



Dissertação de Mestrado

**Aplicação da Transformada Wavelet em
Análise de Texturas de Gráficos de
Recorrência para Detecção de Patologias
Laríngeas**

Taciana Araújo de Souza

Instituto Federal de Educação, Ciência e Tecnologia da Paraíba
Programa de Pós-Graduação em Engenharia Elétrica

Aplicação da Transformada Wavelet em Análise de Texturas de Gráficos de Recorrência para Detecção de Patologias Laríngeas

Taciana Araújo de Souza

Dissertação de Mestrado apresentada à Coordenação do Programa de Pós Graduação em Engenharia Elétrica do Instituto Federal de Educação, Ciência e Tecnologia da Paraíba como requisito necessário para obtenção do grau de Mestre em Ciências no Domínio da Engenharia Elétrica.

Área de Concentração: Processamento de Sinais.

Suzete Élide Nóbrega Correia, D.Sc.
Orientadora

Silvana Luciene do Nascimento Cunha Costa, D.Sc.
Co-Orientadora

João Pessoa, Paraíba, Brasil
13 de Abril de 2015
©Taciana Araújo de Souza

Dados Internacionais de Catalogação na Publicação – CIP
Biblioteca do IFPB, *campus* João Pessoa

S729a

Souza, Taciana Araújo de.

Aplicação da transformada wavelet em análise de texturas de gráficos de recorrência para detecção de patologias laringeas/ Taciana Araújo de souza. – 2015.

79 f. : il.

Dissertação (Engenharia Elétrica) – Instituto Federal de Educação, Ciência e Tecnologia da Paraíba – IFPB, Campus João Pessoa.

Orientadora: Prof.^a Dr.^a Suzete Élide N. Correia .
Co-orientadora: Silvana Luciene N. Cunha Costa.

1. Sinal de voz. 2. Processamento digital. 3.
Transformada wavelet. I. Título.

CDU 621.391

Instituto Federal de Educação, Ciência e Tecnologia da Paraíba
Programa de Pós Graduação em Engenharia Elétrica

**Aplicação da Transformada Wavelet em Análise de Texturas de
Gráficos de Recorrência para Detecção de Patologias
Laringeas**

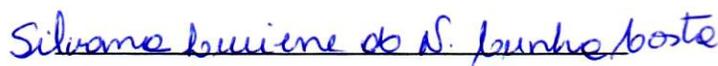
Taciana Araújo de Souza

Dissertação de Mestrado defendida e aprovada em 13/04/15

Banca examinadora:



Prof. Suzete Élide Nóbrega Correia, D.Sc.
Orientadora (IFPB)



Prof. Silvana Luciene do Nascimento Cunha Costa, D.Sc.
Co-Orientadora (IFPB)



Prof. Carlos Danilo Miranda Regis, D.Sc.
Membro da Banca (IFPB)



Prof. José Josemar de Oliveira Júnior, D.Sc.
Membro da Banca (UFRN)



Prof. Washington César de Almeida Costa, D.Sc.
Membro da Banca (IFPB)

Aos meus pais Maria do Socorro Araújo e Gabriel Francisco de Souza.

Agradecimentos

- ▣ Primeiramente agradeço a Deus pelo dom da vida, e aos meus pais, Maria do Socorro Araújo e Gabriel Francisco de Souza, aos quais devo minha vida, minha formação e grande parte de meus princípios;
- ▣ À minha orientadora, Suzete Élide Nóbrega Correia, que foi incumbida da árdua tarefa de me conduzir nesse trabalho, a quem devo muito pela paciência, pela compreensão diante das minhas limitações e dificuldades, pela disponibilidade em me orientar e, além de tudo, pelo exemplo de professora e ser humano, a quem respeito e admiro;
- ▣ À minha co-orientadora, Silvana Luciene do Nascimento Cunha Costa, pela parceria com a professora Suzete nesta orientação, a quem também devo muito, pela paciência e compreensão em diversos momentos, por todos os conselhos, por todo o conhecimento compartilhado e pelo apoio nos momentos de dificuldade durante esta caminhada;
- ▣ Aos professores examinadores, que fizeram parte das minhas bancas de qualificação e/ou de defesa da dissertação, pelas valiosas sugestões, que contribuíram para o enriquecimento deste trabalho de pesquisa;
- ▣ Ao professor Luiz Guedes Caldeira, por todo o apoio, por ter acreditado e me feito acreditar que seria possível chegar até aqui, por todo o conhecimento compartilhado desde a graduação, pelos conselhos e pelas palavras de incentivo durante esta caminhada;
- ▣ A todos os colegas da turma pioneira da Pós-graduação em Engenharia Elétrica do IFPB, com os quais compartilhei durante esse período bons momentos de interação, além do conhecimento. Em especial ao meu amigo Vinícius Vieira, pela amizade e pelas valiosas contribuições na execução deste trabalho. E também ao colega Micael Souza, pelas contribuições no desenvolvimento desta pesquisa;
- ▣ À minha querida avó Antônia Lopes de Souza, que me ensinou o valor da humanidade, respeito, e do carinho, que compreendeu mesmo com saudade minha ausência em alguns momentos durante esta caminhada;
- ▣ Aos meus irmãos Alessandra, Arthur, Ellis e Mateus, pelo amor e carinho, me apoiando em todas as etapas da minha formação;
- ▣ Aos professores da Pós-graduação em Engenharia Elétrica do IFPB, em especial ao professor DSc. Neilor César dos Santos, pelas lições em sala de aula e, principalmente, pelo exemplo de dedicação e compromisso com o ensino;
- ▣ A todas as pessoas que, de alguma forma, contribuíram para a minha formação acadêmica e que compartilham da minha alegria em concluir mais uma etapa.

*“A Matemática é só um monte de números e símbolos sem significado se não humanizamos a situação em que ela é aplicada.”
(Leandro Teixeira)*

Resumo

A análise acústica do sinal de voz, devido à sua natureza não invasiva e ao baixo custo, tem se mostrado uma eficiente ferramenta para auxílio ao diagnóstico das desordens vocais provocadas por patologias na laringe. Os gráficos apresentam padrões de larga e pequena escala, cujas variações em sua textura representam o comportamento do sinal de voz, proporcionando informações acerca do estado de normalidade ou de alteração na qualidade vocal. Os padrões de pequena escala podem ser vistos como características de textura e servem como base para uma análise quantitativa dos gráficos de recorrência. Técnicas de Processamento Digital de Imagens são empregadas para a análise da textura contida nos gráficos de recorrência, baseada na transformada wavelet bidimensional. A fim de discriminar sinais saudáveis de sinais patológicos, são extraídos diversos descritores de texturas dos coeficientes de cada sub-banda obtida pela decomposição wavelet bidimensional. Nesta pesquisa, duas abordagens foram aplicadas, as quais se diferenciam pela forma de extração dos padrões representativos dos sinais: extração dos descritores de textura diretamente das sub-bandas da transformada wavelet; e extração dos descritores de Haralick, a partir da matriz de co-ocorrência. Os sinais de voz foram classificados como saudáveis ou patológicos, como também foi realizada a discriminação entre patologias. Paralisia, edema de Reinke e nódulos nas pregas vocais foram as patologias laríngeas consideradas na pesquisa. Os melhores resultados foram obtidos com os descritores de Haralick, empregando redes neurais MLP (*Multilayer Perceptron*) na classificação, em conjunto com o algoritmo de otimização por enxame de partículas PSO (*Particle Swarm Optimization*) empregado na seleção das características mais representativas. O sistema proposto melhorou significativamente a acurácia na discriminação entre patologias, com resultados superiores aos encontrados na literatura, que empregam a análise de recorrência.

Palavras-Chave: Gráficos de recorrência; Análise de textura; Transformada wavelet; Otimização por enxame de partículas.

Abstract

The acoustic analysis of the speech signal due to its noninvasive nature and its low cost, has proved an effective tool to aid in the diagnosis of vocal disorders caused by diseases in the larynx. In this work, a method based on analysis of texture of recurrence plots is employed, that characterize the dynamics of the speech production system. The recurrence plots have large and small-scale patterns, whose texture variations represent the voice signal behavior, providing information about the state of normalcy or changes in voice quality. Small-scale patterns can be seen as texture features and serve as the basis for a quantitative analysis of the recurrence plots. Digital images processing techniques are used for texture analysis contained in the recurrence plots, based on two-dimensional wavelet transform. In order to discriminate between healthy and pathological signals, several texture descriptors are extracted from each subband coefficients obtained by two-dimensional wavelet decomposition. In this study, two approaches were applied, which differ in the way of extraction of representative patterns of signals: extraction of texture descriptors directly from the sub-bands of wavelet transform; and extracting the Haralick descriptors from the co-occurrence matrix. The speech signals are classified as healthy or pathological, and a discrimination pathologies was also performed. Paralysis, Reinke's edema and nodules on the vocal folds were the laryngeal pathologies considered in the survey. The best results were obtained with Haralick descriptors, using Multilayer Perceptron neural network, in the classification, with the optimization algorithm (Particle Swarm Optimization) used in the selection of the most representative features. The proposed system improved the accuracy significantly in discriminating laryngeal diseases, with results superior to those found in the literature, employing the recurrence analysis.

Key-Words: Recurrence plots; Texture analysis; Wavelet transform; Particle swarm optimization.

Lista de Siglas e Abreviaturas

MLP – *Multilayer Perceptron*

PSO – *Particle Swarm Optimization*

PDS – *Processamento Digital de Sinais*

MNL – *Medidas de análise não linear*

MRQ – *Medidas de quantificação da recorrência*

TWP – *Transformada Wavelet Packet*

SVM – *Support Vector Machine*

RP – *Recurrence Plot*

PDI – *Processamento Digital de Imagens*

DWT – *Discrete Wavelet Transform*

GLCM – *Gray-Level Co-occurrence matrix*

MEEI – *Massachusetts Eye and Ear Infirmary*

VRA – *Visual Recurrence Analysis*

RNAs – *Redes Neurais Artificiais*

RQA – *Recurrence Quantification Analysis*

LPC – *Linear Prediction Coding*

LBP – *Local Binary Pattern*

Lista de Símbolos

- $x(t)$ – série temporal
 t – tempo
 m – dimensão de imersão
 x_i – vetores m -dimensionais de uma série temporal
 t – tempo
 N – número de estados de um sistema dinâmico
 T – comprimento da série temporal
 τ – passo de reconstrução
 ε – raio de vizinhança
 θ – função degrau unitário
 d_0 – dimensão fractal do atrator
 $f(t)$ – função real
 $\psi(t)$ – função wavelet mãe
 $\psi_{a,b}(t)$ – funções de expansão wavelets
 $\psi_{j,k}(t)$ – funções de expansão na escala j e translação k
 $DWT_f(j, k)$ – transformada wavelet discreta
 $\phi(t)$ – função escalonamento
 \langle, \rangle – produto interno
 $L^2(\mathbb{R})$ – espaço das funções quadráticas integráveis
 V_j – subespaço de aproximação na resolução j
 W_j – complemento ortogonal de V_j em V_{j-1}
 $h(n)$ – resposta ao impulso do filtro passa-baixa
 $g(n)$ – resposta ao impulso do filtro passa-alta
 $\phi(x, y)$ – função escalonamento bidimensional
 $\psi^1(x, y), \psi^2(x, y), \psi^3(x, y)$ – wavelets bidimensionais
 I_{j-1} – imagem na resolução $j - 1$
 Sb – sub-banda wavelet
 $Sb(x, y)$ – coordenada (x, y) na sub-banda Sb
 Sb_1^j – sub-banda dos coeficientes de aproximação na resolução j
 Sb_2^j – sub-banda dos coeficientes de detalhes horizontais na resolução j
 Sb_3^j – sub-banda dos coeficientes de detalhes verticais na resolução j
 Sb_4^j – sub-banda dos coeficientes de detalhes diagonais na resolução j
 \oplus – soma direta

- \otimes – operação de convolução
- N_{Sb} – número de coeficientes da sub-banda
- r_n – ordem da norma
- r_m – ordem do momento
- μ – média
- Energia $_{Sb}$ – energia da sub-banda Sb
- Desvio Médio $_{Sb}$ – desvio médio dos coeficientes da sub-banda Sb
- Entropia $_{Sb}$ – entropia dos coeficientes da sub-banda Sb
- Média $_{Sb}$ – média dos coeficientes da sub-banda Sb
- Norma $_{Sb}$ – norma dos coeficientes da sub-banda Sb
- Momento $_{Sb}^r$ – momento dos coeficientes da sub-banda Sb
- Suavidade $_{Sb}$ – suavidade dos coeficientes da sub-banda Sb
- Resíduo Médio $_{Sb}$ – resíduo médio dos coeficientes da sub-banda Sb
- d – distância
- $P(i, j, d, \theta)$ – frequência relativa do *pixel* de referência i com relação a um *pixel* vizinho j , separados por uma distância d na direção θ
- Z_n – níveis de cinza
- $p(Sb(x, y))$ – probabilidade de ocorrer o coeficiente $Sb(x, y)$
- σ^2 – variância dos coeficientes da sub-banda
- H_k – máscaras de Haar
- I^{j+1} – imagem na resolução $j + 1$
- Ng – número de níveis de cinza distintos contidos na imagem
- $p(i, j)$ – (i, j) th é a entrada normalizada da matriz de co-ocorrência
- $p_x(i)$ – i th entrada na matriz de probabilidade marginal obtida pela soma das linhas de $p(i, j)$
- $p_y(j)$ – j th entrada na matriz de probabilidade marginal obtida pela soma das colunas de $p(i, j)$
- HX – entropia de p_x
- HY – entropia de p_y
- μ_x, μ_y – médias de p_x e p_y
- σ_x, σ_y – desvios padrão de p_x e p_y
- $v_i(t)$ – velocidade da partícula i no instante t
- $x_i(t)$ – posição da partícula i
- p_B – melhor posição da partícula
- g_B – melhor posição já alcançada pelos conhecidos da partícula
- $\#(i, j)$ – número de ocorrências do *pixels* i vizinho ao *pixel* j
- N_e – número de neurônios na camada de entrada da rede neural
- N_i – número de neurônios na camada intermediária da rede neural
- N_s – número de neurônios na camada de saída da rede neural

Lista de Figuras

2.1	Esquema dos sistemas envolvidos na produção da fala dos sinais de voz.	6
2.2	Vista anterior e posterior dos principais componentes da laringe.	7
2.3	Nódulos nas pregas vocais.	8
2.4	Edema de Reinke nas pregas vocais	9
2.5	Paralisia unilateral das pregas vocais	9
3.1	Construção de um gráfico de recorrência.	12
3.2	(a) Sinal de voz saudável; (b) Sinal de voz afetado por paralisia; (c) Sinal de voz afetado por edema; e (d) Sinal de voz afetado por nódulo.	13
3.3	Wavelet de Haar.	14
3.4	Representação da relação entre a escala e os subespaços de funções wavelets.	16
3.5	Representação das sub-bandas obtidas pela decomposição wavelet com (a) um nível de decomposição; (b) dois níveis de decomposição.	18
3.6	Representação da decomposição wavelet utilizando de filtros.	18
3.7	Direções utilizadas para o cálculo da matriz de co-ocorrência.	20
3.8	Matriz correspondente a uma imagem com 4 níveis de cinza.	20
3.9	Matriz de co-ocorrência.	21
3.10	Matrizes de co-ocorrência da imagem em níveis de cinza ilustrada pela Figura 3.8 para $d =$ 1: (a) $\theta = 0^\circ$; (b) $\theta = 45^\circ$; (c) $\theta = 90^\circ$; (d) $\theta = 135^\circ$	21
4.1	Diagrama de blocos.	28
4.2	Máscaras de Haar.	28
4.3	Decomposição wavelet de um gráfico de recorrência de um sinal de voz utilizando a wavelet de Haar, com um nível de decomposição.	29
4.4	Modelo de neurônio.	31
4.5	Modelo de uma rede neural MLP.	31
5.1	Acurácia média (%) utilizando apenas redes neurais MLP e utilizando PSO com MLP.	37
5.2	Acurácia média (%) obtidas nos trabalhos de Costa e Vieira em comparação com os melhores resultados obtidos nesta pesquisa utilizando PSO com MLP.	38
5.3	Gráficos de recorrência de sinais de vozes saudáveis: (a) feminino, 26 anos; (b) masculino, 32 anos; (c) feminino, 37 anos; (d) masculino, 30 anos.	39

-
- 5.4 Gráficos de Recorrência de vozes afetadas por edema de Reinke: (a) bilateral, feminino; (b) bilateral, feminino; (c) bilateral, feminino; (d) indefinido, feminino. 40
- 5.5 Gráficos de recorrência de vozes afetados por nódulo: (a) indefinido, feminino; (b) unilateral esquerdo, masculino; (c) bilateral, feminino; (d) bilateral, feminino; 40
- 5.6 Gráficos de recorrência de sinais de vozes afetadas por paralisia: (a) unilateral esquerdo, feminino; (b) unilateral esquerdo, feminino; (c) unilateral esquerdo, masculino; (d) unilateral esquerdo, feminino. 41
- 5.7 Gráficos de recorrência: (a) sinal de voz saudável; (b) sinal de voz afetado por edema de Reinke; (c) sinal de voz afetado por nódulo; e (d) sinal de voz afetado por paralisia. 41

Lista de Tabelas

5.1	Matriz de confusão	34
5.2	Relação entre os casos de classificação entre patologias e valores de sensibilidade e especificidade.	34
5.3	Acurácia média (%) para classificação utilizando descritores extraídos das sub-bandas.	35
5.4	Resumo dos melhores resultados para cada caso de classificação utilizando 13 descritores de Haralick com a rede MLP.	37
5.5	Resultados com MLP X PSO com MLP.	37
A.1	Classificação Saudável X Patológico	49
A.2	Classificação Saudável X Paralisia	49
A.3	Classificação Saudável X Edema	49
A.4	Classificação Saudável X Nódulo	50
A.5	Classificação Edema X Nódulo	50
A.6	Classificação Edema X Paralisia	50
A.7	Classificação Paralisia X Nódulo	50
B.1	Classificação Saudável X Patológico utilizando 13 descritores com a matriz de co-ocorrência na direção $\theta = 0^\circ$	51
B.2	Classificação Saudável X Patológico utilizando 13 descritores com a matriz de co-ocorrência na direção $\theta = 45^\circ$	51
B.3	Classificação Saudável X Patológico utilizando 13 descritores com a matriz de co-ocorrência na direção $\theta = 90^\circ$	52
B.4	Classificação Saudável X Patológico utilizando 13 descritores com a matriz de co-ocorrência na direção $\theta = 135^\circ$	52
B.5	Classificação Saudável X Paralisia utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 0^\circ$	52
B.6	Classificação Saudável X Paralisia utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 45^\circ$	52
B.7	Classificação Saudável X Paralisia utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 90^\circ$	52
B.8	Classificação Saudável X Paralisia utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 135^\circ$	52

B.9	Classificação Saudável X Edema utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 0^\circ$	53
B.10	Classificação Saudável X Edema utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 45^\circ$	53
B.11	Classificação Saudável X Edema utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 90^\circ$	53
B.12	Classificação Saudável X Edema utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 135^\circ$	53
B.13	Classificação Saudável X Nódulo utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 0^\circ$	53
B.14	Classificação Saudável X Nódulo utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 45^\circ$	53
B.15	Classificação Saudável X Nódulo utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 90^\circ$	54
B.16	Classificação Saudável X Nódulo utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 135^\circ$	54
B.17	Classificação Edema X Nódulo utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 0^\circ$	54
B.18	Classificação Edema X Nódulo utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 45^\circ$	54
B.19	Classificação Edema X Nódulo utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 90^\circ$	55
B.20	Classificação Edema X Nódulo utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 135^\circ$	55
B.21	Classificação Edema X Paralisia utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 0^\circ$	56
B.22	Classificação Edema X Paralisia utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 45^\circ$	56
B.23	Classificação Edema X Paralisia utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 90^\circ$	56
B.24	Classificação Edema X Paralisia utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 135^\circ$	57
B.25	Classificação Paralisia X Nódulo utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 0^\circ$	57
B.26	Classificação Paralisia X Nódulo utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 45^\circ$	57
B.27	Classificação Paralisia X Nódulo utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 90^\circ$	58
B.28	Classificação Paralisia X Nódulo utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 135^\circ$	58
C.1	Classificação SDLxPTL	59

C.2	Classificação SDLxPRL	60
C.3	Classificação SDLxEDM	60
C.4	Classificação SDLxNDL	61
C.5	Classificação EDMxNDL	61
C.6	Classificação EDMxPRL	62
C.7	Classificação PRLxNDL	62

Sumário

1	Introdução	1
1.1	Formulação do Problema	1
1.2	Justificativa	1
1.3	Estado da arte	3
1.4	Objetivos	5
1.4.1	Geral	5
1.4.2	Específicos	5
1.5	Organização do Trabalho	5
2	Mecanismo de Produção Vocal	6
2.1	Produção da Fala	6
2.2	Laringe	7
2.3	Patologias da Laringe	8
2.3.1	Nódulo	8
2.3.2	Edema de Reinke	8
2.3.3	Paralisia das pregas vocais	9
3	Fundamentação Teórica	10
3.1	Gráficos de Recorrência	10
3.2	Wavelets	13
3.2.1	Análise Multirresolucional	15
3.2.2	Transformada Wavelet Bidimensional	17
3.3	Texturas	18
3.3.1	Medidas extraídas das sub-bandas Wavelets	19
3.3.2	Matriz de co-ocorrência (<i>Gray-Level Co-occurrence matrix - GLCM</i>)	20
3.3.3	Descritores de Haralick	21
3.3.4	Otimização por enxame de partículas (<i>Particle Swarm Optimization - PSO</i>)	24
4	Materiais e Métodos	27
4.1	Base de dados	27
4.2	Metodologia	27
4.2.1	Gráficos de Recorrência	28
4.2.2	Transformada Wavelet Bidimensional	28

4.2.3	Extração de Características	29
4.3	Seleção de Características	30
4.3.1	Otimização por enxame de partículas (<i>Particle Swarm Optimization - PSO</i>)	30
4.4	Classificação	30
4.4.1	Rede Neural MLP	30
5	Resultados e Discussão	33
5.1	Método das sub-bandas	34
5.2	Método dos descritores	36
6	Considerações Finais	42
6.1	Contribuições da pesquisa	43
6.2	Sugestões para Trabalhos Futuros	43
	Referências Bibliográficas	47
	APÊNDICES	48
A	Método das sub-bandas	49
B	Método dos descritores	51
C	Características selecionadas pelo PSO	59

1.1 – Formulação do Problema

A fala é um dos meios de comunicação mais utilizados pelos seres humanos. O funcionamento da voz pode ter implicações na vida social e profissional de um indivíduo, visto que a voz humana é uma ferramenta importante de comunicação [1]. Os diferentes sons produzidos pelo sistema de produção da fala são resultado da ação de um conjunto de órgãos, tais como: os pulmões, a traqueia, a laringe, a faringe e as cavidades oral e nasal. Portanto, a voz pode sofrer alterações devido a patologias em algum desses órgãos e, particularmente, da laringe.

A laringe é responsável por fornecer o sinal acústico para a fala e tem funções biológicas que protegem a via aérea [2]. Algumas patologias capazes de alterar a voz são nódulos, pólipos, cistos, edemas, que podem ser decorrentes do uso incorreto da fala ou hábitos não saudáveis, e paralisia das pregas vocais, que podem ter causas neurológicas [2-4].

A fim de diagnosticar precocemente tais patologias, médicos otorrinolaringologistas realizam em seus pacientes, além da audição da voz, exames tais como: videolaringoscopia ou videoestroboscopia, considerados invasivos. Além desses exames existem métodos capazes de auxiliar na detecção de tais patologias de forma menos invasiva, visto que desordens na voz são perceptíveis por meio da análise do sinal acústico correspondente e, dessa forma, é possível diagnosticar, de forma objetiva, patologias que provocam alterações vocais.

1.2 – Justificativa

A análise acústica, que emprega técnicas de Processamento Digital de Sinais (PDS), é considerada não invasiva e de baixo custo por utilizar computadores pessoais e microfone para realizar a aquisição e análise do sinal. Esta técnica tem se mostrado uma ferramenta eficiente de auxílio ao diagnóstico das desordens provocadas por patologias na laringe e, pode complementar os métodos baseados na inspeção direta das pregas vocais, diminuindo a regularidade dos exames invasivos [5]. Portanto, pode ser empregada como uma ferramenta de apoio ao médico otorrinolaringologista ou ao fonoaudiólogo na avaliação de disfonias e acompanhamento terapêutico.

Métodos baseados na análise do sinal de voz são utilizados na classificação de doenças da voz e no acompanhamento da evolução de tratamentos, por apresentar as seguintes vantagens em relação aos métodos tradicionais: propiciar um exame mais confortável ao paciente; fornecer uma avaliação

quantitativa da doença; e possibilitar o desenvolvimento de sistemas automáticos de auxílio ao diagnóstico por computador por um baixo custo [3].

A análise acústica dos sinais de voz pode ser feita baseada na análise linear do modelo de produção de fala ou em técnicas de análise dinâmica não linear. Na análise linear, em que é considerado um modelo linear de produção de fala, é possível representar, de forma geral, o sistema de produção de voz como um sistema fonte-filtro, no qual as pregas vocais são consideradas a fonte sonora, o trato vocal é considerado o filtro e o sinal resultante é o sinal acústico da voz [1, 2, 4, 5]. Na análise dinâmica não linear um sinal de voz pode ser considerado como uma observação do sistema de produção da fala, que pode ser usada para revelar e modelar sua dinâmica usando técnicas não lineares [6, 7].

Dentre os métodos de análise não linear utilizados para estimar medidas e propriedades não lineares de processos naturais, os gráficos de recorrência têm mostrado bons resultados na caracterização de sistemas dinâmicos. Tais gráficos foram desenvolvidos por Eckmann *et al.* [8] baseados na ideia de recorrência, proposta pelo físico e filósofo francês J. H. Poincaré (1890) e constituem um método que permite visualizar a dinâmica de sistemas recorrentes [9]. Uma das vantagens dos gráficos de recorrência é que permitem visualizar o comportamento recorrente de sistemas dinâmicos independentemente da dimensionalidade dos mesmos.

Tais gráficos são representados por matrizes bidimensionais que correspondem a imagens digitais com dois níveis de cinza (preto e branco), nas quais os *pixels* podem formar estruturas típicas, como: pontos isolados, linhas diagonais, verticais e horizontais. Essas estruturas caracterizam os sistemas analisados e permitem a abordagem dos gráficos de recorrência do ponto de vista de pequena escala como texturas [10]. Os padrões de pequena escala podem ser vistos como características de textura e são a base de uma análise quantitativa dos gráficos de recorrência proposta por Marwan, através das medidas de quantificação da recorrência [10]. Tais medidas foram desenvolvidas por Zbilut e Webber (1994) e são baseadas na abordagem estatística de textura. No entanto, texturas podem ser analisadas de acordo com uma abordagem estatística, estrutural ou espectral.

Vários métodos têm sido propostos em busca de uma descrição de textura eficiente. As três abordagens comumente empregadas são a estatística, estrutural e espectral. Na abordagem estatística, a análise de textura é realizada utilizando medidas estatísticas baseadas na distribuição dos níveis de cinza. A abordagem estrutural é baseada na ideia de que texturas são compostas por primitivas, que são estruturas geométricas básicas que formam a textura, como linhas paralelas regularmente espaçadas. A abordagem espectral representa a imagem em um espaço de coordenadas cuja interpretação está relacionada com as características da textura (frequência ou tamanho), obtidas através de transformadas, como Fourier, Gabor e wavelet [11, 12].

A transformada wavelet possui propriedades interessantes para a análise de textura, como a representação dos sinais tanto no domínio do tempo quanto da frequência; a possibilidade de se usar a análise multirresolucional, de modo que as texturas possam ser analisadas na escala mais adequada; e o fato de que existem diversas famílias wavelets que podem ser selecionadas de acordo com a aplicação [9]. As características extraídas a partir dos coeficientes da transformada wavelet são numerosas incluindo a energia, medidas estatísticas como média e desvio padrão, medidas obtidas a partir do histograma e de matrizes de co-ocorrência [13].

A transformada wavelet bidimensional decompõe a imagem em quatro sub-bandas de coeficientes, uma contendo uma aproximação da imagem, e três sub-bandas contendo detalhes nas direções: horizontal, vertical e diagonal. Texturas presentes em uma imagem podem ser representadas a partir de medidas extraídas das sub-bandas de coeficientes, uma vez que estes evidenciam padrões texturais presentes na imagem. Para texturas suaves, os coeficientes possuem valores baixos, quando há transições, os coeficientes tendem a possuir maior valor, enfatizando as mudanças [9].

Os gráficos de recorrência de um sistema dinâmico podem apresentar padrões em sua estrutura que caracterizam a evolução do sistema. Analisando sinais de voz do ponto de vista de sistemas dinâmicos é possível observar que os gráficos de recorrências de sinais saudáveis apresentam uma maior presença de linhas diagonais, enquanto que os sinais patológicos podem apresentar pontos isolados e tendem a formar estruturas verticais e horizontais. A transformada wavelet bidimensional permite a extração dos traços diagonais, horizontais e verticais de uma imagem. Portanto, esta técnica pode ser empregada para a análise dos gráficos de recorrência.

Neste trabalho, uma nova abordagem baseada em análise de texturas, através do Processamento Digital de Imagens (PDI), é proposta para avaliar os gráficos de recorrência na detecção das patologias na laringe. A transformada wavelet é aplicada aos gráficos de recorrência dos sinais de voz, em um nível de resolução, empregando a wavelet de Haar. A escolha desta família wavelet deve-se às características dos gráficos de recorrência, que apresentam estruturas horizontais/verticais que se assemelham a forma de onda desta wavelet, considerada a mais simples. Além disso, outro fator relevante é a facilidade de implementação, visto que esta pode ser realizada através de convolução da imagem dos gráficos de recorrência com as máscaras de Haar, descritas com detalhes no Capítulo 4.

Esta pesquisa foi dividida em duas abordagens, utilizando dois métodos para extração das características de textura dos gráficos de recorrência. No primeiro método, denominado método das sub-bandas, foram aplicados 12 descritores de textura sobre cada sub-banda obtida, totalizando para cada sinal de voz 48 características a serem analisadas. No segundo método, denominado método dos descritores, foram extraídos 13 descritores de Haralick para cada sub-banda obtida, totalizando 52 características a serem analisadas. Após a extração de características, em cada caso, foi realizada a classificação utilizando uma rede neural *Multilayer Perceptron* (MLP).

No método dos descritores, além da extração das 52 características dadas pelos descritores de Haralick, foi utilizado um algoritmo de otimização *Particle Swarm Optimization* (PSO), com o objetivo de selecionar o melhor conjunto de características, que resultam nos melhores casos de classificação entre sinais saudáveis e sinais patológicos. Logo, é apresentada uma nova proposta para a análise dos gráficos de recorrência, que pode ser utilizada em diversas situações modeladas por sistemas dinâmicos.

1.3 – Estado da arte

Vieira *et al.* [14] analisam a discriminação de sinais de voz saudáveis e afetados por patologias laringeas utilizando medidas de quantificação de recorrências e redes neurais *Multilayer Perceptron* (MLP). Algumas medidas de quantificação de recorrência utilizadas para analisar o sinal de voz foram Taxa de Recorrência, Determinismo, Comprimento máximo das estruturas diagonais, Entropia de Shannon, Laminaridade, Tempo de Permanência, Comprimento máximo das estruturas verticais, Transitividade e

Divergência. Utilizando as medidas de forma combinada na detecção da presença de patologias chegou-se a uma precisão de 90%. Dentre as medidas avaliadas, a que mais se destacou na discriminação entre sinais saudáveis e sinais patológicos (edemas, nódulos e paralisia) foi o comprimento médio das linhas diagonais. Na classificação entre patologias o melhor resultado foi obtido utilizando o tempo de recorrência do tipo 2.

Costa [6], em sua tese de doutorado, realizou a caracterização e classificação de sinais de vozes saudáveis e vozes afetadas por patologias laringeas por meio da análise dinâmica não linear e também por meio da análise de quantificação de recorrência. A caracterização consiste em investigar o potencial de cada característica para discriminar os sinais de voz considerados. A análise dos sinais de voz é baseada na combinação de características dos conjuntos de medidas de análise não linear (MNL) e das medidas de quantização de recorrência (MRQ). No processo de classificação foi utilizada a técnica de análise discriminante com as funções linear ou quadrática, com validação cruzada, e as médias da taxa de acurácia obtidas variaram nos intervalos de confiança: (95,44% ;100%) para a classificação entre vozes saudáveis e patológicas; 100% para a classificação entre vozes saudáveis e afetadas por paralisia e (94,75% ;100%) entre vozes saudáveis e afetadas por edema, e entre saudáveis e nódulos. Portanto, os resultados mostram que o método utilizado é eficiente para a discriminação entre sinais de vozes saudáveis e patológicas e, portanto, pode ser utilizado para desenvolver uma ferramenta computacional para auxiliar o diagnóstico de patologias laringeas.

Carvalho [3], em sua dissertação de mestrado, utiliza a transformada wavelet para caracterizar os sinais de voz saudáveis e patológicos. Os dados utilizados neste trabalho consistem de 60 amostras de sinais de vozes, contendo vozes saudáveis e três patologias laringeas (nódulo vocal, edema de Reinke e disфонia neurológica). As características extraídas a partir dos coeficientes obtidos pela decomposição wavelet foram: energia wavelet, entropia wavelet e entropia wavelet relativa. Os resultados mostraram que a técnica proposta permite discriminar vozes saudáveis e patológicas com taxas de acerto acima de 90% e, além disso, permite classificar as amostras de voz em um tipo de patologia da laringe com desempenho superior a 86% incluindo a classe de disфонia neurológica. Portanto, o uso de características extraídas pela transformada wavelet mostrou-se eficiente na discriminação entre vozes saudáveis e patológicas.

Soares [15], em sua tese de doutorado, apresenta uma nova metodologia inteligente para análise e classificação de imagens de câncer de pele, baseada em técnicas de processamento digital de imagens para extração de características de cor, forma e textura, utilizando a Transformada Wavelet Packet (TWP) e técnicas de aprendizado de máquina denominada Máquina de Vetor Suporte (*Support Vector Machine - SVM*). Para o banco de dados utilizado neste trabalho, os resultados apresentaram um bom desempenho obtendo um acerto global de 92,73% para melanoma, e 86% para lesões não-melanoma e benigna.

Silva [9], em sua dissertação de mestrado, utiliza descritores de textura sobre os coeficientes obtidos pela transformada wavelet bidimensional para extrair características das imagens. Para isto, um conjunto de 36 bases wavelets separáveis foi analisado com objetivo de selecionar, entre as bases, aquelas que melhor responderam à tarefa de caracterização de textura. Dentre os descritores de textura analisados destacaram-se: energia, desvio médio, entropia, resíduo médio, momento central de ordem 2, normas de ordem 1 e 2, desvio padrão, variância e suavidade. Neste trabalho, as características obtidas foram utilizadas a fim de segmentar imagens e os resultados obtidos pelo método proposto foram considerados satisfatórios.

1.4 – Objetivos

1.4.1 – Geral

Desenvolver um método para a detecção automática de patologias laringeas empregando análise de texturas extraídas a partir da transformada wavelet bidimensional em gráficos de recorrência dos sinais de voz.

1.4.2 – Específicos

- ▣ Relacionar os aspectos teóricos referentes aos gráficos de recorrência, à transformada wavelet bidimensional e à análise de textura.
- ▣ Gerar as imagens dos gráficos de recorrência dos sinais avaliados e aplicar a transformada wavelet bidimensional para obter os coeficientes de aproximação e detalhes da imagem;
- ▣ Analisar os coeficientes obtidos para cada sub-banda através de descritores de textura a fim de classificar um sinal de voz em saudável ou patológico, bem como a discriminação entre patologias.
- ▣ Discriminar os sinais em saudáveis ou patológicos empregando algoritmos de classificação tais como redes neurais MLP.

1.5 – Organização do Trabalho

Os demais capítulos estão organizados conforme está descrito a seguir. O Capítulo 2 apresenta conceitos relativos ao sistema de produção vocal e algumas patologias que atingem a laringe. O Capítulo 3 apresenta a fundamentação teórica necessária para compreensão do método proposto, abordando os conceitos básicos relacionados aos Gráficos de Recorrência, à Transformada Wavelet e à Análise de Textura. Na análise de Textura são abordados dois métodos, um deles baseado na matriz de co-ocorrência e nos descritores de Haralick e, o outro método é baseado em alguns descritores extraídos diretamente dos coeficientes wavelets.

No Capítulo 4 são apresentados os materiais e métodos utilizados neste trabalho, tais como informações sobre a base de dados utilizada nessa pesquisa. Além disso, é apresentado o método utilizado para obtenção dos gráficos de recorrência assim como os métodos de classificação utilizados para discriminação entre sinais saudáveis e sinais afetados por patologias.

No Capítulo 5 são apresentados e discutidos os resultados dos dois métodos propostos. E o Capítulo 6 apresenta as considerações finais, as contribuições deste trabalho e as sugestões para trabalhos futuros.

Mecanismo de Produção Vocal

Neste capítulo será apresentado um breve resumo sobre os órgãos que constituem o processo de produção vocal. Além disso, serão apresentadas algumas patologias que ocorrem no sistema fonador e causam desordens vocais, tais como: edema de Reinke, nódulos e paralisia nas pregas vocais.

2.1 – Produção da Fala

O processo de produção vocal é realizado a partir de movimentos sincronizados e articulados envolvendo a liberação de ar dos pulmões que passa pela laringe, que contém as pregas vocais e, finalmente, passa pelo trato vocal, cuja principal função é modular a onda sonora que vem das pregas vocais [6, 16].

A fonte necessária para a produção da fala, que produz o fluxo de ar, é constituída pelo sistema respiratório, que inclui os pulmões, brônquios e traquéia. O sistema fonatório, constituído pelas pregas vocais, representa a fonte vibratória. O sistema formado pelas cavidades oral e nasal, representa, a fonte de ressonância e, juntamente com o sistema articulatorio, constituído pela língua, lábios, mandíbula, palato e os dentes, constituem o trato vocal [17].

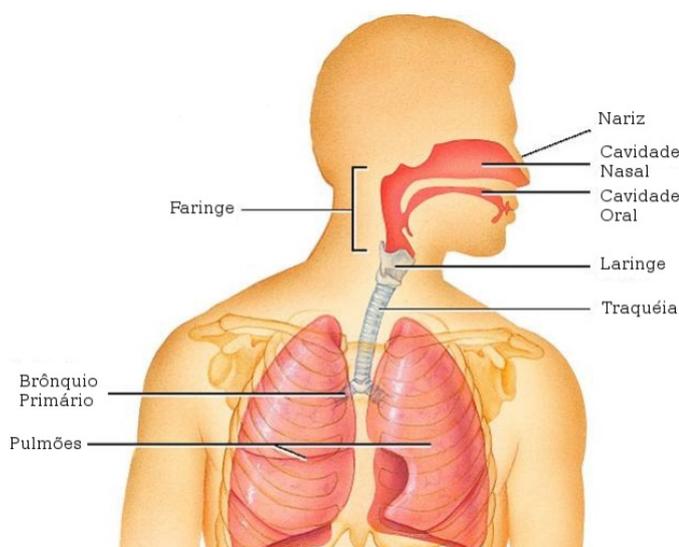


Figura 2.1 – Esquema dos sistemas envolvidos na produção da fala dos sinais de voz [18].

Considerando a estrutura e o funcionamento do sistema de produção da fala, qualquer alteração estrutural causada por alguma patologia no trato vocal, implica na obtenção de um sinal de voz diferente do

normal [19]. Dentre as patologias que podem causar alterações no sinal de voz, destacam-se as patologias que afetam a laringe. Portanto, para compreender essas patologias é necessário entender o funcionamento da laringe.

2.2 – Laringe

A laringe está situada no pescoço ao nível da 3ª e 6ª vértebras cervicais e pode ser dividida em três espaços: supraglote, glote e infraglote. As pregas vocais estão situadas na laringe e o espaço entre as pregas vocais é chamado de glote. O som da voz é produzido na glote, sendo imediatamente acrescido de ressonância na própria supraglote [6].

A maior e mais importante cartilagem da laringe, em respeito à fonação, é a cartilagem tireóidea, que possui o formato de um escudo, sendo composta de duas lâminas laterais, de forma quadrangular, e dois pares de cornos posteriores. O ângulo de união entre as lâminas direita e esquerda, chamado de proeminência laríngea, varia de acordo com o sexo. Esta variação é responsável pela definição do tamanho das pregas vocais e contribui na definição da frequência vocal emitida [20].

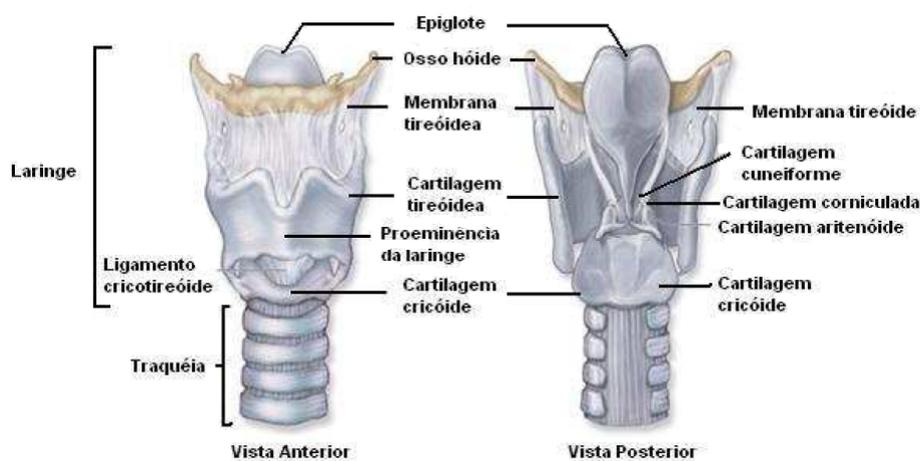


Figura 2.2 – Vista anterior e posterior dos principais componentes da laringe [16].

- ▣ Proteção
- ▣ Respiração
- ▣ Fonação

Com relação à função de proteção, a laringe atua evitando a entrada de qualquer coisa, exceto o ar, ao pulmão. Para a respiração as pregas vocais abduzem ativamente durante o movimento respiratório, contribuindo para regular a troca gasosa com o pulmão e a manutenção do equilíbrio ácido-base. E, com relação à função de fonação, as mudanças de tensão e longitude das pregas vocais, ampliação da abertura glótica e a intensidade do esforço respiratório provocam variações no tom da voz, tom esse que resulta da vibração das pregas vocais [7, 21, 22].

Algumas das patologias que afetam a laringe estão relacionadas a distúrbios orgânicos ou neurológicos. As patologias que afetam a laringe causam alterações na estrutura histológica das pregas vocais, influenciando no seu padrão de vibração, que pode ser afetado por mudanças nas camadas da mucosa, uma vez que a massa, o tamanho, a rigidez e a tensão do mecanismo de vibração estão comprometidos [6]. A seguir, uma breve descrição de algumas patologias encontradas na literatura.

2.3 – Patologias da Laringe

Uma voz normal, sob o ponto de vista perceptivo-auditivo, pode ser definida como uma voz agradável, melódica, audível e coerente, sendo essencial para uma boa comunicação. No entanto, uma voz com alguma patologia pode apresentar aspereza ou rouquidão e causar constrangimento à pessoa que fala [16]. Dentre as patologias que atingem a laringe e afetam a qualidade vocal do indivíduo serão abordadas neste trabalho nódulos, edema de Reinke e paralisia das pregas vocais, as quais são descritas a seguir.

2.3.1 – Nódulo

Nódulos podem ser classificados como lesões de massa bilaterais, de característica esbranquiçada ou levemente avermelhada. Devido às características peculiares da função vocal e de órgãos que participam da produção vocal os nódulos são lesões muito comuns em mulheres adultas jovens, entre 25 e 35 anos de idade, e em crianças de ambos os gêneros, devido às características peculiares da função vocal e de órgãos que participam da produção vocal [16].

Profissionais que utilizam a voz como instrumento de trabalho, tais como professores ou cantores, podem ter a laringe afetada por nódulos. O tamanho do nódulo pode comprometer o fechamento glótico e, além disso, a massa excedente por conta da presença da patologia resulta em um aumento da periodicidade de vibração das pregas vocais, maior nível de perturbação de frequência [7].



Figura 2.3 – Nódulos nas pregas vocais [23].

2.3.2 – Edema de Reinke

O edema corresponde ao acúmulo de fluido que ocorre em algum lugar na prega vocal. Particularmente, quando está situado na camada superficial da lâmina própria, também chamada de

espaço de Reinke, o edema é denominado edema de Reinke [2]. Em geral, são bilaterais e tem tamanho assimétrico [2, 7].

O desenvolvimento do edema de Reinke ocorre gradualmente e os sinais de vozes sofrem uma diminuição da frequência fundamental. Desse modo, a voz torna-se mais áspera e, conseqüentemente, ocorre o aumento do esforço para falar [7]. É considerada uma lesão benigna e o produz mudanças no movimento das pregas vocais, as quais têm efeito acústico de rouquidão na voz do paciente [19].



Figura 2.4 – Edema de Reinke nas pregas vocais [23].

2.3.3 – Paralisia das pregas vocais

Paralisia é um tipo de lesão que implica na imobilidade das pregas vocais, a qual é causada por dano ou disfunção do seu principal nervo: o nervo laríngeo recorrente. A paralisia pode ser unilateral, na qual a prega vocal afetada não pode mover-se para a linha média e ajudar no fechamento da glote, ou bilateral, na qual os dois lados tem espessura com aspecto reduzido e não conseguem mover-se totalmente para a linha média [7].

A paralisia das pregas vocais podem ser causadas pelos seguintes motivos: lesão do nervo durante uma ou mais cirurgias comuns, a pressão sobre o nervo de um tumor que cresce ao lado dele, inflamação que interrompe o nervo de trabalhar e um acidente vascular cerebral (para o caso bilateral). A paralisia de prega vocal unilateral ocorre com maior incidência do que a paralisia bilateral [23].



Figura 2.5 – Paralisia unilateral das pregas vocais [23].

Neste trabalho, a fim de analisar o método de discriminação entre sinais de voz saudáveis e sinais afetados por alguma patologia laríngea, serão utilizados sinais de voz afetados por alguma dentre essas três patologias: nódulos, edema de Reinke e paralisia das pregas vocais.

Este capítulo apresenta os conceitos teóricos necessários para compreender os métodos propostos para discriminação entre sinais de voz saudáveis e patológicos.

3.1 – Gráficos de Recorrência

Diversos fenômenos físicos podem ser caracterizados por um sistema dinâmico, o qual constitui um modelo matemático determinístico, ou seja, uma equação, na qual uma variável desse suposto sistema dinâmico evolui em função do tempo [24]. Nesse sentido, o processo de produção da fala pode ser considerado um fenômeno modelado por um sistema dinâmico autônomo, no qual assume-se que a série temporal é obtida a partir de equações de evolução que não contêm o tempo explicitamente [8].

Um sistema dinâmico pode depender de uma ou mais variáveis de estado, no entanto, nem sempre é possível observar a evolução no tempo de todas essas variáveis. Um exemplo disso é a observação de um sinal de voz, que corresponde a uma série temporal e apresenta a evolução no tempo de apenas uma variável de estado. Isto significa que um sistema autônomo modela a situação em que o comportamento do sistema muda no tempo, mas as características estruturais permanecem as mesmas, com independência do tempo.

Existem diversos métodos clássicos utilizados para calcular parâmetros dinâmicos de séries temporais, tais como: dimensão da informação, entropia, expoentes de Lyapunov, etc [8]. Tais métodos são baseados na teoria do caos e são utilizados para estimar propriedades não lineares de processos naturais. Contudo, a maioria deles exigem séries um tanto longas ou estacionariedade da série, o que implica em algumas limitações para a aplicação desses métodos [6].

A análise de um sistema dinâmico pode ser feita ao observar o seu espaço de fase, que constitui um espaço n -dimensional abstrato no qual se representa o comportamento do sistema e n corresponde ao número de variáveis do sistema [7]. Contudo, quando a quantidade de variáveis do sistema é maior do que três o nosso sistema visual humano só é capaz de visualizar o espaço de fase utilizando uma análise baseada em projeções, visto que somos capazes de visualizar apenas o espaço tridimensional.

Em 1987, Eckmann *et al.* [8] propuseram uma nova ferramenta denominada Gráfico de Recorrência (*Recurrence Plot - RP*), capaz de revelar propriedades dinâmicas de sistemas estacionários e não estacionários. Além disso, os Gráficos de Recorrência permitem uma análise bidimensional do comportamento de sistemas dinâmicos multidimensionais, sendo esta, portanto, uma grande vantagem em relação à análise visual do espaço de fase, que é limitado à representação tridimensional [7]. Nesta

seção são apresentados os conceitos básicos sobre os Gráficos de Recorrência que são necessários para o entendimento do trabalho desenvolvido.

Os Gráficos de Recorrência baseiam-se no Teorema da Recorrência, que foi proposto pelo físico e filósofo francês J. H. Poincaré em 1890. Esta ferramenta permite ilustrar os pontos da trajetória do espaço de fase, que são considerados recorrentes em sistemas dinâmicos dissipativos. Além do mais, os gráficos de recorrência podem ser utilizados para a análise de séries curtas e não-estacionárias, a qual é considerada uma das principais vantagens em comparação com outras técnicas de análise dinâmica não-linear [6, 10].

Na maioria dos experimentos não é possível medir todas as variáveis de estado de um sistema dinâmico e, geralmente, obtém-se a evolução temporal de apenas uma variável observada, $x(t)$, que representa a trajetória do espaço de fase do sistema. Contudo, as variáveis intrínsecas do sistema podem ser projetadas utilizando diversas técnicas de reconstrução do espaço de fase. O método utilizado neste trabalho é baseado no Teorema de Imersão de Takens, cuja ideia é de que a série temporal contém informações sobre as variáveis intrínsecas do sistema dinâmico.

Segundo o teorema da imersão de Takens [25] é possível reconstruir um espaço de dimensão m topologicamente equivalente ao espaço de estados original de dimensão d_0 , desde que $m > 2d_0$. Para isto a série temporal, $\overrightarrow{x}(t)$, de comprimento T , precisa passar por um processo de imersão no espaço de fases, no qual são criadas $m - 1$ cópias defasadas de $\tau, 2\tau, \dots, (m - 1)\tau$ amostras, em que τ é chamado passo de reconstrução. Este processo irá gerar N vetores x_i de dimensão m , em que m é chamada de dimensão da imersão.

Segundo Marwan [10], um gráfico de recorrência é definido matematicamente pela expressão:

$$\overrightarrow{R}_{i,j}^{m,\varepsilon} = \Theta(\varepsilon - \|\overrightarrow{x}_i - \overrightarrow{x}_j\|), \overrightarrow{x}_i \in R^m, i, j = 1, \dots, N \quad (3.1)$$

em que:

- N é o número de estados \overrightarrow{x}_i considerados;
- Θ é a função degrau unitário;
- $\|\cdot\|$ é a norma utilizada (geralmente, a norma euclidiana) ;
- ε é o raio de vizinhança no ponto x_i ;
- m é a dimensão de imersão.

Um Gráfico de Recorrência constitui uma matriz de ordem $N \times N$, que é construída a partir da função degrau unitário, a qual mapeia os pontos recorrentes do sistema da seguinte forma: se um ponto é considerado recorrente, ou seja, está dentro da vizinhança com raio ε , é atribuído o valor 1 (um) ao ponto (i, j) da matriz e, caso contrário, é atribuído o valor 0 (zero). Desse modo obtém-se uma matriz formada por zeros e uns, que corresponde a uma imagem binária.

Na Figura 3.1 é apresentado um exemplo da construção de um gráfico de recorrência. A partir de uma trajetória no espaço de fase de um sistema dinâmico arbitrário (a) e observando um trecho no qual há um retorno dessa trajetória (b), é possível ver que, no raio de vizinhança ε de cada um dos estados consecutivos $\overrightarrow{x}_2, \overrightarrow{x}_3$ e \overrightarrow{x}_4 , há recorrência com os estados $\overrightarrow{x}_7, \overrightarrow{x}_8$ e \overrightarrow{x}_9 , respectivamente. Tal evolução paralela é retratada no gráfico de recorrência (c), no qual são formadas estruturas diagonais, neste caso particular. Pode-se observar que a diagonal principal será sempre totalmente preenchida por pontos pretos

(pontos recorrentes) pois reflete o cálculo da distância entre um estado e ele mesmo, o que sempre será igual a zero e, por consequência, menor que o raio de vizinhança estabelecido. Além disso, o gráfico de recorrência é simétrico, visto que a distância de um estado \vec{x}_i a um estado \vec{x}_j será sempre igual à distância de \vec{x}_j até \vec{x}_i [7].

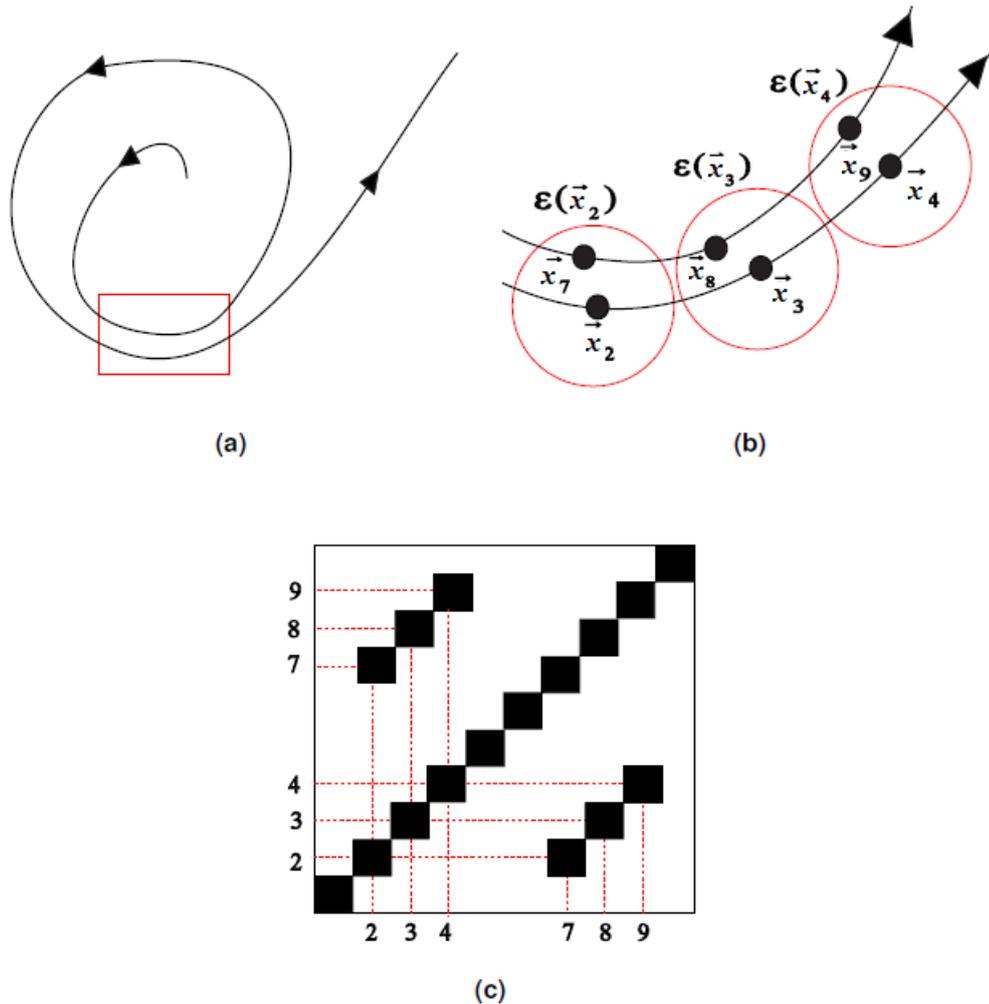


Figura 3.1 – Construção de um gráfico de recorrência. Adaptada de [7].

Na Figura 3.2 são apresentados exemplos de gráficos de recorrência gerados a partir de sinais de voz, sendo um saudável e os demais afetados por patologias (edema, nódulo e paralisia). Conforme foi mencionado anteriormente, pode-se observar que os gráficos de recorrências são matrizes bidimensionais que correspondem a imagens digitais com dois níveis de cinza (preto e branco), nas quais os *pixels* podem formar estruturas típicas tais como: pontos isolados, linhas diagonais, verticais e horizontais. Tais estruturas caracterizam os sistemas analisados e permitem a abordagem dos gráficos de recorrência do ponto de vista de pequena escala como texturas e, portanto, pode-se utilizar o Processamento Digital de Imagens (PDI) sobre esses gráficos para extrair informações relevantes desses sistemas.

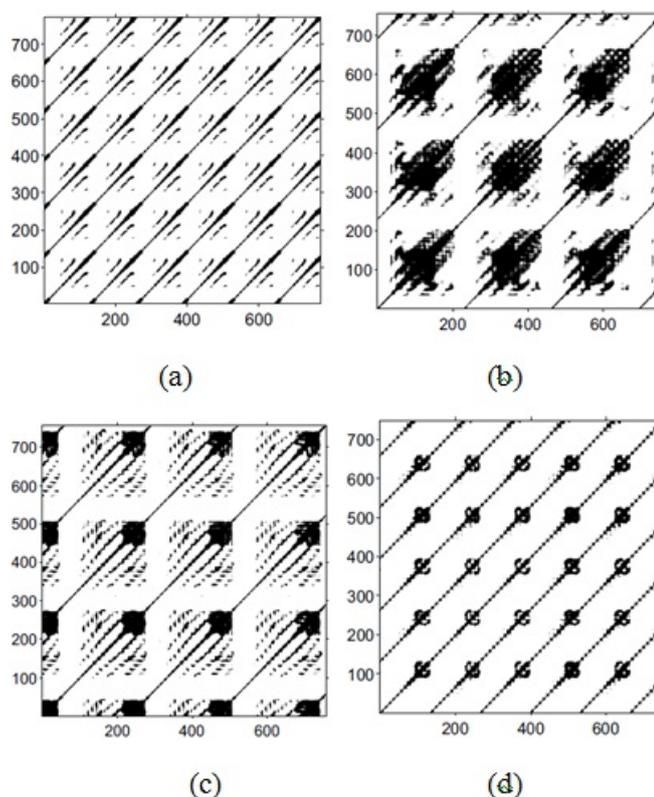


Figura 3.2 – (a) Sinal de voz saudável; (b) Sinal de voz afetado por paralisia; (c) Sinal de voz afetado por edema; e (d) Sinal de voz afetado por nódulo [7].

3.2 – Wavelets

Uma das mais conhecidas técnicas para processamento e análise de sinais é a transformada de Fourier, que consiste em expressar uma função como uma soma infinita de senos e cossenos, provendo uma forma de representar no domínio da frequência um sinal inicialmente expresso no domínio do tempo. As wavelets, diferentemente das funções senoidais, são ondas que possuem sua energia concentrada no tempo, que é um diferencial com relação à outras ondas.

Wavelets são funções de duração limitada, isto é, de suporte compacto. Esta propriedade significa que seu domínio é diferente de zero em uma extensão finita e igual a zero em todo o resto do domínio, o que possibilita a análise de sinais transitórios, não estacionários ou variáveis no tempo. Este fato torna interessante a utilização das wavelets no caso específico da análise de imagens, pois as mudanças de regiões ou bordas podem ser destacadas facilmente [26].

A transformada wavelet constitui um mecanismo para decompor sinais nas suas partes constituintes. Portanto, é uma técnica usada para representar, descrever e processar sinais como uma combinação linear de funções elementares. Para isto, utiliza funções que são versões escalonadas e transladadas de uma função conhecida como wavelet mãe [27].

Em 1910 surge o primeiro sistema ortogonal de funções conhecidos como wavelets de Haar, que é apresentada na Figura 3.3. Por volta de 1980, Yves Meyer e Stephane Mallat [27] aplicou as wavelets na teoria chamada multirresolução, que é uma estratégia de processamento de sinais que permite que objetos

difíceis de serem identificados em uma determinada resolução possam ser identificados a partir de uma resolução mais apropriada, seja esta maior ou menor.

Outro ponto importante em relação ao processamento de imagens é que a Transformada wavelet permite representar uma imagem toda, em uma escala menor, com uma quantidade mínima de valores [9]. Portanto, apresentando economia de processamento em determinadas aplicações como, por exemplo, na compressão de imagens.

Um sinal pode ser representado por uma função $f(t)$ formada pela combinação linear de um conjunto de funções reais $\psi_k(t)$ com coeficientes reais c_k , sendo k um índice inteiro [28].

$$f(t) = \sum_k c_k \psi_k(t) \quad (3.2)$$

As funções reais $\psi_k(t)$ neste caso são denominadas wavelets, que são funções reais definidas no espaço vetorial $L^2(\mathbb{R})$. Uma função pode ser considerada uma wavelet se satisfaz as seguintes propriedades [29]:

- A área total sob a curva da função é 0, ou seja, $\int_{-\infty}^{\infty} \psi(t) dt = 0$;
- A energia da função é finita $\int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty$.

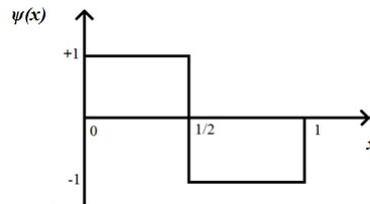


Figura 3.3 – Wavelet de Haar.

A expansão wavelet é obtida a partir da wavelet mãe e por um sistema constituído de dois índices inteiros a que corresponde ao fator escala e b que corresponde ao fator translação, e funções de expansão (ou funções base) $\psi_{a,b}(t)$ que pode ser representada por:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad a, b \in \mathbb{R} (a \neq 0) \quad (3.3)$$

em que o fator $\frac{1}{\sqrt{a}}$ é um fator de normalização que mantém a mesma energia para todas as wavelets independente da escala utilizada.

Segundo Mallat [27] o fator b está associado à localização das wavelets no tempo e o fator a está associado à sua localização no domínio da frequência. Um número discreto de wavelets é gerado quando os fatores a e b são restritos aos valores discretos $a = a_0^j$ e $b = ka_0^j b_0$, em que $a_0 > 1$, $b_0 > 0$ e $j, k \in \mathbb{Z}$. No caso particular em que $a_0 = 2$ e $b_0 = 1$ é possível construir um conjunto de funções $\psi(t)$ ortonormal dado pela Equação 3.4:

$$\psi_{j,k}(t) = 2^{-j/2} \psi(2^{-j}t - k) \quad (3.4)$$

Portanto, a Transformada Wavelet Discreta (*Discrete Wavelet Transform - DWT*) utilizando as bases wavelets diádicas ortonormais é representada pela Equação 3.5 [3]:

$$DWT_f(j, k) = \langle f, \psi_{j,k} \rangle = |2|^{-j/2} \int_{-\infty}^{\infty} f(t) \psi(2^{-j}t - k) dt \quad (3.5)$$

em que o produto interno $\langle f, \psi_{j,k} \rangle$ é chamado coeficiente wavelet discreto. E para reescrever a função $f(t) \in L^2(\mathbb{R})$ utiliza-se a equação:

$$f(t) = \sum_{x,y} \langle f, \psi_{j,k} \rangle \psi_{j,k}(t). \quad (3.6)$$

3.2.1 – Análise Multirresolucional

A análise de multirresolução em wavelets foi desenvolvida na década de 80, no trabalho de Mallat [27]. Esta análise provê um modelo matemático para representar sinais em diferentes escalas ou resoluções utilizando um conjunto de filtros especializados em extrair as informações do sinal. A multirresolução é bastante útil para análise de imagens, pois permite que objetos difíceis de serem identificados em uma determinada resolução possam ser identificados a partir de uma resolução mais apropriada, seja esta mais alta ou mais baixa [9].

Para decompor um sinal por meio da análise multirresolução, é necessário, além da função wavelet mãe, uma função complementar de norma unitária $\phi(t)$, chamada função escala. Essa análise é feita em duas partes distintas do sinal, a principal, obtida utilizando um filtro passa-baixa, e a residual, obtida a partir da aplicação de filtros passa-alta sobre o sinal, correspondente aos detalhes da imagem. A parte principal corresponde a uma aproximação do sinal original e pode ser sucessivamente dividida, resultando num processo de decomposição [27].

Na análise multirresolução um sinal é decomposto em vários níveis de resolução, onde cada nível corresponde a um subespaço $\{V_j\}_{j \in \mathbb{Z}}$ de funções de $L^2(\mathbb{R})$, que satisfazem as seguintes condições:

$$V_j \subset V_{j-1}, j \in \mathbb{Z} \quad (3.7)$$

em que

$$\bigcap_{j=-\infty}^{\infty} V_j = \{0\} \quad , \quad \bigcup_{j=-\infty}^{\infty} V_j = L^2(\mathbb{R}) \quad (3.8)$$

A condição de causalidade imposta na pela Equação 3.7 implica que o subespaço V_j apresenta uma resolução menor do que o subespaço V_{j-1} . Portanto, quanto menor a escala j maior a resolução e, além disso, para cada subespaço V_j existe um subespaço W_j associado, chamado de complemento ortogonal de V_j com relação à V_{j-1} . Este subespaço W_j corresponde aos detalhes necessários para passar do nível de resolução j para o nível $j - 1$ [27].

$$V_{j-1} = V_j \oplus W_j \quad (3.9)$$

Mallat e Daubechies [27, 29] mostraram a relação entre o cálculo de coeficientes em uma transformada wavelet e um banco de filtros, que são sistemas que alteram apenas algumas frequências de um sinal de entrada. Portanto, a análise multirresolução pode ser feita a partir de um banco formado

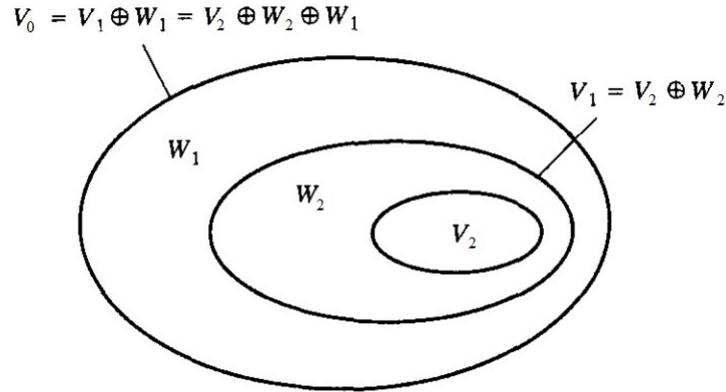


Figura 3.4 – Representação da relação entre a escala e os subespaços de funções wavelets. Adaptada de [30].

por dois filtros: $h(n)$ (passa-baixas) e $g(n)$ (passa-altas), tais que a convolução desses filtros com o sinal permite obter os coeficientes em um determinado nível de resolução.

Sejam V_0 e W_0 subespaços de V_{-1} . Cada wavelet mãe $\psi(t) \in V_0$ tem uma função escala $\phi(t) \in W_0$ associada a ela que podem ser representadas como combinações lineares das funções base de V_{-1} de acordo com as Equações 3.10 e 3.11, as quais correspondem à relação entre a transformada e os filtros correspondentes à análise multirresolução $h(n)$ e $g(n)$.

$$\phi(t) = \sum_n h(n)\phi(2t - n) \quad (3.10)$$

$$\psi(t) = \sum_n g(n)\phi(2t - n) \quad (3.11)$$

A função escalonamento ϕ existe se os coeficientes do filtro passa-baixas satisfaz a seguinte condição:

$$\sum_n h(n) = 1 \quad (3.12)$$

De maneira geral, como $V_{j-1} = V_j \oplus W_j$ uma função $f(t) \in L^2(\mathbb{R})$ que pode ser representada como uma combinação linear das funções base de V_{j-1} , pode ser representada pelas funções base de V_j e W_j da seguinte forma:

$$f(t) = \sum_k c_j(k)2^{-j/2}\phi(2^{-j}t - k) + d_j(k)2^{-j/2}\psi(2^{-j}t - k) \quad (3.13)$$

em que

$$c_j(k) = \sum_m h(m - 2k)c_{j-1}(m), \quad (3.14)$$

$$d_j(k) = \sum_m g(m - 2k)c_{j-1}(m). \quad (3.15)$$

3.2.2 – Transformada Wavelet Bidimensional

A aplicação da transformada bidimensional utilizada em imagens é, geralmente, obtida a partir de transformações unidimensionais aplicadas separadamente nas direções vertical e horizontal, com uma função de escala $\phi(x, y)$ definida como [27]:

$$\phi(x, y) = \phi(x)\phi(y) \quad (3.16)$$

em que $\phi(x)$ é a função de escala unidimensional.

Associando a wavelet unidimensional $\psi(x)$ à função de escala $\phi(x)$, para fins de processamento de imagens, obtém-se três wavelets bidimensionais, obtidas a partir de filtragens passa-alta e passa-baixa aplicadas nas linhas e nas colunas. As três wavelets obtidas são:

$$\psi(x, y)^1 = \phi(x)\psi(y) \quad (3.17)$$

$$\psi(x, y)^2 = \psi(x)\phi(y) \quad (3.18)$$

$$\psi(x, y)^3 = \psi(x)\psi(y) \quad (3.19)$$

A Equação 3.16 gera uma aproximação da imagem original com uma resolução menor, portanto, os coeficientes gerados por ela são chamados de coeficientes de aproximação. A Equação 3.17 mede as variações horizontais. A Equação 3.18 mede as variações verticais. E, a Equação 3.19 corresponde as variações diagonais.

Para obter a decomposição wavelet bidimensional em uma imagem pode ser feita a filtragem individuais das linhas e colunas da imagens utilizando filtros discretos $h(n)$ e $g(n)$ da seguinte forma: inicialmente é realizada a convolução das linhas da imagem I_{j-1} com os filtros $h(n)$ e $g(n)$. Em seguida, as duas imagens restantes passam por um processo chamado de decimação (*downsampling*), que consiste em descartar os termos de uma sequência cujos índices não são múltiplos de um inteiro q ($q = 2$ para a transformada wavelet) [9]. Deste modo, são descartadas as linhas ímpares da imagem e, do mesmo modo, as colunas das imagens decimadas são convoluídas com os filtros passa-baixas e passa-altas e, assim, as colunas ímpares são excluídas no processo de decimação [31].

As Equações 3.16 - 3.19 geram um nível de decomposição wavelet, produzindo quatro sub-bandas, que são: uma sub-banda LL, correspondente à aplicação da função escala sobre o sinal, e as sub-bandas LH, HL e HH, que correspondem, respectivamente, à aplicação das medidas das variações horizontais, verticais e diagonais. Esta decomposição pode ser realizada recursivamente na sub-banda LL, obtendo níveis adicionais de decomposição. Na Figura 3.5 é apresentada uma representação da decomposição wavelet com um nível e com dois níveis de decomposição.

Na Figura 3.6 é apresentado um exemplo da decomposição wavelet, com um nível de decomposição, de um gráfico de recorrência de um sinal de voz, realizada por meio dos filtros $h(n)$ e $g(n)$.

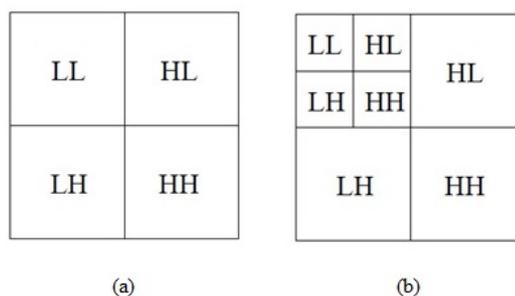


Figura 3.5 – Representação das sub-bandas obtidas pela decomposição wavelet com (a) um nível de decomposição; (b) dois níveis de decomposição.

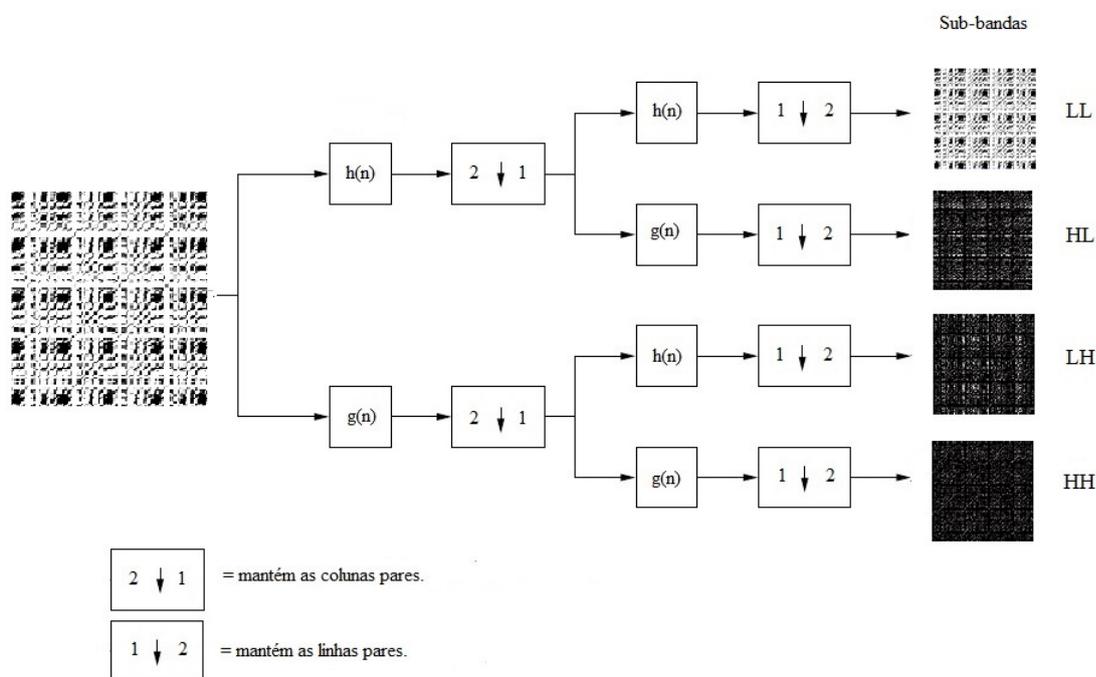


Figura 3.6 – Representação da decomposição wavelet utilizando de filtros.

3.3 – Texturas

O Processamento Digital de Imagens (PDI) compreende a área de reconhecimento de padrões, que tem como objetivo tentar descrever e identificar objetos presentes em uma imagem. A análise de textura é uma área do PDI. Segundo Gonzalez e Woods [30] a quantificação do conteúdo de textura é uma abordagem importante para a descrição de regiões e, esses padrões homogêneos ou arranjos espaciais de *pixels* não podem ser descritos somente por intensidades locais ou cores.

Quando se olha para uma imagem conecta-se facilmente regiões que têm um mesmo nível de cinza ou cor similar e, embora não haja um acordo sobre uma definição formal de textura, uma característica importante de texturas é a repetição de um padrão ou padrões sobre uma região [32]. Soares [15] afirma que a textura refere-se a um padrão visual que possui algumas propriedades de homogeneidade que não

resultam simplesmente de uma cor ou intensidade, pois ao contrário de outras características, como por exemplo o brilho, a textura não pode ser definida em um *pixel*, mas sim a partir de uma região ou conjunto de *pixels*.

A análise de textura pode ser realizada a partir de diversas abordagens, das quais se destacam as abordagens estatística, estrutural e espectral. A abordagem estatística é realizada, em geral, a partir de medidas de primeira ordem, que são as medidas estatísticas de textura calculadas a partir dos valores originais da imagem (por exemplo a variância), e não consideram o relacionamento entre os *pixels* e, utilizando medidas estatísticas de segunda ordem, que consideram o relacionamento entre grupos de dois *pixels* na imagem original [33].

A abordagem estrutural é baseada na ideia de que texturas são compostas por primitivas, que é a estrutura geométrica básica que forma a textura, por exemplo, linhas paralelas regularmente espaçadas. E a abordagem espectral baseia-se em propriedades do espectro de Fourier [34].

Haralick *et al.* [35] utilizaram medidas estatísticas de segunda ordem para análise de textura como uniformidade, densidade, aspereza, regularidade, intensidade, entre outras características da imagem a partir da matriz de co-ocorrência dos *pixels*. Esta matriz é formada pelos valores de probabilidade das diferentes combinações dos valores de intensidade dos níveis de cinza que ocorrem em uma imagem. Neste trabalho foram utilizadas medidas estatísticas de primeira ordem extraídas diretamente das sub-bandas wavelets dos gráficos de recorrência no método 1 e os descritores de Haralick no método 2, os quais são apresentados a seguir.

3.3.1 – Medidas extraídas das sub-bandas Wavelets

Os coeficientes das sub-bandas da transformada wavelet contém informações de textura contidas na imagem original. As regiões com texturas suaves possuem principalmente coeficientes baixos e, nas regiões com transições, os coeficientes tendem a ser maiores, enfatizando as mudanças [9].

Silva [9], em seu trabalho, propôs um conjunto de medidas que são extraídas das sub-bandas Sb , as quais são apresentadas nas Equações 3.20 - 3.27. Desse modo, dada uma imagem $I(x, y)$, tem-se que o coeficiente da sub-banda Sb na coordenada (x, y) será denotado por $Sb(x, y)$, e N_{Sb} é o número de coeficientes da sub-banda.

$$\text{Energia}_{Sb} = \sqrt{\frac{1}{N_{Sb}} \sum_{x,y} (Sb(x, y))^2} \quad (3.20)$$

$$\text{Desvio Médio}_{Sb} = \frac{1}{N_{Sb}} \sum_{x,y} |Sb(x, y)| \quad (3.21)$$

$$\text{Entropia}_{Sb} = -\frac{1}{N_{Sb}} \sum_{x,y} |Sb(x, y)| \log |Sb(x, y)| \quad (3.22)$$

$$\text{Média}_{Sb} = \frac{\sum_{x,y} Sb(x, y)}{N_{Sb}} \quad (3.23)$$

$$\text{Norma}_{Sb} = \left(\sum_{x,y} |Sb(x, y)|^n \right)^{\frac{1}{n}} \quad (3.24)$$

em que r_n é a ordem da norma.

$$\text{Momento}_{Sb}^{r_m} = \sum_{x,y} (Sb(x,y) - \mu)^{r_m} p(Sb(x,y)) \quad (3.25)$$

em que r_m é a ordem do momento, μ é a média dos valores dos coeficientes da sub-banda Sb e $p(Sb(x,y))$ é a probabilidade de ocorrer o coeficiente $Sb(x,y)$.

$$\text{Suavidade}_{Sb} = 1 - \frac{1}{1 + \sigma^2} \quad (3.26)$$

em que σ^2 é a variância dos coeficientes da sub-banda Sb .

$$\text{Resíduo Médio}_{Sb} = \sum (Sb(x,y) - \mu)^2 \quad (3.27)$$

3.3.2 – Matriz de co-ocorrência (*Gray-Level Co-occurrence matrix - GLCM*)

A matriz de co-ocorrência corresponde a uma matriz de frequências relativas $P(i, j, d, \theta)$, construída a partir da relação entre dois *pixels* de uma imagem, de modo que um deles é chamado de pixel referência com intensidade i e o outro de pixel vizinho com intensidade j , separados por uma distância d e na direção θ . Cada pixel dentro da imagem torna-se o *pixel* referência, iniciando no canto superior esquerdo e procedendo até o inferior direito.

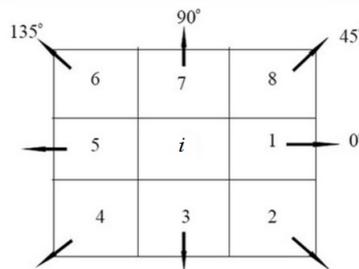


Figura 3.7 – Direções utilizadas para o cálculo da matriz de co-ocorrência [36].

A Figura 3.7 mostra as quatro direções $\theta = 0^\circ, 45^\circ, 90^\circ$ e 135° que podem ser consideradas para a construção da matriz de co-ocorrência. A distância d a ser considerada dependerá do tamanho da imagem. Em texturas grosseiras, a distribuição dos níveis de cinza varia lentamente com a distância, e para texturas finas a distribuição dos níveis de cinza varia rapidamente com a distância [33].

Considere a matriz de uma imagem em níveis de cinza, conforme a Figura 3.8. A imagem apresenta quatro níveis de cinza $Z_n = 0, 1, 2$ e 3 . A Figura 3.9 representa a construção da matriz de co-

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

Figura 3.8 – Matriz correspondente a uma imagem com 4 níveis de cinza.

$i \backslash j$	0	1	2	3
0	#(0,0)	#(0,1)	#(0,2)	#(0,3)
1	#(1,0)	#(1,1)	#(1,2)	#(1,3)
2	#(2,0)	#(2,1)	#(2,2)	#(2,3)
3	#(3,0)	#(3,1)	#(3,2)	#(3,3)

Figura 3.9 – Matriz de co-ocorrência.

$i \backslash j$	0	1	2	3
0	4	2	1	0
1	2	4	0	0
2	1	0	6	1
3	0	0	1	2

(a)

$i \backslash j$	0	1	2	3
0	4	1	0	0
1	1	2	2	0
2	0	2	4	1
3	0	0	1	0

(b)

$i \backslash j$	0	1	2	3
0	6	0	2	0
1	0	4	2	0
2	2	2	2	2
3	0	0	2	0

(c)

$i \backslash j$	0	1	2	3
0	2	1	3	0
1	1	2	1	0
2	3	1	0	2
3	0	0	2	0

(d)

Figura 3.10 – Matrizes de co-ocorrência da imagem em níveis de cinza ilustrada pela Figura 3.8 para $d = 1$: (a) $\theta = 0^\circ$; (b) $\theta = 45^\circ$; (c) $\theta = 90^\circ$; (d) $\theta = 135^\circ$.

ocorrência da uma imagem na Figura 3.8, que contém 4 níveis de cinza. De maneira geral, seja N o número de níveis de cinza, então matriz de co-ocorrência será uma matriz quadrada de ordem N . Considerando dois níveis distintos i e j contidos na imagem, cada entrada da matriz representa o número de ocorrências da combinação (i, j) . Por exemplo, a entrada $P(0, 1) = \#(0, 1)$ corresponde ao número de vezes que ocorre o nível 0 vizinho ao nível 1 para uma determinada distância d na direção θ . No caso da imagem da Figura 3.8 para distância $d = 1$ e $\theta = 0^\circ$ temos $\#(0, 1) = 2$.

Na Figura 3.10 são apresentadas as matrizes de co-ocorrência para a imagem da Figura 3.8 nas direções $\theta = 0^\circ, 45^\circ, 90^\circ$ e 135° considerando a distância $d = 1$. Portanto, para cada direção e, cada distância considerada da matriz de co-ocorrência, é possível extrair medidas capazes de caracterizar uma textura. Tais medidas são denominadas de descritores de Haralick e são apresentados na próxima seção.

3.3.3 – Descritores de Haralick

Haralick *et al.* [35] propuseram medidas capazes de quantificar a textura, as quais são medidas estatísticas de segunda ordem e são extraídas a partir da matriz de co-ocorrência. Diversos trabalhos têm utilizado os descritores de Haralick para a análise de textura obtendo resultados relevantes [33, 37-39].

Dada uma matriz de co-ocorrência com entradas $P(i, j)$. Considere N_g o número de níveis de cinza distintos contidos na imagem e $p(i, j)$ a (i, j) -ésima entrada normalizada da matriz de co-ocorrência.

Os descritores de Haralick utilizados neste trabalho, e a descrição do que cada medida representa, são apresentados nas Equações 3.28 - 3.40 [35, 37-39].

Entropia

É uma medida de aleatoriedade, mede a informação contida em $p(i, j)$. O valor da entropia é alto quando os valores da matriz de co-ocorrência são iguais, e é baixo quando a concentração de valores na diagonal é alta ou quando as entradas da matriz de co-ocorrência possuem valores baixos.

$$\text{Entropia} = - \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} p(i, j) \log(p(i, j)) \quad (3.28)$$

Segundo Momento Angular

É a medida da homogeneidade da imagem. Numa imagem homogênea pode-se dizer que existem poucas transições de níveis de cinza.

$$\text{Segundo momento angular} = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} (p(i, j))^2 \quad (3.29)$$

Contraste

É a variação local entre os *pixels*, ou seja, mede a presença de transição abrupta de níveis de cinza (bordas).

$$\text{Contraste} = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} p(i, j)(i - j)^2 \quad (3.30)$$

Momento da diferença inversa

Denota a homogeneidade da matriz de co-ocorrência. O momento da diferença é máximo quando a concentração dos valores na diagonal da matriz de co-ocorrência for máxima.

$$\text{Momento da diferença inversa} = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} \frac{1}{1 + (i - j)^2} p(i, j) \quad (3.31)$$

Correlação

Corresponde a uma ideia de linearidade de dependências de tons de cinza em uma imagem. O valor da correlação é alto em uma imagem onde existe uma certa ordenação local dos níveis de cinza.

$$\text{Correlação} = \frac{\sum_{i=1}^{Ng} \sum_{j=1}^{Ng} (ij)p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (3.32)$$

Média da soma

É um indicador da média de tonalidades entre pares de *pixels*.

$$\text{Média da soma} = \sum_{i=2}^{2Ng} i \cdot p_{x+y}(i) \quad (3.33)$$

Entropia da soma

Indica quantidade de desorganização entre os *pixels*.

$$\text{Entropia da soma} = - \sum_{i=2}^{2Ng} p_{x+y}(i) \log(p_{x+y}(i)) \quad (3.34)$$

Variância da soma

Indica a variação da tonalidade do fundo da imagem.

$$\text{Variância da soma} = \sum_{i=2}^{2Ng} (i - \text{Entropia da soma})^2 p_{x+y}(i) \quad (3.35)$$

Variância

Mede a dispersão em relação à média, ou seja, indica a variação de tonalidades na imagem.

$$\text{Variância} = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} (i - \mu_x \mu_y)^2 p(i, j) \quad (3.36)$$

Variância da diferença

Similar a variância da soma, mas com a média centrada em zero.

$$\text{Variância da diferença} = \text{variância de } p_{x-y} \quad (3.37)$$

Entropia da diferença

Similar à entropia da soma, é um indicador de quantidade de desorganização entre os *pixels*.

$$\text{Entropia da diferença} = - \sum_{i=0}^{Ng-1} p_{x-y}(i) \log(p_{x-y}(i)) \quad (3.38)$$

Informação de Correlação 1

É um indicador da correlação baseado na medida da entropia para elementos independentes e pares de elementos na imagem.

$$\text{Informação de correlação 1} = \frac{HXY - HXY1}{\max\{HX, HY\}} \quad (3.39)$$

Informação de Correlação 2

É a medida da correlação baseada na similaridade entre entropias de elementos independentes da imagem e pares de elementos da imagem.

$$\text{Informação de correlação } 2 = (1 - \exp[-2.0(HXY2 - HXY)])^{1/2} \quad (3.40)$$

em que, nas expressões anteriores:

$$p_x(i) = \sum_{j=1}^{Ng} p(i, j)$$

$$p_y(j) = \sum_{i=1}^{Ng} p(i, j)$$

$$p_{x+y}(k) = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} p(i, j); k = 2, 3, \dots, 2Ng.; i + j = k$$

$$p_{x-y}(k) = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} p(i, j); k = 0, 1, \dots, Ng - 1.; |i - j| = k$$

$$HXY = - \sum_i \sum_j p(i, j) \log(p(i, j))$$

$$HXY1 = - \sum_i \sum_j p(i, j) \log(p_x(i)p_y(j))$$

$$HXY2 = - \sum_i \sum_j p_x(i)p_y(j) \log(p_x(i)p_y(j))$$

HX e HY são as entropias de p_x e p_y .

μ_x, μ_y, σ_x e σ_y são as médias e os desvios padrão de p_x e p_y .

3.3.4 – Otimização por enxame de partículas (*Particle Swarm Optimization - PSO*)

A Otimização pode ser considerada uma área da matemática cujo objetivo é minimizar ou maximizar uma solução para uma função objetivo. Na área de engenharia a maioria dos problemas reais são não-estruturados e difíceis de modelar. O desenvolvimento de algoritmos de otimização permitem que problemas complexos sejam solucionados, com soluções viáveis, com recursos disponíveis e em um tempo polinomial [40].

Métodos de otimização baseados em conceitos de evolução de indivíduos de uma população, conhecidos como Métodos Evolutivos, podem ser utilizados na busca de uma solução ótima para problemas complexos de engenharia. Tais métodos utilizam uma abordagem baseada na probabilidade e na aleatoriedade para variar os parâmetros de busca [41]. Tais métodos são descritos com mais detalhes em [42].

O método de otimização por Enxame de Partículas (*Particle Swarm Optimization - PSO*) foi desenvolvido por Kennedy e Eberhart [43], baseados em observações do comportamento social dos animais tais como bandos de pássaros e cardumes de peixes. Trata-se de uma técnica de computação evolucionária inspirada nos mecanismos de adaptação de indivíduos em sociedade que considera a prospecção individual combinada à ideia de cooperação entre indivíduos da população.

Desse modo, cada indivíduo ou partícula de uma dada população ou enxame move-se através do espaço de busca em iterações sucessivas, cooperando e competindo com outras partículas. Pode-

se dizer que o um indivíduo "voa" lembrando da melhor posição do espaço de busca que visitou e, em direção ao melhor indivíduo de uma vizinhança topológica [41]. A qualidade das soluções é medida por uma função predefinida de aptidão (*fitness*), que depende do problema. A estrutura teórica do PSO é muito simples, possui codificação e implementação computacional fáceis e, além disso, apresenta baixo custos computacional. Logo, devido a estas características, o PSO tem ganhado muita atenção para diversas aplicações em diferentes campos de pesquisa [44].

Cada partícula, em uma dada iteração, tem uma velocidade e uma posição no espaço. A partícula possui também uma pequena memória que armazena sua melhor posição até então atingida (p_B), e a melhor posição já alcançada pelos conhecidos da partícula (g_B), ou seja, o melhor de todas as partículas pertencentes à vizinhança [43]. O algoritmo básico de PSO é construído então como segue [41]:

- ▣ Inicialização: A população de partículas é inicializada com posições e velocidades aleatórias em L dimensões no espaço de busca.
- ▣ Avaliação: Para cada partícula, a função de aptidão desejada é avaliada em L variáveis.
- ▣ Atualização do melhor local: A aptidão avaliada para a partícula na iteração atual é comparada com o melhor valor de aptidão para a partícula em todas as iterações precedentes (armazenadas em p_B). Se o valor atual é melhor do que p_B , então é atribuído a p_B o valor atual de aptidão, e a p_i é atribuída a posição atual x_i no espaço L -dimensional.
- ▣ Atualização do melhor global: A partícula na vizinhança com maior sucesso até então é identificada, atualizando g_B , que armazena o valor da melhor aptidão para o enxame inteiro.
- ▣ Atualização os parâmetros de movimento das partículas: A velocidade e a posição de cada partícula são atualizadas de acordo com as Equações 3.41 e 3.42.
- ▣ Avaliação do critério de parada: Os passos 2 a 5 são repetidos até que os critérios de parada sejam satisfeitos (geralmente uma aptidão suficientemente boa ou um número máximo de iterações).

Na inicialização do algoritmo, cada partícula é inicializada com uma posição e uma velocidade aleatórias. A velocidade da partícula e posição das partículas podem ser atualizadas, através das Equações 3.41 e 3.42 , respectivamente.

$$v_i(t + 1) = v_i(t) + \varphi_1 \times (p_B - x_i(t)) + \varphi_2 \times (g_B - x_i(t)) \quad (3.41)$$

$$x_i(t + 1) = x_i(t) + v_i(t + 1) \quad (3.42)$$

em que φ_1 e φ_2 são constantes limitadas a um intervalo finito.

No algoritmo de otimização por nuvem de partículas binário cada partícula se move em um espaço de busca restrito a 0 e 1 em cada dimensão. Neste caso, a função da velocidade da partícula é usada como uma distribuição de probabilidade:

$$\text{Se } rand < \frac{1}{1 + e^{v_i(t+1)}} \Rightarrow x_i(t + 1) = 1; \text{ Caso contrário } \Rightarrow x_i(t + 1) = 0; \quad (3.43)$$

$rand()$ é um número aleatório selecionado de uma distribuição uniforme entre [0,1].

Na seleção de características, a versão binária do algoritmo tem como objetivo determinar o subconjunto de características mais representativo do espaço de características. Cada partícula no algoritmo representa uma possível solução. A codificação inicial da posição de cada partícula é feita aleatoriamente. Cada posição recebe um código binário $P = \{F_1 F_2 \dots F_n\}$, $n = 1, 2, \dots, l$; em que l é o comprimento do vetor de características. Se o valor do bit F_i for 1, denota que a característica foi selecionada, caso contrário denota rejeição. A otimização por nuvem de partículas binária realiza uma busca no espaço de 2^l para determinar o subconjunto de características ótimo, ou seja, que fornece a maior separação entre as classes [45].

Neste capítulo, foram apresentados conceitos relativos aos gráficos de recorrência, à transformada wavelet e à análise de textura, necessários para compreensão dos métodos de discriminação entre sinais de voz saudáveis e sinais afetados por patologias na laringe propostos nesta pesquisa. No próximo capítulo são apresentados os materiais utilizados, tais como a base de dados e os *softwares* utilizados, além da metodologia empregada nos dois métodos distintos utilizados.

Neste capítulo são descritos os materiais utilizados para o desenvolvimento dessa pesquisa, assim como a metodologia empregada.

4.1 – Base de dados

Nesta pesquisa são utilizados 53 sinais de voz de pacientes com laringes saudáveis e 112 sinais de voz de pacientes afetados por patologias laríngeas, sendo 51 sinais de voz de laringes afetadas por paralisia nas pregas vocais, 43 sinais de voz de laringes afetadas por edema de Reinke, e 18 sinais de voz de laringes afetadas por nódulos vocais.

Os sinais são da vogal sustentada /a/. Os sinais de voz de laringes saudáveis, originalmente amostrados a 50.000 amostras/s, foram sub-amostrados a 25.000 amostras/s para equiparar à taxa de amostragem dos sinais de voz de laringes patológicas. Tais sinais pertencem à base de dados *Disordered Voice Database, Model 4337*, da *Kay Elemetrics*, gravada pelo *Massachusetts Eye and Ear Infirmary (MEEI) Voice and Speech Lab* [46].

4.2 – Metodologia

O diagrama de blocos mostrado na Figura 4.1 ilustra, de maneira geral, a metodologia utilizada neste trabalho. Inicialmente, são gerados os gráficos de recorrência de cada sinal de voz da base de dados. Em seguida, para cada gráfico, é obtida a decomposição wavelet para um nível de decomposição através da aplicação da transformada wavelet bidimensional, resultando em quatro sub-bandas. A extração de características foi realizada utilizando dois métodos distintos. No método das sub-bandas foram aplicados 12 descritores de textura sobre cada sub-banda obtida, totalizando para cada sinal de voz 48 características a serem analisadas. E, no método dos descritores, foram extraídos 13 descritores de Haralick para cada sub-banda obtida, totalizando 52 características a serem analisadas. Após a extração de características, em cada método, foi realizada a classificação utilizando uma rede neural MLP. Além disso, no método 2, é utilizado um algoritmo de otimização, o *Particle Swarm Optimization - PSO*, com o objetivo de determinar quais combinações das características que resultam nos melhores casos de classificação entre sinais saudáveis e patológicos.

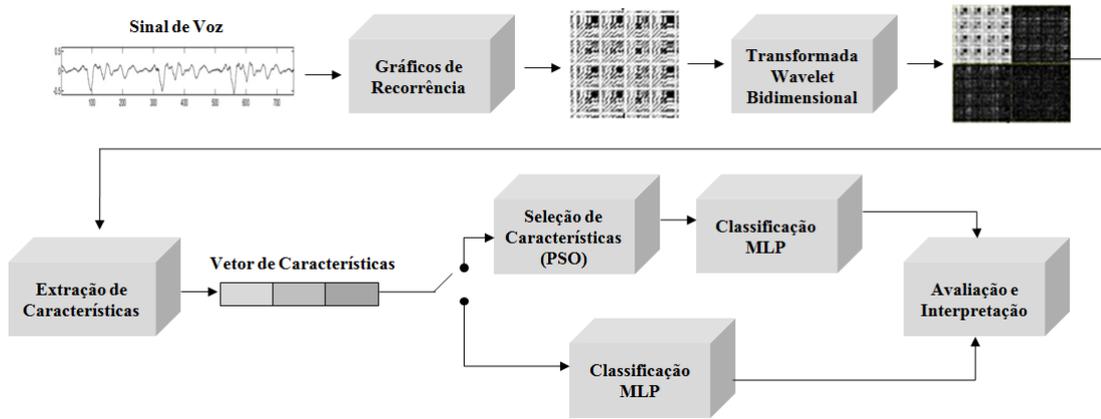


Figura 4.1 – Diagrama de blocos.

4.2.1 – Gráficos de Recorrência

Conforme foi mencionado anteriormente, os gráficos de recorrência permitem a análise de sinais não estacionários. Um sinal de voz é considerado estacionário quando considerados até aproximadamente 32 ms. Neste trabalho, a fim de investigar a eficiência do método considerando a não-estacionaridade dos sinais de voz, foram utilizados 400 ms dos sinais que têm 25000 amostras/s. Portanto, para a construção dos gráficos de recorrência são consideradas 10000 amostras do sinal, ou seja, o número de estados $N = 1000$.

Os gráficos de recorrência são obtidos utilizando o *software* VRA (*Visual Recurrence Analysis*), do qual são extraídos os parâmetros de imersão, τ e m .

4.2.2 – Transformada Wavelet Bidimensional

Após a obtenção dos gráficos de recorrência foi aplicada a decomposição wavelet com um nível de decomposição. Como as estruturas dos gráficos possuem descontinuidades, a família wavelet escolhida foi a de Haar [47]. A decomposição pela wavelet de Haar pode ser realizada através da convolução da imagem com as máscaras de Haar [48], que são ilustradas na Figura 4.2.

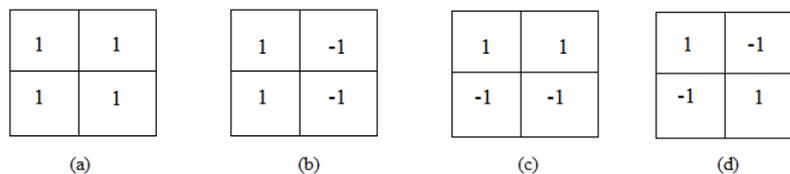


Figura 4.2 – Máscaras de Haar.

Portanto, a decomposição wavelet utilizando a família de Haar é obtida através da Equação 4.1.

$$Sb_k^j = \frac{1}{4} I^{j-1} \otimes H_k; \quad k = 1, 2, 3, 4. \tag{4.1}$$

em que \otimes denota o operador convolução, Sb_k^j é a sub-banda da imagem na resolução menor j , I^{j-1} é a sub-imagem 2×2 na resolução $j-1$, e H_k são as máscaras de Haar definidas na Figura 4.2. Desse modo, são obtidas para cada gráfico de recorrência quatro sub-bandas. A sub-banda Sb_1^j correspondente aos coeficientes de aproximação (LL), a sub-banda Sb_2^j correspondente aos coeficientes de detalhes verticais (LH), a sub-banda Sb_3^j corresponde aos detalhes horizontais (HL) e a sub-banda Sb_4^j corresponde aos detalhes diagonais (HH).

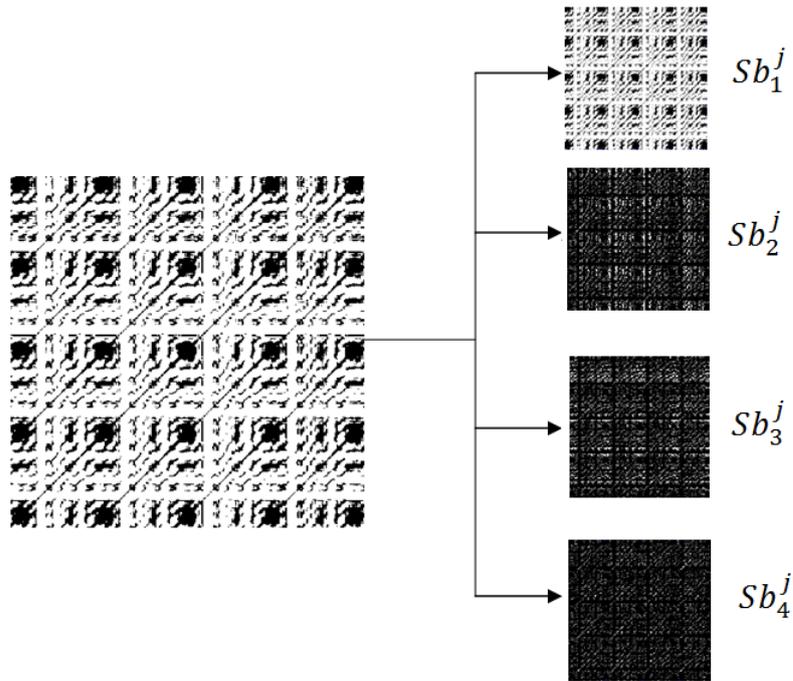


Figura 4.3 – Decomposição wavelet de um gráfico de recorrência de um sinal de voz utilizando a wavelet de Haar, com um nível de decomposição.

4.2.3 – Extração de Características

Para a extração de características foram considerados dois métodos distintos. O método das sub-bandas corresponde à análise da textura utilizando descritores aplicados diretamente nas sub-bandas obtidas pela decomposição wavelet. E o método dos descritores corresponde à análise da textura a partir da matriz de co-ocorrência através dos descritores de Haralick.

No método das sub-bandas, para cada sub-banda obtida a partir dos gráficos de recorrência, são extraídos 12 descritores: energia, desvio médio, entropia, média, norma (de ordem $r = 1$ e $r = 2$), momento (de ordens $r = 1, 2, 3$ e 4), suavidade e resíduo médio. Estas medidas são analisadas individualmente e em conjunto, formando conjuntos com as 12 medidas para cada sub-banda e, além disso, formando um conjunto com todas as medidas de todas as sub-bandas, totalizando 48 medidas.

No método dos descritores, para cada sub-banda wavelet, foram obtidas as matrizes de co-ocorrência com $\theta = 0^\circ, 45^\circ, 90^\circ$ e 135° . Para todos os ângulos analisados foi considerada a distância $d = 1$, totalizando quatro matrizes de co-ocorrência para cada sub-banda. E, para cada matriz de co-ocorrência foram obtidos os 13 descritores de Haralick: entropia, segundo momento angular, contraste,

momento da diferença inversa, correlação, média da soma, entropia da soma, variância da soma, variância, variância da diferença, informação de correlação 1 e informação de correlação 2.

Inicialmente, as medidas foram analisadas em conjunto, sendo divididas por sub-bandas e pelo ângulo considerado para gerar as matrizes de co-ocorrência. Essas medidas foram analisadas em conjunto por sub-bandas e também com todas as medidas juntas, totalizando 52 medidas (13 medidas dos coeficientes de aproximação, 13 medidas dos coeficientes horizontais, 13 medidas dos coeficientes verticais e 13 medidas dos coeficientes diagonais). Por exemplo, para cada sub-banda dos coeficientes de aproximação foram geradas as matrizes de co-ocorrência para $\theta = 0^\circ, 45^\circ, 90^\circ$ e 135° , cada uma considerando a distância $d = 1$, portanto, obtém-se quatro matrizes. Logo, para cada gráfico de recorrência são obtidos quatro sub-bandas e, para cada sub-banda, são obtidas quatro matrizes de co-ocorrência.

4.3 – Seleção de Características

4.3.1 – Otimização por enxame de partículas (*Particle Swarm Optimization - PSO*)

No método dos descritores, são obtidos os descritores de Haralick para cada sub-banda (com 1 nível de resolução) nas quatro direções da matriz de co-ocorrência, gerando um total de 52 características para cada direção. Este fato implica num custo computacional que pode ser otimizado se for utilizado um número menor de características.

Neste método, as características obtidas a partir da direção 45° da matriz de co-ocorrência se destacaram em relação às demais direções. Portanto, neste caso, foi utilizado o algoritmo PSO a fim de selecionar o conjunto de características que resulta nos melhores valores de acurácia para cada caso de classificação. Nesse problema, a função objetivo é definida como sendo a busca pelo melhor valor de acurácia com menor desvio padrão.

4.4 – Classificação

4.4.1 – Rede Neural MLP

Uma rede neural consiste em uma máquina que é busca modelar a maneira como o cérebro humano realiza uma determinada tarefa. Haykin [49] define Redes Neurais Artificiais (RNAs) como sistemas paralelos distribuídos e constituídos de unidades de processamento simples, chamadas "neurônios", que têm capacidade computacional relacionada à aprendizagem e à generalização. Nesse sistema, o conhecimento é adquirido por um processo chamado "treinamento" ou "aprendizagem", que é realizado pelo algoritmo de aprendizagem.

Os neurônios são unidades de processamento fundamentais de uma rede neural. A Figura 4.4 mostra um modelo de neurônio, o qual é constituído por três elementos básicos: sinapses ou elos de conexão, cada um caracterizado por um peso sináptico que é multiplicado por um sinal de entrada x_i ; um somador, que realiza a soma dos sinais de entrada, ponderados pelas respectivas sinapses do neurônio; e uma função de ativação, para restringir a amplitude da saída de um neurônio.

O algoritmo de aprendizagem tem a função de modificar os pesos sinápticos da rede de uma forma ordenada, a fim de realizar uma determinada tarefa. Além disso, uma rede neural é capaz de modificar sua própria topologia [49].

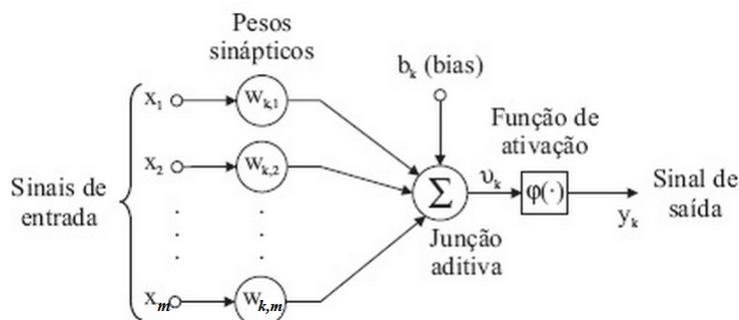


Figura 4.4 – Modelo de neurônio. Adaptado de [50].

Dentre os diversos tipos de RNAs, as redes perceptrons de múltiplas camadas (*Multilayer Perceptron-MLP*) vem obtendo sucesso em problemas de reconhecimento de padrões. As redes MLPs são ditas totalmente conectadas, pois possuem um conjunto de unidades sensoriais que formam a camada de entrada, uma ou mais camadas ocultas ou intermediárias de neurônios computacionais e uma camada de saída. O treinamento de uma MLP é do tipo supervisionado por meio de algoritmo de retropropagação do erro (*error-backpropagation*), detalhes sobre o algoritmo são apresentados em [49].

Uma rede neural MLP é tem três características: cada neurônio da rede possui uma função de ativação não-linear; a rede contém neurônios ocultos, que constituem a camada intermediária; e possui alto grau de conectividade, determinado pelas sinapses da rede. Na Figura 4.5 é apresentado um modelo de uma rede MLP.

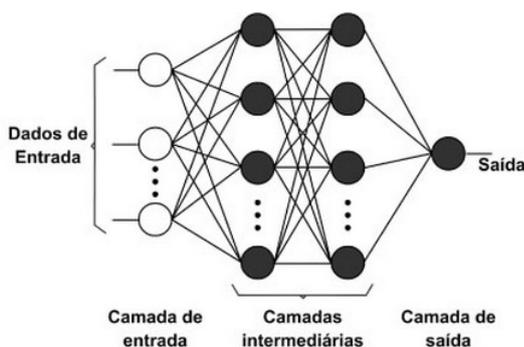


Figura 4.5 – Modelo de uma rede neural MLP. Adaptado de [51].

Neste trabalho, a rede neural MLP foi implementada com auxílio do *software* Matlab. O número de neurônios da camada de entrada (N_e) da rede foi igual ao número de medidas a serem analisadas para cada caso analisado. O número de neurônios da camada de saída utilizado foi $N_s = 2$, correspondentes as duas classes analisadas, por exemplo, saudável *versus* nódulo. Quanto ao número de neurônios da camada oculta N_i o critério utilizado é dado pela Equação 4.2.

$$N_i = \frac{N_e + N_s}{2} \quad (4.2)$$

O algoritmo de treinamento utilizado foi o *backpropagation*, com taxa de aprendizagem igual a 0,3 e taxa de momento igual a 0,2. Além disso, a classificação é realizada com validação cruzada, dividindo-se o conjunto de dados em k subconjuntos, dos quais $k - 1$ são utilizados para treinar o classificador e, o subconjunto restante é utilizado na fase de teste do classificador. Esse processo é repetido k vezes, para que todos os subconjuntos sejam empregados na fase de teste. Neste trabalho foi utilizado $k = 10$, portanto, 90% dos dados eram empregados na fase de treino e 10% na fase de teste, e esse processo foi repetido 10 vezes para cada caso de classificação.

Neste capítulo, foram apresentados os materiais e a metodologia utilizada nesta pesquisa. No capítulo seguinte serão descritos os dois métodos empregados para a classificação de sinais de voz. Para ambos os métodos são discutidos os resultados obtidos e, além disso, é feito um estudo comparativo entre o melhor resultado e alguns resultados obtidos nas pesquisas de Costa [6] e Vieira [7].

Resultados e Discussão

Neste trabalho são propostos dois métodos de extração de características, os quais são denominados método das sub-bandas, correspondente à extração das medidas das sub-bandas wavelets, e o método dos descritores, correspondente a extração dos descritores de Haralick. Os resultados obtidos na discriminação de sinais afetados por patologias da laringe utilizando cada um desses métodos são descritos nesse capítulo. A análise dos resultados será feita a partir das medidas de acurácia média, que pode ser obtida a partir da matriz de confusão ilustrada na Tabela 5.1, que contém os valores obtidos em um teste para identificação da presença ou ausência de uma determinada patologia, ou para identificação de algum tipo de patologia entre duas patologias possíveis.

Na matriz de confusão, um resultado identificado como positivo significa que o indivíduo pode ter a doença (verdadeiro positivo - VP) ou pode não tê-la (falso positivo - FP). Analogamente, quando um resultado for identificado como negativo, o indivíduo pode não ter a doença (verdadeiro negativo - VN) como pode tê-la (falso negativo - FN). As classes a serem analisadas são denominadas C_1 e C_2 , que indicam, respectivamente, a presença e a ausência de uma determinada patologia [6].

A acurácia mede a capacidade do teste de identificar corretamente quando há e quando não há presença da doença, ou no caso de classificação entre patologias, mede qual a taxa de classificação correta entre essas patologias. A Equação 5.1 define a medida da acurácia como o quociente entre todos os sinais corretamente classificados e o número total de sinais avaliados. Todos os casos de classificação foram realizados utilizando a validação cruzada dividindo o conjunto de dados em 10 subconjuntos, portanto, foram obtidas 10 medidas de acurácia e o resultado que consideramos é a média desses valores, a qual denominamos acurácia média.

$$\text{Acurácia} = \frac{VP + VN}{VP + VN + FP + FN} \quad (5.1)$$

Além da medida da acurácia média, também foram consideradas as medidas de sensibilidade e especificidade. A primeira representa a capacidade do teste em identificar corretamente a doença entre aqueles que a possuem. E a especificidade, mede a capacidade do teste em excluir corretamente aqueles que não possuem a doença. Essas medidas são descritas pelas Equações 5.2 e 5.3. Para o caso de classificação que discrimina duas patologias entre si, o valor da sensibilidade e especificidade tem significado de acordo com a Tabela 5.2.

$$\text{Sensibilidade} = \frac{VP}{VP + FN} \quad (5.2)$$

$$\text{Especificidade} = \frac{VN}{VN + FP} \quad (5.3)$$

Tabela 5.1 – Matriz de confusão

Classes reais	Classes preditas	
	C_1	C_2
C_1	Verdadeiro Positivo (VP)	Falso Positivo (FP)
C_2	Falso Negativo (FN)	Verdadeiro Negativo (VN)

Tabela 5.2 – Relação entre os casos de classificação entre patologias e valores de sensibilidade e especificidade.

Casos de classificação	Sensibilidade	Especificidade
EDM x NDL	detecta corretamente edema	detecta corretamente nódulo
EDM x PRL	detecta corretamente edema	detecta corretamente paralisia
PRL x NDL	detecta corretamente paralisia	detecta corretamente nódulo

Nas Tabelas 5.2 - 5.5, EDM corresponde ao conjunto de sinais com edema de Reinke, PRL corresponde ao conjunto de sinais com paralisia, NDL corresponde ao conjunto de sinais com nódulo, PTL corresponde ao conjunto formado por todos os sinais com alguma patologia (edema, nódulo e paralisia) e SDL corresponde ao conjunto de sinais saudáveis.

5.1 – Método das sub-bandas

Nesse método é realizada a extração de características utilizando 12 descritores de textura, os quais foram apresentados no Capítulo 2, na seção 2.3.1, e foram utilizados por Silva [9] em seu trabalho de dissertação. Tais medidas são extraídas a partir de cada sub-banda wavelet obtida pela decomposição dos gráficos de recorrência através da transformada wavelet com um nível de decomposição. Portanto, foram obtidas, para cada gráfico de recorrência, quatro sub-bandas.

A classificação foi feita utilizando uma rede neural MLP com validação cruzada considerando como vetor de característica os descritores de textura extraídos diretamente das sub-bandas wavelet, definidos na Seção 2.3.1.: energia, desvio médio, entropia, média, norma (de ordens $r = 1, 2$), momento (de ordens $r = 1, 2, 3, 4$), suavidade e resíduo médio, totalizando 12 medidas por sub-banda.

Na Tabela 5.3 são apresentados os resultados de cada caso de classificação obtidos quando os descritores foram calculados para cada uma das quatro sub-bandas: LL, correspondente aos coeficientes de aproximação e sub-bandas HL, LH e HH, correspondentes aos coeficientes de detalhes horizontais, verticais e diagonais, respectivamente. E, além disso, apresenta os resultados da classificação realizada com a combinação das 12 características para todas as sub-bandas, totalizando 48 medidas.

Os resultados mostram que o melhor conjunto de dados para classificação entre sinais saudáveis e sinais afetados por uma das patologias analisadas isoladamente é o conjunto formado pela união de todas as medidas de todas as sub-bandas, pois os melhores casos de classificação foram obtidos utilizando as 48

Tabela 5.3 – Acurácia média (%) para classificação utilizando descritores extraídos das sub-bandas.

Casos de classificação	Sub-bandas				
	LL	HL	LH	HH	Todas
SDLxPTL	75,07 ± 3,29*	74,52 ± 2,23	74,01 ± 2,09	73,46 ± 3,53	73,85 ± 3,92
SDLxPRL	68,36 ± 6,59	71,55 ± 5,01	70,36 ± 3,69	64,64 ± 5,34	72,09 ± 4,15*
SDLxEDM	69,89 ± 3,06	71,51 ± 4,86	64,88 ± 3,64	63,00 ± 5,77	74,00 ± 3,46*
SDLxNDL	74,76 ± 2,73	71,67 ± 2,25	71,90 ± 3,63	80,71 ± 4,23	87,62 ± 3,17*
EDMxNDL	68,95 ± 2,84	62,00 ± 3,89	70,48 ± 2,19*	65,48 ± 5,83	60,86 ± 4,24
EDMxPRL	57,44 ± 2,73*	53,22 ± 3,98	49,41 ± 3,41	54,34 ± 3,39	51,43 ± 4,83
PRLxNDL	74,17 ± 1,57*	74,17 ± 2,64	71,07 ± 2,05	70,24 ± 3,22	74,17 ± 2,94*

* melhores resultados de classificação.

medidas em conjunto, chegando a atingir uma acurácia de 87,62 % para o caso de classificação SDLxNDL. Neste caso, conforme pode ser observado no Apêndice A na Tabela A.4, os valores de especificidade e sensibilidade atingem, respectivamente, a média de 96,33 % e 60,00 %. Portanto, o sistema apresentou melhor desempenho em identificar corretamente os sinais saudáveis.

Essa taxa de sensibilidade de apenas 60 % pode ser explicada pela pequena quantidade de sinais afetados por nódulos (18 sinais), da base de dados, em relação ao número de sinais saudáveis (53 sinais), diminuindo o desempenho do classificador. Neste caso, sugere-se repetir a análise com outra base de dados, aumentando a quantidade de sinais para que o classificador possa ser treinado com maior eficiência.

Com relação à classificação entre patologias os melhores valores para a classificação foram obtidos para o caso PRLxNDL, atingindo uma acurácia de 74,17 % para as sub-bandas LL e HL, além de atingir o mesmo valor para o conjunto de todas as sub-bandas, variando apenas o desvio padrão. Desse modo, as sub-banda LL e HL tiveram um melhor desempenho, pelo fato de utilizar apenas 12 medidas para atingir o mesmo valor de acurácia média para o conjunto das sub-bandas, que tem 48 medidas. Neste caso, conforme pode ser observado na Tabela A.7 do Apêndice A, o melhor valor de especificidade foi obtido para a sub-banda HL, atingindo 15 %, e o melhor valor de sensibilidade foi obtido pela sub-banda LL, atingindo 98 %. Portanto, o método proposto é eficiente em identificar corretamente os sinais que tem paralisia. Contudo, é ineficiente na detecção correta de sinais com nódulo. No caso da classificação entre sinais saudáveis e afetados por nódulo, a pequena quantidade de sinais com nódulo (18 sinais) em relação à quantidade de sinais afetados por paralisia (51 sinais) também influenciou no desempenho do classificador.

No caso de classificação entre sinais com EDMxPRL, conforme a Tabela A.6 do Apêndice A, as taxas de acurácia média para todas as sub-bandas atingiram valores na faixa de 50 %. Tabelas contendo as taxas de acurácia, sensibilidade e especificidade, para o método das sub-bandas, são detalhadas no apêndice A. Portanto, o sistema é totalmente ineficiente nesse caso de classificação. A fim de melhorar os valores das taxas de acurácia, sensibilidade e especificidade foi realizado o método dos descritores, no qual são utilizados outros descritores de Haralick para extrair características utilizadas na classificação. Os resultados obtidos utilizando o método dos descritores são detalhados na próxima seção.

5.2 – Método dos descritores

Neste método a extração de características é realizada a partir de 13 descritores de Haralick, descritos na seção 2.3.3. do Capítulo 2, que foram extraídos de cada sub-banda obtida da decomposição wavelet dos gráficos de recorrência. A decomposição wavelet foi feita para um nível de decomposição e, portanto, foram geradas quatro sub-bandas, correspondendo aos coeficientes de aproximação (LL) e de detalhes (HL, LH e HH). A classificação entre sinais saudáveis e patológicos foi realizada utilizando uma rede neural MLP e, o melhor caso de classificação foi analisado também com o algoritmo de otimização PSO, a fim de determinar qual o melhor subconjunto de características que resultam nos melhores resultados para cada caso de classificação.

A extração de características foi realizada a partir da decomposição wavelet dos sinais de voz do banco de dados, para um nível de decomposição. Portanto, foram obtidas as quatro sub-bandas LL, HL, LH e HH. Em seguida, foram calculadas quatro matrizes de co-ocorrência para cada sub-banda, nas direções 0° , 45° , 90° , e 135° , todas considerando a distância $d = 1$. Finalmente, o conjunto de dados foi dividido em quatro subconjuntos de dados, ou seja, para cada direção da matriz de co-ocorrência foi realizada a classificação para as quatro sub-bandas individualmente (com 13 características cada) e incluindo o de medidas com todas as características em conjunto (com 52 características).

A Tabela 5.4 apresenta os melhores resultados obtidos na classificação utilizando como características os descritores de Haralick e, como classificador, uma rede neural MLP. Dentre as possíveis direções da matriz de co-ocorrência analisadas, pode-se observar que a direção $\theta = 45^\circ$ atingiu os melhores resultados em três casos de classificação (SDLxPRL, SDLxEDM e EDMxNDL). No Apêndice B são detalhados os valores de acurácia, sensibilidade e especificidade para todos os casos de classificação.

Além disso, no caso SDLxPTL, conforme Tabela B.2, os valores obtidos com o conjunto de medidas extraídos de todas as sub-bandas atingiram uma acurácia de 82,42 %, que é ligeiramente menor do que o melhor valor para esse caso de classificação, obtido na direção $\theta = 90^\circ$ conforme a Tabela 5.4.

As medidas extraídas a partir da matriz de co-ocorrência na direção $\theta = 45^\circ$ se destacaram na maioria dos casos de classificação. Também é possível observar na Tabela 5.4 que os valores de acurácia, para os casos de classificação entre patologias, apresentaram taxas inferiores a 80 %. E, além disso, os valores de especificidade atingem valores inferiores a 50 % nos casos EDMxNDL e PRLxNDL.

A fim de melhorar os resultados obtidos e otimizar o sistema, reduzindo o número de características a serem utilizadas na classificação, foi realizada a seleção de características utilizando o algoritmo de otimização PSO. Portanto, todas as 52 medidas extraídas de todas as sub-bandas foram inseridas no classificador na fase de treino, a fim de selecionar o conjunto de medidas capaz de atingir as maiores taxas de acurácia.

Na Tabela 5.5 é apresentada uma comparação entre os resultados obtidos com o conjunto de todas as 52 medidas extraídas para a matriz de co-ocorrência na direção $\theta = 45^\circ$ e os resultados obtidos após a seleção de características com o auxílio do PSO. Os valores de sensibilidade são indicados por Sen e, a especificidade por Espec.

No apêndice C são apresentadas as tabelas que contém o conjunto de características selecionadas pelo PSO, nas quais o valor 1 indica que a medida foi selecionada e o valor 0 indica que a medida não foi selecionada. Para cada caso de classificação ocorreu a redução do conjunto de

Tabela 5.4 – Resumo dos melhores resultados para cada caso de classificação utilizando 13 descritores de Haralick com a rede MLP.

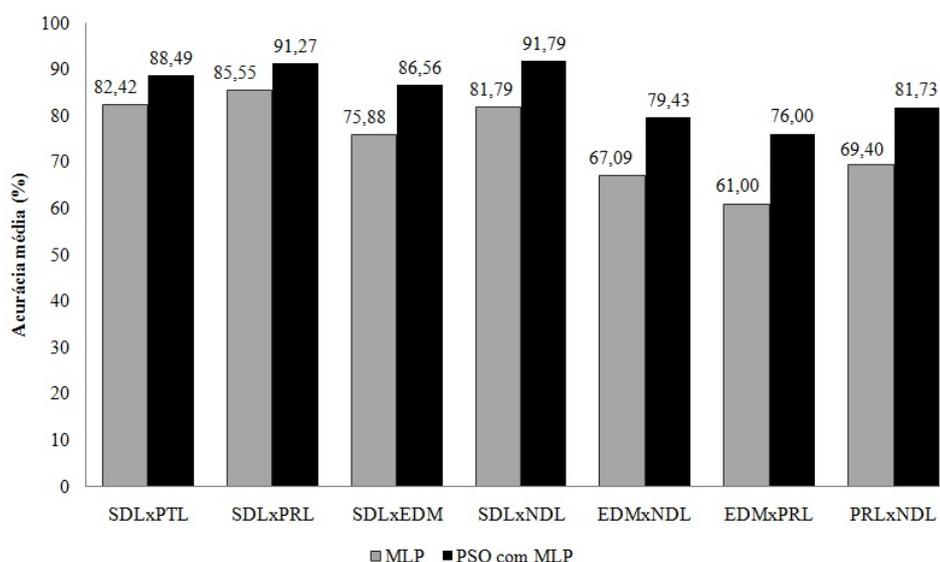
	θ	Sub-banda	Acurácia média	Sensibilidade	Especificidade
SDLxPTL	90°	Todas	82,98 ± 2,41	93,71	59,67
SDLxPRL	45°	Todas	85,55 ± 1,62	86,33	85,00
SDLxEDM	45°	HH	80,33 ± 6,20	72,50	87,00
SDLxNDL	135°	Todas	86,25 ± 3,55	70,00	92,33
EDMxNDL	45°	HL	70,95 ± 4,12	93,00	20,00
EDMxPRL	135°	LH	63,44 ± 5,17	67,00	60,00
PRLxNDL	0°	HL	75,48 ± 2,97	95,00	25,00

características (52 medidas) para um número de características indicado por *Num* na Tabela 5.5, que foram selecionadas pelo PSO.

Tabela 5.5 – Resultados com MLP X PSO com MLP.

	MLP				PSO com MLP				
	Acurácia	Desvio	Sen	Espec	Acurácia	Desvio	Sen	Espec	Num
SDLxPTL	82,42	2,59	92,87	61,00	88,49	5,10	87,58	63,67	23
SDLxPRL	85,55	1,62	86,33	85,00	91,27	4,49	74,00	93,00	30
SDLxEDM	75,88	5,85	62,50	87,33	86,56	4,58	63,00	88,33	31
SDLxNDL	81,79	4,72	50,00	91,33	91,79	3,20	50,00	96,33	28
EDMxNDL	67,09	2,51	88,50	15,00	79,43	2,76	98,00	10,00	27
EDMxPRL	61,00	5,89	66,00	57,67	76,00	4,86	67,50	61,00	29
PRLxNDL	69,40	5,98	86,00	25,00	81,73	2,32	98,00	10,00	27

Na Figura 5.1 é possível observar que os resultados obtidos, utilizando o conjunto de características selecionadas pelo PSO, melhoraram as taxas de acurácia em todos os casos de classificação. Além disso, houve uma redução do custo computacional, visto que houve a redução do número de medidas necessárias para atingir melhores taxas de acurácia.

**Figura 5.1** – Acurácia média (%) utilizando apenas redes neurais MLP e utilizando PSO com MLP.

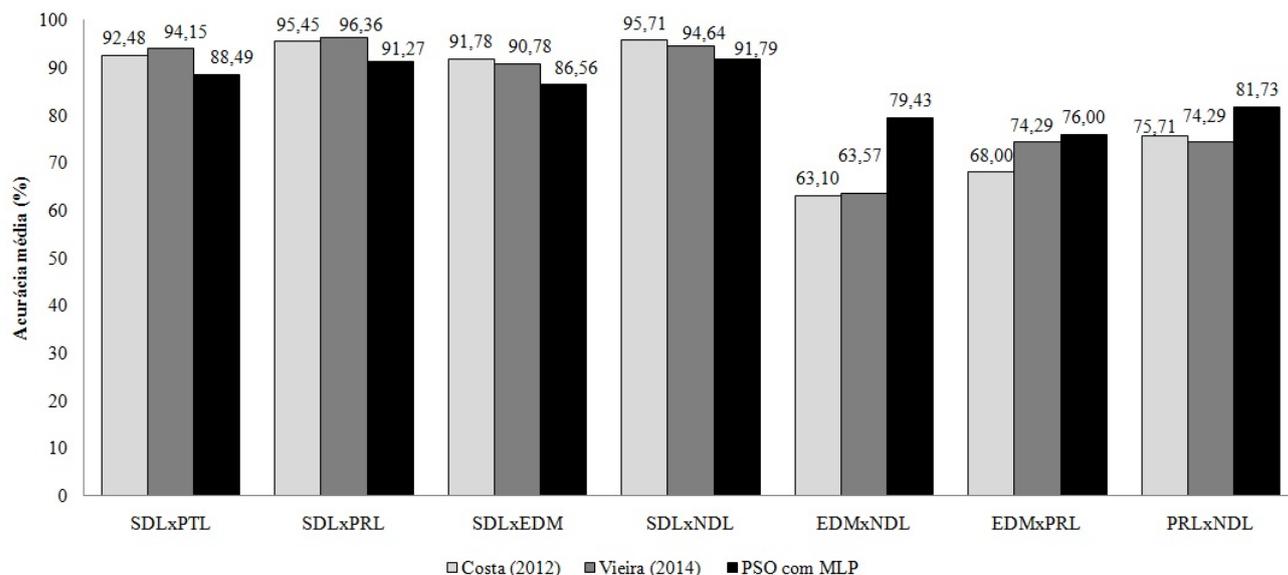


Figura 5.2 – Acurácia média (%) obtidas nos trabalhos de Costa [5] e Vieira [7] em comparação com os melhores resultados obtidos nesta pesquisa utilizando PSO com MLP.

A Figura 5.2 apresenta um estudo comparativo entre os resultados obtidos nessa pesquisa pelo método dos descritores, utilizando o PSO para selecionar características, e os resultados obtidos exclusivamente com as medidas de quantificação da recorrência (MQR) nos trabalhos de Costa [6] e Vieira [7]. No Apêndice C são detalhadas as características selecionada pelo PSO em cada caso de classificação.

Em Vieira [7], os resultados apresentados na Figura 5.2 foram obtidos a partir de sinais de voz sem segmentação, para os quais foram consideradas diversas taxas de recorrência na obtenção do gráfico de recorrência desses sinais. No entanto, em Costa [6], todas as características foram obtidas a partir da segmentação dos sinais. Ambos utilizaram os *software* RQA (*Recurrence Quantification Analysis*), diferentemente desta pesquisa, que utilizou o *software* VRA (*Visual Recurrence Analysis*) para obter os gráficos de recorrência.

Algumas medidas de quantificação de recorrência utilizadas por Vieira [7] para analisar o sinal de voz foram: Taxa de Recorrência, Determinismo, Comprimento máximo das estruturas diagonais, Entropia de Shannon, Laminaridade, Tempo de Permanência, Comprimento máximo das estruturas verticais, Transitividade e Divergência. Utilizando as medidas de forma combinada na detecção da presença de patologias chegou-se a uma precisão de 90%. Dentre as medidas avaliadas, a que mais se destacou na discriminação entre saudáveis e patologias (edemas, nódulos e paralisia) foi o comprimento médio das linhas diagonais. Na classificação entre patologias o melhor resultado foi obtido utilizando o tempo de recorrência do tipo 2.

Costa [6] também realizou a análise dos sinais de voz baseada na combinação das medidas de quantização de recorrência (MRQ), além de outras medidas não-lineares. Portanto, utilizou de vetores híbridos formados por características obtidas utilizando medidas de análise não linear (MNL), MQR e coeficientes extraídos da análise preditiva linear (LPC). No processo de classificação foi utilizada a técnica de análise discriminante com as funções linear ou quadrática, com validação cruzada.

De acordo com os gráficos da Figura 5.2, os valores de acurácia média na classificação entre patologias apresentou um aumento significativo com relação aos trabalhos de Costa [6] e Vieira [7], principalmente nos casos EDMxNDL e PRLxNDL. Nos casos de classificação entre sinais saudáveis e sinais afetados por alguma patologia, os trabalhos de Costa e Vieira apresentam valores superiores de acurácia com relação ao método proposto. No entanto, a classificação entre patologias neste trabalho mostrou-se mais eficiente.

Nas Figuras 5.3 - 5.6 são apresentados exemplos de gráficos de recorrência de sinais de voz da base de dados utilizada nesta pesquisa. Em cada item (a) são apresentados um exemplo de gráfico de recorrência que ilustra o padrão de determinada classe (saudável ou alguma das patologias). Nos demais itens (b)-(d) são apresentados exemplos de sinais de voz de uma mesma classe que apresentam gráficos de recorrência com estruturas diferentes do padrão para cada classe, mostrando a variabilidade dos mesmos. Portanto, pode-se perceber que mesmo pertencente a uma mesma classe de sinais, o gráfico de recorrência pode apresentar estruturas diferentes na imagem.

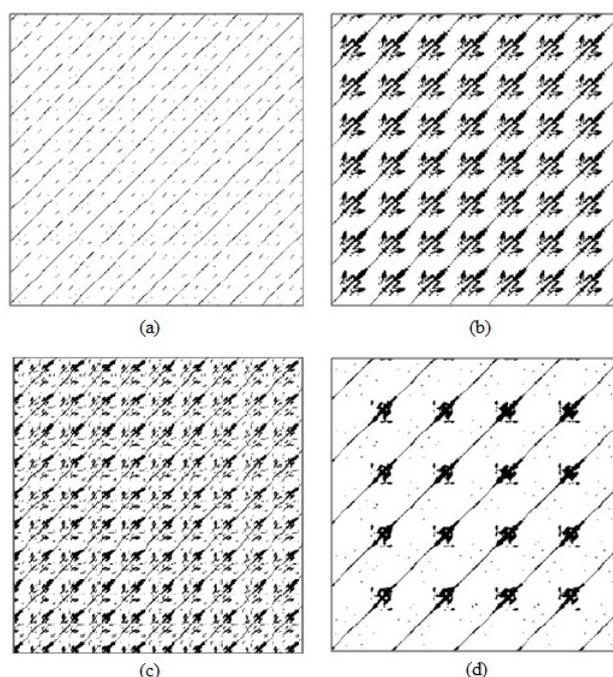


Figura 5.3 – Gráficos de recorrência de sinais de vozes saudáveis: (a) feminino, 26 anos; (b) masculino, 32 anos; (c) feminino, 37 anos; (d) masculino, 30 anos.

Na Figura 5.7, observa-se que as classes consideradas nessa pesquisa (saudável, edema, nódulo e paralisia) podem conter sinais de voz cuja textura dos gráficos de recorrência são semelhantes. Isto implica na dificuldade de classificação entre os sinais. A variabilidade interclasse, assim como a semelhança entre elementos de classes distintas, decorre de diversos fatores. Um fator a ser considerado é o fato da base de dados conter sinais de voz de pessoas do sexo feminino e do sexo masculino, que têm faixas de frequência fundamental distintas. Além disso, fatores como o grau de severidade da patologia também resultam em texturas distintas.

Desse modo, torna-se difícil caracterizar uma textura contida nesses gráficos a partir de uma única medida. Portanto, é necessário um método que utilize mais de uma medida para quantificar a textura

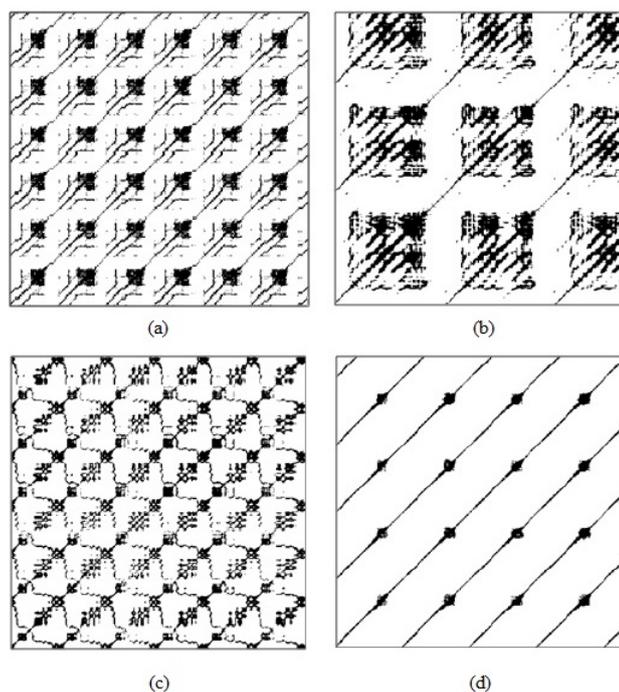


Figura 5.4 – Gráficos de Recorrência de vozes afetadas por edema de Reinke: (a) bilateral, feminino; (b) bilateral, feminino; (c) bilateral, feminino; (d) indefinido, feminino.

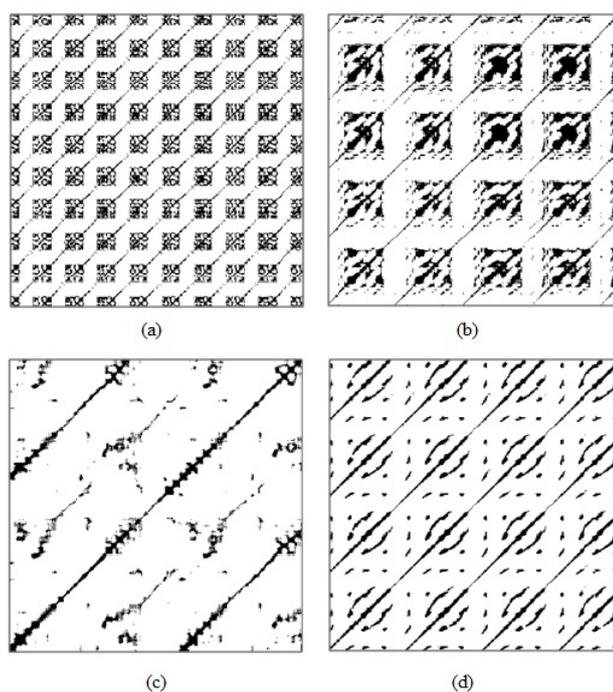


Figura 5.5 – Gráficos de recorrência de vozes afetados por nódulo: (a) indefinido, feminino; (b) unilateral esquerdo, masculino; (c) bilateral, feminino; (d) bilateral, feminino;

contida na imagem. Logo, o método proposto mostrou-se robusto, visto que, mesmo dentro de uma grande variabilidade interclasse e da semelhança entre alguns gráficos de sinais de classes diferentes, foi capaz de atingir taxas de acurácia significativas.

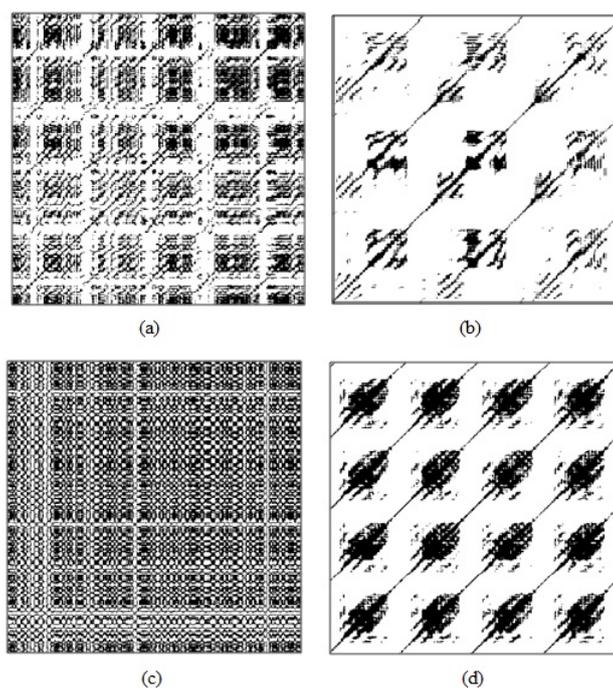


Figura 5.6 – Gráficos de recorrência de sinais de vozes afetadas por paralisia: (a) unilateral esquerdo, feminino; (b) unilateral esquerdo, feminino; (c) unilateral esquerdo, masculino; (d) unilateral esquerdo, feminino.

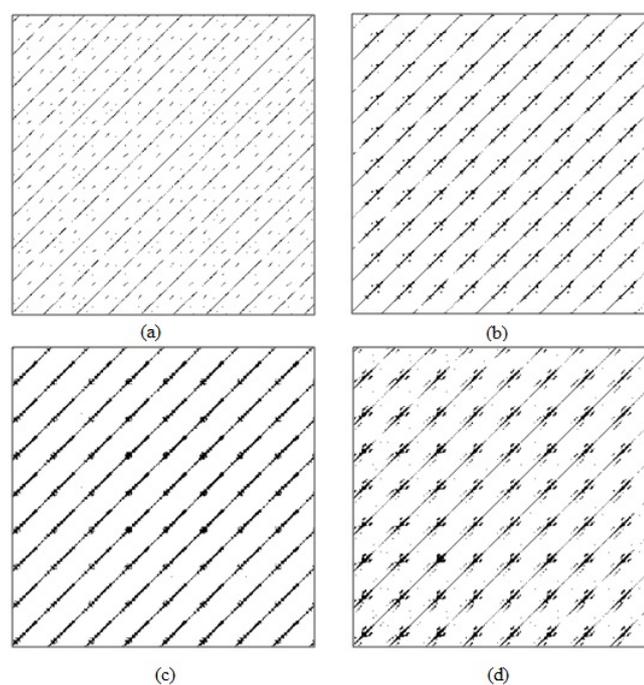


Figura 5.7 – Gráficos de recorrência: (a) sinal de voz saudável; (b) sinal de voz afetado por edema de Reinke; (c) sinal de voz afetado por nódulo; e (d) sinal de voz afetado por paralisia.

Os resultados de maior destaque obtidos nesta pesquisa foram apresentados neste capítulo. Nos apêndices A e B podem ser consultados os demais resultados. A partir da análise e discussão desses resultados são apresentadas, no próximo capítulo, as considerações finais dessa pesquisa.

Considerações Finais

A análise acústica de sinais de voz para discriminação entre sinais saudáveis e patológicos pode ser utilizada para desenvolver sistemas de auxílio ao diagnóstico. Pesquisas mostram que cada patologia pode apresentar características capazes de diferenciá-las, quando comparadas com sinais saudáveis ou com outra patologia.

Neste trabalho foi realizada a análise acústica baseada no modelo não linear de produção da voz a partir da análise dos gráficos de recorrência dos sinais de voz. A partir da observação dos gráficos de recorrência de sinais de voz é possível detectar a presença de características de textura na imagem desses gráficos. A partir disso, surgiu o interesse de aplicar técnicas de PDI, tais como a transformada wavelet e os descritores de Haralick, comumente empregadas na análise de textura, para quantificar a textura contida nesses gráficos.

Os resultados obtidos foram satisfatórios, visto que, comparando com outros trabalhos da literatura apresentaram resultados relevantes. Além de mostrar um bom desempenho nos casos de detecção da presença ou ausência de alguma patologia, o método dos descritores atingiu taxas de classificação significativas na diferenciação entre patologias (EDMxNDL, EDMxPRL e PRLxNDL).

A utilização do algoritmo de otimização por enxame de partículas (PSO) na etapa de seleção de características contribuiu, significativamente, na obtenção dos melhores resultados. Este foi um diferencial com relação aos trabalhos de Costa [6] e Vieira [7], determinando o melhor conjunto de características em cada caso de classificação de uma maneira automática.

Neste trabalho foram empregados dois métodos para a discriminação de sinais de voz saudáveis e afetados por alguma patologia da laringe. No método das sub-bandas foram utilizadas medidas de análise de textura, propostas no trabalho de Silva [9]. Neste método, as medidas são extraídas diretamente das sub-bandas obtidas pela decomposição wavelet com um nível de resolução. Os resultados obtidos utilizando este método não atingiram altas taxas de classificação. Portanto, foi necessária a investigação do método dos descritores, a fim de melhorar os valores de acurácia, sensibilidade e especificidade.

No método dos descritores, foram utilizados os descritores de Haralick para a análise da textura dos gráficos de recorrência. Para cada sub-banda wavelet foram calculadas as matrizes de co-ocorrência em quatro direções distintas. E, na direção $\theta = 45^\circ$, foram obtidos a maioria dos melhores resultados de classificação, indicando que nessa direção estariam as características mais relevantes para caracterizar os sinais.

Por fim, as 52 medidas extraídas nessa direção foram utilizadas pelo PSO a fim de determinar qual o melhor conjunto de características capaz de discriminar entre sinais saudáveis e patológicos. Desse modo, observou-se que os valores de acurácia para classificação entre patologias aumentou

significativamente. Portanto, o método proposto mostrou-se robusto e eficiente, visto que mesmo diante da variabilidade interclasse e da semelhança entre sinais de classes distintas da base de dados, foi capaz de discriminar os sinais de voz. Logo, é apresentada uma nova proposta para a análise dos gráficos de recorrência, que pode ser utilizada em diversas situações modeladas por sistemas dinâmicos.

6.1 – Contribuições da pesquisa

- ▣ Uma nova abordagem para análise dos gráficos de recorrência, baseada em técnicas de PDI, a partir da análise de textura.
- ▣ A técnica demonstrou ser eficiente na classificação entre patologias.

6.2 – Sugestões para Trabalhos Futuros

As sugestões para trabalhos futuros estão elencadas a seguir.

- ▣ Investigar a utilização de outras famílias wavelet, assim como outros níveis de resolução;
- ▣ Investigar a utilização de outros algoritmos para análise de texturas, como o *Local Binary Pattern* (LBP);
- ▣ Utilizar outra base de dados ou separar os sinais pelo gênero e/ou pelo grau de severidade da patologia;
- ▣ Combinar as medidas utilizadas neste trabalho com outras medidas extraídas a partir do modelo linear de produção da voz ou não-linear.

Referências Bibliográficas

- [1] N. C. d. Almeida, "Sistema inteligente para diagnóstico de patologias na laringe utilizando máquinas de vetor de suporte," *Universidade Federal do Rio Grande do Norte. Dissertação de mestrado*, 2010.
- [2] A. Parraga, "Aplicação da transformada wavelet packet na análise e classificação de sinais de vozes patológicas," *Universidade Federal do Rio Grande do Sul. Dissertação de Mestrado*, 2002.
- [3] R. T. S. Carvalho, "Transformada wavelet na detecção de patologias da laringe," *Universidade Federal do Ceará. Dissertação de mestrado*, 2012.
- [4] C. D. P. Crovato, "Classificação de sinais de voz utilizando a transformada wavelet packet e redes neurais artificiais," *Universidade Federal do Rio Grande do Sul. Dissertação de Mestrado*, 2004.
- [5] S. L. d. N. C. Costa, "Análise acústica, baseada no modelo linear de produção da fala, para discriminação de vozes patológicas," *Universidade Federal de Campina Grande. Tese de doutorado*, 2008.
- [6] W. C. d. A. Costa, "Análise dinâmica não linear de sinais de voz para detecção de patologias laríngeas," *Universidade Federal de Campina Grande. Tese de doutorado*, 2012.
- [7] V. J. D. Vieira, "Avaliação de distúrbios da voz por meio de análise de quantificação de recorrência," *Instituto de Educação, Ciência e Tecnologia da Paraíba. Dissertação de mestrado*, 2014.
- [8] J.-P. Eckmann, S. O. Kamphorst, and D. Ruelle, "Recurrence plots of dynamical systems," *Europhys. Lett*, vol. 4, no. 9, pp. 973–977, 1987.
- [9] R. D. d. Silva, "Segmentação de imagens baseada em descritores de textura extraídos por transformadas wavelets," *Universidade Federal do Pará. Dissertação de mestrado*, 2008.
- [10] N. Marwan, "Encounters with neighbours: current developments of concepts based on recurrence plots and their applications," 2003.
- [11] M. L. Alves, "Reconhecimento da rugosidade em imagens monocromáticas por meio de análise de textura. tese de doutorado," *Universidade Federal Fluminense. Tese de doutorado*, 2011.
- [12] A. Materka, M. Strzelecki, *et al.*, "Texture analysis methods—a review," *Technical university of lodz, institute of electronics, COST B11 report, Brussels*, pp. 9–11, 1998.
- [13] P. e. D. D. V. Wouwer, G. V.; Schenunders, "Statistical texture characterization from discrete wavelet representations," *IEEE Transactions on Image Processing, Vol. 8, No. 4.*, 1999.

- [14] V. Vieira, S. Costa, W. Costa, S. Correia, and F. Assis, "Discriminação de sinais de voz com análise de quantificação de recorrência e redes neurais mlp," *Anais do XXXI Simpósio Brasileiro de Telecomunicações (SBrT 2013)*, pp. 1–4, 2013.
- [15] H. B. Soares, "Análise e classificação de imagens de lesões da pele por atributos de cor, forma e textura utilizando máquina de vetor de suporte," *Universidade Federal do Rio Grande do Norte. Tese de doutorado*, 2008.
- [16] F. A. d. A. Sobrinho, "Medida da dispersão da periodicidade de um sinal de voz normal e voz patológica através da seção de poincaré," *Universidade de São Paulo. Tese de doutorado*, 2011.
- [17] M. E. Dajer, "Análise de sinais de voz por padrões visuais de dinâmica vocal," *Universidade de São Paulo. Tese de doutorado*, 2010.
- [18] P. R. Scalassara, "Utilização de medidas de previsibilidade em sinais de voz para discriminação de patologias de laringe," *Universidade de São Paulo. Tese de doutorado*, 2009.
- [19] S. Barbon Júnior, "Identificação de patologias na laringe com base na discriminative paraconsistent machine (dpm)," *Universidade de São Paulo. Tese de doutorado*, 2011.
- [20] M. Behlau, "Voz," *O livro do especialista*, vol. 1, 2001.
- [21] M. d. O. Rosa, "Modelagem da laringe: da biologia ao computador," *Revista das Letras*, Vol. 30, 2011.
- [22] S. M. Zitta, "Análise perceptivo-auditiva e acústica em mulheres com nódulos vocais," *Universidade Tecnológica Federal do Paraná. Dissertação de mestrado.*, 2005.
- [23] M. L. Sulica, "Specialized care for the voice," *Online: <http://voicemedicine.com/>*, Acessado em 12 janeiro de 2015.
- [24] E. G. Souza, "Caracterização de sistemas dinâmicos através de gráficos de recorrência," *Universidade Federal do Paraná. Tese de doutorado*, 2008.
- [25] F. Takens, "Detecting strange attractors in turbulence," *Springer*, 1981.
- [26] C. A. B. Castanón, "Recuperação de imagens por conteúdo através de análise multiresolução por wavelets," *Universidade de São Paulo. Dissertação de mestrado*, 2003.
- [27] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 11, no. 7, pp. 674–693, 1989.
- [28] C. S. Burrus, R. A. Gopinath, and H. Guo, "Introduction to wavelets and wavelet transforms," vol. 998, 1998.
- [29] I. Daubechies, "Ten lectures on wavelets," vol. 61, 1992.
- [30] R. C. Gonzalez and R. E. Woods, "Digital image processing, 2nd," *SL: Prentice Hall*, vol. 2, 2002.
- [31] S. I. N. Correia, "Reconhecimento de caracteres manuscritos usando wavelets," *Universidade Federal de Campina Grande. Tese de doutorado*, 2005.

- [32] J. R. Parker, *Algorithms for image processing and computer vision*. John Wiley & Sons, second edition ed., 2010.
- [33] J. P. R. d. Nascimento, “Análise e classificação de imagens baseadas em características de textura utilizando matrizes de co-ocorrência,” *Universidade Federal do Paraná. Dissertação de mestrado*, 2003.
- [34] P. L. de Paula Filho, “Reconhecimento de espécies florestais através de imagens macroscópicas,” *Universidade Federal do Paraná. Tese de doutorado*, 2012.
- [35] R. M. Haralick, K. Shanmugam, and I. H. Dinstein, “Textural features for image classification,” *Systems, Man and Cybernetics, IEEE Transactions on*, no. 6, pp. 610–621, 1973.
- [36] L. de Paula Batista Benevides, “Detecção de campo de futebol através da análise de texturas.” *Online em <http://slideplayer.com.br/slide/358094/>*, Acessado em 10 de março de 2015.
- [37] D. P. Barbosa, “Avaliação de descritores texturais geoestadísticos de Haralick para o reconhecimento de plantas daninhas,” *Universidade Federal de Viçosa. Dissertação de mestrado*, 2009.
- [38] G. C. F. Medeiros, “Uso de texturas para o acompanhamento da evolução do tratamento de úlceras dermatológicas,” *Universidade de São Paulo. Dissertação de mestrado*, 2001.
- [39] J. A. H. . P. J. R. R. . M. P. M. d. A. Honda, Marcelo Ossamu ; Rodrigues, “Uso de atributos de textura para minimização do *gap* semântico em sistemas de recuperação de imagens baseada em conteúdo (cbir),” *VIII Congresso Brasileiro de Informática em Saúde.*, 2002.
- [40] H. A. d. Silva, “Algoritmo de otimização multinível,” *Pontifícia Universidade Católica do Paraná. Tese de doutorado*, 2012.
- [41] A. A. d. Pina, “Metodologias de análise, síntese e otimização de sistemas para produção de petróleo offshore através de metamodelos e enxame de partículas.” *Universidade Federal do Rio de Janeiro. Tese de doutorado*, 2010.
- [42] T. Back, U. Hammel, and H.-P. Schwefel, “Evolutionary computation: Comments on the history and current state,” *Evolutionary computation, IEEE Transactions on*, vol. 1, no. 1, pp. 3–17, 1997.
- [43] R. C. Eberhart and J. Kennedy, “A new optimizer using particle swarm theory,” vol. 1, pp. 39–43, 1995.
- [44] N. N. Siqueira, “Otimização aplicada ao planejamento de políticas de testes em sistemas nucleares por enxame de partículas. tese de doutorado.” *Universidade Federal do Rio de Janeiro. Dissertação de mestrado.*, 2006.
- [45] C.-J. Tu, L.-Y. Chuang, J.-Y. Chang, C.-H. Yang, *et al.*, “Feature selection using pso-svm,” *IAENG International journal of computer science*, vol. 33, no. 1, pp. 111–116, 2007.
- [46] K. Elemetrics, “Kay elemetrics corp. disordered voice database,” 1994.

- [47] A. Busch, W. W. Boles, and S. Sridharan, "Texture for script identification," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 11, pp. 1720–1732, 2005.
- [48] S.-W. Lee, C.-H. Kim, H. Ma, and Y. Y. Tang, "Multiresolution recognition of unconstrained handwritten numerals with wavelet transform and multilayer cluster neural network," *Pattern Recognition*, vol. 29, no. 12, pp. 1953–1961, 1996.
- [49] H. Simon, *Redes Neurais—Princípios e Prática*. Bookman, 2 ed., 2001.
- [50] S. S. Z. . E. F. S. D. F. de Carvalho; Salassier Bernardo, "Estimação da evapotranspiração de referência no Estado do Rio de Janeiro usando redes neurais artificiais," *Revista Brasileira de Engenharia Agrícola e Ambiental*, vol. 12, no. 2, pp. 174–180, 2008.
- [51] D. V. Fiorin, F. R. Martins, N. J. Schuch, and E. B. Pereira, "Aplicações de redes neurais e previsões de disponibilidade de recursos energéticos solares," *Revista Brasileira de Ensino de Física*, vol. 33, no. 1, p. 1309, 2011.

APÊNDICES

Método das sub-bandas

Neste apêndice serão apresentados os resultados de acurácia, desvio da acurácia, especificidade e sensibilidade obtidos com a utilização do método das sub-bandas para a extração de características. As sub-bandas estão indicadas nas tabelas da seguinte forma: LL, HL, LH e HH correspondem, respectivamente, aos coeficientes de aproximação e aos coeficientes de detalhes horizontais, verticais e diagonais. O conjunto formado por todos os sinais que têm alguma patologia está indicado na tabela como Patológico. E, o conjunto formado por todas as características obtidas em conjunto, ou seja, para todas as sub-bandas, está indicado na tabela por Todas.

Tabela A.1 – Classificação Saudável X Patológico

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	75,07	74,52	74,01	73,46	73,85
Desvio da acurácia (%)	3,29	2,23	2,09	3,53	3,92
Especificidade (%)	58,00	33,33	43,67	37,00	36,67
Sensibilidade (%)	82,95	93,71	88,26	91,06	91,89

Tabela A.2 – Classificação Saudável X Paralisia

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	68,36	71,55	70,36	64,64	72,09
Desvio da acurácia (%)	6,59	5,01	3,69	5,34	4,15
Especificidade (%)	78,00	75,67	78,67	65,33	79,00
Sensibilidade (%)	59,00	67,67	62,33	62,67	64,67

Tabela A.3 – Classificação Saudável X Edema

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	69,89	71,51	64,88	63,00	74,00
Desvio da acurácia (%)	3,06	4,86	3,64	5,77	3,46
Especificidade (%)	84,33	85,67	85,33	69,33	84,00
Sensibilidade (%)	51,50	54,50	40,00	55,00	63,50

Tabela A.4 – Classificação Saudável X Nódulo

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	74,76	71,67	71,90	80,71	87,62
Desvio da acurácia (%)	2,73	2,25	3,63	4,23	3,17
Especificidade (%)	89,33	92,67	86,33	94,00	96,33
Sensibilidade (%)	30,00	10,00	30,00	40,00	60,00

Tabela A.5 – Classificação Edema X Nódulo

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	68,95	62,00	70,48	65,48	60,86
Desvio da acurácia (%)	2,84	3,89	2,19	5,83	4,24
Especificidade (%)	0,00	5,00	5,00	10,00	15,00
Sensibilidade (%)	97,50	85,50	98,00	90,00	83,00

Tabela A.6 – Classificação Edema X Paralisia

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	57,44	53,22	49,41	54,34	51,43
Desvio da acurácia (%)	2,73	3,98	3,41	3,39	4,83
Especificidade (%)	74,67	62,33	61,00	76,00	49,33
Sensibilidade (%)	37,50	42,50	33,50	28,50	53,50

Tabela A.7 – Classificação Paralisia X Nódulo

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	74,17	74,17	71,07	70,24	74,17
Desvio da acurácia (%)	1,57	2,64	2,05	3,22	2,94
Especificidade (%)	5,00	15,00	5,00	0,00	15,00
Sensibilidade (%)	98,00	94,00	94,00	94,67	96,00

Método dos descritores

Neste apêndice serão apresentados os resultados de acurácia, desvio da acurácia, especificidade e sensibilidade obtidos com a utilização do método dos descritores para a extração de características. As sub-bandas estão indicadas nas tabelas da seguinte forma: LL, HL, LH e HH correspondem, respectivamente, aos coeficientes de aproximação e aos coeficientes de detalhes horizontais, verticais e diagonais.

O conjunto formado por todos os sinais que têm alguma patologia está indicado na tabela como Patológico. E, o conjunto formado por todas as características obtidas em conjunto, ou seja, para todas as sub-bandas, está indicado na tabela por Todas. Além disso, os resultados apresentados estão divididos pela direção da matriz de co-ocorrência, ou seja, para cada direção da matriz de co-ocorrência é apresentada uma tabela com os resultados de cada caso de classificação.

Tabela B.1 – Classificação Saudável X Patológico utilizando 13 descritores com a matriz de co-ocorrência na direção $\theta = 0^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	76,40	79,34	81,25	75,88	82,31
Desvio da acurácia (%)	2,40	3,42	2,47	1,78	2,12
Especificidade (%)	58,67	61,33	62,67	37,67	62,00
Sensibilidade (%)	85,00	87,65	90,08	93,94	91,89

Tabela B.2 – Classificação Saudável X Patológico utilizando 13 descritores com a matriz de co-ocorrência na direção $\theta = 45^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	79,23	76,36	76,54	79,49	82,42
Desvio da acurácia (%)	2,42	2,89	3,81	2,15	2,59
Especificidade (%)	62,00	50,67	61,00	58,33	61,00
Sensibilidade (%)	87,50	88,48	84,17	89,39	92,88

Tabela B.3 – Classificação Saudável X Patológico utilizando 13 descritores com a matriz de co-ocorrência na direção $\theta = 90^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	74,45	75,70	80,13	72,54	82,98
Desvio da acurácia (%)	3,35	2,26	3,02	3,94	2,41
Especificidade (%)	40,00	45,67	64,67	36,67	59,67
Sensibilidade (%)	90,30	90,23	87,50	89,55	93,71

Tabela B.4 – Classificação Saudável X Patológico utilizando 13 descritores com a matriz de co-ocorrência na direção $\theta = 135^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	76,54	80,06	74,30	72,84	78,19
Desvio da acurácia (%)	2,75	2,94	4,44	3,03	3,74
Especificidade (%)	52,00	56,67	55,00	32,33	49,00
Sensibilidade (%)	88,33	91,06	83,79	91,97	92,12

Tabela B.5 – Classificação Saudável X Paralisia utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 0^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	77,82	77,64	73,82	58,64	84,64
Desvio da acurácia (%)	4,40	4,52	3,07	5,40	3,62
Especificidade (%)	82,67	81,67	80,67	57,67	90,67
Sensibilidade (%)	72,00	74,33	66,00	59,00	78,67

Tabela B.6 – Classificação Saudável X Paralisia utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 45^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	80,82	78,27	82,64	83,55	85,55
Desvio da acurácia (%)	3,73	4,27	4,02	3,57	1,62
Especificidade (%)	82,67	78,00	83,00	84,33	85,00
Sensibilidade (%)	78,33	78,33	82,00	82,67	86,33

Tabela B.7 – Classificação Saudável X Paralisia utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 90^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	70,71	77,09	79,82	63,45	82,71
Desvio da acurácia (%)	4,64	3,84	3,38	4,58	4,38
Especificidade (%)	75,00	77,33	83,67	55,33	83,00
Sensibilidade (%)	66,33	76,33	76,67	72,33	82,33

Tabela B.8 – Classificação Saudável X Paralisia utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 135^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	73,00	74,18	68,71	70,27	75,00
Desvio médio (%)	3,23	3,77	4,26	4,78	4,09
Especificidade (%)	75,67	81,33	75,00	76,00	78,00
Sensibilidade (%)	71,00	66,00	62,00	64,33	72,00

Tabela B.9 – Classificação Saudável X Edema utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 0^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	73,26	69,57	72,36	64,52	79,72
Desvio da acurácia (%)	4,32	3,50	5,49	4,07	4,64
Especificidade (%)	81,67	77,00	77,00	75,00	82,67
Sensibilidade (%)	63,00	61,00	67,00	50,50	75,50

Tabela B.10 – Classificação Saudável X Edema utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 45^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	75,78	73,89	73,79	80,33	75,89
Desvio da acurácia (%)	4,04	3,92	3,81	6,20	5,85
Especificidade (%)	83,00	83,00	86,33	87,00	87,33
Sensibilidade (%)	68,00	63,00	58,00	72,50	62,50

Tabela B.11 – Classificação Saudável X Edema utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 90^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	66,83	71,87	72,88	63,85	75,63
Desvio da acurácia (%)	4,84	5,32	4,47	7,01	3,46
Especificidade (%)	70,67	77,00	73,00	61,00	78,67
Sensibilidade (%)	63,50	64,00	73,00	66,50	72,00

Tabela B.12 – Classificação Saudável X Edema utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 135^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	70,70	62,67	66,72	60,55	76,18
Desvio da acurácia (%)	4,45	4,41	5,41	5,77	5,20
Especificidade (%)	81,00	75,67	79,00	64,33	83,00
Sensibilidade (%)	58,50	48,00	51,50	56,50	68,00

Tabela B.13 – Classificação Saudável X Nódulo utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 0^\circ$.

Sub-bandas (%)	LL	HL	LH	HH	Todas
Acurácia média (%)	82,80	82,26	68,75	79,29	84,46
Desvio da acurácia (%)	2,87	3,25	5,29	3,57	3,97
Especificidade (%)	92,67	96,67	88,33	93,00	92,33
Sensibilidade (%)	55,00	40,00	15,00	45,00	65,00

Tabela B.14 – Classificação Saudável X Nódulo utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 45^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	78,15	76,19	80,06	77,74	81,79
Desvio da acurácia (%)	3,36	5,58	3,30	2,68	4,72
Especificidade (%)	94,33	94,00	94,00	94,67	91,33
Sensibilidade (%)	30,00	20,00	35,00	25,00	50,00

Tabela B.15 – Classificação Saudável X Nódulo utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 90^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	81,19	76,49	78,15	74,70	85,89
Desvio da acurácia (%)	4,62	4,69	4,27	2,72	4,77
Especificidade (%)	93,00	92,67	92,67	92,33	94,33
Sensibilidade (%)	50,00	25,00	40,00	25,00	65,00

Tabela B.16 – Classificação Saudável X Nódulo utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 135^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	77,14	73,63	74,52	79,05	86,25
Desvio da acurácia (%)	3,81	2,69	5,16	3,68	3,55
Especificidade (%)	92,00	85,67	88,33	94,33	92,33
Sensibilidade (%)	30,00	35,00	40,00	35,00	70,00

Tabela B.17 – Classificação Edema X Nódulo utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 0^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	66,95	65,62	70,67	70,76	67,67
Desvio da acurácia (%)	5,07	3,83	4,37	1,69	4,56
Especificidade (%)	20,00	10,00	15,00	30,00	15,00
Sensibilidade (%)	87,50	90,50	93,50	88,00	88,50

Tabela B.18 – Classificação Edema X Nódulo utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 45^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	68,76	70,95	62,57	69,10	67,10
Desvio da acurácia (%)	1,65	4,12	3,90	2,67	2,51
Especificidade (%)	5,00	20,00	10,00	15,00	15,00
Sensibilidade (%)	95,00	93,00	86,00	90,50	88,50

Tabela B.19 – Classificação Edema X Nódulo utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 90^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	64,33	70,38	67,29	63,19	67,67
Desvio da acurácia (%)	4,50	2,17	4,90	4,43	6,48
Especificidade (%)	0,00	5,00	10,00	5,00	60,00
Sensibilidade (%)	90,50	98,00	90,00	87,50	71,50

Tabela B.20 – Classificação Edema X Nódulo utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 135^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	70,86	70,95	63,81	68,81	67,10
Desvio médio (%)	2,73	4,81	6,03	4,62	3,29
Especificidade (%)	15,00	25,00	5,00	15,00	25,00
Sensibilidade (%)	94,00	90,50	88,00	90,50	86,00

Tabela B.21 – Classificação Edema X Paralisia utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 0^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	52,92	48,67	53,00	56,56	57,44
Desvio da acurácia (%)	3,95	3,81	4,98	6,30	5,73
Especificidade (%)	66,00	61,00	46,33	55,00	68,67
Sensibilidade (%)	38,00	35,00	59,00	58,50	43,00

Tabela B.22 – Classificação Edema X Paralisia utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 45^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	63,78	49,44	59,18	55,56	61,00
Desvio da acurácia (%)	3,97	5,69	6,41	5,23	5,90
Especificidade (%)	70,33	59,67	50,67	56,67	57,67
Sensibilidade (%)	54,50	38,00	69,00	56,00	66,00

Tabela B.23 – Classificação Edema X Paralisia utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 90^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	55,56	54,22	57,05	57,99	60,78
Desvio da acurácia (%)	4,03	6,82	4,62	6,50	6,09
Especificidade (%)	72,67	61,33	58,00	64,33	64,67
Sensibilidade (%)	35,00	45,50	56,50	50,50	58,50

Tabela B.24 – Classificação Edema X Paralisia utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 135^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	59,56	59,78	63,44	48,11	58,17
Desvio médio (%)	4,88	4,16	5,17	3,64	5,90
Especificidade (%)	67,00	56,67	60,00	47,33	56,00
Sensibilidade (%)	52,00	64,50	67,00	50,00	61,00

Tabela B.25 – Classificação Paralisia X Nódulo utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 0^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	74,17	75,48	74,58	72,14	72,62
Desvio da acurácia (%)	1,57	2,97	3,23	3,76	4,91
Especificidade (%)	0,00	25,00	15,00	15,00	15,00
Sensibilidade (%)	100,00	95,00	96,33	92,33	92,00

Tabela B.26 – Classificação Paralisia X Nódulo utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 45^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	70,18	68,63	66,07	72,56	69,40
Desvio da acurácia (%)	6,66	3,16	5,84	5,34	5,99
Especificidade (%)	15,00	5,00	10,00	15,00	25,00
Sensibilidade (%)	90,00	90,33	88,00	92,00	86,00

Tabela B.27 – Classificação Paralisia X Nódulo utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 90^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	74,05	70,06	72,74	66,90	71,49
Desvio da acurácia (%)	4,08	3,59	4,36	5,57	3,29
Especificidade (%)	15,00	10,00	15,00	5,00	20,00
Sensibilidade (%)	94,00	90,33	92,00	89,00	88,33

Tabela B.28 – Classificação Paralisia X Nódulo utilizando 13 descritores para cada sub-banda wavelet com a matriz de co-ocorrência na direção $\theta = 135^\circ$.

Sub-bandas	LL	HL	LH	HH	Todas
Acurácia média (%)	69,88	71,31	66,73	71,01	72,74
Desvio médio (%)	3,13	4,11	3,77	5,32	3,15
Especificidade (%)	0,00	5,00	0,00	20,00	20,00
Sensibilidade (%)	94,00	94,00	90,33	90,67	90,00

Características selecionadas pelo PSO

Neste apêndice são apresentadas as características selecionadas pelo algoritmo de otimização por nuvem de partículas binário (PSO) em cada caso de classificação. Os resultados foram obtidos utilizando o método dos descritores com a matriz de co-ocorrência na direção $\theta = 45^\circ$. Nas tabelas seguintes, o valor 1 indica que a característica foi selecionada e o valor 0 indica que a característica não foi selecionada para a combinação de características que apresenta o melhor resultado da acurácia num determinado caso de classificação.

Tabela C.1 – Classificação SDLxPTL

Descritores de Haralick	Sub-bandas			
	LL	HL	LH	HH
Entropia	0	0	0	1
Segundo momento angular	1	1	0	1
Contraste	1	1	0	0
Momento da diferença inversa	1	0	0	0
Correlação	1	1	0	0
Média da soma	0	0	0	0
Variância da soma	1	1	0	1
Entropia da soma	0	1	1	1
Variância	0	1	1	1
Variância da diferença	0	0	0	1
Entropia da diferença	0	1	1	0
Informação de correlação 1	0	0	0	0
Informação de correlação 2	1	0	0	1

Tabela C.2 – Classificação SDLxPRL

Descritores de Haralick	Sub-bandas			
	LL	HL	LH	HH
Entropia	0	1	0	1
Segundo momento angular	1	0	1	1
Contraste	1	1	1	1
Momento da diferença inversa	1	0	1	1
Correlação	1	0	0	1
Média da soma	0	0	0	1
Variância da soma	1	1	0	1
Entropia da soma	1	0	0	1
Variância	0	0	0	1
Variância da diferença	0	1	0	0
Entropia da diferença	0	1	1	0
Informação de correlação 1	1	1	0	1
Informação de correlação 2	1	0	1	1

Tabela C.3 – Classificação SDLxEDM

Descritores de Haralick	Sub-bandas			
	LL	HL	LH	HH
Entropia	1	1	1	1
Segundo momento angular	0	1	0	1
Contraste	0	0	1	0
Momento da diferença inversa	1	1	1	0
Correlação	1	0	1	1
Média da soma	0	1	1	1
Variância da soma	1	1	0	1
Entropia da soma	0	1	1	1
Variância	1	1	1	0
Variância da diferença	0	1	1	0
Entropia da diferença	1	0	1	0
Informação de correlação 1	1	0	0	0
Informação de correlação 2	0	0	0	1

Tabela C.4 – Classificação SDLxNDL

Descritores de Haralick	Sub-bandas			
	LL	HL	LH	HH
Entropia	0	1	0	1
Segundo momento angular	1	0	0	1
Contraste	0	0	1	1
Momento da diferença inversa	1	0	1	1
Correlação	0	0	0	1
Média da soma	1	0	0	0
Variância da soma	1	1	0	1
Entropia da soma	0	0	1	1
Variância	0	1	1	1
Variância da diferença	0	0	1	0
Entropia da diferença	1	0	1	1
Informação de correlação 1	1	1	0	1
Informação de correlação 2	0	1	0	1

Tabela C.5 – Classificação EDMxNDL

Descritores de Haralick	Sub-bandas			
	LL	HL	LH	HH
Entropia	1	1	1	0
Segundo momento angular	0	0	1	0
Contraste	0	0	1	1
Momento da diferença inversa	0	0	1	0
Correlação	1	1	0	0
Média da soma	0	0	0	1
Variância da soma	1	1	1	1
Entropia da soma	1	0	1	0
Variância	0	0	1	1
Variância da diferença	1	1	1	1
Entropia da diferença	1	1	0	0
Informação de correlação 1	0	0	1	1
Informação de correlação 2	1	0	0	0

Tabela C.6 – Classificação EDMxPRL

Descritores de Haralick	Sub-bandas			
	LL	HL	LH	HH
Entropia	1	1	1	0
Segundo momento angular	0	0	1	0
Contraste	0	0	0	0
Momento da diferença inversa	0	0	1	1
Correlação	1	0	1	1
Média da soma	1	0	1	1
Variância da soma	1	0	1	1
Entropia da soma	1	1	1	0
Variância	1	1	1	1
Variância da diferença	1	0	0	0
Entropia da diferença	0	1	1	0
Informação de correlação 1	0	1	0	0
Informação de correlação 2	0	1	1	1

Tabela C.7 – Classificação PRLxNDL

Descritores de Haralick	Sub-bandas			
	LL	HL	LH	HH
Entropia	1	1	0	0
Segundo momento angular	1	0	1	1
Contraste	0	1	1	0
Momento da diferença inversa	0	1	0	1
Correlação	1	0	0	1
Média da soma	1	1	0	1
Variância da soma	0	0	0	0
Entropia da soma	1	0	1	1
Variância	1	0	1	0
Variância da diferença	0	1	0	0
Entropia da diferença	1	0	0	1
Informação de correlação 1	0	1	0	0
Informação de correlação 2	1	1	1	1