Universidade de São Paulo
Instituto de Física de São Carlos
Departamento de Física e Informática

Leonardo Mendes de Souza

Detecção inteligente de patologias na laringe baseada
em máquinas de vetores de suporte e na transformada
wavelet

São Carlos - SP

2010

Leonardo Mendes de Souza

Detecção inteligente de patologias na laringe
baseada em máquinas de vetores de suporte e
na transformada *wavelet*

Dissertação apresentada ao Programa de
Pós-graduação em Física da Universidade de
São Paulo para obtenção do título de Mestre
em Ciências.

Área de concentração: Física Aplicada -
Opção Computacional.

Orientador: Prof. Dr. Rodrigo Capobianco
Guido

Versão Corrigida

(Versão Original disponível na Universidade que aloja o programa)

São Carlos - SP

2010

AUTORIZO A DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUAL-
QUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA,
DESDE QUE CITADA A FONTE

Ficha catalográfica elaborada pelo Serviço de Biblioteca e Informação IFSC/USP

Souza, Leonardo Mendes de.

Detecção inteligente de patologias na laringe baseada em máquinas de
vetores de suporte e na transformada *wavelet*. / Leonardo Mendes de Sou-
za ; orientador Rodrigo Capobianco Guido. - Versão Corrigida - - São
Carlos, 2010.

102 p.

Dissertação (Mestrado - Programa de Pós-Graduação em Física -
Área de Concentração: Física Aplicada - Opção Física Computacional) -
Instituto de Física de São Carlos da Universidade de São Paulo.

1. Processamento de sinais. 2. Patologias da laringe. 3. Máquinas de
vetores de suporte. 4. Inteligência artificial. 5. *Wavelets*. I. Título.

Dedico a presente dissertação aos meus pais
Jason e Rosineide, a minha namorada Mari-
angela, pelo apoio que me foi dado durante
essa longa caminhada.

Agradecimentos

A Deus, por me ter permitido existir e ter sido sempre presença contante em minha vida me dando motivação, força e sabedoria em todos os momentos. E a Jesus pela amizade e presença certa em todas as horas.

Aos meus pais, Jason e Rosineide, por terem provido todas as condições para eu estudar e por terem me educado para eu me tornar uma pessoa digna.

A minha namorada Mariangela pelo incentivo e apoio em todos os momentos e por me ajudar a tornar-me a cada dia uma pessoa melhor.

Aos meus irmãos Thiago e Karina pelo incentivo recebido. Ao meu amigo Sylvio, pela grande e constante ajuda durante todo o meu percurso academico.

Ao meu orientador, pelo acompanhamento, orientação e suporte deste trabalho, e ao Instituto de Física de São Carlos, por sediar e prover a infra-estrutura necessária para o desenvolvimento do projeto de pesquisa.

Aos senhores Rogério Aquino e Vilma Aquino da empresa Riosoft, pelo apoio e concessão do tempo necessário para realização das disciplinas.

Aos meus amigos e companheiros de mestrado, Lucimar, Fabrício, Paulo Fantinato, João Paulo, Márcio, Regiane e todos os demais.

“Ambição é o caminho para o sucesso. Persistência é o veículo no qual se chega lá.”

Bill Eardley.

Resumo

SOUZA, L.M. *Detecção inteligente de patologias na laringe baseada em máquinas de vetores de suporte e na transformada wavelet*. 2010. 102p. Dissertação (Mestrado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2010.

A detecção de patologias na laringe tem ocorrido basicamente por meio de diagnósticos médicos apoiados em videolaringoscopia, que é considerado um procedimento invasivo e causa certo desconforto ao paciente. Além disso, esse tipo de exame é realizado com solicitação médica e apenas quando as alterações na fala já são marcantes ou estão causando dor. Nesse ponto, muitas vezes a doença já está em grau avançado, dificultando o seu tratamento. Com o objetivo de realizar um pré-diagnóstico de tais patologias, este trabalho propõe uma técnica não invasiva baseada em um novo algoritmo que combina duas Máquinas de Vetores de Suporte, treinadas com o uso de um procedimento de aprendizado semi-supervisionado, alimentadas por um conjunto de parâmetros obtidos com o uso da Transformada *Wavelet* Discreta do sinal de voz do locutor. Os testes realizados com uma base de dados de vozes normais e afetadas por diversas patologias demonstram a eficácia da técnica proposta que pode, inclusive, ser implementada em tempo-real.

Palavras-chave: Processamento de sinais. Patologias da laringe. Máquinas de vetores de suporte. Inteligência artificial. *Wavelets*.

Abstract

SOUZA, L.M. *Intelligent detection of larynx pathologies based on support vector machines and wavelet transform*. 2010. 102p. Dissertação (Mestrado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2010.

Larynx pathology detection is a process that depends basically on medical diagnosis and is based on videolaryngoscopy, which is considered as being an invasive and uncomfortable procedure. Furthermore, this kind of examination depends on a physicist's requirement and is carried out only when speech is considerably modified or causing pain. At that level, the problem is in an advanced stage which difficults its treatment. In order to get a pre-diagnosis of such pathologies, this work proposes a non-invasive technique which is based on a new algorithm that combines two support vector machines, trained with a semi-supervised approach, powered by a set of parameters derived from the discrete wavelet transform of the speaker's voice signal. Tests carried out with the use of a database of normal and pathological voices show the efficacy of the proposed technique which can also be implemented for use in real-time.

Key-words: Signal processing. Larynx pathologies. Support vector machines. Artificial intelligence. *Wavelets*.

Lista de Figuras

2.1	Interpretação física simplificada do sistema bio-gerador de voz (adaptado de (2)) . . .	28
2.2	[esquerda]: visão básica do sistema de produção de voz humana; [direita]: detalhe do trato vocal humano e suas sub-partes (adaptado de (2))	29
2.3	Exemplo de laringe saudável e com patologias comuns nas pregas vocais (adaptado de (3))	31
2.4	Nódulo na prega vocal direita e tumor irritativo à esquerda (adaptado de (3))	32
2.5	Pregas vocais com Edema de Reinke (adaptado de (4))	32
2.6	Pólipos nas pregas vocais (adaptado de (6))	33
2.7	Separação linear das classes W_1 e W_2 com base num hiperplano. Os vetores que definem os limites do hiperplano são conhecidos como vetores de suporte	35
2.8	mapeamento não-linear do espaço de entrada para o espaço de características (adaptado de (11))	36
2.9	arquitetura da SVM (adaptado de (11))	37
2.10	respostas em frequências dos filtros MFCCs (adaptado de (15))	41
2.11	Construção do triângulo de Sierpinski, de acordo com (13)	43
2.12	Funcionamento da DWT, exemplificado para um sinal $s[\cdot]$ de n amostras discretas e máxima frequência π , decomposto até o terceiro nível. É importante observar o espectro de frequências e a quantidade de amostras presentes em cada sub-banda (adaptado de (2)	46

2.13	Relação entre os filtros de análise e síntese	49
2.14	Formato das respostas ao impulso dos filtros <i>wavelet</i> de Haar, Daubechies, Vaidyanathan, Beylkin, Coiflet e Symmlet, respectivamente (2)	52
2.15	Formatos das funções <i>scaling</i> dos filtros <i>wavelet</i> de Haar, Daubechies, Vaidyanathan, Beylkin, Coiflet e Symmlet, respectivamente (2)	52
2.16	Formatos das funções <i>wavelet</i> dos filtros <i>wavelet</i> de Haar, Daubechies, Vaidyanathan, Beylkin, Coiflet e Symmlet, respectivamente (2)	52
3.1	Arquitetura básica da estrutura proposta para treinamento do sistema de detecção de patologias do trato vocal. Legenda: [LPAV]: leitura do próximo arquivo de voz; [CVNV]: consulta para verificação de normalidade da voz; [EP_N]: extração de parâmetros da voz normal; [EP_P]: extração de parâmetros da voz de paciente com patologia; [SVM_N]: SVM das vozes normais; [SVM_P]: SVM das vozes dos indivíduos com patologias; [UAV]: último arquivo de voz.	59
3.2	Arquitetura básica da estrutura proposta para uso do sistema de detecção de patologias do trato vocal, após treinamento. Legenda: [LAV]: leitura do arquivo de voz; [EP]: extração de parâmetros; [SVM_N]: SVM das vozes normais; [SVM_P]: SVM das vozes dos indivíduos com patologias; [EMR]: escolha da máxima resposta.	60
4.1	Testes do grupo 2 realizados com os parâmetros que obtiveram melhor resposta nos testes do grupo 1 para ambas as classes (normal e patológica): SX, DP_M, DP_E e DP_P. O eixo horizontal corresponde à quantidade em percentual de vozes utilizadas para treinamento do classificador, enquanto o eixo vertical corresponde ao percentual de acerto dos testes realizados. A curva representada pela linha rosa corresponde às respostas obtidas utilizando-se a base de testes patológica e a curva representada pela linha azul corresponde às respostas obtidas utilizando-se a base de testes não-patológica.	66

Lista de Tabelas

2.1	Sub-tipos de fonemas (adaptado de (2))	29
2.2	<i>Chunk</i> 1 do formato WAV: componentes e descrições dos 12 bytes integrantes	38
2.3	<i>Chunk</i> 2 do formato WAV: componentes e descrições dos 24 bytes integrantes	39
2.4	<i>Chunk</i> de dados do formato WAV: componentes e descrições dos bytes integrantes	39
2.5	As 25 bandas críticas do sistema auditivo humano que são aproveitadas no presente trabalho para obtenção dos MFCCs	42
2.6	Características das famílias de <i>wavelets</i> utilizadas no presente trabalho, incluindo a quantidade de momentos da função <i>wavelet</i>	51
4.1	Testes do grupo 1 para diversas combinações de dois parâmetros extraídos dos bancos de dados de vozes normais e patológicas. Legenda: D.P. : Desvio Padrão	65
4.2	Testes do grupo 1 para diversas combinações de três parâmetros extraídos dos bancos de dados de vozes normais e patológicas. Legenda: D.P. : Desvio Padrão	65
4.3	Testes do grupo 1 para diversas combinações de quatro parâmetros extraídos dos bancos de dados de vozes normais e patológicas. Legenda: D.P. : Desvio Padrão	65
4.4	Testes do grupo 1 para a combinação dos cinco parâmetros extraídos dos bancos de dados de vozes normais e patológicas. Legenda: D.P. : Desvio Padrão	65

Lista de Abreviaturas

D	Dimensão do Fractal
DSP	<i>Digital Signal Processor</i>
DWT	<i>Discrete Wavelet Transform</i>
FIR	<i>Finite Impulse Response</i>
FPGA	<i>Field Programmable Gate Array</i>
IA	Inteligência Artificial
MFCCs	<i>Mel-frequency cepstral coefficients</i>
MRA	Análise de Multi-Resolução
PDS	Processamento Digital de Sinais
PRFB	<i>Perfect Reconstruction Filter Bank</i>
QMF	<i>Quadrature Mirror Filters</i>
RNA	Redes Neurais Artificiais
SVM	<i>Support vector machine</i>
WAV	<i>Waveform Audio Format</i>

Sumário

1	Introdução e Motivação	23
1.1	Objetivos e Contribuições do trabalho	24
1.2	Organização do Trabalho	25
2	Revisão Bibliográfica e Direcionamento do Trabalho	27
2.1	Estudo da fala humana	28
2.2	Patologias na Laringe	30
2.2.1	Nódulos nas pregas vocais	31
2.2.2	Edema de Reinke	32
2.2.3	Pólipos	33
2.3	Classificadores inteligentes	34
2.4	Conceitos elementares e principais parâmetros oriundos dos sinais de voz . . .	37
2.4.1	Digitalização de sinais e o formato WAV de arquivos de áudio/voz . . .	38
2.4.2	A medida estatística de desvio-padrão	39
2.4.3	O conceito de energia	40
2.4.4	Os <i>Mel frequency cepstral coefficients</i> (MFCCs)	40
2.4.5	O conceito de dimensão fractal (D)	41

2.4.6	A Transformada <i>Wavelet</i> Discreta (DWT)	44
3	Descrição da Técnica Proposta	53
3.1	A Arquitetura do Sistema Proposto	53
3.1.1	Detalhamento do algoritmo proposto para o módulo de treinamento . .	55
3.1.2	Detalhamento do algoritmo proposto para o módulo de aplicação	57
3.2	Implementação	58
4	Testes e Resultados	61
4.1	Materiais e Métodos	61
4.2	Bateria de Testes	63
5	Conclusões e Trabalhos Futuros	67
	Referências	70
	Apêndice I - Coeficientes dos filtros <i>wavelet</i> utilizados nas experiências.	75
	Apêndice II - Código fonte do sistema proposto.	87
	Apêndice III - Publicações durante o mestrado.	101

Capítulo 1

Introdução e Motivação

De acordo com recentes trabalhos desenvolvidos pela equipe de processamento de sinais de voz do câmpus de São Carlos da USP, dentre os quais se destaca (1), é possível afirmar que, do ponto de vista sistêmico, o sinal de voz carrega toda a informação intrínseca do aparelho fonador, possibilitando-se obter a partir desse sinal cada uma das características relacionadas ao trato vocal do locutor. Além disso, é possível utilizar técnicas de processamento digital de sinais (PDS) de voz com o objetivo de detectar propriedades acústicas que diferenciam sinais normais daqueles afetados por alguma patologia no trato vocal de determinado locutor.

Recentemente, modernas técnicas de PDS têm sido usadas para analisar desordens vocais provocadas por patologias na laringe. Tendo em vista que os exames médicos laringoscópicos usados na detecção de tais patologias são técnicas invasivas que causam desconforto ao paciente, as análises acústicas das características temporais e espectrais do sinal de voz podem ser utilizadas como ferramenta auxiliar para o pré-diagnóstico dessas patologias.

1.1 Objetivos e Contribuições do trabalho

Com base nas considerações anteriores, o objetivo específico deste trabalho é o de apresentar uma técnica não-invasiva para auxiliar na detecção de patologias na laringe por meio da análise dos sinais de voz dos locutores. Uma estrutura composta por duas Máquinas de Vetores de Suporte (*Support Vector Machines* - SVMs) treinadas com um procedimento semi-supervisionado e recebendo parâmetros obtidos com base na Transformada *Wavelet* Discreta (*Discrete Wavelet Transform* - DWT) do sinal de voz do paciente constituem a base para o desenvolvimento do projeto. Em particular, espera-se que o sistema desenvolvido possa ser implementado em tempo-real, executando em *background* em *smartphones*, computadores pessoais, entre outros dispositivos, de forma a alertar determinado locutor sobre o fato de que alguma estrutura envolvida na produção de sua voz não está funcionando adequadamente. No que tange às contribuições originais do presente trabalho, destacam-se portanto:

- a obtenção de um novo algoritmo de baixa complexidade para o pré-diagnóstico de patologias no trato vocal;
- a especificação de um sistema binário de classificação com base em uma estrutura composta por duas SVMs, diferentemente das implementações tradicionais;
- a caracterização do melhor conjunto de parâmetros extraídos dos sinais de vozes, para uso com o sistema de classificação proposto, que produz os resultados mais eficientes na distinção entre as vozes normais daquelas afetadas por alguma patologia.

1.2 Organização do Trabalho

Este trabalho está organizado da seguinte forma: o capítulo 2 apresenta uma revisão da literatura, envolvendo os princípios básicos do sistema de fala humana, das SVMs, das patologias na laringe e da DWT, além de outros conceitos importantes. O capítulo 3 apresenta, com detalhes, o algoritmo proposto; já os resultados obtidos com as diversas combinações de parâmetros extraídos por meio da transformada *wavelet* estão descritos no capítulo 4. Tendo em vista os resultados, e apoiado nos conceitos e características estudados durante a revisão de literatura, o capítulo 5 apresenta as conclusões. Por fim, logo após as referências bibliográficas, três apêndices apresentam, respectivamente, uma lista dos coeficientes dos filtros *wavelet* utilizados nas experiências, o código-fonte da implementação em linguagem de programação de alto nível, assim como as publicações obtidas durante o curso de mestrado do autor.

Capítulo 2

Revisão Bibliográfica e Direcionamento do Trabalho

Neste capítulo, uma revisão da literatura é apresentada, abordando basicamente os seguintes tópicos: visão geral do sistema bio-gerador da fala humana e as patologias mais comuns do trato vocal, já que caracterizam o presente trabalho; as SVMs, que constituem os classificadores inteligentes adotados como ponto de partida para formulação do sistema proposto; os conceitos de digitalização de sinais com base no formato WAV de arquivos; de variância estatística; de energia de um sinal; a dimensão fractal e os Mel Frequency Cepstral Coefficients (MFCCs), que caracterizam os parâmetros utilizados para alimentar o classificador proposto e; por fim, a DWT, que é a principal ferramenta utilizada para o cálculo dos parâmetros retrocitados. Todos esses conceitos são necessários para a perfeita compreensão do sistema proposto na presente dissertação.

2.1 Estudo da fala humana

De acordo com (2), a produção de voz pelo corpo humano, embora seja um mecanismo repleto de detalhes, consiste basicamente da propulsão de ar pelos pulmões, seguida de um processo de equalização realizado pelo trato vocal e elementos associados, como ilustram as figuras 2.1 e 2.2. O primeiro detalhe a ser observado na figura 2.1 é que as pregas vocais controlam o fluxo de ar fornecido pelos pulmões, fazendo com que esse sinal de excitação seja periódico, vibrando em determinada frequência, ou aperiódico, similar a um sinal ruidoso. Se o sinal for periódico, esse período é chamado de período de *pitch* e a voz produzida será classificada como vozeada (*voiced speech*), caso contrário, a voz será classificada como não-vozeada (*unvoiced speech*). No primeiro caso encontram-se basicamente as vogais, enquanto no segundo caso estão os demais sons. Dependendo de como agem as estruturas seguintes às pregas vocais, em particular o véu palatino, que controla a passagem do fluxo de ar pelo trato vocal ou nasal, e os próprios tratos vocal e nasal, pode-se ainda refinar essa classificação dos sinais, conforme listado na tabela 2.1.

Figura 2.1 – Interpretação física simplificada do sistema bio-gerador de voz (adaptado de (2))

Qualquer alteração estrutural ou funcional nos elementos da figura 2.2, possivelmente causada por uma patologia, implica a obtenção de um sinal de voz diferente do normal. Quando essa alteração é substancial, o efeito na voz do locutor pode ser percebido mais facilmente por qualquer pessoa, além disso, o profissional da área médica é capaz de arriscar um

Tabela 2.1 – Sub-tipos de fonemas (adaptado de (2))

tipo	descrição	exemplo
<i>fricatives</i>	<i>unvoiced speech</i> com a presença de uma turbulência de ar entre a língua e os dentes superiores	<u>th</u> na palavra <i>thin</i>
<i>plosives</i>	<i>unvoiced speech</i> impulsivo	<u>t</u> na palavra <i>top</i>
<i>whispers</i>	<i>unvoiced speech</i> com uma barreira nas pregas vocais de forma que elas permaneçam parcialmente fechadas	<u>h</u> na palavra <i>he</i>
<i>voiced fricatives</i>	<i>voiced speech</i> misturados com ruído criado na constrição do trato vocal atrás dos dentes e contra o palato	<u>z</u> na palavra <i>zebra</i>
<i>unvoiced fricatives</i>	idem anterior, porém as pregas vocais não vibram simultaneamente com a fricção	<u>h</u> na palavra <i>human</i>
<i>voiced plosives</i>	fonemas <i>voiced</i> misturados com ruído impulsivo criado no trato vocal	<u>g</u> na palavra <i>glass</i>
<i>unvoiced plosives</i>	idem anterior, porém as pregas vocais não vibram simultaneamente com o impulso	<u>b</u> na palavra <i>boat</i>

Figura 2.2 – [esquerda]: visão básica do sistema de produção de voz humana; [direita]: detalhe do trato vocal humano e suas sub-partes (adaptado de (2))

pré-diagnóstico sem uma análise clínica detalhada. Já no caso da alteração não ser significativa, o seu impacto na voz do locutor pode não ser facilmente percebido pelo próprio locutor, por uma outra pessoa qualquer que convive com ele, ou até por um profissional da área médica. Para tais casos, a análise do sinal de voz por um programa de computador é aconselhável, podendo ser utilizada como uma técnica eficaz para alertar o locutor sobre a saúde de sua estrutura vocal, indicando que ele deverá procurar um profissional especializado para realizar uma análise mais detalhada. Conforme mencionado anteriormente, um sistema desse tipo, capaz de fazer a análise com a finalidade de identificar patologias no mecanismo de produção da voz, pode ser bastante útil se estiver sendo executado em *background* em dispositivos tais como telefones celulares, *smarthphones*, *laptops*, entre outros. Neste trabalho, trechos vozeados de sinais de voz, isto é, *voiced speech*, serão utilizados para realizar a análise em questão, já que possuem informações mais propícias para essa finalidade (1).

2.2 Patologias na Laringe

As patologias mais comuns que afetam a voz são os nódulos vocais, também conhecidos como calos nas pregas vocais, os pólipos, cistos e os edemas (inchaço) de Reinke. Outras patologias comuns na laringe, que podem produzir alterações na voz, são: úlcera de contato, granuloma, leucoplasia, além do câncer de laringe, entre outras. Podemos observar na figura 2.3 a anatomia de uma laringe normal e a de laringes com alguns tipos comuns de patologias.

Por parte de um profissional especializado da área médica, a identificação e constatação das patologias na laringe é geralmente realizada por meio de um exame de videolaringoscopia (1), que é relativamente desconfortável. Já do ponto de vista computacional, as patologias podem usualmente ser identificadas a partir da utilização de padrões acústicos perceptuais tais

como a aspereza, a soproidade e a rouquidão (1). Alterações em tais padrões refletem em alguns parâmetros específicos, descritos a seguir, que constituem as entradas do sistema de classificação projetado no presente trabalho. Apenas para fins de conhecimento, três das principais patologias estão brevemente descritas a seguir.

Figura 2.3 – Exemplo de laringe saudável e com patologias comuns nas pregas vocais (adaptado de (3))

2.2.1 Nódulos nas pregas vocais

Segundo (3), o nódulo vocal é uma reação do tecido ao constante estresse, induzido pelo movimento oposicional brusco frequente das pregas vocais. A figura 2.4 ilustra a presença de um nódulo na prega vocal direita e também um tumor irritativo na prega vocal esquerda. O comportamento vocal inadequado e o abuso vocal são indicados como os dois fatores principais do surgimento de nódulos, sendo o álcool, o fumo, os distúrbios hormonais, as infecções das vias aéreas superiores, as alergias e a constituição atlética cultivada, os elementos predisponentes mais significativos. Os indivíduos com maior concentração de fibronectina e com menor concentração de ácido hialurônico podem estar mais predispostos ao desenvolvimento

de nódulos.

Figura 2.4 – Nódulo na prega vocal direita e tumor irritativo à esquerda (adaptado de (3))

Figura 2.5 – Pregas vocais com Edema de Reinke (adaptado de (4))

2.2.2 Edema de Reinke

Em (4), os autores conceituaram o Edema de Reinke como um edema generalizado e bilateral que traduz a presença de uma laringite crônica. Na figura 2.5 é possível observar um

Figura 2.6 – Pólipos nas pregas vocais (adaptado de (6))

exemplo de prega vocal afetada pelo edema. A patologia recebe esse nome por se localizar no espaço anatômico com o nome do anatomista Reinke, que foi o primeiro a investigar a anatomia das pregas vocais. Grande parte dos autores consideram o Edema de Reinke como resultado da agressão do cigarro, associado ao uso excessivo e abusivo da voz.

2.2.3 Pólipos

De acordo com (5) e (6), os pólipos são formações benignas, não cancerosas, com concentração acentuada de sangue e localizadas sobre as pregas vocais. São de massa lisa, gelatinosa e translúcida, atingindo as mais diversas faixas etárias, em indivíduos de ambos os sexos. Desenvolvem-se principalmente devido ao uso abusivo da voz, às reações alérgicas crônicas que afetam a laringe, ao ato de pigarrear constantemente e forçosamente, ou até mesmo à inalação de substâncias irritantes tais como gases industriais ou fumaça de cigarro. Os sintomas mais significativos incluem a rouquidão acompanhada de uma voz entrecortada. Na figura 2.6 é possível observar um exemplo em que dois pólipos nas pregas vocais são claramente visíveis.

2.3 Classificadores inteligentes

De acordo com (7) (8), as redes neurais artificiais (RNAs) constituem uma das mais antigas técnicas de inteligência artificial (IA) em uso. Esse instrumento foi desenvolvido na década de 40, por Walter Pitts e McCulloch, o primeiro, matemático e o segundo, neurofisiologista. A idéia era fazer uma analogia entre neurônios biológicos e circuitos eletrônicos capazes de simular conexões sinápticas pelo uso de resistores variáveis e amplificadores. Nas RNAs, cada neurônio matemático, similarmente ao natural, recebe um ou mais sinais de entrada e devolve um único de saída, que pode ser distribuído como sinal de saída da rede, ou como sinal de entrada para um ou vários outros neurônios de uma camada posterior.

O sistema de classificação apresentado no presente trabalho foi elaborado a partir de uma ferramenta de IA, utilizada para classificação binária, relativamente mais elaborada do que as RNAs. Conhecida como Máquina de Vetores de Suporte (SVM - *Support Vector Machine*) (9), constitui uma técnica de aprendizado supervisionado que vem recebendo crescente atenção da comunidade científica. Os resultados da aplicação dessa técnica são comparáveis, e muitas vezes superiores, aos obtidos por outros algoritmos de aprendizado, como as RNAs. Exemplos de aplicações de sucesso podem ser encontrados em diversos domínios, como na categorização de textos, na análise de imagens, em Bioinformática, entre outras.

De acordo com (10), as SVMs foram desenvolvidas por Vladimir Vapnik em 1998 e constituem uma nova técnica no campo teórico do aprendizado estatístico que se baseia no princípio da Minimização do Risco Estrutural, o qual possui dois objetivos. O primeiro é o de controlar o risco empírico no conjunto de treinamento e o segundo é controlar a capacidade da função de decisão usada para obter esse valor de risco. Dentre as diversas técnicas de classificação e separação de padrões existentes, as SVMs são as que oferecem considerável capacidade de generalização, isto é, novos casos de teste são eficientemente separados com base em um conjunto modesto de exemplos de treinamento.

Figura 2.7 – Separação linear das classes W_1 e W_2 com base num hiperplano. Os vetores que definem os limites do hiperplano são conhecidos como vetores de suporte

Dado um conjunto de vetores de treinamento $\{(\vec{x}_1, y_1); (\vec{x}_2, y_2); \dots; (\vec{x}_n, y_n)\}$ pertencente a duas classes linearmente separáveis, W_1 representada pela saída $y_i = +1$ e W_2 representada pela saída $y_i = -1$, a SVM encontra o hiperplano com a máxima distância do conjunto de treinamento. De acordo com esse princípio, existirá somente um hiperplano com a margem máxima δ , definida como sendo a soma das distâncias do hiperplano até o ponto mais próximo das classes, sendo que esse limiar do classificador é a separação ótima, conforme pode-se observar na figura 2.7.

Nos casos de conjuntos de treinamento não linearmente separáveis (11), como possivelmente ocorre com os dados sob análise no presente trabalho, associa-se ao i -ésimo ponto x_i , ($0 \leq i \leq N$), uma variável σ_i , que representa a magnitude do erro de classificação do mesmo, sendo que a função de penalidade $F(\sigma) = \sum_{i=0}^{N-1} \sigma_i$, representa a soma dos erros de classificação. A solução da SVM pode ser encontrada com base na minimização dos erros de treinamento de $\frac{1}{2} \vec{w}^T \vec{w} + C \sum_{i=0}^{N-1} \sigma_i$, onde $\vec{w}$ é o vetor de pesos e $C \in \mathbb{R}$. Na prática, é possível traduzir o equacionamento anterior em uma operação que compreende duas etapas ilustradas na figura 2.8 (11):

1. o mapeamento não-linear do vetor de entrada, que geralmente possui uma dimensão mais baixa, para um espaço de características de alta dimensão, sendo essa separação tratada

Figura 2.8 – mapeamento não-linear do espaço de entrada para o espaço de características (adaptado de (11))

pelo Teorema de Cover (11) sobre a separabilidade de padrões. A função de mapeamento é chamada de núcleo ou *kernel* da aplicação, geralmente representada por $\phi(\cdot)$. Dentre os *Kernels* mais comuns tem-se: linear, Gaussiano e polinomial.

2. a construção de um hiperplano ótimo para separar as características dos vetores da etapa anterior. Esse hiperplano realiza agora uma separação linear dos padrões de entrada, já que eles foram mapeados no passo anterior.

Com base nas idéias discutidas anteriormente, a arquitetura básica de uma SVM é apresentada na figura 2.9, na qual é possível observar três camadas de elementos. A primeira, à esquerda, é apenas passiva e encaminha os valores de entrada para a camada seguinte. A camada intermediária, ou oculta, é responsável pelo mapeamento não-linear. Já a terceira, à direita, é composta por um único elemento de processamento, associado ao vetor de pesos $\vec{w}$, que separa linearmente as características do mapeamento anterior.

2.4 Conceitos elementares e principais parâmetros oriundos dos sinais de voz

A seguir, são apresentados alguns conceitos elementares utilizados durante o projeto e a implementação do presente trabalho. Além disso, a considerável experiência anterior do grupo de pesquisas no qual se insere o presente trabalho (1) (2) (12) (13) tem mostrado quais os parâmetros que mais traduzem as características acústicas de interesse para a aplicação em questão. Tais parâmetros também são brevemente revistos a seguir.

Figura 2.9 – arquitetura da SVM (adaptado de (11))

2.4.1 Digitalização de sinais e o formato WAV de arquivos de áudio/voz

No presente trabalho, os arquivos utilizados para armazenar os sinais de vozes para o treinamento do classificador apresentado estão todos gravados no formato WAV (14), tendo em vista que ele é o mais simples existente para fins de armazenamento das amostras digitalizadas de tais sinais. Basicamente, um arquivo de áudio/voz no formato WAV é composto por um cabeçalho sub-dividido em duas partes chamadas *chunk 1* e *chunk 2*, seguidas pelos dados brutos, que fazem parte de uma porção identificada como *chunk* de dados. De acordo com a especificação do formato, no caso de sinais amostrados com uma quantização de 16 bits, que são os utilizados aqui, a faixa de valores para cada amostra vai de -32768 até 32767 , e ainda, no padrão *little ending*. Adicionalmente, os arquivos utilizados nas experiências do presente trabalho foram gravados em um único canal (mono) sem nenhum tipo de compressão. A especificação de cada um dos *chunks* mencionados consta nas tabelas 2.2, 2.3 e 2.4. Um trecho de código-fonte escrito em linguagem C++, que se encontra na parte final do presente trabalho, foi elaborado para extrair os “dados brutos” dos arquivos WAV e permitir seu uso nos experimentos.

Tabela 2.2 – *Chunk 1* do formato WAV: componentes e descrições dos 12 bytes integrantes

Bytes	Descrição
0 a 3	<i>string</i> ascii “RIFF”
4 a 7	comprimento do <i>chunk</i>
8 a 12	<i>string</i> ascii “WAV”

Tabela 2.3 – *Chunk* 2 do formato WAV: componentes e descrições dos 24 bytes integrantes

Bytes	Descrição
0 a 3	<i>string</i> ascii “FMT”
4 a 7	comprimento do <i>chunk</i>
8 a 9	0 para mono (um canal) e 1 para estéreo (dois canais)
10 a 11	número de canais (1 para mono e 2 para estéreo)
12 a 15	taxa de amostragem em Hz
16 a 19	bytes por segundo
20 a 21	bytes por amostra
22 a 23	bits por amostra

Tabela 2.4 – *Chunk* de dados do formato WAV: componentes e descrições dos bytes integrantes

Bytes	Descrição
0 a 3	<i>string</i> ascii “data”
4 a 7	comprimento do <i>chunk</i>
8 até o fim	“dados brutos”

2.4.2 A medida estatística de desvio-padrão

A literatura tem mostrado cada vez mais que a classificação de um padrão em um sinal de voz deve concentrar-se em alguma medida de variação nos valores dos parâmetros ao longo do tempo, ao invés de nos valores propriamente ditos. Dessa forma, o presente trabalho adotará, diversas vezes, como entrada para o sistema de classificação proposto, medidas estatísticas da variância de certas variáveis. Para um vetor genérico $\vec{x}$, de comprimento N , a sua variância pode ser definida como $Var(\vec{x}) = \frac{1}{N} \sum_{i=0}^{N-1} x_i^2 - \left(\frac{1}{N} \sum_{i=0}^{N-1} x_i \right)^2$.

2.4.3 O conceito de energia

A energia de um sinal x_i de N amostras discretas é um escalar definido como $E = \sum_{i=0}^{N-1} x_i^2$. No presente trabalho, as energias de várias sub-bandas de frequências de um sinal serão utilizadas para alimentar o sistema de classificação adotado.

2.4.4 Os *Mel frequency cepstral coefficients* (MFCCs)

De acordo com (15), a análise cepstral tem sido largamente empregada em processamento digital de sinais de voz. A implementação mais popular combina o *cepstrum* (16) com uma dilatação não-linear na frequência de acordo com a escala Bark ou mel (12), incorporando, portanto, aspectos da audição. Dessa forma, os coeficientes MFCCs podem ser encarados como uma representação mais conveniente dos componentes espectrais importantes de um sinal de voz, em comparação com outras técnicas de análise.

Para obtenção dos MFCCs, o espectro de potência S de cada *frame* sucessivo do sinal de voz é efetivamente “deformado” tanto em frequência, com o uso da escala Bark ou mel (14), quanto em amplitude, com base em uma escala logaritmica. A seguir, os M coeficientes iniciais c_i da Transformada Discreta de Fourier (15) são obtidos, sendo que geralmente $8 \leq M \leq 14$. A equação 2.1 resume o processo, onde os X_k correspondem aos sinais filtrados de acordo com a escala Bark mencionada anteriormente, sendo que o conjunto de respostas em frequências dos 20 filtros utilizados é exibido na figura 2.10.

$$c_i = \sum_{k=1}^{20} X_k \cos\left(i\left(k - \frac{1}{2}\right)\frac{\pi}{20}\right) \quad , \quad \text{onde} \quad i = 0, 1, \dots, M \quad . \quad (2.1)$$

Figura 2.10 – respostas em frequências dos filtros MFCCs (adaptado de (15))

A escala Bark mencionada acima permite sub-dividir a faixa de frequências audíveis pelo ouvido humano em 25 intervalos chamados de bandas críticas. Quando dois ou mais sons, de frequências diferentes mas pertencentes à mesma banda crítica, estão sendo captados simultaneamente, o de maior amplitude mascara o(s) outro(s), portanto, o de menor amplitude não é percebido. A tabela 2.5 mostra as bandas críticas, na escala de frequências em Hertz (Hz) e na escala Bark (z), que é uma escala adotada para linearizar o intervalo entre as bandas. A conversão entre Hertz e Bark é dada pela equação $z = \frac{26.81f}{1960+f} - 0.53$, onde f representa a frequência em Hertz. Nos casos onde $z < 2$ então aplica-se a correção $z \leftarrow z + 0.15(2 - z)$ e onde $z > 20$ então faz-se $z \leftarrow z + 0.22(z - 20.1)$.

2.4.5 O conceito de dimensão fractal (D)

De acordo com a discussão apresentada por (13) e também com o conceito de dimensão Euclidiana, esta última denominando-se usualmente de dimensão topológica, um ponto possui dimensão 0, uma reta possui dimensão 1, uma superfície possui dimensão 2 e uma porção de espaço possui dimensão 3. Esses exemplos indicam que, para determinar a di-

Tabela 2.5 – As 25 bandas críticas do sistema auditivo humano que são aproveitadas no presente trabalho para obtenção dos MFCCs

<i>Bark (z)</i>	<i>Hertz (Hz)</i>	<i>Bark (z)</i>	<i>Hertz (Hz)</i>
0	0 - 100	12	1720 - 2000
1	100 - 200	13	2000 - 2320
2	200 - 300	14	2320 - 2700
3	300 - 400	15	2700 - 3150
4	400 - 510	16	3150 - 3700
5	510 - 630	17	3700 - 4400
6	630 - 770	18	4400 - 5300
7	770 - 920	19	5300 - 6400
8	920 - 1080	20	6400 - 7700
9	1080 - 1270	21	7700 - 9500
10	1270 - 1480	22	9500 - 12000
11	1480 - 1720	23	12000 - 15500
		24	15500 - 22050

mensão topológica de um objeto, faz-se uma correspondência unívoca desse objeto com um ente geométrico elementar. Diferentemente, a dimensão fractal (D) refere-se à dimensão espacial, ou seja ao espaço ocupado por um sinal ou figura, podendo o valor correspondente ser um número fracionário. A dimensão do fractal, portanto, consiste em uma medida para o nível de irregularidade ou auto-similaridade desse sinal ou figura (13). No presente trabalho, a variância de D será utilizada como medida de irregularidade dos sinais sob análise (17).

Ainda de acordo com (13), é possível definir D da seguinte forma:

$$D = \frac{\log(N)}{\log(\frac{1}{T})} \quad (2.2)$$

o que significa que a regra para a formação do fractal consiste em substituir um segmento, parte ou objeto, por N segmentos de tamanho T correspondentes a uma fração do tamanho anterior. Um exemplo é o da construção do triângulo de Sierpinski (17) que, conforme ilustra a figura 2.11, possui dimensão tal que:

$$D = \frac{\log(N)}{\log(\frac{1}{T})} = \frac{\log(3)}{\log(\frac{1}{4})} = \frac{\log(3)}{\log(4)} \approx 0.79 \quad ,$$

tendo em vista que a cada etapa da construção, um triângulo é substituído por 3 triângulos de

área igual a $\frac{1}{4}$ da área do anterior.

Figura 2.11 – Construção do triângulo de Sierpinski, de acordo com (13)

O triângulo de Sierpinski constitui um caso particular em que a similaridade é facilmente observável por meio da simples observação da figura, possuindo caráter não estatístico. Outras classes de sinais ou figuras, como é o caso dos sinais de voz, possuem outro tipo de similaridade, de caráter estatístico. Independentemente do tipo de similaridade, os sinais unidimensionais de voz são tais que $1 \leq D \leq 2$, de acordo com (17) e (13). Este mesmo autor observa que, durante o processo de construção de um fractal, com lei de formação bem definida, tal como a do triângulo de Sierpinski, é muito simples obter sua dimensão em função dos parâmetros N e T . Entretanto, nos casos em que já se possui o sinal, ou figura, e não fica clara a sua lei de formação por meio de uma simples inspeção visual, um método deve ser utilizado para realizar a medição dessa dimensão. Os métodos mais tradicionais são *walking-divider*, *box-counting*, *prism* e *power spectrum*.

Em sua dissertação de mestrado, Fantinato, P.C. (13) discutiu os quatro métodos mencionados acima, tendo concluído que a técnica do *power spectrum*, que é relativamente diferente das demais, é a mais precisa e portanto será utilizada no presente trabalho. Essa técnica está baseada no uso do espectro de potência do sinal fractal, obtido via Transformada de Fourier ou Transformada *Wavelet* (18), ao invés do sinal propriamente dito e apresenta resultados relativamente precisos para toda a faixa $1 \leq D \leq 2$. O algoritmo consiste em obter $D = 2 - |H|$, onde $H = \frac{\beta-1}{2}$ é o expoente de Hurst (17), β é o coeficiente angular da reta média, obtida via regressão linear por Mínimos Quadrados, que atende ao conjunto de pontos da forma $\log_2(b_i)$ versus $\log_2(a_i)$, sendo b_i o valor normalizado do espectro de potência da banda i que

possui comprimento a_i , para um total de N sub-bandas:

$$\beta = \frac{N \sum_{k=0}^{N-1} \log_2(a_k) \log_2(b_k) - \sum_{k=0}^{N-1} \log_2(b_k) \sum_{k=0}^{N-1} \log_2(a_k)}{N \sum_{k=0}^{N-1} \log_2(a_k)^2 - \left(\sum_{k=0}^{N-1} \log_2(a_k) \right)^2} . \quad (2.3)$$

2.4.6 A Transformada *Wavelet* Discreta (DWT)

Barbon, S.J. (2) observa em sua dissertação que a transformada *wavelet* discreta (19) (20) consiste numa alternativa mais eficiente do que a Transformada de Fourier de Tempo Reduzido (18) para realizar a análise tempo-frequência de um sinal, bem como realizar filtragens e separações de um sinal em sub-bandas do seu espectro de frequências. Duas abordagens básicas existem para compreender a DWT. A primeira está baseada em conceitos da álgebra linear, possuindo cunho mais algébrico, e a segunda está mais próximas dos conceitos de processamento de sinais. Devido a sua praticidade e aplicabilidade, a segunda abordagem será adotada neste trabalho. De acordo com ela, e ainda com base na explicação dada por (2) em sua dissertação, da qual alguns parágrafos e figuras seguintes estão sendo reproduzidos com a devida permissão, constata-se que a DWT funciona como um par de filtros, sendo um deles passa-baixas ($h[\cdot]$) e o outro passa-altas ($g[\cdot]$), em geral, com frequência de corte em metade da máxima frequência presente no sinal de entrada. Dado um sinal discreto, ele é submetido a ambos os filtros via convolução(18). Cada vez que esse processo é aplicado, diz-se que se tem um nível de decomposição e obtêm-se dois novos sinais, um dos quais contém as frequências abaixo da metade da máxima frequência original do sinal, e o outro contém as frequências acima desse limiar. Em particular, os termos chamados coeficientes de detalhamento designam o sinal

obtido quando da passagem do sinal original pelo filtro passa-altas, e os termos chamados coeficientes de aproximação designam o sinal obtido quando da passagem do sinal original pelo filtro passa-baixas. Após aplicar um nível de decomposição no sinal, apenas o novo sinal obtido pela aplicação do filtro passa-baixas é usado para continuar o processo recursivo de decomposição.

Um detalhe fundamental a ser notado é que, cada vez que um nível da transformação é realizado, os dois novos sinais obtidos são sub-amostrados por 2, pois eles contêm apenas metade da faixa de frequências do sinal original, de acordo com o Teorema da Amostragem (18) e como ilustra a figura 2.12. Um sinal de n amostras tem a sua transformada *wavelet* com a mesma quantidade de amostras, composta por uma sequência de coeficientes, iniciando-se com aqueles provenientes da aplicação do filtro passa-baixas no último nível, seguidos pelos coeficientes resultantes da aplicação dos filtros passa-altas nos níveis intermediários, e terminando com os coeficientes resultantes da aplicação do filtro passa-altas do primeiro nível de decomposição.

Para realizar a decomposição até o último nível possível, é necessário que o sinal discreto tenha comprimento equivalente a uma potência de 2, sendo possível realizar $\frac{\log(n)}{\log(2)}$ decomposições para um sinal de comprimento n . Um fator muito importante para que um filtro digital seja considerado um filtro *wavelet* é que a resposta em frequência do filtro passa-baixas seja 0 em $\omega = \pi$.

O processo conjunto de filtragem e sub-amostragem por 2, realizado nos sinais da transformada *wavelet* em cada nível, pode ser representado por uma convolução modificada da seguinte forma:

$$y[\cdot] = x[\cdot] * t[\cdot] = \sum_{k=0}^{n-1} t_k x_{2n-k} \quad , \quad (2.4)$$

ou, mais especificamente:

$$y_{passa-baixas}[\cdot] = x[\cdot] * h[\cdot] = \sum_{k=0}^{n-1} h_k x_{2n-k} \quad (2.5)$$

e

$$y_{passa-altas}[\cdot] = x[\cdot] * g[\cdot] = \sum_{k=0}^{n-1} g_k x_{2n-k} \quad . \quad (2.6)$$

•

•

•

Figura 2.12 – Funcionamento da DWT, exemplificado para um sinal $s[\cdot]$ de n amostras discretas e máxima frequência π , decomposto até o terceiro nível. É importante observar o espectro de frequências e a quantidade de amostras presentes em cada sub-banda (adaptado de (2))

A DWT está diretamente relacionada à análise de multi-resolução (*Multi-Resolution Analysis* - MRA), proposta por Mallat, Meyer, Stromberg e outros (21) (22), que consiste em decompor um vetor (sinal sob análise) $\vec{f}$ em uma soma de outros vetores pertencentes a uma sequência de sub-espços vetoriais (23). Em outras palavras, isso significa representar um sinal em vários níveis de resolução. Então, de acordo com a MRA, para um vetor $\vec{f}$ de n pontos tem-se:

$$\vec{f} = \vec{A} + \vec{D} \quad (2.7)$$

onde

$$\vec{A} = \sum_{k=0}^{\frac{n}{2}-1} \langle \vec{f}, \vec{v}_k \rangle \vec{v}_k \quad ,$$

e

$$\vec{D} = \sum_{k=0}^{\frac{n}{2}-1} \langle \vec{f}, \vec{w}_k \rangle \vec{w}_k \quad ,$$

ou seja:

- $\vec{A}$ é a projeção de $\vec{f}$ num sub-espço V , com uma base de $\frac{n}{2}$ vetores;
- $\vec{D}$ é a projeção de $\vec{f}$ num sub-espço W , com uma base de $\frac{n}{2}$ vetores;
- $V \perp W \leftrightarrow \vec{A} \perp \vec{D}$;
- $\vec{v}_i \perp \vec{w}_i \leftrightarrow \langle \vec{v}_i, \vec{w}_i \rangle = 0$.

O processo acima consiste na decomposição em nível 1. Numa transformada *wavelet* de nível 2, o vetor A é novamente decomposto na soma de dois outros vetores ortogonais, podendo esse processo ser repetido $\frac{\log(n)}{\log(2)}$ vezes, conforme já foi mencionado. Dessa forma, generalizando, para uma decomposição de nível j , temos:

$$\vec{f} = \vec{A}_j + \sum_{i=1}^j \vec{D}_i. \quad (2.8)$$

sendo que:

- $\vec{A}_j$ é a projeção de $\vec{f}$ num sub-espço V_j , com uma base contendo $\frac{n}{2^j}$ vetores;

- $\vec{D}_i$ é a projeção de $\vec{f}$ num sub-espço W_i , com uma base contendo $\frac{n}{2^i}$ vetores;
- $V_j \perp W_j \leftrightarrow \vec{A}_j \perp \vec{D}_j$;
- $v_{i,j}^{\vec{}} \perp w_{i,j}^{\vec{}} \leftrightarrow \langle v_{i,j}^{\vec{}}, w_{i,j}^{\vec{}} \rangle = 0$.

O processo anterior equivale a escrever (24):

$$f[n] = \sum_{k=0}^{\frac{n}{2^j}-1} H_{j,k}[n] \phi_{j,k}[n] + \sum_{t=1}^j \sum_{k=0}^{\frac{n}{2^j}-1} G_{t,k}[n] \psi_{t,k}[n] \quad (2.9)$$

onde

- $\phi[n]$ e $\psi[n]$ formam uma base de Riesz (24) para escrever $\vec{f}$;
- $\phi[n] = \sum_k h_n \phi[2n - k]$, definida recursivamente por dilatações e translações de si mesma é chamada função *scaling* (24);
- $\psi[n] = \sum_k g_n \phi[2n - k]$, também definida recursivamente, é chamada função *wavelet* e é ortogonal a função *scaling*;
- $H_{j,k}[n] = \langle f, \phi_{j,k}[n] \rangle$;
- $G_{t,k}[n] = \langle f, \psi_{t,k}[n] \rangle$;
- $\{0\} \leftarrow \dots \subset V_{-1} \subset V_0 \subset V_1 \subset \dots \rightarrow L^2$;
- se $f[n] \in V_j \rightarrow f[2n] \in V_{j+1}$;
- $V_{j+1} = V_j \oplus W_j$;
- os coeficientes h_k correspondem ao filtro passa-baixas;
- os coeficientes g_k correspondem ao filtro passa-altas;
- $h[\cdot]$ e $g[\cdot]$, que são chamados filtros de análise, formam um par de *Quadrature Mirror Filters* - QMF;
- um filtro com k coeficientes é dito filtro de suporte k .

Cada par de filtros de análise, $h[\cdot]$ e $g[\cdot]$, possui uma única função *scaling* (ϕ) e uma única função *wavelet* (ψ) associadas. A forma de obtenção dessas funções a partir dos filtros, e vice-versa, está documentada com detalhes em (20), não sendo apresentada aqui por estar fora do escopo do trabalho.

É muito importante também o fato de $h[\cdot]$ e $g[\cdot]$ possuírem outros filtros associados, chamados filtros de síntese, representados por $\bar{h}[\cdot]$ e $\bar{g}[\cdot]$, que são utilizados para inverter a transformada, recuperando o sinal original a partir do transformado. Tais filtros obedecem às relações das equações 2.10, 2.11 e 2.12, para $k = 0, \dots, n-1$, que ficam mais claras através do exemplo na figura 2.13, para filtros de suporte 4.

$$g_k = (-1)^k h_{N-k-1} \quad , \quad (2.10)$$

$$\bar{h}_k = h_{N-k-1} \quad , \quad (2.11)$$

$$\bar{g}_k = (-1)^{k+1} h_k \quad . \quad (2.12)$$

Quando $h[\cdot]$, $g[\cdot]$, $\bar{h}[\cdot]$, e $\bar{g}[\cdot]$ mantêm as relações acima, eles constituem um banco de filtros de reconstrução perfeita (*perfect reconstruction filter bank* - PRFB) (21) (22), ou seja, as condições de *anti-aliasing* e *no-distortion*, no domínio Z, representadas nas equações 2.13 e 2.14, respectivamente, são satisfeitas. No presente trabalho a inversão da DWT não se faz necessária, entretanto, é desejável que o algoritmo proposto utilize apenas PRFBs.

Figura 2.13 – Relação entre os filtros de análise e síntese

$$\bar{H}[z] = G[-z] \quad , \quad \bar{G}[z] = -H[-z] \quad . \quad (2.13)$$

$$\bar{H}[z]H[z] + \bar{G}[z]G[z] = 2z^{-N+1} \quad . \quad (2.14)$$

Cálculo da DWT:

Para o cálculo da DWT de um sinal, aplica-se o algoritmo de Mallat, que está minuciosamente descrito em (20). É importante observar que apenas os filtros $h[\cdot]$ e $g[\cdot]$ são utilizados, não sendo necessária a utilização de ϕ e ψ . O procedimento de cálculo envolve apenas a multiplicação de duas matrizes para cada nível de transformação. Se $A[\cdot][\cdot]$ é a matriz de coeficientes dos filtros e $B[\cdot]$ é o sinal original, então $C[\cdot] = A[\cdot][\cdot]B[\cdot]$ corresponde ao sinal transformado, sendo que a disposição dos coeficientes nas matrizes é a seguinte:

$$A[\cdot][\cdot] = \begin{pmatrix} h_0 & h_1 & h_2 & \dots & \dots & \dots & h_{n-1} & 0 & 0 & 0 & 0 & \dots & \dots & 0 & 0 \\ g_0 & g_1 & g_2 & \dots & \dots & \dots & g_{n-1} & 0 & 0 & 0 & 0 & \dots & \dots & 0 & 0 \\ 0 & 0 & h_0 & h_1 & h_2 & \dots & \dots & h_{n-1} & 0 & 0 & 0 & \dots & \dots & 0 & 0 \\ 0 & 0 & g_0 & g_1 & g_2 & \dots & \dots & g_{n-1} & 0 & 0 & 0 & \dots & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \vdots & & & & & & & & & & & & & & \vdots \\ h_{n-1} & 0 & 0 & 0 & \dots & \dots & \dots & 0 & 0 & h_0 & h_1 & \dots & \dots & h_{n-3} & h_{n-2} \\ g_{n-1} & 0 & 0 & 0 & \dots & \dots & \dots & 0 & 0 & g_0 & g_1 & \dots & \dots & g_{n-3} & g_{n-2} \end{pmatrix} \quad ,$$

$$B[\cdot] = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ \dots \\ \vdots \\ b_{n-2} \\ b_{n-1} \end{pmatrix} \quad e \quad C[\cdot] = \begin{pmatrix} c_0 \\ c_{\frac{n}{2}} \\ c_1 \\ c_{\frac{n}{2}+1} \\ \dots \\ \vdots \\ c_{n-1} \\ c_{\frac{n}{2}-1} \end{pmatrix} \quad .$$

Pode-se notar na matriz $A[\cdot][\cdot]$ que dois procedimentos estão embutidos no algoritmo de Mallat: *downsampling* e *wrap-around* (20). O primeiro já foi mencionado anteriormente, e o segundo consiste em fazer com que os últimos coeficientes dos filtros ocupem as posições iniciais de cada linha. Isso faz com que a DWT tenha sempre o mesmo número de elementos do sinal original.

Famílias de Transformadas *Wavelet*:

Conforme ilustra a tabela 2.6 e as figuras 2.14, 2.15 e 2.16, as diversas famílias de filtros *wavelet* existentes (20) diferem no suporte dos filtros, assim como nas características de resposta em frequência e fase dos mesmos, o que faz com que as funções ϕ e ψ também sofram reflexo de tais diferenças. Foram utilizadas neste trabalho, para fins de separação dos sinais de vozes em sub-bandas de frequências, as *wavelets* de Haar, Daubechies, Symmlets, Coiflets, Vaidyanathan e Beylkin, com diversos suportes, todas constituindo filtros com respostas finitas ao impulso (*Finite Impulse Response* - FIR), isto é, limitadas (18), sendo que as respostas em frequência se aproximam das ideais à medida que o suporte cresce.

Tabela 2.6 – Características das famílias de *wavelets* utilizadas no presente trabalho, incluindo a quantidade de momentos da função *wavelet*

Família	Suporte(n)	Fase	Observação	Momentos
Haar	2	linear	é a mais simples das <i>wavelets</i> , criada por Alfred Haar (25)(20)	1
Daubechies	par, maior que 4	não linear	resposta ao impulso <i>maximally flat</i> , criada por Ingrid Daubechies (25)(20)	$\frac{n}{2}$
Symmlets	par, múltiplo de 8	não linear	resposta ao impulso mais simétrica(25)(20)	$\frac{n}{2} - 2$
Coiflets	par, múltiplo de 6	quase linear	resposta ao impulso quase simétrica, criada por Ronald Coifman (25)(20)	$\frac{n}{2} - 1$
Vaidyanathan	24	não linear	otimizada para voz, criada por P. P. Vaidyanathan (25)(20)	—
Beylkin	18	não linear	otimizada para áudio em geral (25)(20)	$\frac{n}{2} - 2$

Figura 2.14 – Formato das respostas ao impulso dos filtros *wavelet* de Haar, Daubechies, Vaidyanathan, Beylkin, Coiflet e Symmlet, respectivamente (2)

Figura 2.15 – Formatos das funções *scaling* dos filtros *wavelet* de Haar, Daubechies, Vaidyanathan, Beylkin, Coiflet e Symmlet, respectivamente (2)

Figura 2.16 – Formatos das funções *wavelet* dos filtros *wavelet* de Haar, Daubechies, Vaidyanathan, Beylkin, Coiflet e Symmlet, respectivamente (2)

Capítulo 3

Descrição da Técnica Proposta

Este capítulo descreve o sistema proposto, incluindo sua arquitetura e cada passo dos algoritmos associados para treinamento e uso do mesmo, entre outros detalhes.

3.1 A Arquitetura do Sistema Proposto

O sistema proposto foi dividido em dois módulos distintos, ilustrados nas figuras 3.1 e 3.2, localizadas ao final do capítulo. O primeiro módulo é responsável pelo treinamento do sistema de classificação com base em um banco de vozes do Hospital das Clínicas de Ribeirão Preto da USP, que está detalhado na seção de Materiais e Métodos. Já o segundo corresponde ao módulo de aplicação, que funciona com base no sistema previamente treinado, utilizado para classificar uma voz desconhecida como pertencente a um indivíduo normal ou com alguma patologia no trato vocal.

Conforme pode-se observar na figura 3.1, o módulo de treinamento é responsável por treinar, de maneira isolada e com base numa abordagem semi-supervisionada, duas SVMs,

sendo essa uma das características diferenciais do presente trabalho em comparação aos esquemas tradicionais encontrados na literatura. Particularmente, o módulo LPAV lê o arquivo de voz sendo processado para treinamento e, em seguida, o módulo CVNV faz uma consulta na base de dados para obter a informação se o sinal de voz sob análise corresponde ao de um paciente normal ou portador de alguma patologia. No primeiro caso, o módulo EP_N é acionado para extrair os parâmetros necessários para o treinamento. Já no segundo caso, o módulo EP_P é acionado para extrair os mesmos parâmetros do caso anterior. No caso de normalidade da voz, o módulo SVM_N entra em funcionamento para treinar uma das SVMs, chamada SVM_N, fixando o valor 1 em sua saída. Por outro lado, no caso de uma voz oriunda de um portador de alguma patologia no trato vocal, o módulo SVM_P é acionado para treinar a outra SVM, chamada SVM_P, fixando o valor -1 em sua saída. Finalmente, caso o módulo UAV detecte que ainda existem mais arquivos de vozes para treinamento, o sistema reinicia, lendo o próximo arquivo. Caso contrário, o treinamento está concluído.

Conforme mencionado anteriormente, o esquema adotado para o treinamento das SVMs é semi-supervisionado. Isso significa que, num primeiro momento, apenas os parâmetros de entrada são utilizados, sem qualquer relação com o valor de saída desejado (1 ou -1). Em particular, cada uma das SVMs eleva a dimensão do sinal de entrada de P para T , que representam, respectivamente, o número de parâmetros utilizados para cada exemplo de treinamento, e o número de exemplos de treinamento. Com o uso desse artifício, é sabido que existirá sempre uma solução direta para a respectiva formulação. A função *kernel* utilizada nessa etapa é a Gaussiana, $\phi(v) = \frac{1}{1+e^{-v}}$, e cada um dos T -ésimos elementos de processamento da camada intermediária é ajustado para apresentar saída máxima (1) somente para o T -ésimo exemplo de treinamento. O leitor pode recorrer à figura 2.9 do capítulo anterior para acompanhar essa explicação.

O módulo de aplicação, ilustrado na figura 3.2, lê o arquivo de voz desconhecido, extrai os parâmetros correspondentes, que são os mesmos extraídos no módulo de treinamento, injetando-os nas entradas de ambas as SVMs (SVM_N e SVM_P). Finalmente, é medida a diferença entre o valor de saída apresentado pela SVM_N e o valor 1. Da mesma forma, é

medida a diferença entre o valor de saída apresentado pela SVM_P e o valor -1 . A menor diferença corresponde à classificação emitida como resposta do sistema, isto é, voz normal ou pertencente a um indivíduo com alguma patologia no trato vocal, respectivamente.

Uma observação interessante é a de que o sistema de classificação planejado, por treinar cada uma das SVMs levando em conta apenas os elementos de uma classe particular a ser identificada, é relativamente distinto dos esquemas tradicionais de classificação. Em particular, cada uma das SVMs, isoladamente, ao invés de criar uma margem de separação entre elementos de classes distintas, minimiza as distâncias de cada ponto do espaço até o ponto 1 ou -1 , conforme o caso. A rotina EMR do módulo de aplicação forma, então, em conjunto com as duas SVMs, um sistema de classificação que leva em conta dois procedimentos de minimização, um de cada SVM, escolhendo a melhor solução. A combinação das SVMs com a estrutura de decisão da rotina EMR é uma das chaves do presente trabalho, que permitiu a obtenção de resultados promissores, conforme documentado no próximo capítulo.

As próximas sub-seções detalham os algoritmos utilizados nos módulos de treinamento e de aplicação, respectivamente. Já os parâmetros mencionados acima, tanto para o treinamento quanto para a aplicação do classificador, serão discutidos no próximo capítulo.

3.1.1 Detalhamento do algoritmo proposto para o módulo de treinamento

Como complemento às explicações apresentadas anteriormente, segue a descrição detalhada das etapas do algoritmo proposto para o módulo de treinamento do sistema.

- **INÍCIO**

- **Repita:**

- **PASSO 1:** (rotina LPAV): procurar na base de dados de treinamento o próximo arquivo WAV disponível, extraindo os “dados brutos” (amostras digitalizadas) correspondentes a um mesmo locutor;
- **PASSO 2:** (rotina CVNV): procurar na base de dados de treinamento a indicação textual que caracteriza o arquivo WAV de voz em questão como pertencente à classe normal ou da classe patológica;
- **PASSO 3:** (rotina CVNV): se a voz sob análise pertence a classe normal, então:
  - * **INÍCIO**
  - * **PASSO 3.1.A:** para cada janela de 1024 amostras do sinal, considerando 50% de sobreposição entre janelas consecutivas, extrair o conjunto de parâmetros a serem utilizados para o treinamento;
  - * **PASSO 3.2.A:** (rotina SVM_N): utilizar os parâmetros para treinar a SVM_N, conforme descrito anteriormente, objetivando o valor 1 na saída;
  - * **FIM;**
- senão:
  - * **INÍCIO**
  - * **PASSO 3.1.B:** para cada janela de 1024 amostras do sinal, considerando 50% de sobreposição entre janelas consecutivas, extrair o conjunto de parâmetros a serem utilizados para o treinamento;
  - * **PASSO 3.2.B:** (rotina SVM_P): utilizar os parâmetros para treinar a SVM_P, conforme descrito anteriormente, objetivando o valor -1 na saída;

* **FIM;**

- **enquanto** ainda existem arquivos na base de treinamento (rotina UAV);
- **FIM.**

3.1.2 Detalhamento do algoritmo proposto para o módulo de aplicação

- **INÍCIO**
- **PASSO 1:** (rotina LAV): ler o arquivo WAV a ser classificado, extraíndo os seus “dados brutos”;
- **PASSO 2:** (rotina EP): para cada janela de 1024 amostras do sinal, considerando 50% de sobreposição entre janelas consecutivas, extrair o conjunto de parâmetros a serem utilizados para o teste;
- **PASSO 3:** (rotina SVM_N): injetar os parâmetros extraídos na SVM_N, obtendo o valor de saída ($S(SVM\_N)$);
- **PASSO 4:** (rotina SVM_P): injetar os mesmos parâmetros extraídos na SVM_P, obtendo o valor de saída ($S(SVM\_P)$);

- **PASSO 5:** (rotina EMR): se $|1 - (S(SVM_N))| < |1 - (S(SVM_P))|$ então:
  - **INÍCIO**
  - **PASSO 5.1.A:** classificar a voz como normal;
  - **FIM;**
- senão:
  - **INÍCIO**
  - **PASSO 5.1.B:** classificar a voz como pertencente a um indivíduo portador de patologia no trato vocal;
  - **FIM;**
- **FIM.**

3.2 Implementação

Em ambos os módulos de treinamento e de aplicação, os trechos de código referentes à extração dos parâmetros do sinal de voz foram implementados em linguagem C++ sob ambiente LINUX (26). Já os trechos referentes às etapas de processamento em si, de ambos os módulos, foram implementados com o uso da plataforma Java também sob ambiente LINUX. Os códigos-fonte respectivos encontram-se no apêndice II. O próximo capítulo descreve os testes e resultados obtidos com o algoritmo proposto.

Figura 3.1 – Arquitetura básica da estrutura proposta para treinamento do sistema de detecção de patologias do trato vocal. Legenda: [LPAV]: leitura do próximo arquivo de voz; [CVNV]: consulta para verificação de normalidade da voz; [EP_N]: extração de parâmetros da voz normal; [EP_P]: extração de parâmetros da voz de paciente com patologia; [SVM_N]: SVM das vozes normais; [SVM_P]: SVM das vozes dos indivíduos com patologias; [UAV]: último arquivo de voz.

Figura 3.2 – Arquitetura básica da estrutura proposta para uso do sistema de detecção de patologias do trato vocal, após treinamento. Legenda: [LAV]: leitura do arquivo de voz; [EP]: extração de parâmetros; [SVM_N]: SVM das vozes normais; [SVM_P]: SVM das vozes dos indivíduos com patologias; [EMR]: escolha da máxima resposta.

Capítulo 4

Testes e Resultados

Neste capítulo encontram-se os testes realizados, que foram divididos em dois grupos, assim como os respectivos resultados obtidos. No primeiro grupo de testes, diversas combinações de parâmetros extraídos dos fonemas vozeados /a/, /e/ e /i/, sustentados em média por 5 segundos, foram experimentados. No segundo grupo de testes utilizou-se a combinação de parâmetros com melhor desempenho no primeiro grupo para analisar a qualidade da resposta expedida pelo classificador proposto, variando-se os percentuais de vozes utilizadas para treinamento e para os testes.

4.1 Materiais e Métodos

Todos os sinais de vozes utilizados correspondem aos fonemas /a/, /e/, e /i/ sustentadas, que foram extraídas do banco de vozes do Hospital das Clínicas de Ribeirão Preto da

USP. Essa base de dados possui 50 vozes com características normais e outras 50 pertencentes a indivíduos com algumas patologias na laringe, tais como nódulo nas pregas vocais, edema de Reinke, entre outras, em diversos níveis. Todos os indivíduos foram previamente examinados por profissionais da área médica para confirmar seu estado saudável ou patológico. A base está separada por pastas nomeadas com o nome dos pacientes examinados. Em cada uma das pastas existem arquivos WAV, amostrados a uma taxa de 16000 Hz, 16-bits, com os fonemas retrocitados.

Os parâmetros utilizados nas experiências estão listados a seguir. Para todos eles, com exceção do último, foi extraído o desvio padrão com base em janelas de 1024 amostras, considerando 50% de sobreposição entre janelas consecutivas, conforme mencionado no capítulo anterior.

Parâmetros:

1. desvio padrão da dimensão do fractal (DP_D): para cada arquivo de voz correspondente a um fonema sustentado por um locutor, o desvio padrão dos valores das dimensões fractais obtidas em cada janela foi computado;
2. desvio padrão no período de *pitch* (DP_P): para cada arquivo de voz correspondente a um fonema sustentado por um locutor, o desvio padrão dos valores dos períodos de *pitch* obtidos em cada janela foi computado;
3. desvios padrões nas energias das sub-bandas (DP_E): para cada arquivo de voz correspondente a um fonema sustentado por um locutor, os 21 desvios padrões dos valores das energias de cada uma das primeiras 21 bandas críticas da escala Bark obtidas em cada janela foram computados;
4. desvios padrões nos coeficientes MFCCs (DP_M): para cada arquivo de voz correspondente a um fonema sustentado por um locutor, as 14 variâncias dos 14 primeiros coeficientes MFCCs obtidos em cada janela foram computadas. Os desvios padrões na primeira e na segunda derivadas dos coeficientes MFCC (DP_M_Δ₁ e DP_M_Δ₂, respectivamente) também foram obtidos e utilizados todas as vezes que os coeficientes DP_M foram consi-

derados. Para cada um dos 14 DP_Ms obtidos, foram calculadas as 13 derivadas primeiras e 12 derivadas segundas. Tendo em vista que os DP_Ms são valores discretos, as derivadas ficam definidas como os módulos das diferenças entre os valores vizinhos, isto é:

$$DP\_M\_Δ_{1,i} = |DP\_M_{i+1} - DP\_M_i|, \quad (0 \leq i \leq 12) \quad ;$$

$$DP\_M\_Δ_{2,j} = |DP\_M\_Δ_{1,j+1} - DP\_M\_Δ_{1,j}|, \quad (0 \leq j \leq 11).$$

5. sexo (SX): como tentativa do uso de um parâmetro não-acústico, foi experimentado atribuir os valores numéricos 1 e -1 para representar as vozes como pertencentes a indivíduos do sexo masculino e feminino, respectivamente. Não foi possível utilizar outros parâmetros não-acústicos, tais como altura, peso, índice de massa corporal, fumante/não-fumante, alcoólatra/não-alcoólatra, entre outros, uma vez que eles não estavam disponíveis.

4.2 Bateria de Testes

A bateria completa dos testes realizados segue, incluindo tabelas e gráficos, assim como os resultados obtidos, cuja discussão e interpretação, que levam às conclusões finais, encontram-se no próximo capítulo.

O primeiro grupo de testes, que objetivou verificar quais são os melhores parâmetros, dentre os citados anteriormente, para uso em conjunto com o classificador proposto, foi implementado com base nas diferentes combinações possíveis desses parâmetros que foram extraídos do banco de dados de vozes normais e patológicas, conforme consta nas tabelas 4.1, 4.2, 4.3 e 4.4, que reportam todas as possíveis combinações dois-a-dois, três-a-três, quatro-a-quatro e com os cinco conjuntos de parâmetros listados acima. Testes e resultados considerando apenas um parâmetro como entrada para o classificador não foram satisfatórios, como era de se esperar,

e portanto, não estão sendo reportados. Neste primeiro grupo de testes foram utilizados 50% das vozes normais e 50% das patológicas, escolhidas ao acaso, para treinamento do sistema. O restante das vozes foi utilizado para os testes.

O segundo grupo de testes consistiu em fixar a melhor combinação de parâmetros obtida no teste anterior, considerando as taxas de acertos para **ambas** as classes (normal e patológico), que corresponde ao conjunto SX, DP_M, DP_E e DP_P, e variar a porcentagem de vozes utilizadas para o treinamento do classificador, observando até que ponto a eficácia do sistema continuava satisfatória. Os resultados obtidos estão plotados no gráfico da figura 4.1.

As tabelas 4.1, 4.2, 4.3 e 4.4, assim como o gráfico da figura 4.1, fazem referência aos resultados obtidos com o uso do fonema /a/ sustentado de cada um dos indivíduos. Os fonemas /e/ e /i/ também foram experimentados, tendo apresentado eficácia de 20 a 30% menor, em geral. Esse resultado já foi também constatado anteriormente (1) com base em outros esquemas de classificação.

Para obtenção dos parâmetros DP_E, as filtragens foram implementadas com base na DWT. Vários filtros *wavelet* de diversos suportes foram experimentados, conforme listado no apêndice I. As variações dos filtros e dos suportes causaram alterações pouco significantes nos resultados, o que mostrou que o classificador projetado se adaptou bem a tais mudanças. De qualquer forma, considerando tais alterações mínimas, foi observado que os melhores resultados nas classificações foram obtidos com o uso dos filtros das famílias Coiflets e Symmlets de suporte por volta de 24 ou 30. Os resultados reportados nas tabelas 4.1, 4.2, 4.3 e 4.4, assim como o gráfico da figura 4.1, consideram tais filtros.

O próximo capítulo apresenta a discussão detalhada dos resultados apresentados nas tabelas e gráficos citados, assim como as conclusões e trabalhos futuros.

Tabela 4.1 – Testes do **grupo 1** para diversas combinações de dois parâmetros extraídos dos bancos de dados de vozes normais e patológicas. Legenda: D.P. : Desvio Padrão

<i>Parâmetros</i>	<i>% de acertos - normais</i>	<i>% de acertos - patológicas</i>
SX + DP_D	12%	94,11%
SX + DP_P	88%	97,05%
SX + DP_E	80%	32,35%
SX + DP_M	84%	75,47%
DP_E + DP_D	48%	73,52%
DP_E + DP_P	60%	82,35%
DP_E + DP_M	12%	85,29%
DP_P + DP_M	80%	94,11%
DP_P + DP_D	96%	88,23%
DP_D + DP_M	20%	91,17%

Tabela 4.2 – Testes do **grupo 1** para diversas combinações de três parâmetros extraídos dos bancos de dados de vozes normais e patológicas. Legenda: D.P. : Desvio Padrão

<i>Parâmetros</i>	<i>% de acertos - normais</i>	<i>% de acertos - patológicas</i>
SX + DP_D + DP_P	92%	88,23%
SX + DP_D + DP_M	96%	76,47%
SX + DP_D + DP_E	72%	55,88%
DP_D + DP_E + DP_M	12%	85,29%
DP_D + DP_E + DP_P	56%	88,23%
DP_E + DP_M + DP_P	84%	91,17%

Tabela 4.3 – Testes do **grupo 1** para diversas combinações de quatro parâmetros extraídos dos bancos de dados de vozes normais e patológicas. Legenda: D.P. : Desvio Padrão

<i>Parâmetros</i>	<i>% de acertos - normais</i>	<i>% de acertos - patológicas</i>
SX + DP_D + DP_P + DP_E	84%	91,17%
SX + DP_D + DP_P + DP_M	92%	88,23%
SX + DP_D + DP_M + DP_E	92%	76,47%
SX + DP_M + DP_E + DP_P	96%	94,11%
DP_D + DP_M + DP_E + DP_P	84%	94,11%

Tabela 4.4 – Testes do **grupo 1** para a combinação dos cinco parâmetros extraídos dos bancos de dados de vozes normais e patológicas. Legenda: D.P. : Desvio Padrão

<i>Parâmetros</i>	<i>% de acertos - normais</i>	<i>% de acertos - patológicas</i>
SX + DP_D + DP_M + DP_E + DP_P	92%	94,11%

Figura 4.1 – Testes do **grupo 2** realizados com os parâmetros que obtiveram melhor resposta nos testes do grupo 1 para **ambas** as classes (normal e patológica): SX, DP_M, DP_E e DP_P. O eixo horizontal corresponde à quantidade em percentual de vozes utilizadas para treinamento do classificador, enquanto o eixo vertical corresponde ao percentual de acerto dos testes realizados. A curva representada pela linha rosa corresponde às respostas obtidas utilizando-se a base de testes patológica e a curva representada pela linha azul corresponde às respostas obtidas utilizando-se a base de testes não-patológica.

Capítulo 5

Conclusões e Trabalhos Futuros

Após uma revisão bibliográfica de importantes conceitos, o presente trabalho apresentou um sistema não-invasivo para pré-diagnóstico de patologias na laringe, baseado na análise de parâmetros extraídos do sinal de voz de um locutor. O sistema foi projetado com base em uma estrutura composta por duas SVMs com *kernels* Gaussianos, associadas com uma rotina de decisão. Nessa abordagem proposta, cada uma das duas classes consideradas, normal e patológica, ficam a cargo de uma única SVM, isto é, cada SVM é treinada levando em conta apenas uma das classes. Essa concepção faz com que a idéia original por trás da SVM seja modificada, pois, ao invés de determinar um hiperplano ótimo para separar as classes, o treinamento da forma como foi realizado faz com que um processo de minimização seja implementado, buscando reduzir ao máximo a distância dos pontos no espaço até o ponto objetivo, que foi definido como 1 no caso das vozes normais e -1 no caso das patológicas. Por fim, no momento da aplicação do sistema para realizar uma classificação, a rotina de decisão retrocitada cuida de verificar qual das SVMs tem sua função objetivo mais próxima dos parâmetros de entrada analisados.

Os testes realizados com os parâmetros extraídos de janelas de 1024 amostras, com 50% de sobreposição, mostraram-se eficientes no pré-diagnóstico de diversas patologias na laringe. Os resultados mais promissores para a identificação correta de ambas as classes utilizaram sexo (SX), desvio-padrões dos MFCCs (DP_M) e suas derivadas, desvios padrões

das energias das sub-bandas de frequências (DP_E) e desvio-padrão do *pitch* (DP_P) como parâmetros de entrada. Já para identificação correta das vozes normais apenas, os parâmetros desvio-padrão do *pitch* (DP_P) e da dimensão fractal (DP_D) foram suficientes para produzir quase que a totalidade dos resultados corretos. Por sua vez, para identificação das vozes patológicas, o uso da combinação de sexo (SX) com o desvio-padrão do *pitch* (DP_P) forneceu resultados consideráveis. Tais resultados permitiram observar que o único parâmetro não-acústico utilizado foi relevante nas experiências.

Embora não tenham sido reportados com mais detalhes, foram realizados diversos testes com as várias combinações de parâmetros mencionados utilizando uma única SVM para separar as duas classes em questão. Essa abordagem tradicional não ofereceu, entretanto, resultados nem sequer próximos aos melhores resultados obtidos com a abordagem proposta. Ainda na abordagem proposta, a variação dos *kernels* das SVMs utilizadas não forneceu resultados melhores do que o *kernel* Gaussiano.

Durante os testes, foi possível observar que os fonemas /e/ e /i/ apresentaram respostas insatisfatórias em comparação com os resultados obtidos com o fonema /a/. Observou-se também que, conforme reportado anteriormente, a variação dos filtros *wavelet* no cálculo das energias das sub-bandas críticas de frequência causou alterações pouco significantes nos resultados, o que mostrou que o classificador projetado se adaptou bem a tais mudanças. De qualquer forma, foi observado que os melhores resultados nas classificações foram obtidos com o uso dos filtros das famílias Coiflets e Symmlets de suporte por volta de 24 ou 30. Isso significa que as respostas em frequência próximas das ideais e as respostas em fase quase lineares, que são as características dessas famílias de filtros, são importantes, isto é, os conteúdos de frequências teoricamente especificados para as bandas críticas devem ser respeitados e, além disso, o atraso constante inerente ao processo de filtragem com filtros de fase linear também é fator importante. A faixa adotada para o tamanho do suporte dos filtros, que equilibra resolução no tempo e na frequência, deve ser levada em conta.

De maneira geral, foi possível observar uma considerável capacidade de generalização do sistema de classificação proposto. Com base no grafico 4.1 do capítulo anterior, é

possível notar que o uso de 35% a 40% da base de vozes para treinar o sistema já foi suficiente para a obtenção de um alto índice de acertos nas classificações, para ambas as classes.

Em termos de trabalhos futuros, pretende-se obter outras bases de dados para implementar o uso de outros parâmetros não-acústicos, conforme mencionado anteriormente, analisando as suas viabilidades para o pré-diagnóstico de patologias na laringe. Implementações em tempo-real do classificador proposto, com base em FPGAs e DSPs (27), também estão programadas. Por fim, pretende-se incrementar o sistema projetado para classificar e distinguir entre as diversas patologias mencionadas, com abordagens do tipo *one-to-many* (8), por exemplo.

Referências

- 1 FONSECA, E.S. *Wavelets, predição linear e LS-SVM aplicados na análise e classificação de sinais de vozes patológicas*. 2008. 107p. Tese (Doutorado). Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2008.
- 2 BARBON JR, S. *Dynamic time warping baseado na transformada wavelet*. 2007. 113p. Dissertação (Mestrado). Instituto de Física de São Carlos da Universidade de São Paulo, São Carlos, 2007.
- 3 BRAGA, J. C.; OLIVEIRA, D. S. F.; ATHERINO, C. C. T.; SCHOTT, T. C. A.; SILVA, J. C. Vocal nodules: functional and anatomical analysis. *Revista CEFAC*, São Paulo, v.8, n.2, p.223-9, abr-jun, 2006.
- 4 BEHLAU, M. *Avaliação e tratamento das disfonias*. São Paulo: Lovise, 1995.
- 5 PREGAS vocais com Edema de Reinke. Disponível em: <http://www.msdbrazil.com/msdbrazil/patients/manual_Merck/mm_sec19_214.html#section_7>. Acesso em: Dezembro 2009.
- 6 PÓLIPOS nas pregas vocais. Disponível em: <<http://www.brasilecola.com/doencas/polipos-nas-cordas-vocais.htm>>. Acesso em: Dezembro 2009.
- 7 LUDWIG JR, O.; MONTGOMERY, E. *Redes Neurais - fundamentos e aplicações com programas em C*. Rio de Janeiro: Ciência Moderna. 2007, p. 11-14.
- 8 RUSSEL, S. *Inteligência artificial*. São Paulo: Campus, 2003.
- 9 LORENA, A. C.; CARVALHO, A. C. P. L. F. Uma introdução às Support Vector Machines. *RITA*, São Paulo, p.1-5, v.14, n.2, 2007.
- 10 BARANOSKI, F. L.; JUSTINO, E. J. R.; BORTOLOZZI, F. Identificação da autoria em documentos manuscritos usando SVM. In: CONGRESSO DA SOCIEDADE BRASILEIRA DE COMPUTAÇÃO (SBC). 25; 2005. *Anais...*, São Leopoldo: SBC, 2005, p. 543-553.

- 11 HAYKIN, S. *Redes neurais: princípios e aplicações*. 2a ed. Porto Alegre: Bookman, 2004.
- 12 VIEIRA, L. S. *Conversão de voz baseada na transformada wavelet*. 2008. 91p. Dissertação (Mestrado). Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2008.
- 13 FANTINATO, P. C. *Segmentação de voz baseada na análise fractal e na transformada wavelet*. 2009. 123p. Dissertação (Mestrado). Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2009.
- 14 BOSI, M; GOLDBERG, R. E. *Introduction to digital audio coding and standards*. Boston-USA: Kruwer Academic Press, 2003.
- 15 DENG, L.; O'SHAUGHNESSY, O. *Speech processing: a dynamic and optimization-oriented approach*. New York: Marcel Dekker, 2003.
- 16 SANCHEZ, F. L. *Análise cepstral baseada em diferentes famílias de transformadas wavelet*. 2008. 96p. Dissertação (Mestrado). Interunidades em Bioengenharia. Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2008.
- 17 AL-AKAIDI, M. *Fractal speech processing*. Cambridge-UK: Cambridge University Press, 2004.
- 18 HAYKIN, S.; VEEN, B. V. *Sinais e sistemas*. Porto Alegre: Bookman, 2002.
- 19 ADDISON, P. S. *The illustrated Wavelet transform handbook: introductory theory and applications in science, engineering, medicine and finance*. Edinburg-UK: Institute of Physics Publishing, 2002.
- 20 JENSEN, A. et al. *Ripples in mathematics: the discrete Wavelet transform*. New York-USA: Springer-Verlag, 2000.
- 21 AKAY, M. *Time-frequency and Wavelets in biomedical signal processing*. New York: IEEE Press in Biomedical Engineering, 1998.
- 22 WILLIAMS, J. R.; AMARATUNGA, K. Introduction to wavelets in engineering. *International Journal for Numerical Methods in Engineering*, v. 37, n.1, p. 2365-2388, 1994.
- 23 LIPSCHUTZ, S. *Álgebra linear*. 3a ed. São Paulo: Makron Books, 1994.
- 24 STRANG, G.; NGUYEN, T. *Wavelets and filter-banks*. Wellesley: Wellesley-Cambridge Academic Press, 1997.

- 25 WALKER, J. S. *A primer on Wavelets and their scientific applications*. Whashington-USA: Chapman and Hall/CRC, 1999.
- 26 SCHILDT, H. *C++: the complete reference*. 4th ed. New York: McGraw-Hill Osborne Media, 2002.
- 27 OLIVIEIRA, A. S.; ANDRADE, F. S. *Sistemas embarcados: hardware e firmware na prática*. São Paulo: Érica, 2006.

Apêndice I - Coeficientes dos filtros *wavelet* utilizados nas experiências.

Coeficientes referentes aos filtros $h[\cdot]$ (passa-baixas de análise).

- Haar :
{0.7071067, 0.7071067};
- Daub4 :
{ 4.82962913144e-01, 8.36516303737e-01, 2.24143868042e-01, -1.29409522551e-01};
- Daub6 :
{ 3.32670552950e-01, 8.06891509311e-01, 4.59877502118e-01, -1.35011020010e-01, -8.54412738820e-02, 3.52262918857e-02};
- Daub8 :
{ 2.30377813308e-01, 7.14846570552e-01, 6.30880767929e-01, -2.79837694168e-02, -1.87034811719e-01, 3.08413818355e-02, 3.28830116668e-02, -1.05974017850e-02};
- Daub10 :
{ 1.60102397974e-01, 6.03829269797e-01, 7.24308528437e-01, 1.38428145901e-01, -2.42294887066e-01, -3.22448695846e-02, 7.75714938400e-02, -6.24149021279e-03, -1.25807519990e-02, 3.33572528547e-03};
- Daub12 :
{ 1.11540743350e-01, 4.94623890398e-01, 7.51133908021e-01, 3.15250351709e-01, -2.26264693965e-01, -1.29766867567e-01, 9.75016055873e-02, 2.75228655303e-02, -3.15820393174e-02, 5.53842201161e-04, 4.77725751094e-03, -1.07730108530e-03};
- Daub14 :
{ 7.78520540850e-02, 3.96539319481e-01, 7.29132090846e-01, 4.69782287405e-01, -1.43906003928e-01, -2.24036184993e-01, 7.13092192668e-02, 8.06126091510e-02, -3.80299369350e-02, -1.65745416306e-02, 1.25509985560e-02, 4.29577972921e-04, -1.80164070404e-03, 3.53713799974e-04};
- Daub16 :
{ 5.44158422431e-02, 3.12871590914e-01, 6.75630736297e-01, 5.85354683654e-01, -1.58291052563e-02, -2.84015542961e-01, 4.72484573913e-04, 1.28747426620e-01, -1.73693010018e-02, -4.40882539307e-02, 1.25509985560e-02, 4.29577972921e-04, -1.80164070404e-03, 3.53713799974e-04};

02, 1.39810279173e-02, 8.74609404740e-03, -4.87035299345e-03, -3.91740373376e-04, 6.75449406450e-04, -1.17476784124e-04};

- Daub18 :

{ 3.80779473638e-02, 2.43834674612e-01, 6.04823123690e-01, 6.57288078051e-01, 1.33197385825e-01, -2.93273783279e-01, -9.68407832229e-02, 1.48540749338e-01, 3.07256814793e-02, -6.76328290613e-02, 2.50947114831e-04, 2.23616621236e-02, -4.72320475775e-03, -4.28150368246e-03, 1.84764688305e-03, 2.30385763523e-04, -2.51963188942e-04, 3.93473203162e-05};

- Daub20 :

{ 2.66700579005e-02, 1.88176800077e-01, 5.27201188931e-01, 6.88459039453e-01, 2.81172343660e-01, -2.49846424327e-01, -1.95946274377e-01, 1.27369340335e-01, 9.30573646035e-02, -7.13941471663e-02, -2.94575368218e-02, 3.32126740593e-02, 3.60655356695e-03, -1.07331754833e-02, 1.39535174705e-03, 1.99240529518e-03, -6.85856694959e-04, -1.16466855129e-04, 9.35886703200e-05, -1.32642028945e-05};

- Daub22 :

{ 1.86942977614e-02, 1.44067021150e-01, 4.49899764356e-01, 6.85686774916e-01, 4.11964368947e-01, -1.62275245027e-01, -2.74230846817e-01, 6.60435881966e-02, 1.49812012466e-01, -4.64799551166e-02, -6.64387856950e-02, 3.13350902190e-02, 2.08409043601e-02, -1.53648209062e-02, -3.34085887301e-03, 4.92841765605e-03, -3.08592858815e-04, -8.93023250666e-04, 2.49152523552e-04, 5.44390746993e-05, -3.46349841869e-05, 4.49427427723e-06};

- Daub24 :

{ 1.31122579572e-02, 1.09566272821e-01, 3.77355135214e-01, 6.57198722579e-01, 5.15886478427e-01, -4.47638856537e-02, -3.16178453752e-01, -2.37792572560e-02, 1.82478605927e-01, 5.35956967435e-03, -9.64321200965e-02, 1.08491302558e-02, 4.15462774950e-02, -1.22186490697e-02, -1.28408251983e-02, 6.71149900879e-03, 2.24860724099e-03, -2.17950361862e-03, 6.54512821250e-06, 3.88653062820e-04, -8.85041092082e-05, -2.42415457570e-05, 1.27769522193e-05, -1.52907175806e-06};

- Daub26 :

{ 9.20213353896e-03, 8.28612438729e-02, 3.11996322160e-01, 6.11055851158e-01, 5.88889570431e-01, 8.69857261796e-02, -3.14972907711e-01, -1.24576730750e-01, 1.79476079429e-01, 7.29489336567e-02, -1.05807618187e-01, -2.64884064753e-02, 5.61394771002e-02, 2.37997225405e-03, -2.38314207103e-02, 3.92394144879e-03, 7.25558940161e-03, -2.76191123465e-03, -1.31567391189e-03, 9.32326130867e-04, 4.92515251262e-05, -1.65128988556e-04, 3.06785375793e-05, 1.04419305714e-05, -4.70041647936e-06, 5.22003509845e-07};

- Daub28 :

{ 6.46115346008e-03, 6.23647588493e-02, 2.54850267792e-01, 5.54305617940e-01, 6.31187849104e-01, 2.18670687758e-01, -2.71688552278e-01, -2.18033529993e-01, 1.38395213864e-01, 1.39989016584e-01, -8.67484115681e-02, -7.15489555040e-02, 5.52371262592e-02, 2.69814083079e-02, -3.01853515403e-02, -5.61504953035e-03, 1.27894932663e-02, -7.46218989268e-04, -3.84963886802e-03, 1.06169108560e-03, 7.08021154235e-04, -3.86831947312e-04, -4.17772457703e-05, 6.87550425269e-05, -1.03372091845e-05, -4.38970490178e-06, 1.72499467536e-06, -1.78713996831e-07};

- Daub30 :

{ 4.53853736157e-03, 4.67433948927e-02, 2.06023863986e-01, 4.92631771708e-01, 6.45813140357e-01, 3.39002535454e-01, -1.93204139609e-01, -2.88882596566e-01, 6.52829528487e-02, 1.90146714007e-01, -3.96661765557e-02, -1.11120936037e-01, 3.38771439235e-02, 5.47805505845e-02, -2.57670073284e-02, -2.08100501696e-02, 1.50839180278e-02, 5.10100036040e-03, -6.48773456031e-03, -2.41756490761e-04, 1.94332398038e-03, -3.73482354137e-04, -3.59565244362e-04, 1.55896489920e-04, 2.57926991553e-05, -2.81332962660e-05, 3.36298718173e-06, 1.81127040794e-06, -6.31688232588e-07, 6.13335991330e-08};

- Daub32 :

{ 3.18922092534e-03, 3.49077143236e-02, 1.65064283488e-01, 4.30312722846e-01, 6.37356332083e-01, 4.40290256886e-01, -8.97510894024e-02, -3.27063310527e-01, -2.79182081330e-02, 2.11190693947e-01, 2.73402637527e-02, -1.32388305563e-01, -6.23972275247e-03, 7.59242360442e-02, -7.58897436885e-03, -3.68883976917e-02, 1.02976596409e-02, 1.39937688598e-02, -6.99001456341e-03, -3.64427962149e-03, 3.12802338120e-03, 4.07896980849e-04, -9.41021749359e-04, 1.14241520038e-04, 1.74787245225e-04, -6.10359662141e-05, -1.39456689882e-05, 1.13366086612e-05, -1.04357134231e-06, -7.36365678545e-07, 2.30878408685e-07, -2.10933963010e-08};

- Daub34 :

{ 2.24180700103e-03, 2.59853937036e-02, 1.31214903307e-01, 3.70350724152e-01, 6.10996615684e-01, 5.18315764056e-01, 2.73149704032e-02, -3.28320748363e-01, -1.26599752215e-01, 1.97310589565e-01, 1.01135489177e-01, -1.26815691778e-01, -5.70914196316e-02, 8.11059866541e-02, 2.23123361781e-02, -4.69224383892e-02, -3.27095553581e-03, 2.27336765839e-02, -3.04298998135e-03, -8.60292152032e-03, 2.96799669152e-03, 2.30120524215e-03, -1.43684530480e-03, -3.28132519409e-04, 4.39465427768e-04, -2.56101095665e-05, -8.20480320245e-05, 2.31868137987e-05, 6.99060098507e-06, -4.50594247722e-06, 3.01654960999e-07, 2.95770093331e-07, -8.42394844600e-08, 7.26749296856e-09};

- Daub36 :

{ 1.57631021844e-03, 1.92885317241e-02, 1.03588465822e-01, 3.14678941337e-01, 5.71826807766e-

01, 5.71801654888e-01, 1.47223111969e-01, -2.93654040736e-01, -2.16480934005e-01, 1.49533975565e-01, 1.67081312763e-01, -9.23318841508e-02, -1.06752246659e-01, 6.48872162119e-02, 5.70512477385e-02, -4.45261419029e-02, -2.37332103958e-02, 2.66707059264e-02, 6.26216795430e-03, -1.30514809466e-02, 1.18630033858e-04, 4.94334360546e-03, -1.11873266699e-03, -1.34059629833e-03, 6.28465682965e-04, 2.13581561910e-04, -1.98648552311e-04, -1.53591712353e-07, 3.74123788074e-05, -8.52060253744e-06, -3.33263447888e-06, 1.76871298362e-06, -7.69163268988e-08, -1.17609876702e-07, 3.06883586304e-08, -2.50793445494e-09};

- Daub38 :

{ 1.10866976318e-03, 1.42810984507e-02, 8.12781132654e-02, 2.64388431740e-01, 5.24436377464e-01, 6.01704549127e-01, 2.60894952651e-01, -2.28091394215e-01, -2.85838631755e-01, 7.46522697081e-02, 2.12349743306e-01, -3.35185419023e-02, -1.42785695038e-01, 2.75843506256e-02, 8.69067555558e-02, -2.65012362501e-02, -4.56742262772e-02, 2.16237674095e-02, 1.93755498891e-02, -1.39883886785e-02, -5.86692228101e-03, 7.04074736710e-03, 7.68954359257e-04, -2.68755180070e-03, 3.41808653458e-04, 7.35802520505e-04, -2.60676135678e-04, -1.24600791734e-04, 8.71127046721e-05, 5.10595048707e-06, -1.66401762971e-05, 3.01096431629e-06, 1.53193147669e-06, -6.86275565776e-07, 1.44708829879e-08, 4.63693777578e-08, -1.11640206703e-08, 8.66684883899e-10};

- Daub40 :

{ 7.79953613666e-04, 1.05493946249e-02, 6.34237804590e-02, 2.19942113551e-01, 4.72696185310e-01, 6.10493238938e-01, 3.61502298739e-01, -1.39212088011e-01, -3.26786800434e-01, -1.67270883090e-02, 2.28291050819e-01, 3.98502464577e-02, -1.55458750707e-01, -2.47168273386e-02, 1.02291719174e-01, 5.63224685730e-03, -6.17228996246e-02, 5.87468181181e-03, 3.22942995307e-02, -8.78932492390e-03, -1.38105261371e-02, 6.72162730225e-03, 4.42054238704e-03, -3.58149425960e-03, -8.31562172822e-04, 1.39255961932e-03, -5.34975984399e-05, -3.85104748699e-04, 1.01532889736e-04, 6.77428082837e-05, -3.71058618339e-05, -4.37614386218e-06, 7.24124828767e-06, -1.01199401001e-06, -6.84707959700e-07, 2.63392422627e-07, 2.01432202355e-10, -1.81484324829e-08, 4.05612705555e-09, -2.99883648961e-10};

- Daub42 :

{ 5.48822509852e-04, 7.77663905235e-03, 4.92477715381e-02, 1.81359625440e-01, 4.19687944939e-01, 6.01506094935e-01, 4.44590451927e-01, -3.57229196172e-02, -3.35664089530e-01, -1.12397071568e-01, 2.11564527680e-01, 1.15233298439e-01, -1.39940424932e-01, -8.17759429808e-02, 9.66003903237e-02, 4.57234057492e-02, -6.49775048937e-02, -1.86538592021e-02, 3.97268354278e-02, 3.35775639033e-03, -2.08920536779e-02, 2.40347092080e-03, 8.98882438197e-03, -2.89133434858e-03, -2.95837403893e-03, 1.71660704063e-03, 6.39418500512e-04, -6.90671117082e-04, -3.19640627768e-05, 1.93664650416e-04, -3.63552025008e-05, -3.49966598498e-05, 1.53548250927e-05, 2.79033053981e-06, -3.09001716454e-

06, 3.16609544236e-07, 2.99213663046e-07, -1.00040087903e-07, -2.25401497467e-09, 7.05803354123e-09, -1.47195419765e-09, 1.03880557102e-10};

- Daub44 :

{ 3.86263231491e-04, 5.72185463133e-03, 3.80699372364e-02, 1.48367540890e-01, 3.67728683446e-01, 5.78432731009e-01, 5.07901090622e-01, 7.37245011836e-02, -3.12726580428e-01, -2.00568406104e-01, 1.64093188106e-01, 1.79973187992e-01, -9.71107984091e-02, -1.31768137686e-01, 6.80763143927e-02, 8.45573763668e-02, -5.13642542974e-02, -4.65308118275e-02, 3.69708466206e-02, 2.05867076275e-02, -2.34800013444e-02, -6.21378284936e-03, 1.25647252183e-02, 3.00137398507e-04, -5.45569198615e-03, 1.04426073918e-03, 1.82701049565e-03, -7.70690988123e-04, -4.23787399839e-04, 3.28609414213e-04, 4.34589990453e-05, -9.40522363481e-05, 1.13743496621e-05, 1.73737569575e-05, -6.16672931646e-06, -1.56517913199e-06, 1.29518205731e-06, -8.77987987336e-08, -1.28333622875e-07, 3.76122874933e-08, 1.68017140492e-09, -2.72962314663e-09, 5.33593882166e-10, -3.60211348433e-11};

- Daub46 :

{ 2.71904194128e-04, 4.20274889318e-03, 2.93100036578e-02, 1.20515531783e-01, 3.18450813852e-01, 5.44931147873e-01, 5.51018517241e-01, 1.81392625363e-01, -2.61392148030e-01, -2.71402098607e-01, 9.21254070824e-02, 2.23573658242e-01, -3.30374470942e-02, -1.64011321531e-01, 2.02830745756e-02, 1.12297043618e-01, -2.11262123562e-02, -7.02073915749e-02, 2.17658568344e-02, 3.84953325225e-02, -1.85235136501e-02, -1.75371010030e-02, 1.27519439315e-02, 6.03184065002e-03, -7.07531927370e-03, -1.13486547335e-03, 3.12287644981e-03, -2.46501400516e-04, -1.06123122888e-03, 3.19420492709e-04, 2.56762452007e-04, -1.50021850349e-04, -3.37889483412e-05, 4.42607120310e-05, -2.63520788924e-06, -8.34787556785e-06, 2.39756954684e-06, 8.14757483477e-07, -5.33900540520e-07, 1.85309178563e-08, 5.41754917953e-08, -1.39993549543e-08, -9.47288590181e-10, 1.05044645369e-09, -1.93240511131e-10, 1.25020330235e-11};

- Daub48 :

{ 1.91435800947e-04, 3.08208171490e-03, 2.24823399497e-02, 9.72622358336e-02, 2.72908916067e-01, 5.04371040839e-01, 5.74939221095e-01, 2.80985553233e-01, -1.87271406885e-01, -3.17943078999e-01, 4.77661368434e-03, 2.39237388780e-01, 4.25287296414e-02, -1.71175351370e-01, -3.87771735779e-02, 1.21016303469e-01, 2.09801137091e-02, -8.21616542080e-02, -4.57843624181e-03, 5.13016200399e-02, -4.94470942812e-03, -2.82131070949e-02, 7.66172188164e-03, 1.30499708710e-02, -6.29143537001e-03, -4.74656878632e-03, 3.73604617828e-03, 1.15376493683e-03, -1.69645681897e-03, -4.41618485614e-05, 5.86127059318e-04, -1.18123323796e-04, -1.46007981776e-04, 6.55938863930e-05, 2.18324146046e-05, -2.02288829261e-05, 1.34115775080e-08, 3.90110033859e-06, -8.98025314393e-07, -4.03250775687e-07, 2.16633965327e-07, -5.05764541979e-10, -2.25574038817e-08, 5.15777678967e-09, 4.74837582425e-10, -4.02465864458e-10, 6.99180115763e-11, -4.34278250380e-12};

- Daub50 :

{ 1.34802979347e-04, 2.25695959185e-03, 1.71867412540e-02, 7.80358628721e-02, 2.31693507886e-01, 4.59683415146e-01, 5.81636896746e-01, 3.67885074802e-01, -9.71746409646e-02, -3.36473079641e-01, -8.75876145876e-02, 2.24537819745e-01, 1.18155286719e-01, -1.50560213750e-01, -9.85086152899e-02, 1.06633805018e-01, 6.67521644940e-02, -7.70841110565e-02, -3.71739628611e-02, 5.36179093987e-02, 1.55426059291e-02, -3.40423204606e-02, -3.07983679484e-03, 1.89228044766e-02, -1.98942578220e-03, -8.86070261804e-03, 2.72693625873e-03, 3.32270777397e-03, -1.84248429020e-03, -8.99977423746e-04, 8.77258193674e-04, 1.15321244046e-04, -3.09880099098e-04, 3.54371452327e-05, 7.90464000396e-05, -2.73304811996e-05, -1.27719529319e-05, 8.99066139306e-06, 5.23282770815e-07, -1.77920133265e-06, 3.21203751886e-07, 1.92280679014e-07, -8.65694173227e-08, -2.61159855611e-09, 9.27922448008e-09, -1.88041575506e-09, -2.22847491022e-10, 1.53590157016e-10, -2.52762516346e-11, 1.50969208282e-12};

- Daub52 :

{ 9.49379575071e-05, 1.65052023353e-03, 1.30975542925e-02, 6.22747440251e-02, 1.95039438716e-01, 4.13292962278e-01, 5.73669043034e-01, 4.39158311789e-01, 1.77407678098e-03, -3.26384593691e-01, -1.74839961289e-01, 1.81291832311e-01, 1.82755409589e-01, -1.04323900285e-01, -1.47977193275e-01, 6.98231861132e-02, 1.06482405249e-01, -5.34485616814e-02, -6.86547596040e-02, 4.22321857963e-02, 3.85357159711e-02, -3.13781103630e-02, -1.77609035683e-02, 2.07349201799e-02, 5.82958055531e-03, -1.17854979061e-02, -5.28738399262e-04, 5.60194723942e-03, -9.39058250473e-04, -2.14553028156e-03, 8.38348805654e-04, 6.16138220457e-04, -4.31955707426e-04, -1.06057474828e-04, 1.57479523860e-04, -5.27779549303e-06, -4.10967399639e-05, 1.07422154087e-05, 7.00007868296e-06, -3.88740016185e-06, -4.65046322064e-07, 7.93921063370e-07, -1.07900423757e-07, -8.90446637016e-08, 3.40779562129e-08, 2.16932825985e-09, -3.77601047853e-09, 6.78004724582e-10, 1.00230319104e-10, -5.84040818534e-11, 9.13051001637e-12, -5.25187122424e-13};

- Daub54 :

{ 6.68713138543e-05, 1.20553123167e-03, 9.95258878087e-03, 4.94525999829e-02, 1.62922027502e-01, 3.67110214125e-01, 5.53849860990e-01, 4.93406122677e-01, 1.02840855061e-01, -2.89716803314e-01, -2.48264581903e-01, 1.14823019517e-01, 2.27273288414e-01, -3.87864186318e-02, -1.78031740959e-01, 1.57993974602e-02, 1.31197971717e-01, -1.40627515558e-02, -9.10229065295e-02, 1.73110182654e-02, 5.79694057347e-02, -1.85124935619e-02, -3.27390666310e-02, 1.61469669223e-02, 1.56655956489e-02, -1.15771864589e-02, -5.86209634546e-03, 6.85663560968e-03, 1.34262687730e-03, -3.33285446952e-03, 1.45752962593e-04, 1.30117745024e-03, -3.41835122691e-04, -3.87901857410e-04, 2.01971987969e-04, 7.66005838706e-05, -7.71114551779e-05, -3.51748361490e-06, 2.06344264773e-05, -3.90116407063e-06, -3.65750090818e-06, 1.63436962472e-06, 3.05088068625e-07, -3.47246814739e-07, 3.28655896805e-

08, 4.02625505286e-08, -1.32133227399e-08, -1.30946560685e-09, 1.52161498477e-09, -2.41552692801e-10, -4.37498622429e-11, 2.21366208806e-11, -3.29579012247e-12, 1.82818835288e-13};

- Daub56 :

{ 4.71080777501e-05, 8.79498515984e-04, 7.54265037764e-03, 3.90926081154e-02, 1.35137914253e-01, 3.22563361285e-01, 5.24998231630e-01, 5.30516293441e-01, 2.00176144045e-01, -2.30498954047e-01, -3.01327809532e-01, 3.28578791633e-02, 2.45808151373e-01, 3.69068853157e-02, -1.82877330732e-01, -4.68382337445e-02, 1.34627567910e-01, 3.44786312750e-02, -9.76853558056e-02, -1.73419228313e-02, 6.77478955019e-02, 3.44801895554e-03, -4.33333686160e-02, 4.43173291006e-03, 2.46880600101e-02, -6.81554976455e-03, -1.20635919682e-02, 5.83881662774e-03, 4.78486311245e-03, -3.72546124707e-03, -1.36037384563e-03, 1.87599866820e-03, 1.41567239314e-04, -7.48674955911e-04, 1.15465606365e-04, 2.29579098223e-04, -8.90390149004e-05, -4.90771341619e-05, 3.64140121105e-05, 4.63866498139e-06, -1.00432604133e-05, 1.24790031757e-06, 1.84036373451e-06, -6.67021547995e-07, -1.75746117320e-07, 1.49066001353e-07, -8.26238731562e-09, -1.78413869087e-08, 5.04404705638e-09, 6.94454032894e-10, -6.07704124722e-10, 8.49222001105e-11, 1.86736726378e-11, -8.36549047125e-12, 1.18885053340e-12, -6.36777235471e-14};

- Daub58 :

{ 3.31896627984e-05, 6.40951680304e-04, 5.70212651777e-03, 3.07735802214e-02, 1.11370116951e-01, 2.80653455970e-01, 4.89758804762e-01, 5.51374432758e-01, 2.89105238335e-01, -1.54028734459e-01, -3.30040948917e-01, -5.57068000729e-02, 2.36105236153e-01, 1.12419174873e-01, -1.60877988594e-01, -1.07845949938e-01, 1.14472295893e-01, 8.32207471624e-02, -8.51254926156e-02, -5.50274895253e-02, 6.34791645842e-02, 3.05315432727e-02, -4.51879812777e-02, -1.29171425542e-02, 2.94704318717e-02, 2.64832730767e-03, -1.70412245736e-02, 1.73788033272e-03, 8.46972549356e-03, -2.55080712778e-03, -3.47379898968e-03, 1.87712092572e-03, 1.08705394222e-03, -1.00077832708e-03, -2.00071136307e-04, 4.11128345474e-04, -2.29201804121e-05, -1.29304484008e-04, 3.64502606856e-05, 2.91334475016e-05, -1.65732839530e-05, -3.59364480402e-06, 4.75060924645e-06, -3.02905459205e-07, -8.97570175063e-07, 2.63389838699e-07, 9.38719741109e-08, -6.28615692201e-08, 1.07659190661e-09, 7.76897885477e-09, -1.89399538617e-09, -3.42680086326e-10, 2.40709945350e-10, -2.94058925076e-11, -7.83250973362e-12, 3.15276241337e-12, -4.28565487006e-13, 2.21919131158e-14};

- Daub60 :

{ 2.33861617273e-05, 4.66637950428e-04, 4.30079716504e-03, 2.41308326715e-02, 9.12383040670e-02, 2.42020670940e-01, 4.50487821853e-01, 5.57572232912e-01, 3.66242683371e-01, -6.61836707759e-02, -3.32966975020e-01, -1.41968513330e-01, 1.99462121580e-01, 1.77829873244e-01, -1.14558219432e-01, -1.57236817959e-01, 7.27786589703e-02, 1.22747746045e-01, -5.38064654582e-02, -8.76586900363e-02, 4.38016646714e-02, 5.67123657447e-02, -3.56733974967e-02, -3.22637589193e-02, 2.70786195952e-

02, 1.52879607698e-02, -1.83997438681e-02, -5.29685966613e-03, 1.09156316583e-02, 6.19671756497e-04, -5.53073014819e-03, 8.43384586662e-04, 2.32452009406e-03, -8.60927696811e-04, -7.67878250438e-04, 5.05094823903e-04, 1.72482584235e-04, -2.16171830116e-04, -8.54830546758e-06, 6.98200837080e-05, -1.33971686329e-05, -1.63615247872e-05, 7.25214553589e-06, 2.32754909849e-06, -2.18726767699e-06, 1.09947433852e-08, 4.26166232601e-07, -1.00041468235e-07, -4.76437996513e-08, 2.60544275497e-08, 5.55339786139e-10, -3.33110568046e-09, 6.98486269183e-10, 1.61362297827e-10, -9.46138799727e-11, 1.00010513139e-11, 3.23942863853e-12, -1.18523759210e-12, 1.54399757084e-13, -7.73794263095e-15);

- Daub62 :

{ 1.64801338645e-05, 3.39412203776e-04, 3.23688406862e-03, 1.88536916129e-02, 7.43360930116e-02, 2.07012874485e-01, 4.09192200037e-01, 5.51139840914e-01, 4.29468808206e-01, 2.71692124973e-02, -3.10955118319e-01, -2.17978485523e-01, 1.40178288765e-01, 2.24966711473e-01, -4.99263491604e-02, -1.86962360895e-01, 1.54369884294e-02, 1.45089500931e-01, -8.13983227346e-03, -1.07612773323e-01, 1.09412974523e-02, 7.53536117432e-02, -1.48800266181e-02, -4.86190754648e-02, 1.61541715659e-02, 2.80476193667e-02, -1.42762752777e-02, -1.39005529392e-02, 1.05176394873e-02, 5.51616357331e-03, -6.52085237587e-03, -1.42826422321e-03, 3.39306677671e-03, -6.39790110601e-05, -1.45904174198e-03, 3.43139829690e-04, 4.99881617563e-04, -2.39658346940e-04, -1.24341161725e-04, 1.08958435041e-04, 1.50133572744e-05, -3.63125515786e-05, 4.03452023518e-06, 8.79530134269e-06, -3.03514236589e-06, -1.36906023094e-06, 9.81001542204e-07, 5.32725065697e-08, -1.97592512917e-07, 3.61682651733e-08, 2.32830971382e-08, -1.06152960215e-08, -6.47431168795e-10, 1.40856815102e-09, -2.52404395415e-10, -7.34893003248e-11, 3.69210880887e-11, -3.32700896712e-12, -1.32433491724e-12, 4.44546709629e-13, -5.55944205057e-14, 2.69938287976e-15);

- Daub64 :

{ 1.16146330213e-05, 2.46656690638e-04, 2.43126191957e-03, 1.46810463814e-02, 6.02574991203e-02, 1.75750783639e-01, 3.67509628597e-01, 5.34317919340e-01, 4.77809163733e-01, 1.20630538265e-01, -2.66698181476e-01, -2.77421581558e-01, 6.47133548055e-02, 2.48310642356e-01, 2.46624448396e-02, -1.92102344708e-01, -4.89951171846e-02, 1.45232079475e-01, 4.44049081999e-02, -1.09456113116e-01, -2.96278725084e-02, 8.08741406384e-02, 1.41061515161e-02, -5.69263140624e-02, -2.38026446493e-03, 3.70514579235e-02, -4.14590766082e-03, -2.16628228363e-02, 6.16752731068e-03, 1.10174007154e-02, -5.41156825727e-03, -4.64921675118e-03, 3.62722464068e-03, 1.46895510046e-03, -1.96474055582e-03, -2.21167872957e-04, 8.67305851845e-04, -1.02453731060e-04, -3.05965442382e-04, 1.05391546173e-04, 8.10367832913e-05, -5.25980928268e-05, -1.29404577940e-05, 1.82426840198e-05, -6.36178153226e-07, -4.55830957626e-06, 1.20288903632e-06, 7.56004762559e-07, -4.28597069315e-07, -5.00336186874e-08, 8.96596631195e-08, -1.21992435948e-08, -1.10438302172e-08, 4.25042231198e-09, 4.38438779994e-

10, -5.88109146263e-10, 8.90472379622e-11, 3.26327074133e-11, -1.43091876516e-11, 1.07561065350e-12, 5.36148222961e-13, -1.66380048943e-13, 2.00071530381e-14, -9.42101913953e-16};

- Daub66 :

{ 8.18635831417e-06, 1.79101615370e-04, 1.82270943516e-03, 1.13959433745e-02, 4.86146665317e-02, 1.48186313180e-01, 3.26718130117e-01, 5.09376172514e-01, 5.11254770583e-01, 2.09582350713e-01, -2.04202622398e-01, -3.15997410766e-01, -1.92783394369e-02, 2.45420612119e-01, 9.98515586803e-02, -1.71428099051e-01, -1.10844133116e-01, 1.21967856403e-01, 9.47880880506e-02, -9.11469683513e-02, -7.03024850540e-02, 7.01911439409e-02, 4.57345618938e-02, -5.34712513358e-02, -2.52485829774e-02, 3.86870607602e-02, 1.07032658200e-02, -2.57287617547e-02, -2.16775861735e-03, 1.53169541158e-02, -1.59428878241e-03, -7.95354038705e-03, 2.38906240816e-03, 3.48080095340e-03, -1.86071821445e-03, -1.20430925760e-03, 1.07438069635e-03, 2.72730584733e-04, -4.90832900759e-04, 4.39316625176e-06, 1.78043189825e-04, -4.16043851627e-05, -4.92956442341e-05, 2.42333539881e-05, 9.07080575782e-06, -8.86612136675e-06, -3.60751610287e-07, 2.28837127614e-06, -4.42692340795e-07, -3.98579129198e-07, 1.82244333257e-07, 3.37797270373e-08, -3.98783819851e-08, 3.67286357683e-09, 5.11121185734e-09, -1.67139267725e-09, -2.49640210524e-10, 2.42683310230e-10, -3.04957445394e-11, -1.42023685988e-11, 5.50941472076e-12, -3.34348121895e-13, -2.15248838683e-13, 6.21474024717e-14, -7.19651054536e-15, 3.28937367841e-16};

- Daub68 :

{ 5.77051063273e-06, 1.29947620067e-04, 1.36406139005e-03, 8.81988940388e-03, 3.90488413517e-02, 1.24152482111e-01, 2.87765059233e-01, 4.78478746279e-01, 5.30555099656e-01, 2.90366329507e-01, -1.28246842174e-01, -3.31525301508e-01, -1.03891915515e-01, 2.16907220187e-01, 1.66601750412e-01, -1.27337358223e-01, -1.60924927177e-01, 7.79918469379e-02, 1.34125960271e-01, -5.44829680641e-02, -1.02947596992e-01, 4.35760946496e-02, 7.31852354367e-02, -3.70128384178e-02, -4.74385596452e-02, 3.07397465739e-02, 2.72283507563e-02, -2.36717379228e-02, -1.31439800166e-02, 1.64093741998e-02, 4.71364926099e-03, -1.00455067083e-02, -6.19474884515e-04, 5.33495076875e-03, -7.69212797506e-04, -2.39945394353e-03, 8.58995987436e-04, 8.75199906407e-04, -5.52735576214e-04, -2.32673214023e-04, 2.65077239755e-04, 2.66005001845e-05, -9.91469777078e-05, 1.35311722724e-05, 2.84495141969e-05, -1.05765749425e-05, -5.71082651099e-06, 4.16987175854e-06, 4.97971810142e-07, -1.11630653481e-06, 1.44819570833e-07, 2.02599066666e-07, -7.52670174041e-08, -1.99034650153e-08, 1.74042333293e-08, -8.66574426136e-10, -2.31650194699e-09, 6.44637821032e-10, 1.30041031860e-10, -9.90477453763e-11, 1.00420873546e-11, 6.08012535400e-12, -2.10787910891e-12, 9.79945115821e-14, 8.57919405179e-14, -2.31708370390e-14, 2.58733838193e-15, -1.14894475448e-16};

- Daub70 :

{ 4.06793406114e-06, 9.42146947557e-05, 1.01912268037e-03, 6.80729288431e-03, 3.12362885114e-

02, 1.03404455861e-01, 2.51307378994e-01, 4.43592739224e-01, 5.37008427509e-01, 3.60345640518e-01, -4.38838818739e-02, -3.23822864912e-01, -1.81786976766e-01, 1.66041357490e-01, 2.17299289321e-01, -6.52628713106e-02, -1.91919589298e-01, 1.93095446660e-02, 1.55292480396e-01, -4.75268083411e-03, -1.20585522643e-01, 4.73422917264e-03, 8.99135475707e-02, -9.31855894990e-03, -6.33560374404e-02, 1.32285495850e-02, 4.12546930647e-02, -1.43668397842e-02, -2.41694978016e-02, 1.27664567156e-02, 1.22894360081e-02, -9.57779789923e-03, -5.08599164923e-03, 6.13775458674e-03, 1.42808879407e-03, -3.35764438092e-03, 7.61596943517e-06, 1.54963746970e-03, -3.34669216425e-04, -5.86481031899e-04, 2.64832881996e-04, 1.70001228366e-04, -1.36588307226e-04, -2.97699596284e-05, 5.30414312291e-05, -2.43700152682e-06, -1.57244207727e-05, 4.30804786171e-06, 3.35334586287e-06, -1.89592961769e-06, -3.90393173328e-07, 5.30236861690e-07, -3.70030837820e-08, -9.99039694453e-08, 3.00818865071e-08, 1.08490273378e-08, -7.45811655289e-09, 5.89795131038e-11, 1.03082334548e-09, -2.43354557375e-10, -6.40793825650e-11, 4.00053662725e-11, -3.12563935710e-12, -2.56706547615e-12, 8.01508853368e-13, -2.59795432889e-14, -3.39772085679e-14, 8.62403743472e-15, -9.29801252932e-16, 4.01462871233e-17};

- Daub72 :

{ 2.86792518275e-06, 6.82602867854e-05, 7.60215109966e-04, 5.24029737740e-03, 2.48905656448e-02, 8.56520925952e-02, 2.17756953097e-01, 4.06433697708e-01, 5.32266895260e-01, 4.17875335600e-01, 4.39751975293e-02, -2.94421039589e-01, -2.46807036978e-01, 9.81142041631e-02, 2.46537277608e-01, 7.27851509579e-03, -1.99337205608e-01, -4.58614007463e-02, 1.54106236627e-01, 5.02761800735e-02, -1.18803754310e-01, -3.98808535755e-02, 9.11567822580e-02, 2.50387214495e-02, -6.82090166368e-02, -1.13191003168e-02, 4.85130835478e-02, 1.42497266176e-03, -3.19807206776e-02, 3.98404019871e-03, 1.90635947806e-02, -5.65781324505e-03, -9.99026347328e-03, 5.02298910666e-03, 4.41348483535e-03, -3.48454144540e-03, -1.50307406629e-03, 1.99079377185e-03, 2.77681279571e-04, -9.46340382326e-04, 8.61456575899e-05, 3.69350728496e-04, -1.15511889584e-04, -1.13189946808e-04, 6.69474119693e-05, 2.37510668366e-05, -2.73139082465e-05, -1.18347105998e-06, 8.37221819816e-06, -1.58614578243e-06, -1.87081160285e-06, 8.31142127970e-07, 2.54842352255e-07, -2.45537765843e-07, 2.75324907333e-09, 4.79904346545e-08, -1.15609368881e-08, -5.61278434332e-09, 3.13884169578e-09, 1.09081555371e-10, -4.51254577856e-10, 8.96241820385e-11, 3.03742909811e-11, -1.59971668926e-11, 8.87684628721e-13, 1.07096935711e-12, -3.02928502697e-13, 5.54226318263e-15, 1.33807138629e-14, -3.20462854340e-15, 3.33997198481e-16, -1.40327417537e-17};

- Daub74 :

{ 2.02206086249e-06, 4.94234375062e-05, 5.66241837706e-04, 4.02414036825e-03, 1.97622861538e-02, 7.05848259771e-02, 1.87326331862e-01, 3.68440972400e-01, 5.18167040855e-01, 4.62207553661e-01, 1.30878963233e-01, -2.46180429761e-01, -2.94375915262e-01, 1.96715004523e-02, 2.51523254360e-

01, 8.18060283872e-02, -1.81962291778e-01, -1.08451713823e-01, 1.29929646959e-01, 1.01780296838e-01, -9.66075406166e-02, -8.23302119065e-02, 7.50476199483e-02, 5.95674108715e-02, -5.92568156326e-02, -3.82538294793e-02, 4.58079441512e-02, 2.09728005925e-02, -3.35235840641e-02, -8.83349389041e-03, 2.26186515445e-02, 1.69047238348e-03, -1.37639819628e-02, 1.51930577883e-03, 7.38775745285e-03, -2.24805318700e-03, -3.39452327640e-03, 1.81687134380e-03, 1.26393425811e-03, -1.11148486531e-03, -3.28078847088e-04, 5.49053277337e-04, 1.53443902319e-05, -2.20894403245e-04, 4.33672612594e-05, 7.05513878206e-05, -3.09866292761e-05, -1.63916249616e-05, 1.35432771841e-05, 1.84994500311e-06, -4.30994155659e-06, 4.85473139699e-07, 1.00212139929e-06, -3.49494860344e-07, -1.50988538867e-07, 1.10903123221e-07, 5.35065751546e-09, -2.25219383672e-08, 4.22448570636e-09, 2.79397446595e-09, -1.29720500146e-09, -1.03141112909e-10, 1.94616489408e-10, -3.20339824412e-11, -1.39841571553e-11, 6.33495544097e-12, -2.09636319423e-13, -4.42161240987e-13, 1.13805283092e-13, -4.51888960746e-16, -5.24302569188e-15, 1.18901238750e-15, -1.19928033585e-16, 4.90661506493e-18};

- Daub76 :

{ 1.42577664167e-06, 3.57625199426e-05, 4.21170266472e-04, 3.08308811925e-03, 1.56372493475e-02, 5.78899436128e-02, 1.60071993564e-01, 3.30775781411e-01, 4.96591175311e-01, 4.93356078517e-01, 2.13050571355e-01, -1.82867667708e-01, -3.21675637808e-01, -6.22665060478e-02, 2.32125963835e-01, 1.49985119618e-01, -1.41795685973e-01, -1.59912565158e-01, 8.56381215561e-02, 1.41414734073e-01, -5.65864586307e-02, -1.14731170710e-01, 4.30958954330e-02, 8.72043982620e-02, -3.66051034028e-02, -6.17662087084e-02, 3.19898775315e-02, 4.00549811051e-02, -2.68914938808e-02, -2.31141340205e-02, 2.09046452556e-02, 1.12904972786e-02, -1.47018820653e-02, -4.13130665603e-03, 9.21478503219e-03, 5.62571574840e-04, -5.07131450921e-03, 7.16982182106e-04, 2.40069778189e-03, -8.44862666553e-04, -9.42461407722e-04, 5.81075975053e-04, 2.81763925038e-04, -3.03102046072e-04, -4.55568269666e-05, 1.26204335016e-04, -1.15540910383e-05, -4.17514164854e-05, 1.33417614992e-05, 1.03735918404e-05, -6.45673042846e-06, -1.55084435011e-06, 2.14996026993e-06, -8.48708758607e-08, -5.18773373887e-07, 1.39637754550e-07, 8.40035104689e-08, -4.88475793745e-08, -5.42427480028e-09, 1.03470453927e-08, -1.43632948779e-09, -1.34919775398e-09, 5.26113255735e-10, 6.73233649018e-11, -8.27825652253e-11, 1.10169293459e-11, 6.29153731703e-12, -2.48478923756e-12, 2.62649650406e-14, 1.80866123627e-13, -4.24981781957e-14, -4.56339716212e-16, 2.04509967678e-15, -4.40530704248e-16, 4.30459683955e-17, -1.71615245108e-18};

- sym8:

{0.032223100604, -0.012603967262, -0.099219543577, 0.297857795606, 0.803738751807, 0.497618667633, -0.029635527646, -0.075765714789};

- sym16:

{0.001889950333, -0.000302920515, -0.014952258337, 0.003808752014, 0.049137179674, -0.027219029917,

-0.051945838108, 0.364441894835, 0.777185751701, 0.481359651258, -0.061273359068, -0.143294238351, 0.007607487325, 0.031695087811, -0.000542132332, -0.003382415951};

- coif6:

{-0.072732619513, 0.337897662458, 0.852572020212, 0.384864846864, -0.072732619513, -0.015655728135};

- coif12:

{0.016387336464, -0.041464936782, -0.067372554722, 0.386110066823, 0.812723635450, 0.417005184424, -0.076488599079, -0.059434418647, 0.023680171946, 0.005611434819, -0.001823208871, -0.000720549445};

- coif18:

{-0.003793512864, 0.007782596427, 0.023452696142, -0.065771911282, -0.061123390003, 0.405176902410, 0.793777222626, 0.428483476378, -0.071799821619, -0.082301927107, 0.034555027573, 0.015880544864, -0.009007976137, -0.002574517689, 0.001117518771, 0.000466216960, -0.000070983303, -0.000034599773};

- coif24:

{0.000892313669, -0.001629492013, -0.007346166328, 0.016068943965, 0.026682300156, -0.081266699681, -0.056077313317, 0.415308407030, 0.782238930921, 0.434386056491, -0.066627474263, -0.096220442034, 0.039334427123, 0.025082261845, -0.015211731528, -0.005658286687, 0.003751436157, 0.001266561929, -0.000589020756, -0.000259974552, 0.000062339034, 0.000031229876, -0.000003259680, -0.000001784985};

- coif30:

{-0.000212080840, 0.000358589688, 0.002178236358, -0.004159358782, -0.010131117521, 0.023408156788, 0.028168028974, -0.091920010569, -0.052043163181, 0.421566206733, 0.774289603730, 0.437991626216, -0.062035963969, -0.105574208714, 0.041289208754, 0.032683574270, -0.019761778945, -0.009164231163, 0.006764185449, 0.002433373213, -0.001662863702, -0.000638131343, 0.000302259582, 0.000140541150, -0.000041340432, -0.000021315027, 0.000003734655, 0.000002063762, -0.000000167443, -0.000000095177};

- Beylkin18 :

{0.099305765374353, 0.424215360812961, 0.699825214056600, 0.449718251149468, -0.110927598348234, -0.264497231446384, 0.026900308803690, 0.155538731877093, -0.017520746266529, -0.088543630622924, 0.019679866044322, 0.042916387274192, -0.017460408696028, -0.014365807968852, 0.010040411844631, 0.0014842347824723, -0.002736031626258, 0.0006404853285212};

- Vaidyanathan24:

{-0.000062906118, 0.000343631905, -0.000453956620, -0.000944897136, 0.002843834547, 0.000708137504, -0.008839103409, 0.003153847056, 0.019687215010, -0.014853448005, -0.035470398607, 0.038742619293, 0.055892523691, -0.077709750902, -0.083928884366, 0.131971661417, 0.135084227129, -0.194450471766, -0.263494802488, 0.201612161775, 0.635601059872, 0.572797793211, 0.250184129505, 0.045799334111}.

Apêndice II - Código fonte do sistema proposto.

Apêndice III - Publicações durante o mestrado.

Durante o curso de mestrado do autor, os seguintes artigos foram publicados:

- SOUZA, L.M.; VIEIRA, L.S.; GUIDO, R.C. Predição linear e transformada wavelet discreta para identificação de patologias na laringe. *Revista de Informática, Inovação e Pesquisa da Faculdade de Tecnologia de São José do Rio Preto*; 2010. - ISSN 19846563, no prelo.
- FANTINATO, P. C. ; GUIDO, R. C. ; CHEN, S. H. ; SANTOS, B. L. S. ; VIEIRA, L. S. ; BARBON JR, S. ; RODRIGUES, L. C. ; SANCHEZ, F. L. ; ESCOLA, J. P. L. ; SOUZA, L. M. ; MACIEL, C. D. ; SCALASSARA, P. R. ; PEREIRA, J. C. . A Fractal-based Approach for Speech Segmentation. In: IEEE INTERNATIONAL SYMPOSIUM ON MULTIMEDIA (ISM2008), 10th; 2008, Berkeley, CA, USA. *Proceedings...*, Berkeley: IEEE, 2008 v.1, p. 551-555.
- GUIDO, R. C., BARBON JÚNIOR, S., VIEIRA, L. S., RODRIGUES, L. C., FANTINATO, P. C., ESCOLA, J. P. L., SOUZA, L. M., ALMEIDA, L. O. B., SLAETS, J. F. W. Autenticação biométrica por comando de voz com implementação em hardware dedicado. In: WORKSHOP NO AMBITO DO PROJECTO - AMBIENTES PARA CO-PROJETO DE HARDWARE/SOFTWARE EM PLATAFORMAS FPGAS COM APLICAÇÃO EM ROBÓTICA MÓVEL, 2; 2008, *Anais...*, São Carlos: 2008. CD-ROM
- GUIDO, R.C., BARBON JÚNIOR, S., VIEIRA, L.S., FANTINATO, P, SANCHEZ, F.L., ESCOLA, J.P.; BOLOGNINI, E., SOUZA, L., ZULATO, P., RIBEIRO, J., LACERDA, M. Speech compression based on the Discrete Shapelet Transform. In: CNMAC - CONGRESSO NACIONAL DE MATEMÁTICA APLICADA E COMPUTACIONAL. 31; 2008. *Anais...*, Recife-PE: CNMAC, 2008. CD ROM.
- VIEIRA, L.S., GUIDO, R.C., BARBON JÚNIOR, S., SANCHEZ, F.L., SERGIO, K.I.C., GUILHERME, M.B.A., SOUZA, L.M., FANTINATO, P.C., BASSI, R.D.S., AMADIO, B.C.G., FONSECA, E.S. On the use of wavelets for voice morphing. In: CNMAC - CONGRESSO NACIONAL DE MATEMÁTICA APLICADA E COMPUTACIONAL. 30; 2007. *Anais...*, Florianópolis-SC: 2007. CD ROM.