de que com a aplicação dos filtros, algumas componentes de alta frequência que não eram percebidas foram realçadas, assim como algumas componentes de baixa frequência que eram anteriormente mascaradas acabaram reaparecendo, conforme já foi apresentado na Figura 5.2.

## 5.4 Avaliação em grupo

Nesta sessão são apresentados os resultados obtidos a partir das medidas TA, TAV e TANV, conforme descrito no Capítulo 3, para diferentes conjuntos de parâmetros característicos. A TA é a relação entre o número de sinais classificados de forma correta e o

Parâmetro	Sinal original	Sinal com pré-ênfase de primeira ordem	Sinal com pré-ênfase de segunda ordem
Energia de curto prazo	64.78	64.89	64.89
Taxa de cruzamento por zero	64.14	81.70	84.04
Centróide espectral	77.87	83.08	81.91
Largura de espectro	85.31	84.89	87.55
Fluxo espectral	83.61	83.29	76.70
Ponto de roll-off	79.89	83.29	84.14
Taxa de espalhamento	65.63	65.31	68.08
Modelo AR ordem=2			
Distância máxima de Itakura	66.13	61.46	75.33
Distância média de Itakura	66.13	64.13	80.66
Diferença máxima espectral	61.33	62.80	70.13
Diferença média espectral	64.80	64.93	76.26
Modelo PLP ordem=2			
Distância máxima de Itakura	62.26	70.93	69.33
Distância média de Itakura	63.20	64.93	67.06
Diferença máxima espectral	63.33	69.33	67.46
Diferença média espectral	60.00	65.06	67.20

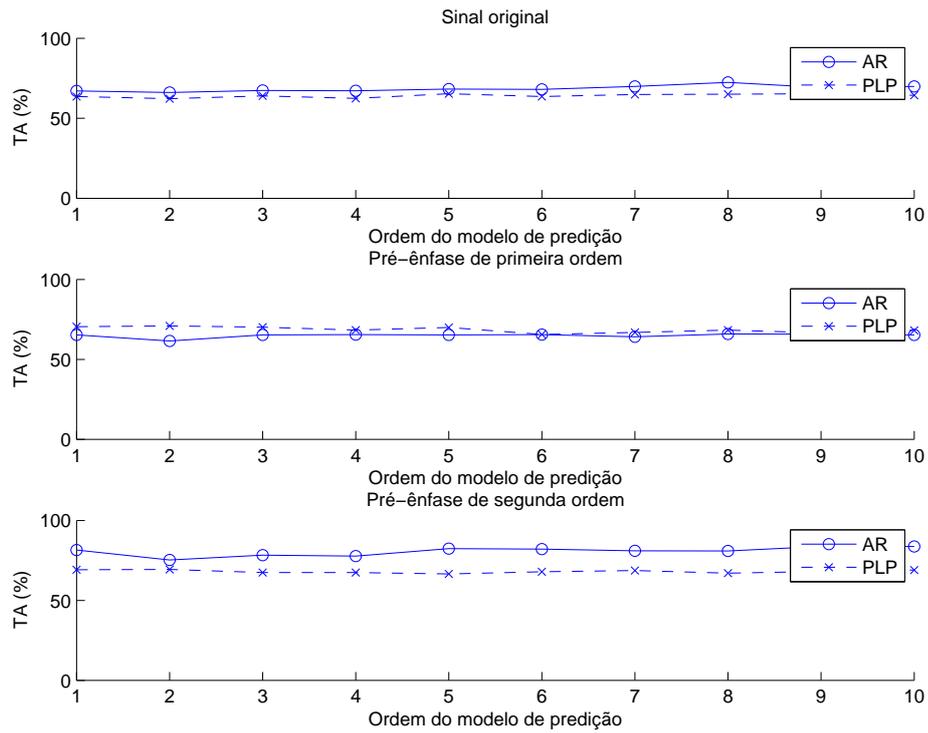
**Tabela 5.5:** *Melhores resultados da TA para o sinal original e os sinais com pré-ênfase.*

número total de sinais. A TAV é a relação entre o número de sinais de vazamento classificados corretamente e o número total de sinais de vazamento. Por fim a TANV é a relação entre o número de sinais de não vazamento classificados corretamente e o número total de sinais de não vazamento. Com estas três medidas podemos avaliar se os parâmetros conseguem discernir ambas as classes com taxa de erro baixa.

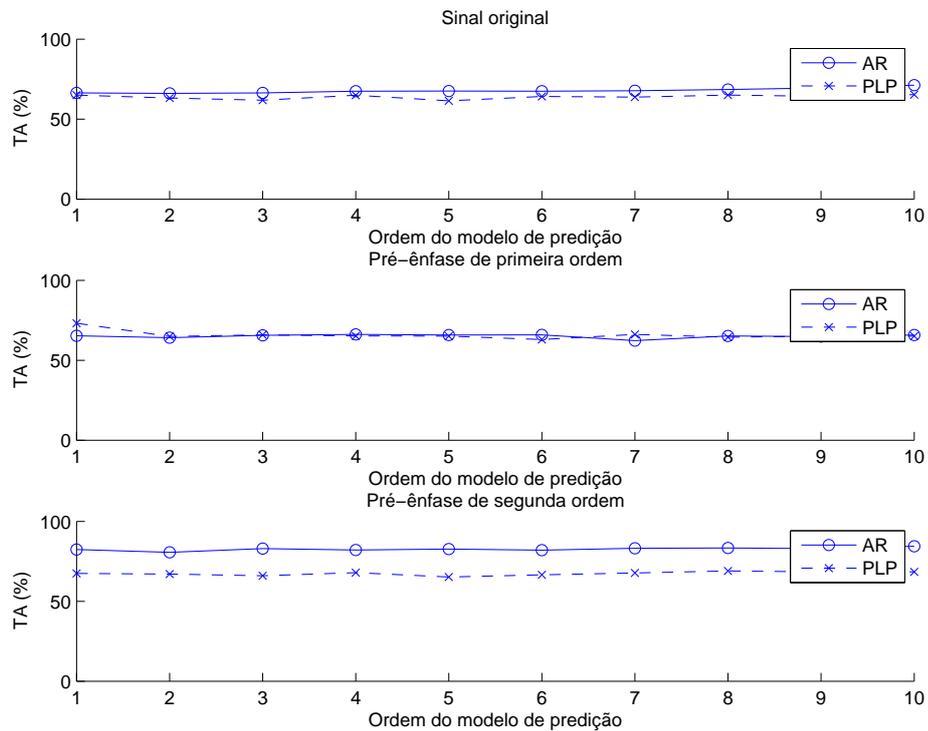
#### 5.4.1 Classificação de acordo com a ordem do modelo de predição linear

Primeiramente foram avaliados os resultados de classificação para os parâmetros característicos obtidos a partir de medidas de distância dos modelo de predição linear de teste com relação à base de modelos de referência de sinais de vazamento. Esta avaliação foi feita com o intuito de estimar a ordem do modelo de predição que melhor auxilia na classificação. Foram avaliados tanto os modelos de predição linear quanto os modelos de predição linear perceptual. Nas Figuras 5.18 e 5.19 são apresentados os resultados para os parâmetros característicos máxima e distância média de Itakura, respectivamente. Para estes parâmetros característicos a ordem que melhor auxilia na TA está entre 2 e 6.

Nas Figuras 5.20 e 5.21 são apresentados os resultados para os parâmetros característicos de diferença máxima e média espectral. Também para este parâmetro característico a ordem do modelo que melhor auxilia na TA está entre 2 e 6.



**Figura 5.18:** Taxa de acerto da distância máxima de Itakura por ordem de predição.



**Figura 5.19:** Taxa de acerto da distância média de Itakura por ordem de predição.

De acordo com as TAs obtidas para os diferentes tipos de medidas dos modelos de predição, fica evidenciado que os modelos de predição podem ser estimados a partir de ordem 2 ao contrário do foi apresentado nos experimentos anteriores, que eram de ordem 10.

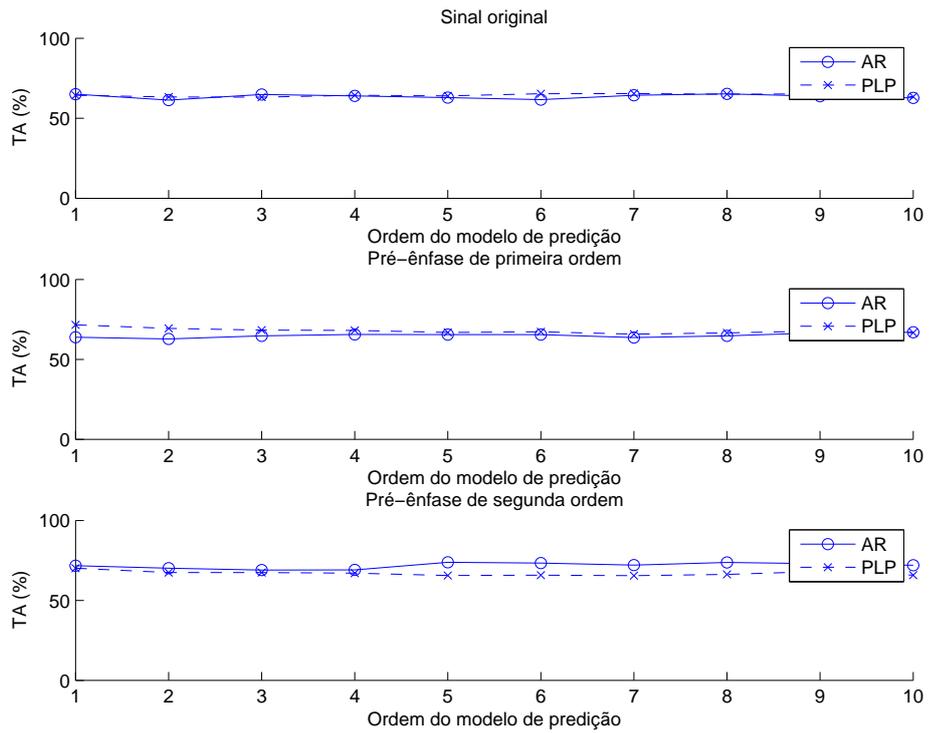


Figura 5.20: Taxa de acerto da distância máxima espectral por ordem de predição.

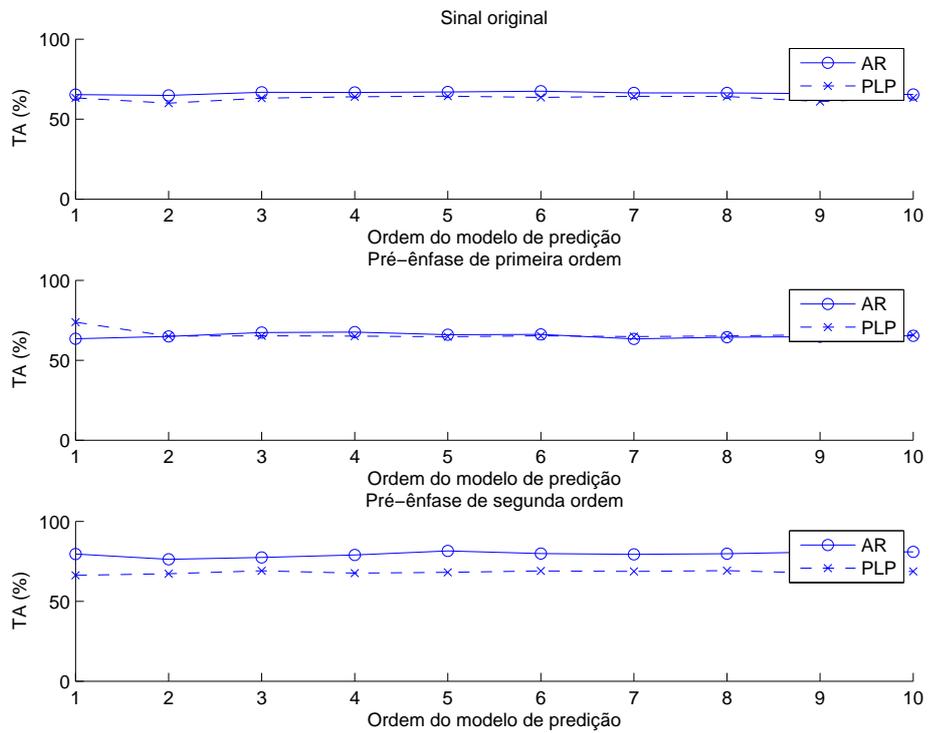


Figura 5.21: Taxa de acerto da diferença média espectral por ordem de predição.

## 5.4.2 Classificação para grupos de diversas combinações de parâmetros

A partir dos resultados obtidos na avaliação individual dos parâmetros, foram selecionados os parâmetros que apresentaram melhores resultados de estatística-M. Portanto, o grupo de parâmetros selecionados para o sinal original (sem pré-ênfase) e para os demais sinais modificados (com pré-ênfase de primeira e segunda ordem) é apresentado na Tabela 5.6, juntamente com suas respectivas taxas de acerto do classificador.

Sinal de entrada	Parâmetros característicos	TA(%)	TANV(%)	TAV(%)
Original	Centróide espectral	86.80	79.23	90.81
	Largura de espectro			
	Fluxo espectral			
	Ponto de roll-off			
Pré-ênfase de primeira ordem	Taxa de cruzamento por zero	88.13	86.15	89.18
	Centróide espectral			
	Largura de espectro			
	Fluxo espectral			
Pré-ênfase de segunda ordem	Taxa de cruzamento por zero	88.93	80.77	93.26
	Centróide espectral			
	Ponto de roll-off			

**Tabela 5.6:** Taxa de acerto dos parâmetros característicos para o sinal original e com pré-ênfase.

Avaliando os resultados apresentados na Tabela 5.6 observa-se que os resultados obtidos para o sinal original foram os piores. Com a utilização da pré-ênfase de primeira ordem e também da adição do parâmetro Taxa de cruzamento por zero, obteve-se uma melhora no desempenho do classificador em torno de 1.5% na TA e em torno de 8% absoluto para a TANV, porém uma reduzindo em torno de 1.5% a TAV. Já com a utilização da pré-ênfase de segunda ordem e removendo os parâmetros Largura de espectro e Fluxo espectral, a TA manteve-se. Por outro lado a TANV reduziu praticamente 6% de forma absoluta para o sinal com pré-ênfase de segunda ordem.

Um segundo experimento foi o de calcular o desempenho do classificador somente utilizando os parâmetros baseados em predição linear, equivalente ao apresentado em [SM07], porém desta vez para modelos de predição com ordem 2, os resultados obtidos para o sinal original são apresentados na Tabela 5.7 e para os sinais com pré-ênfase de primeira ordem, na Tabela 5.8 e os de segunda ordem na Tabela 5.9.

Comparando os resultados obtidos a partir do sinal com pré-ênfase e o sinal original, observa-se que o desempenho do classificador para os parâmetros relacionados ao modelo AR piorou porém o desempenho para os parâmetros relacionados ao modelo PLP obtiveram melhora em termos de TA e TAV.

Parâmetros característicos	TA(%)	TANV(%)	TAV(%)
Modelo AR de ordem $p = 2$			
Distância máxima e média de Itakura	65.47	57.69	97.14
Diferença máxima e média espectral	66.67	33.85	84.08
Distância máxima e média de Itakura e Diferença máxima e média espectral	67.60	35.76	84.48
Modelo PLP de ordem $p = 2$			
Distância máxima e média de Itakura	62.13	19.61	84.69
Diferença máxima e média espectral	63.73	38.46	77.14
Distância máxima e média de Itakura e Diferença máxima e média espectral	66.13	16.92	92.24

**Tabela 5.7:** Resultado do classificador para o sinal original com os parâmetros baseados em modelos AR e PLP.

Parâmetros característicos	TA(%)	TANV(%)	TAV(%)
Modelo AR de ordem $p = 2$			
Distância máxima e média de Itakura	63.60	34.23	79.18
Diferença máxima e média espectral	63.47	40.77	75.51
Distância máxima e média de Itakura e Diferença máxima e média espectral	63.47	29.61	81.42
Modelo PLP de ordem $p = 2$			
Diferença máxima e média de Itakura	68.13	27.69	89.59
Diferença máxima e média espectral	68.80	26.15	91.43
Distância máxima e média de Itakura e Diferença máxima e média espectral	71.87	45.38	85.91

**Tabela 5.8:** Resultado do classificador para o sinal com pré-ênfase de primeira ordem com os parâmetros baseados em modelos AR e PLP.

Parâmetros característicos	TA(%)	TANV(%)	TAV(%)
Modelo AR de ordem $p = 2$			
Distância máxima e média de Itakura	79.33	73.85	82.24
Diferença máxima e média espectral	79.87	75.00	82.40
Distância máxima e média de Itakura e Diferença máxima e média espectral	83.33	79.23	85.51
Modelo PLP de ordem $p = 2$			
Dist. máxima e média de Itakura	70.53	35.00	89.39
Diferença máxima e média espectral	69.87	35.00	88.37
Distância máxima e média de Itakura e Diferença máxima e média espectral	71.20	43.07	86.12

**Tabela 5.9:** Resultado do classificador para o sinal com pré-ênfase de segunda ordem com os parâmetros baseados em modelos AR e PLP.

Analisando os resultados do classificador para sinais com pré-ênfase de segunda ordem, verifica-se que em termos de TA e TANV os resultados foram melhores se comparados com o do sinal original. Porém em termos de TAV somente os parâmetros relacionados ao modelo AR que obtiveram melhora, os demais, relacionados aos modelos PLP não apresentaram mudanças significativa.

O último experimento realizado foi de avaliar os resultados obtidos com a combinação dos parâmetros mais relevantes, descritos na Tabela 5.6, em conjunto com os parâmetros relacionados aos modelos AR e PLP.

Parâmetros característicos	TA(%)	TANV(%)	TAV(%)
Centróide espectral			
Largura de espectro	87.60	82.31	90.41
Fluxo espectral			
Ponto de roll-off			
Modelo AR de ordem $p = 2$			
+ Distância máxima e média de Itakura	83.47	72.69	89.18
+ Diferença máxima e média espectral	85.33	79.61	88.37
+ Diferença máxima e média de Itakura e Diferença máxima e média espectral	84.53	73.08	90.61
Modelo PLP de ordem $p = 2$			
+ Distância máxima e média de Itakura	84.80	73.46	90.82
+ Diferença máxima e média espectral	86.13	75.77	91.63
+ Distância máxima e média de Itakura e Diferença máxima e média espectral	84.27	68.85	92.45

**Tabela 5.10:** Resultado do classificador para o sinal original

Na Tabela 5.10 são apresentados os melhores resultados obtidos dentre as diversas combinações avaliadas para os sinais originais, seguido pelos resultados para sinais com pré-ênfase de primeira ordem e de segunda, Tabelas 5.11 e 5.12.

Consolidando os resultados obtidos para as diversas combinações de parâmetros característicos do sinal original, com a inclusão dos parâmetros estatísticos o resultado final do classificador foi deteriorado. Já no caso do sinal com pré-ênfase de primeira ordem, a adição dos parâmetros de diferença espectral máxima e média do modelo PLP obteve-se uma melhora moderada no desempenho do classificador, basicamente com relação à classificação dos sinais de vazamento.

Por fim, os resultados do classificador para o sinal com pré-ênfase de segunda ordem foram melhorados com a inclusão dos parâmetros relacionados à diferença espectral do modelo PLP, onde a TA mudou de 91.07% para 92.53%, representando um ganho de 1.6%. Este ganho na TA foi basicamente relacionado ao aumento na TNAV.

Uma fato curioso percebido foi de que, apesar das TAs terem sido piores para os parâmetros característicos relacionados à predição linear perceptual, quando estes parâmetros

Parâmetros característicos	TA(%)	TANV(%)	TAV(%)
Centróide espectral			
Largura de espectro	89.87	85.77	92.04
Fluxo espectral			
Ponto de roll-off			
Modelo AR de ordem $p = 2$			
+ Distância máxima e média de Itakura	89.33	84.23	92.04
+ Diferença máxima e média espectral	89.87	84.23	92.86
+ Distância máxima e média de Itakura e Diferença máxima e média espectral	89.87	81.92	94.08
Modelo PLP de ordem $p = 2$			
+ Distância máxima e média de Itakura	88.93	87.31	89.80
+ Diferença máxima e média espectral	90.40	85.00	93.26
+ Distância máxima e média de Itakura e Diferença máxima e média espectral	86.53	82.69	88.57

**Tabela 5.11:** Resultado do classificador para o sinal com pré-ênfase de primeira ordem

Parâmetros característicos	TA(%)	TANV(%)	TAV(%)
Centróide espectral			
Largura de espectro	91.07	83.46	95.10
Fluxo espectral			
Ponto de roll-off			
Modelo AR de ordem $p = 2$			
+ Distância máxima e média de Itakura	90.80	85.00	93.88
+ Diferença máxima e média espectral	91.60	85.00	95.10
+ Distância máxima e média de Itakura e Diferença máxima e média espectral	90.53	85.00	93.47
Modelo PLP de ordem $p = 2$			
+ Distância máxima e média de Itakura	91.33	83.46	95.51
+ Diferença máxima e média espectral	92.53	88.46	94.70
+ Distância máxima e média de Itakura e Diferença máxima e média espectral	90.40	85.38	93.06

**Tabela 5.12:** Resultado do classificador para o sinal com pré-ênfase de segunda ordem

foram combinados com alguns dos parâmetros característicos individuais do sinal, a TA resultante obteve o melhor de classificação do sistema. Possivelmente porque as informações que os parâmetros característicos, relacionados ao PLP, carregam não estão sendo consideradas pelos outros parâmetros e que apesar de estes parâmetros não alcançarem bons resultados de TA, ao serem combinados com os parâmetros individuais os mesmos agregam novas componentes que auxiliam na tarefa de classificação, implicando diretamente no aumento da TA.

# Capítulo 6

## Conclusões e Propostas de Trabalhos

### 6.1 Conclusões

A partir dos estudos e dos experimentos realizados durante o desenvolvimento desta tese de doutorado podemos concluir, de uma forma geral, que a detecção automática de vazamento de água a partir da análise do sinal sonoro por ele gerado é viável.

Tomando como referência um dos métodos encontrados na literatura, o que classifica, utilizando máquina de vetores de suporte, os parâmetros obtidos a partir dos modelos de predição linear, chegou-se à conclusão de que os resultados apresentados no artigo não se mantêm quando da utilização de dados reais, que foram coletados em campo por técnicos especializados.

Os melhores resultados de classificação em grupo foram obtidos a partir da extração dos parâmetros característicos pré-processados com pré-ênfase de segunda ordem.

A melhor ordem dos modelos de predição linear normal e perceptual é menor que a ordem utilizada em experimentos anteriores [SM07]. Em função disto cabe salientar que esta contribuição é relevante uma vez que a ordem do modelo de predição mudou de 10 para 2.

Dentre os parâmetros característicos que apresentaram os melhores desempenhos na detecção, a grande maioria está relacionada diretamente com a informação espectral do sinal, como, por exemplo, os parâmetros característicos largura de banda e centróide espectral. Já os parâmetros extraídos no domínio do tempo não se mostraram muito relevantes.

A técnica que foi muito importante para a obtenção destes resultados foi a aplicação de filtros de pré-ênfase no sinal. Estes filtros aumentam o ganho das altas frequências e atenuam o das baixas. Com o aumento do ganho das altas frequências algumas componentes do sinal original que até então não eram percebidas em função da sua baixa potência puderam então ser destacadas. Até mesmo até componentes com baixa potência nas frequências mais baixas que eram possivelmente mascaradas por suas componentes vi-

zinhas de maior potência. Com este destaque destas componentes os sinais de vazamento e não vazamento puderam ser mais bem caracterizados. Também em função das pré-ênfases o modelo de predição pode estimar melhor as variações mais suaves do espectro uma vez que o espectro tornou-se mais plano.

Uma vez que com a utilização da pré-ênfase de primeira ordem não foi possível obter um espectro mais plano, foi então justificado a utilização da pré-ênfase de segunda ordem.

Ficou evidente que com a utilização das técnicas de pré-ênfase foi quando se obteve resultados relevantes para os modelos de predição. Sem a pré-ênfase os resultados foram não satisfatórios.

Quando da combinação dos parâmetros característicos individuais do sinal com os estatísticos, somente houve melhora quando o sinal foi pré-enfatizado, caso contrário o resultado foi pior que o obtido somente com os parâmetros individuais do sinal.

Por fim, a proposta de utilização de parâmetros característicos baseados no modelo de percepção auditiva humana, quando combinados com os parâmetros característicos centróide espectral, largura de espectro, fluxo espectral e ponto de *roll-off*, foi a combinação que apresentou os melhores resultados onde a taxa de acerto do classificador obtida foi em torno de 92%.

## 6.2 Propostas de Trabalhos Futuros

A sugestão principal de trabalhos futuros para este tema é a avaliação dos parâmetros propostos através de outros métodos de classificação, como redes neurais ou modelo de mistura de Gaussianas. Uma segunda sugestão consiste na aplicação de transformadas wavelets para a extração de parâmetros. Alguns artigos encontrados na literatura apresentaram esta proposta, mas os resultados ainda não foram animadores.

Uma das maiores dificuldades de desenvolver pesquisa sobre sinais acústicos de vazamento está na criação da base de dados de treinamento e teste. Não existe nenhuma base específica para este tipo de sinais de áudio e assim ocorreu nesta pesquisa a necessidade da criação de uma base a partir do zero. Em função de não se ter esta base pronta, torna-se mais complicada a avaliação das diversas propostas. Não há uma referência com a qual comparar. Por isto, uma das tarefas futuras mais interessantes envolve a criação desta base de dados de referência para avaliações de propostas futuras.

# Referências Bibliográficas

- [ABE04] ABENDI. Detecção de vazamentos não visíveis de líquidos sob pressão em tubulações enterradas procedimento PR-051, 2004. 4
- [ASZ<sup>+</sup>06] Changsheng Ai, Xuan Sun, Honghua Zhao, Rujian Ma e Xueren Dong. Pipeline damage and leak detection based on sound spectrum lpcc and hmm. *Proceedings of 6<sup>th</sup> International Conference on Intelligent System Design and Application*, páginas 829–833, 2006. 13
- [BR10] Liselene de Abreu Borges e Miguel Arjona Ramírez. Acoustic water leak detection. *Proceedings of ITS 2010 International Telecommunications Symposium.*, 1:1–3, 2010. 2, 37
- [Bro11] Mike Brookes. VOICEBOX: Speech processing toolbox for matlab, 2011. online web resource. 36
- [Bur98] Christopher J.C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2:121–167, 1998. 17
- [CV95] Corinna Cortes e Vladimir Vapnik. Support-vector networks. *Machine Learning*, 20:273–297, 1995. 10.1007/BF00994018. 17, 19
- [Ell05] Daniel P. W. Ellis. PLP and RASTA (and MFCC, and inversion) in Matlab, 2005. online web resource. 36
- [GST<sup>+</sup>99] T. R. Golub, D. K. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, J. P. Mesirov, H. Coller, M. L. Loh, J. R. Downing, M. A. Caligiuri e C. D. Bloomfield. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science*, 286:531–537, 1999. 20
- [Guy03] Isabelle Guyon. An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3:1157–1182, 2003. 19
- [Her90] Hynek Hermansky. Perceptual linear prediction (PLP) analysis for speech. *Journal of the Acoustical Society of America*, 87(4):1738–1752, 1990. 29
- [Hun00] Osama Hunaidi. Detecting leaks in water-distribution pipes. *Construction Technology Update No. 40*, 2000. 8
- [Ita75] Fumitada Itakura. Minimum prediction residual principle applied to speech recognition. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 23(1):67–72, 1975. 33
- [Mat11] Mathworks. MATLAB - the language of technical computing, 2011. 36

- [Mer09] J. Mercer. Functions of Positive and Negative Type, and their Connection with the Theory of Integral Equations. *Royal Society of London Philosophical Transactions Series A*, 209:415–446, 1909. 18
- [OSB99] Alan V. Oppenheim, Ronald W. Schafer e John R. Buck. *Discrete-Time Signal Processing*. Prentice-Hall, 2nd edição, 1999. 10, 16
- [PM06] John G. Proakis e Dimitris K. Manolakis. *Digital Signal Processing (4th Edition)*. Prentice Hall, March 2006. 28
- [PN03] Paulo Sérgio Padilha e Osvaldo Ioshio Niida. Formação, treinamento, qualificação e certificação de pessoal em estanqueidade - detecção de vazamentos não visíveis em redes pressurizadas. *Pan American Conference on Non Destructive Testing*, 2003. viii, 4, 5, 6
- [Ram06] Miguel Arjona Ramírez. Predição linear harmônica para processamento espectral e temporal de sinais de voz., 2006. 33
- [Rob99] Malcolm J. Robinson, David J. M.; Hawksford. Time-domain auditory model for the assessment of high-quality coded audio. Em *Audio Engineering Society Convention 107*, 9 1999. 29, 30
- [RS78] Lawrence R. Rabiner e Ronald W. Schafer. *Digital Processing of Speech Signals*. Prentice-Hall, 1st edição, 1978. 15, 24, 25
- [SAB05] SABESP. Perdas em sistemas de abastecimento de Água. <http://www.sabesp.com.br>, 2005. viii, 3
- [SAB10] SABESP. Informações sobre o controle de vazamentos da SABESP. <http://site.sabesp.com.br/>, Maio 2010. Último acesso em 9/5/2010. 1
- [Sav95] Michael Savic. Detection of leaks in pipelines. 9 out. 1992, 16 maio 1995, 1995. 13
- [SM07] Toshitaka Sato e Akira Mita. Leak detection using the pattern of sound signals in water supply system. *Proceedings of SPIE*, 6529:65292K.1–65292K.9, 2007. 13, 35, 51, 56
- [SR 10] SR Detecção de Vazamentos Ltda. Pesquisa e detecção de vazamentos não visíveis em vias públicas. <http://www.srdetecvaz.com.br/vias.htm>, Novembro 2010. Último acesso em 18/11/2010. viii, 5, 6, 7
- [SS97] E. Scheirer e M. Slaney. Construction and evaluation of a robust multifeature speech/music discriminator. Em *Acoustics, Speech, and Signal Processing, 1997. ICASSP-97., 1997 IEEE International Conference on*, volume 2, páginas 1331–1334 vol.2, apr 1997. 25, 27
- [TC02] G. Tzanetakis e P. Cook. Musical genre classification of audio signals. *IEEE Transactions on Speech and Audio Processing*, 10(5):293–302, Julho 2002. 25, 27
- [TM08] Yuriko Terao e Akira Mita. Robust water leakage detection approach using the sound signals and pattern recognition. *Proceedings SPIE*, 6932:69322D.1–69322D.9, 2008. 2, 13

- [Wee10] David Weenink. Speech signal processing with praat. <http://www.fon.hum.uva.nl/david/sspbook/sspbook.pdf>, Novembro 2010. Último acesso em 15/01/2011. viii, 17
- [YWL08] Jin Yang, Yumei Wen e Ping Li. Leak acoustic detection in water distribution pipelines. *Proceedings of the 7<sup>th</sup> World Congress on Intelligent Control and Automation*, páginas 3057–3061, 2008. 13
- [Zan09] Nilton Zaniboni. Equipamentos e metodologias para o controle e redução de perdas reais em sistemas de abastecimento de água. Dissertação de Mestrado, Escola Politécnica, Universidade de São Paulo, Brazil, Jun 2009. 7