

Serviço de Pós-Graduação EESC/USP

EXEMPLAR REVISADO

Data de entrada no Serviço.....09/.....08/.....05.....

Ass.:.....*lesan*.....

Perfis Acústicos para Caracterização de Patologias da Laringe

Ricardo Naoki Mori

Tese apresentada à Escola de Engenharia de
São Carlos da Universidade de São Paulo ,
como parte dos requisitos para obtenção do
título de Doutor em Engenharia Elétrica

Orientador: Prof. Dr. José Carlos Pereira

São Carlos
2005



DEDALUS - Acervo - EESC

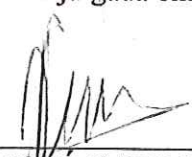


31100052248

FOLHA DE JULGAMENTO

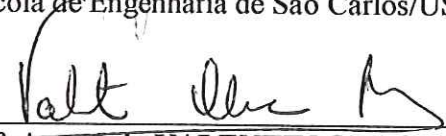
Candidato: Engenheiro **RICARDO NAOKI MORI**

Tese defendida e julgada em 16-06-2005 perante a Comissão Julgadora:



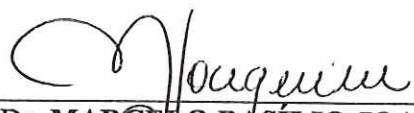
Prof. Titular **JOSE CARLOS PEREIRA (Orientador)**
(Escola de Engenharia de São Carlos/USP)

APROVADO



Prof. Associado **VALENTIN OBAC RODA**
(Escola de Engenharia de São Carlos/USP)

APROVADO




Prof. Dr. **MARCELO BASÍLIO JOAQUIM**
(Escola de Engenharia de São Carlos/USP)

aprovado




Dr. **ÁLVARO MACEDO DA SILVA**
(EMBRAPA)

aprovado




Prof. Dr. **ARTHUR OSCAR SCHELP**
(UNESP/Botucatu)

aprovado



Prof. Associado **HOMERO SCHIABEL**
Vice-Coordenador em Exercício do Programa de Pós-Graduação
em Engenharia Elétrica



Profa. Titular **MARIA DO CARMO CALJURI**
Presidente da Comissão de Pós-Graduação

Aos meus pais que sempre me apoiaram
em todos os momentos difíceis.

Agradecimentos

Ao Professor Dr. José Carlos Pereira pela orientação, conselhos, ensinamentos, presteza e pela grande paciência que foram de suma importância para o desenvolvimento do trabalho.

Ao Professor Dr. Marcos Grellet pelo auxílio no desenvolvimento da tese.

Ao amigo Marcelo Rosa pela auxílio inestimável dado ao trabalho.

Aos meus amigos: Parê, Granato, Espanhol, Eugênia, Farelo, Everthon, Elton, Karen, Thais e Yog pelo auxílio ao trabalho e amizade

Aos Professores do Departamento de Engenharia Elétrica pela formação e auxílio.

Aos Funcionários do Departamento de Engenharia Elétrica pela presteza.

A todos os meus colegas de pós-graduação pelo companheirismo e amizade.

À FAPESP pelo auxílio financeiro para o desenvolvimento da tese.

SUMÁRIO

LISTA DE FIGURAS.....	iii
LISTA DE TABELAS.....	iv
LISTA DE SIGLAS.....	v
RESUMO.....	vi
ABSTRACT.....	vii
1 INTRODUÇÃO	1
1.1 MEDIDAS ACÚSTICAS NO DOMÍNIO DO TEMPO	3
1.2 MEDIDAS ACÚSTICAS NO DOMÍNIO DA FREQUÊNCIA.....	4
1.3 OBJETIVO DO TRABALHO.....	5
1.4 ORGANIZAÇÃO DA TESE	6
2 FISIOLOGIA DO APARATO VOCAL E PATOLOGIAS ASSOCIADAS	7
2.1 A VOZ E O TRATO VOCAL.....	7
2.2 PATOLOGIAS E DISFONIAS DO APARATO VOCAL.....	9
2.2.1 Edema de Reinke	11
2.2.2 Diplofonia.....	11
2.2.3 Disfonia espasmódica.....	11
2.2.4 Disfonia ventricular.....	12
2.2.5 Espessamento das pregas vocais	12
2.2.6 Falsete	12
2.2.7 Laringite Traumática.....	13
2.2.8 Nódulos.....	13
2.2.9 Pólipos.....	13
2.2.10 Quebra de Altura.....	14
2.2.11 Quebra de vocalização	14
2.2.12 Úlceras de contato.....	15
2.2.13 Carcinoma	15
2.2.14 Cisto.....	15
2.2.15 Disartria	16
2.2.16 Glanuloma	16
2.2.17 Hiperkeratose	17
2.2.18 Laringite infecciosa	17
2.2.19 Leucoplasia.....	17
2.2.20 Membrana laringeana	18
2.2.21 Mudanças endócrinas.....	18
2.2.22 Paralisia das pregas vocais.....	18
3 MODELOS E PARÂMETROS ACÚSTICOS	20
3.1 MODELOS DIGITAIS PARA SINAIS DE VOZ.....	20
3.1.1 Modelo do Trato Vocal.....	20
3.1.2 Modelo de Radiação	21
3.1.3 Modelo da Excitação	21
3.1.4 Modelo da glote.....	21
3.2 PARÂMETROS ACÚSTICOS	22
3.2.1 Fator de Perturbação de Frequência (PFF - Period Frequency Factor)	23
3.2.2 Fator de Perturbação Direcional (DPF - Directional Perturbation Factor).....	23
3.2.3 Perturbação Média Relativa (RAP - Relative Average Perturbation)	24
3.2.4 Jitter Médio (JM).....	25
3.2.5 Razão de Jitter (JR)	25
3.2.6 Fator de Jitter (JF - Jitter Factor)	26
3.2.7 Índices de Variabilidade.....	26

3.2.8	<i>Perturbações de Ordem Superior</i>	27
3.2.9	<i>Quociente de Perturbação de Amplitude (APO)</i>	28
3.2.10	<i>Medidas Espectrais de Ruído</i>	29
3.2.11	<i>Relação de soproidade (Br)</i>	30
3.2.12	<i>Nível Espectral de Ruído(SNL)</i>	31
3.2.13	<i>Energia de Ruído Normalizada (NNE – Normalized Noise Energy)</i>	31
3.2.14	<i>Parâmetros de Cepstrum</i>	32
3.2.15	<i>Avaliação por Resíduos</i>	32
3.2.16	<i>Coefficiente de excesso</i>	32
3.2.17	<i>Nivelamento Espectral</i>	33
3.2.18	<i>Medidas Realizadas sobre o pulso glotal</i>	33
3.2.19	<i>H-Index</i>	35
3.2.20	<i>Relação Excitação Glotal/Ruído (GNE – Glottal to Noise Excitation)</i>	35
3.2.21	<i>Índice do Ruído Turbulento (TNI – Turbulent Noise Index)</i>	36
4	DETETORES E ESTIMADORES DE PITCH	39
4.1	ESTIMADOR.....	39
4.1.1	<i>Método AR</i>	40
4.1.2	<i>Método SIFT</i>	40
4.1.3	<i>Método AMDF (Average Magnitude Difference Function)</i>	40
4.1.4	<i>Método do Domínio Cepstral</i>	41
4.1.5	<i>Estimadores baseados na auto-correlação</i>	41
4.1.6	<i>Técnicas Perceptuais</i>	41
4.1.7	<i>Técnicas que utilizam o domínio Wavelet</i>	42
4.2	DETETORES DE PITCH	42
4.2.1	<i>Detetores baseados na análise de cruzamento de zero</i>	42
4.2.2	<i>Detetores baseados em mínimos ou máximos locais</i>	43
4.3	CORRETORES DE ERRO.....	43
5	MATERIAIS E MÉTODOS	44
5.1	DADOS UTILIZADOS.....	44
5.1.1	<i>Procedimento de análise</i>	45
6	RESULTADOS	48
6.1	DEFINIÇÃO DE UM ALGORITMO DE DETECÇÃO DE PERÍODO DE PITCH	48
6.2	VOZES DIPLOFÔNICAS	51
6.3	MEDIDAS TEMPORAIS	53
6.3.1	<i>Jitter (JR)</i>	53
6.3.2	<i>Shimmer</i>	53
6.4	MEDIDAS DE RUÍDO	53
6.5	PERFIS ACÚSTICOS DOS PARÂMETROS PARA AS DIVERSAS DOENÇAS	54
6.5.1	<i>Perfil acústico para fenda nas pregas vocais</i>	54
6.5.2	<i>Perfil acústico para edemas</i>	57
6.5.3	<i>Perfil Acústico para Nódulo</i>	59
6.5.4	<i>Perfil acústico para pólipos</i>	60
6.5.5	<i>Perfis acústicos para Paralisia Unilateral das pregas vocais</i>	62
6.5.6	<i>Patologias com número insuficiente de vozes para cálculo de estatísticas</i>	63
6.6	AVALIAÇÃO DOS RESULTADOS OBTIDOS	63
7	CONCLUSÕES	65
7.1	AVALIAÇÃO DOS RESULTADOS OBTIDOS	65
7.2	SUGESTÕES SOBRE TRABALHOS FUTUROS	66
8	REFERÊNCIAS BIBLIOGRÁFICAS	67
9	BIBLIOGRAFIA CONSULTADA	70

LISTA DE FIGURAS

Figura 1 – Trato vocal	9
Figura 2 – Edema de Reinke.....	11
Figura 3 - Nódulos.....	13
Figura 4 – Cisto	16
Figura 5 - glanuloma.....	17
Figura 6 - Modelo digital para sinais de voz	20
Figura 7 - Pulso glotal	21
Figura 8 - Correlação cruzada entre as janelas de 500 e 1kHz (GNE).....	36
Figura 9 - Modelo para um medidor do período de Pitch	39
Figura 10 - $A_c(p)$	50
Figura 11 - Sinal composto por 2 senóides separadas por $3/2(F_0)$	51
Figura 12 - sinal F_0 e o sinal composto.....	52

LISTA DE TABELAS

Tabela 1 - Patologias e suas Origens	10
Tabela 2 – parâmetros acústicos para fenda e vogal /a/	56
Tabela 3 - parâmetros acústicos para fenda e vogal /e/	56
Tabela 4 – parâmetros acústicos para Edema utilizado a vogal /a/	58
Tabela 5 – Parâmetros acústicos para Edema utilizando a vogal /e/	58
Tabela 6 – parâmetros acústicos para nódulo utilizando a vogal /a/	59
Tabela 7 – Parâmetros acústicos para nódulo utilizando a vogal /e/	60
Tabela 8 – parâmetros acústicos para pólipos e vogal /a/	61
Tabela 9 – parâmetros acústicos para pólipos e vogal /e/	61
Tabela 10 – Parâmetros acústicos para paralisia unilateral	63
Tabela 11 – Coeficiente de dispersão dos parâmetros para as diversas patologias	64

LISTA DE SIGLAS

AMDF	Average Magnitude Difference Function
APQ	Amplitude Perturbation Quotient
AR	Auto-Regressivo
AVI	Amplitude Variability Index
Br	Breathiness relation
CPP	Cepstral Peak Proeminence
DPF	Direction Perturbation Factor
FFT	Fast Fourier Transform
FPQ	Frequency Perturbation Quotient
GNE	Glottal to Noise Ratio
JF	Jitter Factor
JR	Jitter Ratio
JM	Jitter Mean
HNR	Harmonic to Noise Ratio
HFPR	High Frequency Power Ratio
Hr	Harmonics relative
NNE	Normalized Noise Energy
N/S	Noise/Signal
PFF	Perturbation Frequency Factor
PPQ	Pitch Perturbation Quotient
PVI	Period Variability Index
Ra	Relative Harmonic
RAP	Relative Average Perturbation
RPK	Pearson r at autocorrelation Peak
SIFT	Simple Inverse Filter Tracking
SFF	Spectral Flatness of Filter
SFR	Spectral Flatness of Residue
SNL	Spectral Noise Level
SI	Speed Index
SQ	Speed Quotient
To	Time open
Tc	Time closed

Resumo

A caracterização das patologias da laringe até o presente momento é predominantemente realizada de forma qualitativa. Alguns trabalhos recentes têm quantificados parâmetros acústicos no intuito de discriminar as diversas patologias. Este trabalho verifica os valores médios das diversas medidas acústicas da voz e as relaciona com as diferentes patologias. Desta forma, futuramente será possível utilizar estes valores para quantificar e acompanhar diferentes estágios do tratamento da patologia. O conjunto de valores médios compõem o que chamamos de perfil acústico.

Palavras chaves: processamento digital de sinais; análise acústica da voz; doenças da laringe

Abstract

Until now the characterization of larynx pathologies has been done in a subjective manner. Recently some researches have been conducted in order to quantify acoustic parameters mainly to discriminate the larynx pathologies. In this work several acoustic parameters of voice was evaluated and their correlation with the pathologies were established. In the future, would be possible to use these parameters to quantify and to follow the therapy procedures. These values constitute the acoustic profile of the pathologies.

Keywords: digital signal processing; acoustical analysis of voice; larynx diseases

Capítulo 1

Introdução

Alteração da voz (disfonia) é um evento constante na vida das pessoas. Infecções respiratórias ou uso excessivo da voz em ambientes adversos, geralmente produzem rouquidão na voz. Geralmente os sintomas desaparecem com o alívio da infecção ou após um descanso da voz. No entanto, há situações onde a voz é alterada permanentemente mesmo após a doença ter sido curada.

A voz é resultado de fenômenos fisiológicos que determinam sua emissão acústica. Os sons são classificados basicamente em três tipos:

- Sons vocálicos, que são sons produzidos por pulsos de ar quase periódicos gerados na laringe através da oscilação das pregas vocais e equalizados pelo trato vocal.
- Sons fricativos, produzidos por passagem turbulenta de ar através de alguma constrição formada pelo trato vocal.
- Sons plosivos, que são sons produzidos pelo fechamento completo do trato vocal com conseqüente aumento da pressão anterior à obstrução e subseqüente liberação abrupta.

Dos três tipos de sons citados, os sons vocálicos são o que apresentam vibração das pregas vocais, sendo, portanto, os sons utilizados na análise da laringe (pregas vocais) com o objetivo de detecção de patologias.

A produção da voz humana requer uma unidade funcional obtida por meio da ação integrada e harmoniosa de uma série de estruturas que não participam de uma mesma unidade orgânica, por não existir um único órgão especialmente desenvolvido para a produção da fala.

Na maioria das vezes a ocorrência de disfonias (alteração no comportamento vocal) provenientes de alteração orgânica ocorre na laringe, especificamente nas pregas vocais. Entre as diversas patologias inerentes as pregas vocais encontramos: edemas, úlceras, pólipos, granuloma e leucoplasias. Não raramente, a intervenção cirúrgica torna-se necessária para a melhoria da qualidade vocal ou mesmo como medida profilática ao agravamento da patologia.

O mecanismo de funcionamento da laringe é bastante complexo. Diversas teorias foram propostas para explicar seu funcionamento tais como: Teoria mioelástica – aerodinâmica, das cordas vibrantes, do caos determinístico, entre outras. A Teoria mais aceita é a Teoria mioelástica – aerodinâmica. O correto movimento de vibrações da mucosa depende do perfeito equilíbrio de todas as estruturas que constitui a laringe. Este equilíbrio é crítico, pois qualquer alteração em uma das partes desse conjunto é suficiente para comprometer a vibração regular das pregas vocais.

Conforme já mencionado, os pulsos de ar oriundos da laringe, devidamente equalizado pelo trato vocal produzem a voz humana (sons vocálicos).

A amplitude, frequência e forma de onda são as três características que descrevem um sinal. Observando que estão contidos na forma, as frequências harmônicas e ruído inerentes ao sinal, uma patologia que cause alguma disфонia, inevitavelmente ocasionará alterações em uma ou mais características da voz. Uma laringe normal produz vibrações quase periódicas de amplitude constante para os sons vocálicos.

As técnicas acústicas para análise da voz baseiam-se na fonte de sinal e nos métodos de análise. Uma fonte é o sinal direto, a pressão de som irradiada pela boca ou sinal de contato com a garganta. Essa pressão é o sinal mais imediato para a análise acústica, porém sua utilidade para avaliar as condições da laringe é limitada. A produção de uma vogal sustentada é realizada pela fonte glotal, que será afetada pela patologia da laringe, e pela estrutura supra-glotal, cujas características ressonantes possivelmente não serão afetadas pela patologia. As

possíveis mudanças no pulso glotal devido a patologia alterariam o formato de onda da pressão irradiada, podendo atenuar ou reforçar componentes ou até mesmo eliminá-los.

Outra fonte é o sinal indireto, o sinal glotal ou residual são obtidos por técnicas de filtragem inversa. O som vocálico é modelado como uma fonte (glote) excitando um tubo acústico. A filtragem inversa retira os efeitos do tubo e o sinal resultante aproxima-se do sinal periódico da fonte. Caso a estrutura supra-glotal não é afetada pela patologia e a causa da mudança da voz esta na fonte, então a aproximação do sinal periódico da fonte contém informação suficiente para se analisar os efeitos acústicos da patologia.

O uso da instrumentação eletrônica na análise objetiva para quantificar e fornecer medidas ou parâmetros acurados das características do sinal de voz. Para cada uma dessas fontes, os sinais podem ser analisados nos domínios do tempo e da frequência, dependendo das vantagens e limitações que apresentarem.

1.1 Medidas acústicas no domínio do tempo

A vibração das pregas vocais produz um pulso de ar glotal quase periódico. A utilização de medidas de perturbação do período de *Pitch* do sinal acústico discrimina casos patológicos de normais. Baseado nesse fato, várias medidas na perturbação de “pitch” e amplitude são propostas na literatura. A primeira foi o fator de perturbação de frequência (PFF) determinado por Lierberman (1963), sendo definido como a frequência relativa das perturbações de período de “pitch” maiores que 0,5 mS ocorrendo em uma vogal sustentada. É definida, portanto, como a diferença de tempo entre as durações de períodos sucessivos no sinal de voz.

Uma segunda medida de perturbação é a perturbação média relativa (RAP), proposta por Koike, Takashi e Calterra (1977) e analogamente ao PFF mede a flutuação dos períodos de pitch (usando janela de três amostras).

Utilizando a frequência para substituir o período e variando a janela de dados significativos, pesquisadores propuseram diversas medidas de perturbação. Dessa forma, tem-se na literatura o quociente de perturbação de frequência (FPQ) proposto por Takahashi e Koike (1975), o quociente de perturbação de período de “pitch” (PPQ), razão de *Jitter* (JR) que fornece a relação entre a média das perturbações pelo período médio, proposto para destacar a redução do *Jitter* para frequências fundamentais relativamente altas. Além da razão de *Jitter*, existe também o fator de *Jitter*, análogo ao JR, porém definido no domínio da frequência. A nomenclatura atual engloba todos esses fatores de perturbação de período (frequência) numa única denominação: *Jitter*.

Por outro lado, quando a perturbação é calculada sobre a amplitude do sinal, a maioria das medidas descritas acima têm seu correspondente e a denominação é *Schimmer*. Uma outra característica do sinal a ser considerada além das perturbações de período e amplitude é a forma de onda do sinal. A relação sinal ruído fornece informações para a discriminação de vozes patológicas e normais.

1.2 Medidas acústicas no domínio da frequência

A análise no domínio da frequência possibilita um outro conjunto de medidas acústicas da voz. O espectrógrafo é um instrumento utilizado na análise de frequência da voz, o qual verifica a distribuição espectral de energia de um determinado trecho da voz, sustentado por um banco de filtros. Como resultado dessa análise tem-se um gráfico tempo x frequência, sendo a densidade de energia indicada pela intensidade. Além disso, a frequência fundamental e os formantes das vogais sustentadas são visualizados no espectrograma.

Diversos estudos comprovam que existem diferenças significativas entre espectros de vozes normais e patológicas. Isso pelo fato do fluxo de ar turbulento originário do fechamento

incompleto da glote ou por vibração “não balanceada” das pregas vocais, as componentes espectrais do ruído podem ser originadas.

Essas componentes são distribuídas ao longo do espectro de forma variada e a extensão da distribuição depende da gravidade da patologia. Além disso, a presença do ruído contribui para a rouquidão, que é o primeiro sintoma para inúmeras patologias.

Proposto por Shoji et al. (1992a) a relação de potências em alta frequência (HFPR) calcula a média do logaritmo da densidade espectral de potência em intervalos de frequência bem definidos. Algumas outras medidas definidas no domínio da frequência são: relação harmônicas de ruído (HNR); intensidade relativa de harmônicas (Hr) determinado somando-se todas as intensidades das harmônicas (da segunda harmônica em diante) normalizadas pela intensidade total do espectro de voz (fundamental, harmônicas e ruído), sendo que o limite estabelecido para pessoas normais é de 67,5%; harmônicas relativas (Ra) que é a relação entre as componentes harmônicas e não harmônicas obtidas em um segmento de três períodos; relação de soproidade (Br) que estabelece a razão entre a energia em altas frequências e a energia total do sinal; nível espectral de ruído (SNL) que analisa as intensidades espectrais em intervalos de frequência e por último a energia de ruído normalizada (NNE) que considera o sinal composto de componentes periódicas mais ruído aditivo.

Os primeiros estudos associam uma patologia com um único parâmetro acústico. Nota-se que as patologias geralmente não estão associadas a um único efeito acústico, como perturbação em frequência ou adição de ruído.

1.3 Objetivo do trabalho

Os estudos sobre parâmetros acústicos concentram-se principalmente na sua utilização como avalistas do diagnóstico. Desta forma é muito importante verificar quais são as medidas mais importantes para uma determinada patologia. Nesse sentido é proposto o conceito de

perfis acústicos, que procuram identificar estatisticamente quais são os parâmetros mais significativos para cada patologia.

1.4 Organização da tese

Este trabalho apresenta no capítulo 2 a fundamentação teórica que visa abordar aspectos anatômicos e funcionais do aparato vocal.

No capítulo 3 são abordadas a fisiologia do aparato vocal, as principais patologias e disfonias do trato vocal.

O capítulo 4 apresenta uma revisão sobre estimadores de pitch e a metodologia utilizada no trabalho.

No capítulo 5 são especificados os materiais utilizados e a forma de captura dos dados.

O capítulo 6 são apresentados os resultados obtidos e no capítulo 7 são realizadas as conclusões, discussões e propostas de trabalhos futuros.

Capítulo 2

2 Fisiologia do aparato vocal e patologias associadas

Neste capítulo serão apresentadas a fisiologia do aparato vocal e a apresentação das principais patologias e disfonias associadas ao trato vocal.

2.1 *A voz e o trato vocal (Guyton, 1985)*

A estrutura do aparelho fonador é dividida em três subsistemas básicos: os pulmões, a laringe e o trato vocal.

O pulmão é uma estrutura elástica que sofre colapso à semelhança de um balão e expulsa todo seu ar pela traquéia toda vez que não houver uma força para mantê-lo insuflado. Estes podem sofrer expansão e retração por duas maneiras, primeiramente pelo movimento do diafragma para baixo e para cima, a fim de aumentar ou diminuir a altura da cavidade torácica, e pela elevação e abaixamento das costelas para aumentar e diminuir o diâmetro antero-posterior da cavidade torácica. No momento da fala normal a vazão de ar ocorre de modo passivo.

A laringe está especialmente adaptada para atuar como um elemento vibrador. O elemento vibrátil é constituído pelas pregas vocais, que fazem protusão das paredes laterais da laringe para o centro da glote; são estriadas e posicionadas por vários músculos específicos da própria laringe.

Durante a respiração normal, as pregas se abrem para permitir a fácil passagem de ar. Durante a fonação, as pregas vocais se fecham de modo que a passagem de ar entre elas provoque vibração. A altura da vibração (frequência de vibração) é determinada principalmente pelo grau de estiramento das pregas, mas também pelo grau de aproximação das pregas entre si e pela massa de suas bordas.

Imediatamente no interior de cada prega vocal existe um forte ligamento elástico, denominado ligamento vocal. Esse ligamento fixa-se anteriormente à grande cartilagem tireóideia, que faz protusão na superfície anterior do pescoço, sendo conhecido como epiglote. Posteriormente, o ligamento vocal liga-se aos processos vocais das duas cartilagens aritenóides. Tanto a cartilagem tireóide quanto as cartilagens aritenóides articulam-se, por sua vez com a cartilagem cricóide.

As pregas vocais podem ser estiradas por rotação da cartilagem tireóide para frente ou por rotação posterior das cartilagens aritenóides, ativadas por músculos que estendem das cartilagens tireóide e aritenóide para a cartilagem cricóide. Os músculos localizados no interior das pregas vocais, lateralmente aos ligamentos vocais, os músculos tireoaritenóides, podem tracionar as cartilagens aritenóides em direção à cartilagem tireóide, afrouxando as cordas vocais. Além disso podem modificar a forma e a massa das bordas das pregas vocais. Esticando-as para emitir sons agudos e relaxando-as para sons mais baixos.

Por fim, vários outros grupos de músculos laríngeos muito pequenos que situam-se entre as cartilagens aritenóides e a cartilagem cricóide podem determinar a rotação dessas cartilagens para dentro ou para fora ou tracionar suas bases, estabelecendo as várias configurações possíveis das pregas vocais.

Os três órgãos principais da articulação são os lábios, a língua e o palato mole sendo responsáveis pela modulação do sinal glotal.

Os ressonadores incluem a boca, o nariz e os seios nasais associados, a faringe e até mesmo a própria cavidade torácica. Eles podem ser representados como um conjunto de tubos de ressonância que enfatizam ou de-enfatizam o sinal glotal.

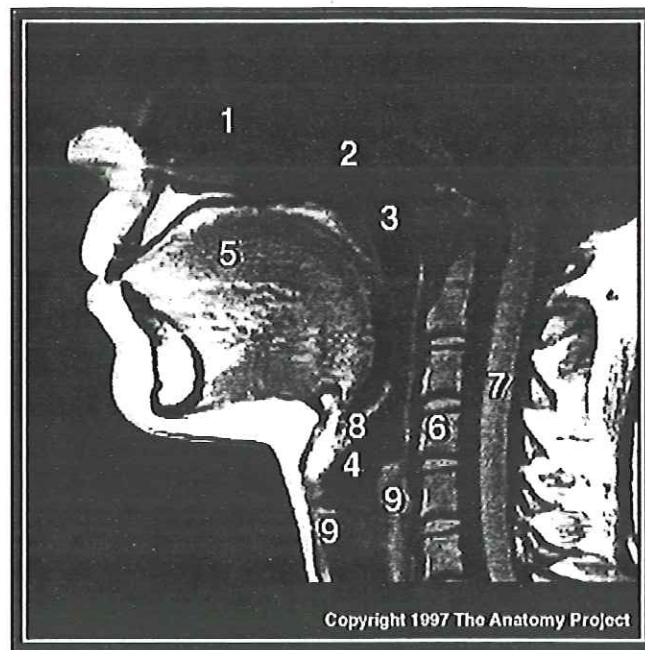


Figura 1 – Trato vocal

Legenda:

- | | |
|-------------------|------------------------|
| 1. Cavidade Nasal | 6. Coluna cervical |
| 2. Naso Faringe | 7. Medula espinhal |
| 3. Faringe | 8. Epiglote |
| 4. Laringe | 9. Cartilagem cricóide |
| 5. Língua | |

A figura 1 mostra uma ressonância magnética computadorizada que mostra os principais órgãos do aparelho fonador.

2.2 Patologias e Disfonias do aparato vocal

As patologias do aparato vocal podem ocorrer em qualquer dos três subsistemas do trato vocal, sendo convencionalmente classificados quanto à origem (Pinho, 1998) por:

- a) origem funcional, onde por um mal uso ou funcionamento de parte do trato vocal, origina-se uma fala anormal.
- b) origem orgânica, em que a patologia é proveniente de anomalias físicas no sistema de produção da fala.

A análise subjetiva utilizada por especialistas classifica as vozes patológicas em vozes ásperas (vozes que apresentam um grande estiramento das pregas vocais), vozes soprosas (vozes que não apresentam um fechamento total da glote) e vozes roucas (vozes que apresentam uma modificação da frequência fundamental de oscilação das pregas vocais e uma oscilação errática em alguns instantes, geralmente provocado pelo aumento de massa das pregas).

As patologias do sub-sistema glotal são apresentadas na Tabela 1 conforme a sua classificação etiológica.

Origem funcional	Origem orgânica
Edema de Reinke	Carcinoma
Diplofonia	Cistos
Disfonia Espasmódica	Disartria
Disfonia Ventricular	Granuloma
Espessamento	Hiperkeratose
Falsete	Laringite infecciosa
Laringite Traumática	Leucoplasia
Nódulos	Membrana Laringiana
Pólipos	Mutações endócrinas
Quebra de Altura	Paralisia de Pregas Vocais
Quebra de Vocalização	
Úlceras de contato	

Tabela 1 - Patologias e suas Origens

Um breve relato das patologias e disfunções do sub-sistema glotal é realizado a seguir buscando atrelar suas origens aos correlatos acústicos decorrentes.

2.2.1 *Edema de Reinke*

Conhecido também como degeneração polipóide, é resultado do aumento da massa bilateral das pregas vocais causado por um edema no corpo gelatinoso de Reinke, logo abaixo do epitélio da prega vocal, na camada superficial da lâmina própria. A Figura 2 mostra o caso de edema de Reinke.

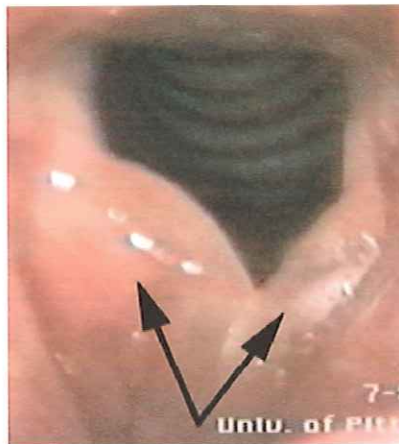


Figura 2 – Edema de Reinke

2.2.2 *Diplofonia*

A diplofonia é caracterizada pelo aparecimento de um sinal com duas frequências distintas. Esta ocorre devido ao aparecimento de uma falsa constrição modulando o sinal do pulso glotal original.

2.2.3 *Disfonia espasmódica*

Esta patologia se caracteriza pela presença da voz normal com súbitos momentos de afonia.

2.2.4 *Disfonia ventricular*

Definida como uma adução das pregas ventriculares (falsas pregas) que vibram com a passagem do ar oriundo do pulmão. As pregas verdadeiras estão ligeiramente mais abertas durante a fase de vocalização.

A voz distingue-se por ter pouca variação de altura e ser de baixa frequência, motivada pela maior massa destas pregas em relação às pregas vocais verdadeiras. Algumas vezes pode-se perceber um padrão diplofônico, uma vez que ambas pregas vibram.

2.2.5 *Espessamento das pregas vocais*

Caracteriza-se pelo aumento da massa ao longo das pregas vocais, usualmente resultante do abuso contínuo da voz. Podendo propiciar o aparecimento de nódulos. Perceptualmente, a voz apresenta-se como rouca, devido a modificação da massa das pregas vocais.

2.2.6 *Falsete*

Caracteriza-se perceptualmente por uma frequência fundamental acima da gama de alturas atingidas por pessoas normais e geralmente produzida por mal uso.

Em termos acústicos a voz de falsete apresenta elevada altura, associada à escapes de ar, desta forma caracterizando soprosidade. Ocorrem também quebras de frequência para baixo.

2.2.7 *Laringite Traumática*

Nesta doença as pregas vocais tem suas bordas inchadas e espessas, acompanhado por maior concentração sangüínea. Com o aumento da massa e do volume das pregas vocais, sua vibração torna-se inadequada, provocando assim uma voz rouca, de baixa intensidade.

2.2.8 *Nódulos*

Produzidos por uso contínuo e abusivo das pregas vocais, são caracterizados por uma protuberância esbranquiçada na borda glótica das pregas vocais, geralmente bilateral. Localizados na porção média da membrana vocal. Por estarem posicionados em uma camada muito superficial das pregas vocais, interferem com a vibração da mucosa, produzindo aperiodicidade no sinal glotal. A Figura 3 mostra um caso com um nódulo bilateral.

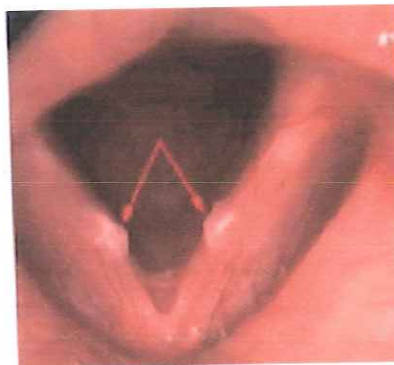


Figura 3 - Nódulos

2.2.9 *Pólipos*

Relacionados com hiperfuncionamento da vocalização resultam de um único evento vocal (não contínuo), podendo ser identificado como uma haste peduncular ou séssil (base

ampla) na borda interna de uma das pregas vocais, pouco irritando a mucosa devido a sua maciez.

A voz de um paciente com esta patologia é soprosa (devido a adução não ser total) e, dependendo do tamanho e localização, pode abafar a vibração da mucosa, produzindo assim rouquidão.

2.2.10 Quebra de Altura

Duas são as origens deste distúrbio: a primeira é observada na puberdade quando o indivíduo sofre modificações hormonais e as estruturas laríngeas sofrem marcante desenvolvimento. A segunda causa advém do uso intensivo da voz por um período prolongado, geralmente quando a mucosa vibra em uma frequência inapropriada, à qual está acostumada.

Analisando o sinal de voz, verifica-se que esta apresenta uma mudança repentina na frequência fundamental em aproximadamente uma oitava, elevando-a ou reduzindo-a de acordo com o nível apropriado, de vibração glotal.

2.2.11 Quebra de vocalização

Perda da voz por um curto período de tempo. O indivíduo fala normalmente, e repentinamente tem uma cessação completa da voz. Geralmente caracterizado por uma hiperfuncionalidade das pregas vocais, ou por perda de pressão subglótica e até mesmo pela presença excessiva de muco nas paredes das pregas vocais.

Em termos sonoros, caracteriza-se pelo aparecimento de um trecho não vocálico, como uma rouquidão intensa, durante a emissão de um sinal vocálico.

2.2.12 Úlceras de contato

Consistem em crateras com uma estrutura escamosa muito espessa sobre o tecido conectivo, aliada a alguma inflamação. Utilizando-se a laringoscopia podem-se notar ulcerações em ambas pregas vocais com grande acúmulo de tecido granuloso ao longo das bordas de aproximação na porção posterior das pregas.

A origem deste distúrbio é relacionada primeiramente com o uso freqüente de limpeza de garganta e tosse. Associada à utilização de freqüência fundamental baixa (som grave) e com ataque vocal excessivamente brusco. A segunda origem é o refluxo gastroesofágico, que irrita as paredes das estruturas laríngeas, e a terceira etiologia é a entubação do paciente em cirurgias.

2.2.13 Carcinoma

Também denominado de câncer, pode afetar qualquer porção do trato vocal, produzindo diversos efeitos sobre a voz do paciente. O carcinoma pode ser classificado com relação ao tipo: T1 e T2 (tumores precoces), T3 e T4 (tumores avançados), com relação à localização e à extensão. Podem causar como sintomas desde uma sopro leve até o aparecimento de uma severa rouquidão.

2.2.14 Cisto

Consiste em uma lesão benigna resultante do acúmulo de fluidos em uma estrutura tipo saco. Como as pregas vocais estão alinhadas com pequenas glândulas de secreção de muco (que auxilia a vibração das pregas vocais), existe a possibilidade de supressão de secreção com eventual acúmulo de líquidos, produzindo uma protuberância na prega vocal. O

cisto, diferentemente do nódulo, ocorre na maioria das vezes somente em uma das pregas vocais, e em qualquer posição.

Devido a sua própria etiologia, impede a total adução, permitindo um escape de ar turbulento, dando um aspecto soproso à voz do paciente. A Figura 4 mostra um caso de Cisto.

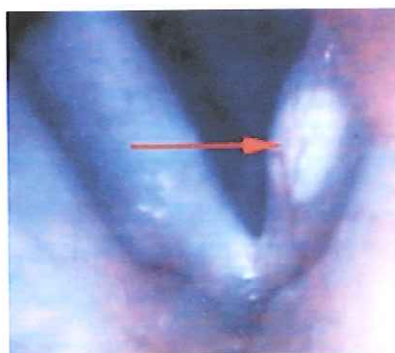


Figura 4 – Cisto

2.2.15 Disartria

Problema motor da fala decorrente de dano em alguma parte do sistema nervoso central ou periférico, podendo resultar em paralisia unilateral ou bilateral das pregas vocais, ou musculatura anexa, que são acionados pelo nervo recorrente.

O descontrole nervoso do trato vocal pode caracterizar problemas de prosódia, coarticulação e alteração na intensidade da voz assim como altura e qualidade de vocalização.

2.2.16 Glanuloma

Caracterizado por altos níveis de rouquidão e soproside, são o resultado de traumatismos nas estruturas da laringe. A mucosa apresenta hiperplasia e granulação de aspecto rígido. A Figura 5 mostra um caso de Glanuloma.



Figura 5 - glanuloma

2.2.17 Hiperkeratose

Resultante da irritação contínua do tecido laríngeo, esta doença caracteriza-se por placas rosadas, que surgem sobre as pregas vocais, instalando-se geralmente sobre a comissura anterior (acidente anatômico).

Os efeitos sobre a voz podem ser de desprezíveis a severos, com o aparecimento de níveis de sopro e rouquidão.

2.2.18 Laringite infecciosa

Pacientes advindos de laringite funcional que apresentam abuso ou mau uso das estruturas de fonação acabam experimentando esta patologia tão logo tenham resfriado.

A voz é caracterizada por uma intensa sopro e rouquidão uma vez que as pregas vocais, apesar de inchadas, apresentam vibração de mucosa.

2.2.19 Leucoplasia

Lesões de acréscimo sobre a membrana superficial da mucosa, apresentando coloração esbranquiçada, sendo consideradas pré-cancerosas.

Com o acréscimo de massa, reduz-se a altura da voz provocando redução da frequência fundamental e de intensidade fraca. Em várias ocasiões as pregas vocais tornam-se assimétricas, gerando diplofonia.

2.2.20 Membrana laringeana

Crescimento de uma membrana na glote, nas duas pregas vocais, impedindo a vibração correta destas, podendo causar respiração curta, tendo origem congênita ou resultado de um trauma nas bordas internas das pregas vocais.

A voz é descrita perceptualmente como áspera, com alta frequência fundamental, uma vez que esta membrana tende a reduzir o espaço glotal, geralmente iniciando pela comissura inferior.

2.2.21 Mudanças endócrinas

Alterações hormonais tem relevância sobre as características vocais das pessoas, causando modificações sobre a frequência fundamental, que pode ser extremamente grave ou aguda.

Alterações na produção hormonal da tireóide podem produzir voz perceptualmente rouca, com altura extremamente baixa. Altos níveis de estrógeno e progesterona nas mulheres podem produzir leve espessamento das pregas, causando redução de altura e rouquidão. Hormônios androgênicos na menopausa tornam as paredes das pregas vocais mais espessas, podendo torná-las áspera.

2.2.22 Paralisia das pregas vocais

Qualquer dano sobre o nervo laríngeo superior ou no nervo laríngeo recorrente, que controlam as atividades musculares nas estruturas de vocalização, resultam em uma paralisia (unilateral ou bilateral) das pregas vocais, impedindo a produção de um sinal glotal adequado. No caso unilateral, uma das pregas vocais fica na posição abduzora, cabendo à outra a função de aproximar-se, compensando a inatividade da primeira.

A voz apresenta-se extremamente soproza, uma vez que as pregas vocais não conseguem aduzirem-se à mínima abertura para prover vibração da mucosa. No caso unilateral, havendo compensação da prega vocal normal, pode-se obter um padrão diplofônico, com percepção de soprozidade.

Capítulo 3

3 Modelos e Parâmetros acústicos

Neste capítulo serão apresentados o modelo digital de produção da voz e os principais parâmetros acústicos para diagnóstico de patologias da laringe.

3.1 Modelos Digitais para sinais de voz

A Figura 6 - Modelo digital para sinais de voz - 6 mostra um modelo para a produção de voz:

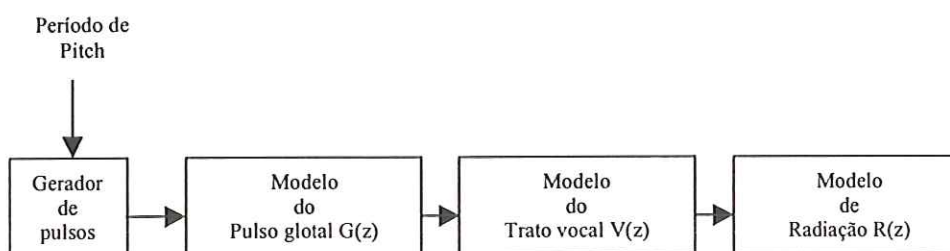


Figura 6 - Modelo digital para sinais de voz

3.1.1 Modelo do Trato Vocal

As ressonâncias (formantes) da fala corresponde aos pólos da função de transferência $V(z)$ (tubos acústicos). Um modelo tudo pólo é uma boa representação dos efeitos do trato vocal para a maioria dos sons da fala. Entretanto a teoria acústica mostra que em casos de sons anasalados e fricativos são necessárias tanto ressonâncias como anti-ressonâncias (pólos e zeros). As ressonâncias são atribuídas ao trato vocal e as anti-ressonâncias à cavidade nasal.

3.1.2 Modelo de Radiação

O modelo de radiação procura representar a pressão nos lábios. Assim procura simular o efeito da impedância de radiação. Esta causa um efeito semelhante a um filtro passa-altas. Uma aproximação razoável para os efeitos da radiação é dada por (Rabiner e Schaffer, 1978):

$$R(z) = R_0(1 - Z^{-1}) \quad (1)$$

onde:

R_0 : representa um ganho.

3.1.3 Modelo da Excitação

O modelo de Excitação corresponde à entrada do sistema do trato vocal. No caso em questão é utilizado um modelo gerador de impulsos unitários separados por um período de Pitch.

3.1.4 Modelo da glote

A saída do pulso glotal é ilustrado na Figura 7,

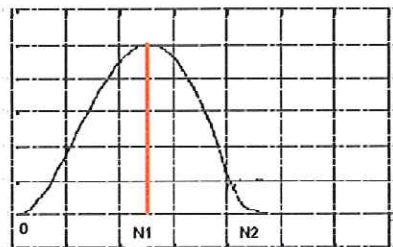


Figura 7 - Pulso glotal

e pode ser representado pela equação 11 (Rosenberg, 1971):

$$g(n) = \begin{cases} 1 - \cos(\pi n/N_1) & 0 \leq n \leq N_1 \\ \cos(\pi(n - N_1)/2N_2) & N_1 \leq n \leq N_1 + N_2 \\ 0 & \text{para outros casos} \end{cases} \quad (2)$$

onde 0 a N_1 : é a fase de abertura
 de N_1 a N_2 : é a fase de fechamento
 $> N_2$: é a fase fechada

Uma boa aproximação bastante utilizada para a equação 20 é um filtro composto somente por pólos dada pela equação 3.

$$H(z) = \frac{G}{1 - \sum_{k=0}^p a_k z^{-k}} \quad (3)$$

onde G : representa um ganho.

3.2 Parâmetros Acústicos

A popularização do computador tornou programas como o Dr. Speech (Tiger DRS Inc.) alvo de grande interesse de profissionais da área. Principalmente da área de fonoaudiologia. Estes programas são baseados principalmente no cálculo de parâmetros acústicos dos sinais vocálicos, que guardam estreita relação com os aspectos perceptivos-auditivos da voz sendo de grande valia na quantificação da análise, possibilitando o discernimento quanto à origem desses aspectos.

Muitas vezes esses parâmetros não são correlacionados diretamente com um diagnóstico (como nódulo ou pólipos), o que provoca uma dificuldade para o correto entendimento do real significado dos valores medidos. Desta forma apesar de alguns desses parâmetros possuírem quase 40 anos ainda existe interesse por informações sobre o seu

significado e suas aplicações como demonstrado pelo artigo que descreve os principais parâmetros publicados por Andrade et al. (2002)

Nesta seção, uma revisão dos parâmetros acústicos da voz é apresentada considerando parâmetros que nunca foram estudados com vozes brasileiras como o GNE e o TNI. Os parâmetros de perturbação quando aplicados ao período (ou frequência) são denominados Jitter e quando aplicados a amplitude recebem a denominação de Shimmer.

3.2.1 Fator de Perturbação de Frequência (PFF - Period Frequency Factor)

Este parâmetro representa a frequência relativa das perturbações do período de “Pitch” maiores que 0.5 ms. Outros nomes relacionados com este parâmetro são: Fator de Perturbação de “Pitch” e Fator de Perturbação de Período.

Lieberman (1963) determinou que vozes patológicas possuem um valor maior de PFF que vozes normais e que o PFF é sensível ao tamanho e localização da anomalia da laringe. A fórmula deste parâmetro é :

$$PFF = \frac{\sum_{i=2}^N |P_i - P_{i+1}| > 0.5mS}{\sum_{i=1}^N P_i} \quad (4)$$

Onde : P_i corresponde ao período de Pitch i

N é o número de períodos

3.2.2 Fator de Perturbação Direcional (DPF - Directional Perturbation Factor)

Este parâmetro sugerido por Hecker e Krueel (1970), considera a direção das diferenças dos períodos, isto é, a porcentagem de alteração do sinal algébrico(aumento ou

diminuição do período do pulso glotal). Este estudo utilizou segmentos de vozes provenientes de frases que foi demonstrado inadequado por Sorensen e Horii (1984), o qual sugerem a utilização de vogais sustentadas, verificando a mudança da direção do Jitter para 11 ciclos dos sinais.

A contagem da mudança dos ciclos é realizada seguindo-se o seguinte procedimento:

O primeiro período é considerado como período referência, se o segundo ciclo apresentar um período menor lhe é atribuído um sinal negativo caso contrário um sinal positivo. A partir deste ponto considera-se o segundo sinal como referência e verifica-se o comprimento do posterior. O processo é repetido até o último ciclo.

A partir desse conjunto de sinais é verificado a quantidade de mudanças de sinais e este é computado pela seguinte fórmula:

$$DPF = \frac{NMS}{n-1} \times 100\% \quad (5)$$

onde :

NMS : Número de mudanças algébricas do sinal
n : número de ciclos contínuos

3.2.3 *Perturbação Média Relativa (RAP - Relative Average Perturbation)*

Este parâmetro foi proposto por Koike, Takashi e Calterra (1977) e analogamente ao PFF mede a flutuação dos períodos de “pitch”, porém de uma forma mais sofisticada (janela de três períodos)

$$RAP = \frac{\frac{1}{N-2} \sum_{i=2}^{N-1} \left| \frac{P_{i-1} + P_i + P_{i+1}}{3} - P_i \right|}{\frac{1}{N} \sum_{i=1}^N P_i} \quad (6)$$

Onde : P_i corresponde ao período de Pitch i

N é o número de períodos

3.2.4 Jitter Médio (JM)

Este foi o parâmetro procura verificar o valor médio do jitter, sendo proposto por Horii (1979).

$$JM = \frac{\sum_{i=1}^{N-1} |P_i - P_{i+1}|}{N-1} \quad (7)$$

Onde : P_i corresponde ao período de Pitch i

N é o número de períodos

3.2.5 Razão de Jitter (JR - Jitter Ratio)

Este parâmetro fornece a relação entre a média das perturbações pelo período médio. A razão de *Jitter* ou *Jitter Ratio* (JR) foi proposto para realçar a redução do Jitter para frequências fundamentais relativamente altas. Este parâmetro foi proposto por Horii (1979). Algebricamente esta relação é dada por:

$$JR = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |P_i - P_{i+1}|}{\frac{1}{N} \sum P_i} \cdot 10^3 \quad (8)$$

Onde : P_i é o período de Pitch i

N é o número de períodos

3.2.6 Fator de Jitter (JF – Jitter Factor)

Este parâmetro proposto por Murry e Doherty (1980) é análogo ao JR, porém é definido no domínio da frequência:

$$JF = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |F_i - F_{i+1}|}{\frac{1}{N} \sum F_i} \cdot 10^2 \quad (9)$$

Onde : F_i corresponde frequência correspondente ao período de Pitch i

N é o número de períodos

3.2.7 Índices de Variabilidade

Basicamente estes, parâmetros derivam do coeficiente de variação aplicado a todas às amostras do sinal:

3.2.7.1 Índice de variabilidade de período (PVI - *Period Variability Index*)

Este parâmetro mede o coeficiente de variação para o período a sua equação é dada por:

$$PVI = \frac{\frac{1}{N} \sum_{i=1}^n (P_i - \bar{P})^2}{\bar{P}^2} \quad (10)$$

onde: A_i representa o período de *pitch* i

\bar{P} representa a média dos períodos

3.2.7.2 Índice de Variabilidade de Amplitude (AVI - *Amplitude Variability Index*)

Este parâmetro definido por Deal e Emanuel (1978) mede a variação do sinal com relação à média.

$$AVI = \frac{\frac{1}{N} \sum_{i=1}^n (A_i - \bar{A})^2}{\bar{A}^2} \quad (11)$$

onde: A_i representa a amplitude no período de *pitch* i

\bar{A} representa a média das amplitudes na janela de cálculo

3.2.8 *Perturbações de Ordem Superior*

Assimetria - este parâmetro é baseado no momento de terceira ordem e permite avaliar a tendência lateral da curva de distribuição das perturbações. A assimetria é descrita por:

Curtose – determina o grau de achatamento da distribuição das perturbações, sendo definida pela expressão :

$$\text{curtose} = \frac{\sum_{i=1}^N (p_i^k)^4}{N \sigma_{p^k}^3} - 3 \quad (13)$$

Esta medida é semelhante ao parâmetro EX (*Excess coeficient*) definido por Davis (1979), onde as disfonias são caracterizadas por curvas de distribuição platicurticas (curtose positiva), enquanto as vozes normais apresentam distribuições leptocurticas (curtose negativa).

3.2.9 Quociente de Perturbação de Amplitude (APQ - Amplitude Perturbation Quotient)

Este parâmetro foi proposto por Koike, Takahashi e Calterra (1977) e Davis (1979) que em seus estudos determinaram que a melhor medida seria realizada em cinco períodos subsequentes do sinal vocálico. A sua equação é dada por:

$$APQ = \frac{\frac{1}{N-4} \sum_{i=3}^{N-2} \left| \frac{A(i-2) + A(i-1) + A(i) + A(i+1) + A(i+2)}{5} - A(i) \right|}{\frac{1}{N} \sum_{i=1}^N A(i)} \quad (14)$$

3.2.10 Medidas Espectrais de Ruído

3.2.10.1 Espectro Médio de Tempo Longo (LTAS, *Long Time Average Spectral*)

Basicamente se constitui na densidade espectral de potência para um grande número de amostras.

3.2.10.2 Relação de Potência em Altas Frequências (HFPR - *High Frequency Power Ratio*)

Proposto por Shoji et al. (1992a), este parâmetro em última análise é semelhante ao LTAS porém calcula a média do logaritmo da densidade espectral de potência em intervalos de frequência bem definidos. Este parâmetro é calculado por:

$$HPFR_i = \sum_{w=f_i}^{f_h} \overline{PSD_w} \quad (15)$$

onde PSD é a densidade espectral de potência em um determinado intervalo de frequência.

3.2.10.3 Relação Harmônicas/Ruído (HNR – *Harmonic to Noise Ratio*)

Esta medida foi desenvolvida por Yumoto, Gould e Baer (1982) e é definida como:

$$HNR = \frac{N \sum_{t=1}^{T_a} S_a^2(t)}{\sum_{i=1}^N \sum_{t=1}^{T_i} [s_i(t) - s_a(t)]^2} \quad (16)$$

onde:

$s_a(t)$ é a estimativa livre de ruído

$s_i(t)$ é o i -ésimo trecho do sinal

T_a é o comprimento da estimativa livre de ruído

T_i é o comprimento do i -ésimo trecho do sinal

3.2.10.4 Intensidade Relativa de Harmônicas

Este parâmetro mede a energia contida na Segunda harmônica relativa à energia total do sinal. O limite estabelecido para pessoas normais por este parâmetro é de 67,5%.

3.2.10.5 Harmônicas Relativas (Ra - *Relative Harmonic*)

É a relação entre as componentes Harmônicas e as não harmônicas calculadas em um segmento de três períodos. Ela foi definida por Kojima *et al* (1980):

$$Ra = 10 \log \left[\frac{\sum_{i=1}^N S_i}{\sum_{i=1}^N N_i} \right] \quad (17)$$

Onde :

S_i são as componentes harmônicas e

N_i são as componentes não harmônicas

3.2.11 Relação de soprosidade (Br - *Breathiness relation*)

Fukuzawa, el-Assuooty e Honjo (1988) procura estabelecer a razão entre a energia em altas frequências e a energia total do sinal. Esta medida diferentemente da maior parte das

medidas de ruído não necessita da determinação correta do período de pitch o que representa uma vantagem se este parâmetro for comprometido pela doença.

O numerador na realidade é a segunda derivada do sinal, o que corresponde a um filtro passa altas.

$$Br = \frac{\sum_{i=1}^{N-2} [s(t_i) - 2s(t_{i-1}) + s(t_{i-2})]^2}{\sum_{i=1}^N [s(t_i)]^2} \quad (18)$$

3.2.12 *Nível Espectral de Ruído (SNL- Spectral Noise Level)*

Este parâmetro analisa as intensidades espectrais em intervalos de frequência, sendo útil na quantificação do som.

3.2.13 *Energia de Ruído Normalizada (NNE – Normalized Noise Energy)*

Este parâmetro proposto por Kasuya et al. (1988) considera o sinal composto de componentes periódicas mais um ruído aditivo. É definido como:

$$NNE = 10 \log \left[\frac{\frac{1}{L} \sum_{w=N_L}^{N_H} \sum_{m=1}^L |\hat{N}_m(w)|^2}{\frac{1}{L} \sum_{w=N_L}^{N_H} \sum_{m=1}^L |S_m(w)|^2} \right] \quad (19)$$

onde:

\hat{N} é a estimativa espectral do ruído.

$S_m(\omega)$ é a densidade de potência espectral do sinal de voz

L é a quantidade de janelas de dados que são utilizadas para análise

N_L e N_H são os limites inferiores e superiores, em amostras, sobre o qual a medida é desenvolvida.

3.2.14 Parâmetros de Cepstrum

- a) Relação sinal/ruído (N/S – *Noise/Signal*)
- b) Proeminência de pico Cepstral (CPP - *Cepstral Peak Proeminence*)
- c) Pearson r no pico da auto correlação (RPK - *Pearson r at autocorrelation Peak*)
- d) Relação harmônicas ruído (HNR Cepstral)

3.2.15 Avaliação por Resíduos

Uma fonte modificada de informação é o produto da filtragem inversa residual do sinal. Esta filtragem tenta remover as componentes relativas do sistema vocal presente no sinal de voz. Neste enfoque pode-se utilizar vários dos parâmetros já definidos anteriormente como, por exemplo, PPQ, APQ, P/A, SNR além de outros sugeridos como:

3.2.16 Coeficiente de excesso

Este parâmetro fornece a forma de distribuição das amplitudes. É a relação entre o quarto e o quadrado do segundo momentos centrais. Sendo positivo a distribuição é concentrada e por outro lado se for negativo a distribuição é dispersa. Em suma, este parâmetro fornece um julgamento da relação sinal/ruído. O coeficiente de excesso é calculado por:

$$EX = \frac{E\left[(x - \bar{x})^4\right]}{\left\{E\left[(x - \bar{x})^2\right]\right\}^2} - 3 \quad (20)$$

3.2.17 Nivelamento Espectral

O nivelamento espectral é uma medida da variação do sinal, ou em outras palavras, esta técnica analisa as amplitudes das componentes espectrais do sinal. O nivelamento espectral é calculado pela relação entre a média geométrica e a média aritmética do espectro isto é:

$$SF = 10 \log \left[\frac{(x_1 \cdot x_2 \cdot x_3 \cdot \dots \cdot x_n)^{1/n}}{\frac{1}{N}(x_1 + x_2 + \dots + x_n)} \right] \quad (21)$$

Caso o espectro apresente componentes em altas frequências (sinal ruidoso), o SF aproxima-se de zero. Por outro lado, se o espectro é decrescente o parâmetro apresenta valores negativos com grande módulo.

Dois parâmetros baseados nesta técnica têm sido propostos para o processamento de sinal de voz. O primeiro calcula o nivelamento espectral do resíduo (SFR - *Spectral Flatness of Residue*) e o segundo analisa as componentes espectrais dos lábios, da glote e do trato vocal através do nivelamento espectral do filtro inverso (SFF - *Spectral Flatness of Filter*).

3.2.18 Medidas Realizadas sobre o pulso glotal

O pulso glotal estimado pela filtragem inversa (anulação dos efeitos do trato supra glotal) pode ser analisado por vários parâmetros para se determinar sua forma.

3.2.18.1 Quociente de Abertura e de Fechamento

Basicamente o quociente de abertura (OQ – *Open quotient*) mede a relação entre o tempo de abertura (T_o) da glote pelo tempo total (período). Similarmente o quociente de fechamento (CQ – *Closed Quotient*) é a razão entre o tempo no qual a glote permanece fechada e o período. A relação entre ambos é:

$$OQ = 1 - CQ;$$

Onde OQ : quociente de abertura

CQ : quociente de fechamento

3.2.18.2 Quociente de Velocidade (SQ – *Speed Quotient*)

Este quociente determina a razão entre o tempo no qual a glote permanece aberta (T_o - *Time open*) e aquele no qual permanece fechada (T_c - *Time closed*).

$$SQ = \frac{T_o}{T_c}$$

3.2.18.3 Índice de velocidade (SI – *Speed Index*)

O índice de velocidade é uma medida normalizada entre os tempos de abertura e fechamento. Seu valor situa-se no intervalo $[-1,1]$. Por exemplo, se a glote não apresenta fechamento total. O SI é +1. Por outro lado se a glote permanece sempre fechada (situação impossível na vocalização), o SI é -1.

$$SI = \frac{T_o - T_c}{T_o + T_c} = \frac{SQ - 1}{SQ + 1} \quad (22)$$

3.2.19 Índice H (HI - H-Index)

Proposto por Shoji et al. (1992b) procura verificar a eficiência glotal através da análise da palavra “HI” /hai/ através da análise do pico de energia entre a abertura glotal realizada durante o início do fricativo glotal /h/ e o seu fechamento realizado durante a vogal /a/ em decibéis demonstrando a correlação entre o H-index e a relação entre o fluxo AC/DC proposto por Isshiki (1989).

3.2.20 Relação Excitação Glotal/Ruído (GNE – Glottal to Noise Excitation)

Proposto por Michaelis, Gramss e Strube (1997), este parâmetro é baseado na correlação entre os coeficientes de envelopes de Hilbert para diferentes bandas de frequência. Esta medida indica quanto do sinal vocálico provém da vibração das pregas vocais e quanto provém da turbulência provocada pela rouquidão.

O processo descrito por Michaelis consiste no seguinte algoritmo:

- 1) Alterar a frequência de amostragem para 10Khz;
- 2) Realizar a filtragem inversa do sinal de voz através de um filtro LPC de 13^a ordem com a utilização de uma janela de Hanning de 30ms de comprimento de deslocamento de 10ms;
- 3) Calcular os envelopes de Hilbert (correspondente ao valor absoluto do sinal analítico derivado da transformada de Hilbert) para 3 frequências de 1KHz de largura de banda (500, 1500 e 2000 Hz);
- 4) Calcular a função de correlação cruzada entre os diversos envelopes;
- 5) Escolher o maior valor entre as funções de correlação cruzada;
- 6) Escolher o máximo entre os valores obtidos em (5).

Para o cálculo das correlações cruzadas a normalização utilizada foi inicialmente subtrair a média de cada um dos envelopes calculados, e em seguida dividir cada elemento pela norma do vetor.

A Figura 8 mostra um exemplo de uma voz normal da correlação cruzada entre os envelopes de Hilbert de 500 e 1000Hz para uma janela de 60ms .

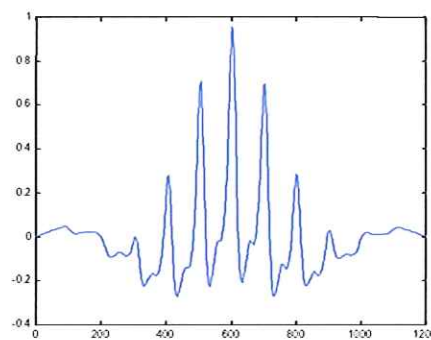


Figura 8 - Correlação cruzada entre as janelas de 500 e 1kHz (GNE)

Para um paciente normal o valor de GNE se aproxima de 1.

3.3 Índice do Ruído Turbulento (TNI – Turbulent Noise Index)

O Índice do Ruído Turbulento foi proposto por Mitev e Hadjitodorov (2000) de forma a mensurar o fator de ruído em um sinal de voz. Este trabalho é baseado em um medidor de frequência fundamental proposto por Milenkovic (1987). Diferentemente dos parâmetros HNR e NNE que são os mais utilizados para a medição da soproscopicidade da voz, o TNI mede o ruído através da diferença entre sinais subsequentes.

A equação que calcula o TNI é:

$$TNI = 100 \left(1 - \frac{1}{N-1} \sum_{n=1}^{N-1} R(t_n, T_n) \right) \quad (23)$$

onde :

$R(t_n, T_n)$ é a função de auto-correlação normalizada pela energia

$$R(t_o, T) = \frac{\sum_{i=0}^{T_w-1} v(t_o + i)v(t_o + i + T)}{\sqrt{\sum_{i=0}^{T_w-1} v^2(t_o + i) \sum_{i=0}^T v^2(t_o + i + T)}} \quad (24)$$

n : ciclo a ser calculado

t_n : início do ciclo glotal

T : Período instantâneo do ciclo

T_w = mínimo período de duração do pitch na janela a ser considerada

$$R(t_n, T_n) = \max[R(T, T)] \quad T_{\min} \leq T \leq T_{\max}$$

$$t_n - \delta \leq t \leq t_n + \delta$$

δ : corresponde a máxima variação do período de pitch a ser considerado.

No caso de periodicidade do sinal, o valor de $R(t_n, T_n)$ será próximo a 1 e o valor de TNI será baixo. No caso de sinal ruidoso o somatório terá um valor menor que 1 e o valor do TNI será elevado.

Considerações citadas pelos autores:

a) Valores de TNI não sofrem grande influência se o valor do jitter estiver abaixo de 10% ou seja o $\delta = 0.1T_0$

b) De forma a diminuir o erro gerado pela discretização do sinal existe a sugestão dos autores em utilizar uma frequência de amostragem superior a 20Khz (o que nem sempre é possível).

Capítulo 4

4 Detetores e Estimadores de Pitch

A determinação do período glotal em vozes normais já é bem conhecido e muito utilizada em sistemas de comunicação (Hess, 1983), mas somente há algum tempo o estudo se concentrou nos efeitos de ruído e jitter para o caso de vozes patológicas como o estudo realizado por Manfredi et al. (2000).

Através de um estudo preliminar realizado no laboratório e em colaboração ao trabalho de Razera (2004), verificou-se que praticamente todos os determinadores e estimadores de período de *pitch* são sensíveis a altos níveis de ruído e *Jitter*. Sendo assim deve-se utilizar métodos modificados ou mistos de forma a aumentar a precisão da medida.

O modelo mais utilizado para a medida de período de pitch é mostrado na Figura 9.

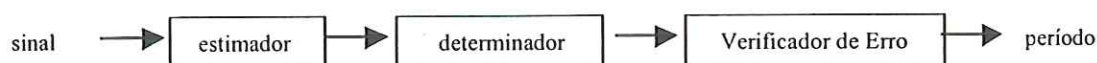


Figura 9 - Modelo para um medidor do período de Pitch

4.1 Estimador

O estimador é responsável pela caracterização da média do valor do período de pitch no sinal. Existem diversos tipos de estimadores, a seguir serão descritos os mais importantes para o caso de vozes patológicas.

4.1.1 Método AR (Auto regressive)

O modelo auto-regressivo (AR) permite uma estimação espectral simples com grande resolução e robustez com relação ao ruído. Existe uma relação próxima com análise de Predição Linear. Mantendo relação também com o método SIFT (*Simple Inverse Filter Tracking*).

4.1.2 Método SIFT (*Simple Inverse Filter Tracking*)

O método SIFT é baseado na análise do sinal obtido através da filtragem inversa do modelo de predição linear. Em vozes normais, filtros de baixa ordem são suficientes para sua caracterização ($p \approx 4$) não trazendo informações mais do que 2 formantes. Mas no caso de vozes patológicas existe a necessidade de filtros de ordem superior ou até de ordem variável devido a alta taxa de ruído inserido no sinal

4.1.3 Método AMDF (*Average Magnitude Difference Function*)

O método de AMDF pode ser utilizado para detectar mudanças rápidas e lentas da frequência fundamental e é baseado na diferença entre dois ciclos sucessivos. O AMDF é dado pela seguinte fórmula:

$$AMDF(\eta) = \frac{1}{M - \eta} \sum_{i=1}^{M-\eta} |x(i) - x(i + \eta)|, \eta = 0, \dots, M - 1 \quad (25)$$

onde: $x(i)$ = sinal do período atual

$x(i+\eta)$ = sinal do período posterior

M = corresponde ao período de pitch

Note que se $AMDF(0) = 0$ e $AMDF(T)=0$, T corresponde a um sinal de perfeita periodicidade.

4.1.4 Método do Domínio Cepstral

O Cepstrum que é o inverso da Transformada do logaritmo de uma transformação espectral (Transformada de Fourier) separa o sinal vocal em duas componentes básicas: A função de transferência do trato vocal e o sinal glotal, o qual apresenta um pico T_0 representando uma medida temporal do período. Este método só pode ser aplicado quando não há superposição espectrais, como no caso da voz.

Estimadores baseados na auto-correlação

Este tipo de detetor de período é baseado na grande semelhança do sinal de vogais sustentadas ciclo a ciclo. A verossimilhança é grande em períodos subsequentes, com isso é possível verificar o período através da detecção de máximos locais. Os inconvenientes deste tipo de procedimento são o alto custo computacional do cálculo da auto correlação e a diminuição da amplitude dos máximos locais com o distanciamento do ponto zero da função de auto-correlação.

Neste tipo de detetor podem ainda ser encontradas variações com pré-processamento através de funções não lineares e transformadas lineares.

4.1.5 Técnicas Perceptuais

Estas técnicas baseiam-se no fato de vogais sustentadas apresentarem um sinal quase repetitivo. Existem várias técnicas de implementação

- a) Método através do correlograma da saída de um banco de filtros;
- b) Casamento de Harmônicas (Martin, 1981), que se baseia na correlação um pente espectral de valor unitário e a FFT do sinal vocálico. Esse método foi utilizado e será descrito detalhadamente.

4.1.6 Técnicas que utilizam o domínio Wavelet

Já é muito conhecido o fato de Transformadas Wavelet serem capazes de detectar temporalmente eventos transientes no sinal. Através da saída da decomposição Wavelet é possível detectar o fechamento do pulso glotal como proposto por Kadambe e Boudreaux-Bartels (1992) utilizando a “Wavelet cubic spline dyadic”.

4.2 Detetores de pitch

O detetor de pitch é responsável pela determinação do ponto que irá ser considerado para a medida do período do sinal.

4.2.1 Detetores baseados na análise de cruzamento de zero

O princípio básico de funcionamento é avaliar o período através do número de cruzamentos de zeros que o sinal promove. Para tanto deve-se filtrar fortemente o sinal através de filtros tipo passa-baixa.

Existem alguns modos para aumentar o desempenho deste tipo de detetor:

- a) Utilização de elementos não lineares para aumentar a influência da primeira harmônica no sinal a ser avaliado.
- b) Utilização de filtros ajustáveis ou adaptativos.

4.2.2 Detetores baseados em mínimos ou máximos locais

Este detetor procura encontrar um máximo ou mínimo local de uma janela com a máxima probabilidade de se encontrar o próximo ciclo do sinal vocal. Esta janela é determinada pelo estimador de pitch. No caso de vozes não patológicas é utilizado um valor de comparação e um filtro não linear que zera todos os valores abaixo do valor de comparação com a finalidade de facilitar a detecção do máximo ou mínimo.

Este método demonstrou-se ineficaz no caso de vozes provenientes de vozes patológicas devido ao alto nível de Shimmer encontrado nesses sinais. Muitas vezes existem dois ou mais valores máximos no mesmo ciclo com amplitudes muito próximas. Para a realização do cálculo do período de pitch neste trabalho foi considerado uma modificação do método utilizado em vozes nomais: Considera-se todos os pontos de máximo ou mínimo locais como sendo pontos prováveis. O responsável pela determinação do valor é realizado pelo corretor de erro.

4.3 Corretores de Erro

O modo mais comum para realizar a correção de erros é aproximar os valores obtidos ao valor da mediana dos períodos. Mas, no caso de vozes patológicas, esse tipo de corretor altera as características de jitter do sinal original, invalidando as medidas de período realizadas. Sendo assim, até o momento na pesquisa bibliográfica realizada, não foi encontrado nenhum corretor de erro específico para vozes patológicas.

Capítulo 5

5 Materiais e métodos

Neste capítulo serão apresentados os materiais utilizados, o método de determinação de pitch e o procedimento de análise dos valores obtidos.

5.1 Dados utilizados

Para o estudo foram utilizadas um conjunto de vozes adquiridas no Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto em 1999. Inicialmente considerou-se a possibilidade de se avaliar cerca de 100 pacientes com um total aproximado de 300 vozes gravadas. Estas foram utilizadas para o teste das subrotinas utilizadas nos cálculos. Inicialmente cerca de 1% das vozes foram descartadas por motivo de afonia durante a locução.

Para a aquisição foi utilizado um computador ligado a uma placa de *som Sound Blaster* de 16 bits com a taxa de amostragem de 22,5 kHz. O microfone utilizado é do modelo Leson SM58, unidirecional e dinâmico.

Para o cálculo dos parâmetros, do período de *pitch* foram confeccionadas subrotinas na linguagem do software MatLab e os resultados foram gravados em arquivos texto para análise posterior.

5.2 Procedimento de análise

Como praticamente não existem trabalhos que analisam os parâmetros acústicos e as diferentes patologias foi necessário criar um meio de avaliar os dados obtidos. São utilizados os valores considerados normais encontrados em outros trabalhos para efeito de comparação com os valores obtidos.

Os valores limites para os casos considerados normais e patológicos para as diversas medidas acústicas são apresentados abaixo:

a) JF, JM e JR

O valor de JF para casos patológicos compatíveis com o níveis de sopro e rouquidão considerados por Martin et al. (1995) é de 2,5% e de aspereza em torno de 1,4% . O valor JR é altamente correlacionado com o valor de JF, pois um utiliza a frequência (JF) como base de cálculo e o outro o período (JR).

O valor de JM para vozes normais está abaixo de 0,5 ms porem como este parâmetro não é normalizado podem ocorrer variações devido a mudança da frequência fundamental.

b) Jitter DPF e Shimmer DPF

O estudo realizado por Sorenson e Horii (1984) para Jitter DPF considera valores de Jitter DPF e Shimmer DPF para vozes normais. Existem diferenças entre os valores considerados normais para cada uma das vogais e o sexo do paciente. Para facilitar a análise foi considerado valores próximos a 50% como normais e valores próximos a 67% como patológicos.

c) RAP

Preciado e Ferdandez (1998) consideram os valores normais para valores abaixo de 0,4% e para pessoas com rouquidão com valores acima de 85%. O valor obtido para o paciente F0 foi de 0,075%, valor para compatível para paciente normal.

d) Shimmer Factor

Os valores considerados normais de acordo com o estudo de Mitev e Hadjitodorov (2000) são abaixo de 2% para pessoas normais e acima de 3% para pessoas com patologias. Em ordem crescente estão os valores para soproidade, aspereza e rouquidão.

e) PVI e AVI

Uma voz normal foi utilizada como parâmetro de comparação com o intuito de verificar a coerência dos dados obtidos. O valor do paciente F0 (normal) para PVI é de $2,6e-5$, AVI inferior é de $9,3e-3$ e para AVI superior de $2,7e-3$.

f) HNR

O estudo de Rodrigues et al. (1994) foi utilizado para comparação com os valores obtidos. Para indivíduos normais o HNR é da ordem de 10 db e segundo Hadjitodorov e Mitev (2002) os valores para casos patológicos é da ordem de 17 db.

g) NNE

Para esta medida os valores considerados até -11 db são consideradas vozes normais e acima de 17db como vozes patológicas de acordo com o estudo de Hadjitodorov e Mitev (2002).

h) TNI

Mitev e Hadjitodorov (2000) em seu parâmetro TNI classificam vozes normais com valores abaixo de 0,5 e vozes patológicas com valores superiores a 1,5. Cabe salientar que não foram encontrados trabalhos avaliando os valores obtidos, nem um estudo realizado com vozes brasileiras.

i) BR

O estudo de Wolfe et al. (2000) considera valores próximos a $-0,12$ db como pacientes normais e próximos de $0,50$ db como pacientes com alguma característica perceptual observável. Para o paciente F0 (normal) o valor encontrado foi de $04,7e-2$ (-47db) esse fato pode indicar um erro de cálculo ou de escala no programa utilizado.

j) SFF, SFR e EX

Neste caso foi utilizado um paciente F0 (normal) para avaliação dos resultados. O valor para SFF normal considerado é -15 , para SFR normal é $-2,5$ e para EX é 49.

l) GNE

Para o caso de GNE valores considerados normais se encontram próximos a 1 e casos patológicos abaixo de 0,8. Para o paciente F0 (normal) o valor de GNE é de 0,98.

Capítulo 6

6 Resultados

Neste Capítulo serão apresentados os principais resultados obtidos.

6.1 Definição de um algoritmo de detecção de período de pitch

Foram realizadas diversas experiências com detectores de período de pitch. Inicialmente sem a utilização de um estimador robusto. O sistema não conseguia mensurar o período devido à grande diferença de amplitude entre um ciclo e o próximo. O próximo passo lógico foi de utilizar um sistema de estimação da F0 de forma a restringir a busca por um cruzamento por zero ou máximo local a uma região.

Após vários testes, verificou-se que:

- l) Algumas patologias geram harmônicas espúrias que podem afetar algoritmos baseados em auto-correlação (AR e SIFT)
- m) Todos os estimadores tem seu desempenho degradado pelo ruído.

Ou seja, é necessário manter uma relação entre a capacidade de rejeitar o ruído e a capacidade de desconsiderar harmônicos espúrios durante o cálculo da estimação. Através de testes práticos com vozes patológicas, verificou-se que o método perceptual de casamento de harmônicas apresenta o melhor desempenho sem nenhuma modificação extra.

Além disso é possível realizar um detetor de período de pitch utilizando a auto-correlação do sinal e obter um detector de erros grosseiros ao mesmo tempo. Isto é facilmente demonstrado a seguir:

Considerando se que a resolução do detector do período de pitch depende do tamanho da janela escolhida, pode-se utilizar uma janela de sinal com 5 a 10 períodos de pitch (aproximadamente 4096 pontos) com preenchimento por zeros até a resolução desejada. No caso dos experimentos realizados foram acrescentados zeros até que a resolução estivesse em 1Hz. A partir desse momento utiliza-se um pente espectral $P(n,q)$ onde:

$$P(n, p) = \begin{cases} 1 & n = k * q \quad k = 1, 2, 3, \dots, Np \\ 0 & \text{caso contrário} \end{cases} \quad (26)$$

onde: q é a frequência escolhida para o pente

k são as harmônicas relativas a q

Np é a ordem do pente

A função de estimação de harmônicas é dados por:

$$A_c(p) = \sum_{k=1}^{N/2p} A(kq)P(kq, q) \quad (27)$$

onde: $A(.)$ é a transformada de Fourier do sinal a ser avaliado.

A Figura 10 mostra um exemplo dos valores de $A_c(p)$ (equação 27).

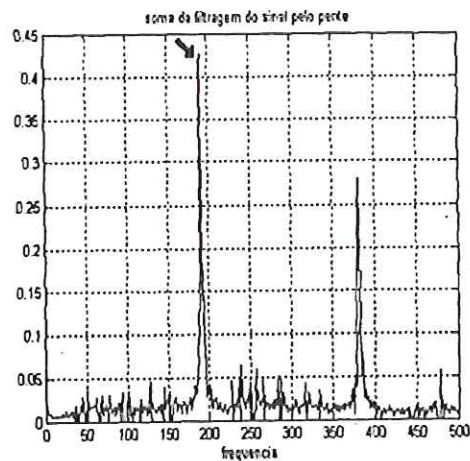


Figura 10 - $A_c(p)$

Em caso onde existam poucas formantes (vogal /i/) pode ocorrer a presença de um erro de estimação gerando uma F_0 que pode representar uma sub-harmônica do sinal. Para solucionar esse efeito, $A_c(p)$ só será computada se seu valor for menor que 10% do valor máximo da transformada de Fourier do sinal.

Obtido o valor da estimação deve-se ajustar o tamanho da janela de forma que a discretização esteja próxima a 5% do valor do F_0 estimado considerando uma janela 2^n pontos.

Desta forma pode-se calcular a auto-correlação através da FFT e IFFT do sinal janelado e ao mesmo tempo obter uma estimativa grosseira de F_0 . Pois a auto-correlação de um sinal pode ser calculada por:

$$A(t) = \text{real} \left\{ F^{-1} \left[F(s)_k F^*(s)_k \right] \right\} \quad (28)$$

Onde:

s é a janela de tamanho k

F é a transformada de Fourier

e o $*$ representa o complexo conjugado



e de $F(s)$ é possível realizar o método de casamento de harmônicas discreta.

6.2 Vozes diplofônicas

As vozes diplofônicas encontradas no banco de dados causaram a maior parte das discrepâncias entre os valores medidos e valores reais de período. Isto ocorreu porque todos os algoritmos de determinação de Pitch utilizam a consideração que existe uma forte correlação ciclo a ciclo do sinal. Aparentemente pacientes que possuem esse tipo de patologia durante a fala tentam diminuir os efeitos acústicos da diplofonia, através de dois sinais distintos que sejam mais agradáveis ao ouvido humano. Este tipo de sinal tem a característica de os dois sinais aparentarem ser um único sinal em um intervalo de tempo curto.

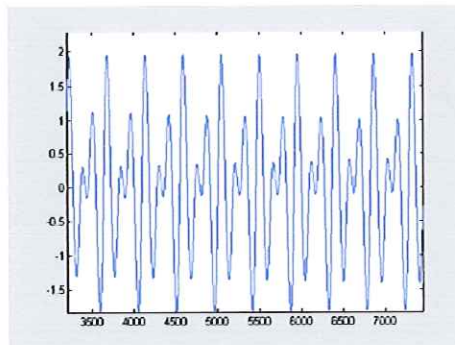


Figura 11 - Sinal composto por 2 senóides separadas por $3/2(F_0)$

Para ilustrar o tipo de problema foi gerado um sinal diplofônico, advindo de duas senóides (440Hz e 659Hz) conforme mostrado na Figura 11. É possível observar que a inserção de uma segunda frequência no sinal vocálico causa uma modulação que pode diminuir a correlação ciclo a ciclo da frequência fundamental (440Hz).

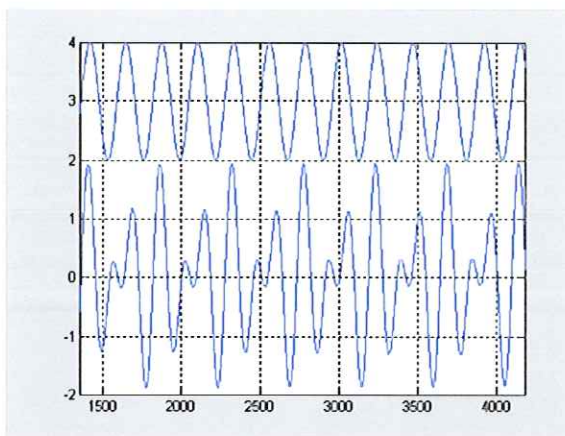


Figura 12 - sinal F0 e o sinal composto

Na Figura 12 pode-se observar que não existe correlação entre os pontos de mínimos locais e a frequência fundamental representada pela senoide superior. Desta forma a correlação ciclo-a-ciclo diminui aumentando por consequente o erro do jitter.

Em termos práticos os sinais que apresentaram diplofonia, advindos de pacientes com paralisia unilateral e nódulo, apresentaram os maiores valores de Jitter (acima de 8%). Desta forma pode-se presumir que valores elevados de Jitter podem ocorrer devido a falhas no cálculo de parâmetros.

No teste realizado com paciente com o maior valor de JR= 17%, utilizou-se um filtro Butterworth de ordem 8 sintonizado em uma frequência ligeiramente acima da estimativa de F0 de forma a diminuir a diplofonia do sinal. Através desse processo o jitter (JR) passou para 6%.

Como esse tipo de fenômeno não está previsto no modelo de determinação de Pitch nem no modelo de fala, os algoritmos testados não foram capazes de eliminar o erro causado pela diplofonia.

Segundo os resultados obtidos, o efeito causado pela diplofonia é mais significativa em vogais /e/ e /i/ do que na vogal /a/, para as medidas de Jitter e Shimmer.

Aparentemente o conteúdo harmônico da vogal /a/ não sofre a mesma degradação que ocorre com as outras vogais. Desta forma a correlação entre um período e o subsequente é menos susceptível ao efeito diplofônico.

6.3 Medidas Temporais

6.3.1 Jitter (JR)

Essa medida é influenciada pela vogal utilizada para o teste. No banco de dado a vogal /i/ apresentou uma distribuição mais concentrada (menor variância) apresentando um valor máximo 8%, a vogal /e/ 9% e a vogal /a/ 12%.

6.3.2 Shimmer

De forma semelhante à medida anterior, os valores máximos também foram influenciados pela vogal utilizada. As vogais /e/ e /i/ apresentaram uma distribuição mais concentrada entre 0 e 15% de *Shimmer*, enquanto a vogal /a/ apresenta valores entre 0 a 20%.

Outra verificação encontrada é a diferença entre os valores de *Shimmer* considerando a parte positiva ou superior e negativa ou inferior do sinal. Os valores superiores possuem em geral valores maiores que os inferiores.

6.4 Medidas de Ruído

A presença de valores elevados de *Jitter* podem comprometer os valores obtidos pelos parâmetros de ruído como o HNR e NNE. Este tipo de perturbação indica que um

período de pitch é ligeiramente diferente do próximo devido a mudança da frequência fundamental, desta forma provocando distorção harmônica do sinal vocálico.

O NNE procura encontrar um sinal médio que representa o sinal sem ruído. Se o sinal apresenta um baixo valor de *Jitter* essa consideração é verdadeira, uma vez que a média do ruído deve ser igual a zero. Mas se o sinal apresentar diferenças no período de *Pitch*, a distorção harmônica resultante será erroneamente considerada como ruído, elevando o valor do parâmetro.

O parâmetro HNR também é susceptível ao *Jitter*, pois este gera um conjunto de harmônicos espúrios que podem incrementar o erro de determinação do ruído.

É importante salientar que para fins de diagnóstico essas mudanças não sejam significativas, uma vez que possuem o mesmo gradiente com relação a presença da patologia, dificultam o processo de análise dos efeitos acústicos.

6.5 Perfis acústicos dos parâmetros para as diversas doenças

Foi realizado um estudo estatístico dos valores dos parâmetros acústicos para as vogais sustentadas /a/, /e/ e /i/, utilizando-se a média, mediana e desvio padrão.

Cabe salientar que em muitas das vozes utilizadas para o processo de cálculo notou-se mais de uma patologia, como o nódulo com uma fenda nas pregas vocais por exemplo.

6.5.1 Perfil acústico para fenda nas pregas vocais

No banco de dados utilizado para o estudo foram encontrados diversos tipos de fenda, além da fusiforme. Desta forma, quando indicou um tipo de fenda nas pregas vocais, esta voz

foi inserida na população de estudo. Foram utilizadas 17 vozes para o estudo estatístico da vogal /a/ e /e/. As tabelas 2 e 3 indicam os resultados dos parâmetros.

Foi considerado que o aumento da soproidade (ruído) foi o principal fator a ser considerado.

Os valores de JF se encontraram com valores entre de 2% a 3% para ambas as vogais, indicando a presença da patologia. Estes estão em um patamar intermediário, indicando que a presença da fenda aumenta o valor do *Jitter* através do aumento do ruído mas, não é o fator mais significativo para a caracterização da doença.

O RAP apresentou valores muito abaixo do esperado 0,14% indicando provavelmente erro no cálculo da medida.

Os valores de *Jitter* DPF e *Shimmer* DPF indicam valores coerentes aos considerados na literatura.

O valor de *Shimmer Factor* se encontra em um valor intermediário das doenças estudadas 5,2% para *Shimmer* Superior e 6% para *Shimmer* inferior. Sendo compatível com os valores esperados.

Os valores de PVI (CV *Jitter*) e de AVI (CV *Shimmer*) encontram-se com valores de aproximadamente 10 vezes do paciente F0.

O valor médio de HNR indica a presença de patologia 21db apresentado um dos valores mais elevados. Enquanto NNE possui um a média em torno de -2,4db e de mediana de 2,5db provavelmente indicando a presença de uma voz pouco ruidosa no banco de dados que causou um desvio padrão. O valor de TNI se encontra acima de 3% indicando alto valor de ruído no sinal mas, novamente a grande diferença entre a média e a mediana, e o alto valor do desvio padrão indicam a possibilidade de uma voz pouco ruidosa no banco de dados afetando na estatística dos dados.

O valor de SFR e SFF estão próximos aos valores do paciente F0, indicando a possibilidade de não serem significativos na avaliação da doença.

O coeficiente EX apresentou uma boa separação entre o paciente normal e os dados calculados apesar de não apresentar valores significativos entre as diversas doenças. O valor de GNE se encontra com uma boa separação entre o normal e o patológico mas não apresenta diferença significativa entre as diversas patologias.

Tabela 2 – parâmetros acústicos para fenda e vogal /a/

	JF	JM	JR	Jitter_DPF	Shimmer_DPF	RAP	Shimmer Factor Inferior
Média	2.8243	1.4035e-4	2.8140e+1	6.8072e+1	6.9754e+1	1.4219e-03	5.2178
Mediana	2.2362	1.1271e-4	2.2162e+1	7.1481e+1	7.3121e+1	8.1830e-4	4.3381
S	2.3754	1.2947e-4	2.3662e+1	7.8876	1.9073e+1	2.5465e-3	5.2610

	Shimmer Factor superior	CV_Jitter	CV Shimmer Factor inferior	CV Shimmer Factor Superior	HNR	NNE
Media	5.9030	1.8807e-4	2.9489e-2	2.7890e-2	2.1620e+1	-2.4095e+001
Mediana	4.9935	1.4734e-4	1.6091e-2	2.1969e-2	1.1895e+1	2.4512e+001
S	3.3053	1.3926e-4	4.2166e-2	2.5540e-2	3.2161e+1	5.4558

	TNI	BR	SFF	SFR	EX	GNE
Média	4.2823	1.5028e-2	-1.3559e+1	-1.3559e+1	8.0490	6.4446e-001
Mediana	2.6922	1.3471e-2	-1.3964e+1	-2.9324	6.4412	6.3726e-001
S	3.5411	9.3725e-3	1.8736	8.1370e-1	5.5665	2.8262e-002

Tabela 3 - parâmetros acústicos para fenda e vogal /e/

	JF	JM	JR	Jitter_DPF	Shimmer_DPF	RAP	Shimmer Factor Inferior
Média	2.1848e+000	1.0620e-004	2.1828e+001	6.5453e+001	6.8984e+001	6.5297e-004	3.2669e+000
Mediana	2.3772e+000	1.0513e-004	2.3701e+001	6.7773e+001	7.2309e+001	4.9082e-004	2.7437e+000
S	1.0851e+000	6.2413e-005	1.0814e+001	6.3304e+000	1.5194e+001	4.9620e-004	1.9986e+000

	Shimmer Factor superior	CV_Jitter	CV Shimmer Factor inferior	CV Shimmer Factor Superior	HNR	NNE
Media	4.0343e+000	2.3661e-004	2.9596e-002	3.3460e-002	2.1697e+001	-2.7119e+001
Mediana	3.4505e+000	1.0581e-004	1.9205e-002	2.9151e-002	1.7690e+001	-2.7808e+001
S	2.2086e+000	2.7889e-004	2.8700e-002	2.3802e-002	1.6820e+001	3.9600e+000

	TNI	BR	SFF	SFR	EX	GNE
Média	3.3993e+000	1.7139e-002	-1.3044e+001	-3.4324e+000	9.5903e+000	
Mediana	2.7613e+000	1.0833e-002	-1.2161e+001	-3.2991e+000	5.7094e+000	
S	2.5818e+000	1.6758e-002	2.7465e+000	8.9304e-001	7.8919e+000	

6.5.2 Perfil acústico para edemas

De forma semelhante ao caso anterior, além do edema de Reinke, verificou-se a existência de outros tipos de edema, classificados segundo o tamanho. Em geral nos estudos clínicos são utilizados de 30 a 40 pacientes, no banco de dados não existia número suficiente para cada tipo de edema, optou-se traçar o perfil acústico de um edema sem nenhuma classificação. Para o cálculo dos parâmetros acústicos, foram utilizadas 14 vozes para a vogal /a/ (Tabela 4) e 10 para a vogal /e/ (Tabela 5).

O fator bastante associado à presença de edemas é o aumento do jitter associado principalmente à rouquidão e ao aumento da soproidade.

O valor de Jitter foi superior ao associado à fenda sendo o aumento de aproximadamente 15% para a vogal /a/ e de 50% para a vogal /e/, mas os valores próximos das medianas pode indicar a presença de dados com valores de jitter muito superiores a média, modificando desta forma o resultado final. Essa explicação é muito plausível pois os níveis de jitter associados à soproidade e rouquidão são ligeiramente semelhantes.

Os parâmetros direcionais Jitter DPF e Shimmer DPF encontram-se com valores semelhantes aos valores encontrados nas outras patologias com um baixo valor de desvio padrão.

Novamente o valor do RAP apresentou valores abaixo do esperado 0,14%.

Os valores de Shimmer superior e inferior estão ligeiramente mais elevados do que o caso com fenda; esse fato mostra a alteração causada pela associação da rouquidão e soproidade no sinal vocálico.

Os valores de PVI e AVI são ligeiramente superiores aos valores encontrados nos casos com fenda.

O valor de HNR é ligeiramente menor que o valor para a fenda indicando a presença maior de sinal com relação ao ruído de fundo, enquanto o valor de NNE se encontra próximo ao encontrado no caso da fenda. Neste caso indicando uma diferença no modo de avaliação do ruído.

O valor de TNI é superior ao encontrado no caso da fenda por volta de 180% indicando a possibilidade da utilização desse parâmetro de forma a separar essas duas doenças.

Novamente os valores de SFF e SFR estão próximos aos valores encontrados no paciente F0.

Tabela 4 – parâmetros acústicos para Edema utilizando a vogal /a/

	JF(%)	JM	JR	Jitter_DPF	Shimmer DPF	RAP	Shimmer Factor_i
Média	3.1677	1.7758e-4	3.1734e+1	6.8665e+1	7.5558e+1	1.4550e-3	6.8415
Mediana	2.1856	1.5451e-4	2.1856e+1	6.9398e+1	7.4370e+1	8.5596e-4	4.8546
S	2.3827	1.2051e-4	2.3816e+1	5.0057	3.3392	1.7388e-3	5.4538

	Shimmer_ Factor_s	CV_Jitter	CV CV Shimmer Factor_i	CV Shimmer Factor_s	HNR	NNE
Media	7.5459	6.1244e-4	3.8455e-2	6.9144e-2	1.1514e+1	-2.1943e+1
Mediana	6.5532	4.1962e-4	2.4538e-2	2.9444e-2	8.5989	-2.2438e+1
S	4.4552	7.4473e-4	6.1717e-2	9.3830e-2	9.0723	4.3224

	TNI	BR	SFF	SFR	EX	GNE
Média	6.0946	2.1910e-2	-1.3484e+1	-2.9648	9.3806	6.4490e-1
Mediana	4.8288	1.9398e-2	-1.3703e+1	-2.8320	6.7243	6.3726e-1
S	4.2015	1.5432e-2	2.3435	6.4923e-1	7.3179	2.3557e-2

Tabela 5 – Parâmetros acústicos para Edema utilizando a vogal /e/

	JF	JM	JR	Jitter_DPF	Shimmer_DPF	RAP	Shimmer Factor Inferior
Média	3.2755e+000	1.9100e-004	3.2578e+001	7.0092e+001	7.3557e+001	9.5903e-004	6.8722e+000
Mediana	2.6224e+000	1.4721e-004	2.6026e+001	7.0707e+001	7.1996e+001	6.5266e-004	3.9436e+000
S	1.8454e+000	1.3115e-004	1.8073e+001	4.1567e+000	3.3571e+000	8.8945e-004	7.3795e+000

	Shimmer Factor superior	CV_Jitter	CV Shimmer Factor inferior	CV Shimmer Factor Superior	HNR	NNE
Média	5.6825e+000	3.0271e-004	2.3924e-002	2.2689e-002	2.5141e+001	-2.4656e+001
Mediana	4.4469e+000	1.7489e-004	1.4172e-002	1.6553e-002	2.2368e+001	-2.5457e+001
S	4.2171e+000	2.9846e-004	1.8070e-002	1.9403e-002	1.6733e+001	4.9151e+000

	TNI	BR	SFF	SFR	EX	
Média	3.9195e+000	1.5911e-002	-1.4119e+001	-3.0995e+000	5.8012e+000	
Mediana	2.6597e+000	8.5733e-003	-1.4229e+001	-3.1367e+000	5.7360e+000	
S	5.0822e+000	1.8072e-002	1.6238e+000	5.2057e-001	3.0896e+000	

6.5.3 Perfil Acústico para Nódulo

O nódulo foi o tipo de doença mais comum encontrada no banco de dados com 21 vozes para a vogal /a/ e 18 para a vogal /e/. A tabela 6 mostra o cálculo dos parâmetros acústicos para a vogal /a/ e a tabela 7 para a vogal /e/. Esta doença está caracterizada principalmente pelo aumento da aspereza da voz e de soproidade.

O jitter apresentou valores ligeiramente menores que os valores apresentados para nódulo e edema, geralmente associado à aspereza da voz. Os valores de Shimmer tanto superior como inferior estão entre a voz com fenda e com edema. O PVI e o AVI possuem valores médios intermediários entre os valores da fenda e do edema.

O HNR, NNE, TNI e GNE apresentaram valores muito elevados para a condição de aspereza da voz, mas compatíveis com a presença de fenda nas pregas vocais que em geral é associado à presença do nódulo.

Tabela 6 – parâmetros acústicos para nódulo utilizando a vogal /a/

	JF(%)	JM	JR	Jitter DPF	Shimmer DPF	RAP	Shimmer Factor Inferior
Média	2.2539	9.8582e-5	2.2453e+1	6.8916e+1	7.2192e+1	7.0251e-4	5.5465
Mediana	1.9191	8.7232e-5	1.9155e+1	7.0143e+1	7.3698e+1	5.5273e-4	4.5249
S	1.2128	5.6136e-5	1.1988e+1	4.7723	5.1174	4.2469e-4	2.4441

	Shimmer Factor Superior	CV Jitter	CV Shimmer Factor Inferior	CV Shimmer Factor Superior	HNR	NNE
Media	5.9418e+000	4.2813e-004	4.6814e-002	7.9990e-002	1.6808e+001	-2.3619e+1
Mediana	5.1949e+000	1.4524e-004	3.5554e-002	5.9011e-002	1.1648e+001	-2.4534e+1
S	2.1603e+000	5.3135e-004	4.3789e-002	7.1685e-002	1.6756e+001	3.7966

Média	TNI	BR	SFF	SFR	EX	GNE
Média	5.8631e+000	3.0218e-002	-1.3962e+001	-3.0655e+000	5.8867	6.5146e-1
Mediana	2.8901e+000	2.3881e-002	-1.4450e+001	-3.0946e+000	4.0545	6.3726e-1
S	6.7283e+000	2.2253e-002	3.3657e+000	7.4553e-001	6.3078	2.8262e-2

Tabela 7 – Parâmetros acústicos para nódulo utilizando a vogal /e/

	JF(%)	JM	JR	Jitter DPF	Shimmer DPF	RAP	Shimmer Factor Inferior
Média	2.2174e+000	9.9724e-005	2.2109e+001	6.4416e+001	7.0887e+001	5.0478e-004	3.9055e+000
Mediana	2.2000e+000	9.0027e-005	2.1995e+001	6.4482e+001	7.2321e+001	4.4544e-004	3.2255e+000
S	1.1278e+000	5.2821e-005	1.1231e+001	6.8933e+000	8.3689e+000	2.7000e-004	1.8628e+000

	Shimmer Factor Superior	CV Jitter	CV Shimmer Factor Inferior	CV Shimmer Factor superior	HNR	NNE
Media	4.2100e+000	4.0849e-004	2.5292e-002	7.1602e-002	3.3486e+001	-2.5923e+001
Mediana	3.9424e+000	1.5674e-004	1.8381e-002	3.5612e-002	2.1454e+001	-2.6847e+001
S	1.8041e+000	9.4158e-004	1.9626e-002	1.4791e-001	3.7909e+001	3.7466e+000

	TNI	BR	SFF	SFR	EX	
Média	2.1205e+000	3.6844e-002	-1.3336e+001	-3.0482e+000	9.2636e+000	
Mediana	2.0673e+000	1.8875e-002	-1.3415e+001	-2.8501e+000	6.3141e+000	
S	1.3640e+000	4.2238e-002	2.5817e+000	6.1914e-001	1.0619e+001	

6.5.4 Perfil acústico para pólipos

O Pólipo foi a doença que apresentou a maior assimetria estatística (diferença entre a média e a mediana) para os valores de Jitter. Este fato pode ser explicado por existirem

diferentes tipos de pólipos, por seu crescimento ser normalmente unilateral e por utilizar uma amostra pequena. Foram utilizados 14 vozes, a tabela 8 mostra os parâmetros acústicos para a vogal /a/ e a tabela 9 para a vogal /e/.

Essa doença apresenta valores intermediários de aspereza com baixo valor de soproidade.

O jitter apresentou valores intermediários com alta dispersão e diferença entre a média e a mediana. Esse fenômeno é explicado pela grande variedade de tamanhos de pólipos. Os valores de Shimmer são ligeiramente menor que o nódulo considerando-se a mediana. O valor de HNR apresenta valores menores que o valor do nódulo indicando um melhor fechamento das pregas vocais.

Tabela 8 – parâmetros acústicos para pólipos e vogal /a/

	JF(%)	JM	JR	Jitter_DPF	Shimmer_DPF	RAP	Shimmer Factor Inferior
Média	4.0065	2.7876e-4	4.0214e+1	6.9678e+1	7.1902e+1	1.2104e-3	3.9288
Mediana	1.3007	8.1453e-5	1.2973e+1	6.7793e+1	7.0303e+1	1.1222e-3	2.9722
S	5.4733	4.0957e-4	5.5054e+1	7.3266	8.1921	2.9284e-4	2.3401

	Shimmer Factor Superior	CV_Jitter	CV Shimmer Factor Inferior	CV Shimmer Factor superior	HNR	NNE
Media	4.7654e+000	4.4440e-004	1.7081e-002	3.2117e-002	1.0072e+001	-2.5769e+1
Mediana	4.1449e+000	3.6134e-004	1.6313e-002	2.8729e-002	8.4010e+000	-2.6804e+1
S	1.8600e+000	2.8145e-004	1.2051e-002	1.6890e-002	5.7009e+000	4.0705

	TNI	BR	SFF	SFR	EX	GNE
Média	2.6764e+000	4.9629e-003	-1.5655e+001	-2.5861e+000	1.2823e+001	6.7348e-001
Mediana	2.3037e+000	4.6475e-003	-1.5770e+001	-2.5262e+000	8.8729e+000	6.7348e-001
S	1.1371e+000	1.5554e-003	9.4271e-001	2.5460e-001	1.2472e+001	1.4219e-002

Tabela 9 – parâmetros acústicos para pólipos e vogal /e/

	JF(%)	JM	JR	Jitter_DPF	Shimmer_DPF	RAP	Shimmer Factor Inferior
Média	5.8819e+000	4.0847e-004	5.8881e+001	6.6965e+001	6.8844e+001	3.0521e-003	7.6566e+000
Mediana	1.5543e+000	8.4527e-005	1.5366e+001	6.1126e+001	7.0537e+001	1.1988e-003	2.6924e+000
S	8.7990e+000	6.5830e-004	8.8312e+001	1.7440e+001	1.3572e+001	4.0579e-003	1.0717e+001

	Shimmer Factor Superior	CV_Jitter	CV Shimmer Factor Inferior	CV Shimmer Factor superior	HNR	NNE
Média	1.1231e+001	5.4131e-004	1.6496e-002	2.7829e-002	1.3567e+001	-2.4782e+001
Mediana	2.9895e+000	3.5414e-004	1.4674e-002	2.4509e-002	1.4292e+001	-2.7824e+001
S	1.7274e+001	5.1696e-004	1.0606e-002	2.1163e-002	3.1510e+000	7.7761e+000

	TNI	BR	SFF	SFR	EX	
Média	3.7385e+000	5.7633e-003	-1.4626e+001	-2.9965e+000	1.0810e+001	
Mediana	3.6831e+000	5.4107e-003	-1.4760e+001	-3.0773e+000	1.0918e+001	
S	2.5898e+000	3.4825e-003	2.6000e+000	4.3616e-001	5.4944e+000	

6.5.5 Perfis acústicos para Paralisia Unilateral das pregas vocais

A Paralisia Unilateral apresentou um alto grau de JR, o que pode ser justificável pela falta de inervação de uma das pregas vocais, causando um efeito diplofônico que dificulta o processo de determinação do período de pitch. No banco de dados foram encontradas 4 vezes com esse tipo de patologia. A tabela 10 mostra o cálculo dos parâmetros acústicos para a vogal /a/.

A principal característica da paralisia unilateral é o alto valor de sopro e um valor entre médio para elevado de aspereza.

O valor de *jitter* nas amostras avaliadas foi de 4% muito acima do valor encontrado por Zhang et al. (no prelo) que se encontra em torno de 1,4%. Esse fato não é significativo pois o número de amostras utilizadas foi pequena em comparação com utilizados pelo trabalho citado (67 vezes). O valor de Shimmer também foi ligeiramente mais elevado 6% contra 3.6% encontrado no artigo citado. Os valores de HNR e TNI possuem valores elevados sendo altamente coerente com o aumento significativo da sopro e aspereza.

Tabela 10 – Parâmetros acústicos para paralisia unilateral para vogal /a/

	JF(%)	JM	JR	Jitter DPF	Shimmer DPF	RAP	Shimmer Factor Inferior
Média	4.1806	2.1275e-4	4.1352e+1	6.9672e+1	7.5584e+1	3.4022e-3	5.9817
Mediana	4.4042	2.4476e-4	4.3208e+1	6.7888e+1	7.2554e+1	9.6640e-4	4.2124
S	2.3990	2.3990	2.3860e+1	4.9572	7.6219	5.1377e-3	4.4542

	Shimmer Factor Superior	CV Jitter	CV Shimmer Factor Inferior	CV Shimmer Factor superior	HNR	NNE
Media	5.8862	4.3078e-4	5.9493e-2	7.2499e-2	2.9708e+1	-2.2549e+1
Mediana	5.9152	4.0042e-4	1.7856e-2	3.1224e-2	2.0571e+1	-2.2243e+1
S	2.6899	3.0922e-4	9.0050e-2	9.8815e-2	3.2984e+1	5.2656

	TNI	BR	SFF	SFR	EX	GNE
Média	4.2828	2.7907e-2	-1.4764e+1	-2.8284	1.8873e+1	6.5533e-1
Mediana	2.3339	3.4290e-2	-1.4251e+1	-2.8459	1.1501e+1	6.5533e-1
S	5.0851	1.8071e-2	1.4233	5.0092e-1	2.1280e+1	2.0870e-2

A tabela da vogal /e/ não foi mostrada devido ao insuficiente número de vozes (1 voz).

6.5.6 *Patologias com número insuficiente de vozes para cálculo de estatísticas*

No banco de dados, duas patologias apresentam somente um sinal de voz. Neste caso, optou-se por não realizar a estatística.

Estas duas patologias foram as seguintes:

- a) Espessamento de pregas vocais
- b) Cisto

6.6 *Avaliação dos resultados obtidos*

Para avaliar os resultados foi calculado o coeficiente de dispersão para os parâmetros e as diversas patologias que é mostrado na equação 29.

$$CD = \frac{\bar{x}}{s} \quad (29)$$

Onde: \bar{x} é a média dos valores
 s é o desvio padrão

O resultado obtido é mostrado na tabela 11, sendo que os valores que estão em branco na tabela corresponde a valores muito pequenos que não foram considerados.

Tabela 11 – Coeficiente de dispersão do parâmetros para as diversas patologias

Parâmetros	Fenda		Edema		Nódulo		Pólipo		Paralisia unilateral
	/a/	/e/	/a/	/e/	/a/	/e/	/a/	/e/	/a/
Jitter DPF	8,63	10,34	13,73	16,89	14,45	9,35	9,51	3,84	14,04
Shimmer DPF			22,69	21,89	14,13	8,47	8,78	5,07	9,92
NNE	4,41	6,85	5,08	5,01	6,22	6,91	6,33	3,19	4,28
SFF	7,25	4,74	4,74	8,71	4,14	4,92	16,66	5,63	10,39
SFR	16,74	3,85	4,55	5,96	4,09		10,36	6,82	5,66
GNE	23		32		21,66		47,37		31,35
HNR								4,3	

Entre fenda e edema para a vogal /a/ o Shimmer DPF as diferencia pois os parâmetros significativos são iguais com exceção de valor de Shimmer DPF. Também para fenda e nódulo verifica-se situação idêntica, sendo que a discriminação é realizada através do SFR quando a vogal /e/ é utilizada.

Quando temos um pólipo presente nas pregas vocais o parâmetro HNR torna-se substancial para a vogal /e/ não aparecendo significativamente em nenhuma outra patologia anteriormente citada.

Para o caso de paralisia unilateral os resultados obtidos não diferem dos resultados das patologia fenda e edema, encontrados para a vogal /a/. Os valores dos parâmetros parâmetros para a vogal /e/ não foram calculados devido a amostra insuficiente de dados.

Capítulo 7

7 Conclusões

Neste capítulo serão apresentadas as principais ponderações e conclusões sobre o trabalho desenvolvido e sugestões de trabalhos futuros

7.1 Avaliação dos resultados obtidos

A utilização de perfis acústicos para a análise de vozes patológicas torna-se conveniente e necessária pois através dessa ferramenta é possível verificar a variável de separação entre as diversas patologias.

Para avaliação dos resultados obtidos é necessário verificar a influência das características perceptivas das diferentes patologias e os valores típicos dos diversos parâmetros.

A paralisia unilateral quando causa o efeito diplofônico deve ser estudada de forma específica pois as hipóteses utilizadas para a extração do período de pitch consideram o sinal sendo formado por um único valor.

A grande maioria dos parâmetros acústicos tem grande sensibilidade aos erros de determinação do período de pitch, portanto é necessário o estudo de algoritmos mais adequados para sua extração de forma a aumentar a confiabilidade do cálculo dos parâmetros.

A maioria dos parâmetros apresentou uma grande dispersão inviabilizando a sua utilização como parâmetro do perfil acústico.

7.2 Ponderações sobre os resultados

As medidas acústicas de ruído, em especial o NNE, são profundamente dependentes da determinação correta do período de *pitch*. Aqui resalta-se a dificuldade em se estimar corretamente o pitch principalmente por 2 fatores:

- 1) a presença de ruído no sinal geralmente ocasiona um posicionamento inadequado da janela do sinal utilizado para a determinação do ruído.
- 2) A presença de jitter no sinal ocasiona um efeito de modulação em frequência computado como ruído na determinação do ruído.

Portanto medidas independentes da determinação do período de *pitch*. Apresentaram geralmente resultados melhores em termos de separação de patologias da laringe.

Não existe padronização do método de extração do período de *pitch* dentre os programas comerciais. O que pode dificultar estudos comparativos entre os valores obtidos.

Sugestões sobre trabalhos futuros

- Acréscimo de outras medidas acústicas para avaliação do perfil acústico.
- Estudo da influência do Jitter sobre as medidas de ruído, principalmente NNE e HNR.
- Avaliação de um método que compense o efeito do Jitter na estimação do ruído.

Capítulo 8

Referências Bibliográficas

ANDRADE, L. B. de O.; VIERA J.M.; RAZEIRA D. E.; PEREIRA J. C. Medidas de perturbação da voz; um novo enfoque. *Fonoaudiologia Brasil*, v.2, n. 2, p.39-46, 2002.

BHUTA T.; PATRICK L.; GARNETT J.D. Perceptual evaluation of voice quality and its correlation with acoustic measurements. *Journal of Voice*. n.18, v.3, p.299-304, 2004.

DAVIS, S. B. Acoustic characteristics of normal and pathological voices. In *Lass, N. J. (Ed) Speech and Language: Advances in Basic Research and Practice*, v.1, New York Academic Press, p.271-335, 1979.

DEAL, R. E.; EMANUEL, F. W. Some waveforms and spectral features of vowel roughness. *Journal of Speech and Hearing Research*, v.21, p.250-64. 1978.

EMANUEL, F. W.; LIVELY, M. A.; McCOY, J. F. Spectral noise levels and roughness ratings for vowels produced by males and females. *Folia Phoniatrica*, n.25, p.110-20, 1973.

FUKAZAWA, T.; EL-ASSUOOTY, A.; HONJO, I. A new index for evaluation of turbulent noise in pathological voice. *Journal of the Acoustical Society of America*. v.83, p.163-7, 1988.

GOLD, B.; MORGAN, N. Speech and audio signal processing – Processing and perception of speech and music. *John Wiley & Sons inc.*, New York, 2000.

GUYTON, A. C. Fisiologia Humana. *Interamericana*, Rio de Janeiro, 1985

HADJITODOROV S.; MITEV P. A computer system for acoustic analysis of pathological voices and laryngeal diseases screening. *Medical Engineering e Physics*, v.24, p.419-29, 2002.

HECKER M. H. L.; KRUEL, E. J. Description of the speech of patients with câncer of the vocal folds . Part I: Measures of fundamental frequency. *Journal of the Acoustical Society of America*, v.49, n.4, parte 2, p.1275-82, 1970.

HESS W. Pitch Determination of Speech Signals. *Springer-Verlag*, Berlin Heidelberg, 1983.

HORII, Y. Fundamental Frequency perturbation observed in sustained phonation. *Journal of Speech and Hearing Research*, v.25, p.12-4, 1979.

ISSHIKI, N.; OKAMURA, K.; KOJIMA, H.; HARITA, Y. Phonosurgery: Theory and Practice. *Springer-Verlag*, p.46-8, Tokyo, 1989.

KADAMBE S.; BOUDREAUX-BARTELS. G. F. Application of Wavelet Transform for Pitch Detection of Speech Signals. *IEEE Transactions on Information Theory*, v.38, n.2, p.917-24, 1992.

- KASUYA, H.; OGAWA, S.; MASHIMA K.; EBIHARA, S. Normalized Noise Energy as a acoustic measure to evaluate pathologic voice. *Journal of the Acoustical Society of America*, v.80, p.1329-34, 1988.
- KOIKE, Y.; TAKAHASHI, H.; CALCATERRA, T. C. Acoustic measures for detecting laryngeal pathology. *Journal of Phonetics*, v.14, p.517-24, 1977.
- KOJIMA, H.; GOULD, W. J.; LAMBIASE, A.; ISSHIKI, N. Computer analysis of hoarseness. *Acta Otolaryngologica*, v.89, p.547-54, 1980.
- LIEBERMAN, P. Some acoustics measures of the fundamental periodicity of normal and pathologic larynges. *Journal of the Acoustical Society of America*, v.35, p.344-53, 1963.
- MANFREDI, C.; D'ANIELLO, M.; BUSCAGLIONI, P.; ISMAELLI, A. A comparative analysis of fundamental frequency estimation methods with application to pathological voices. *Medical Engineering & Physics*, v.22, p.135-47, 2000.
- MARTIN, D.; FITCH, J.; WOLFE, V. Pathologic Voice Type and the Acoustic Prediction of Severity. *Journal of Speech and Hearing Research*, v.38, p.765-71, 1995.
- MARTIN, P.H. Detection de F0 par intercorrélacion avec une fonction peigne, *Journées d'Etude sur la Parole*, v.12, p.221-32, 1981
- MICHAELIS, D.; GRAMSS, T.; STRUBE, H. W. Glottal to noise excitation ratio - a new measure for describing pathological voices. *Acustica / acta acustica* v.83, p.700-6, 1997.
- MILENKOVIC, P. Least Mean Square Measures of Voice Perturbation. *Journal of Speech and Hearing Research*, v.30, p.529-38, 1987.
- MITEV, P.; HADJITODOROV, S. A method for turbulent noise estimation in voiced signals. *Medical Biological Engineering and computing*, v.38, 625-31, 2000.
- MURRY, T.; DOHERTY E. T. Selected Acoustics Characteristics of Pathologic and Normal Speakers. *Journal of Speech and Hearing Research*, v.23, p.361-9, 1980.
- PINHO, S. Avaliação e tratamento da voz. In PINHO S (Org). *Fundamentos em Fonoaudiologia Tratando os Distúrbios da Voz*. Rio de Janeiro: Guanabara-Koogan, 1998, p.2-48.
- PRECIADO, J. A.; FERNANDEZ, S. Digital analysis of the acoustic signal in vocal pathology diagnosis. Sensitivity and specificity of shimmer and jitter measurements. *Acta Otorrinolaringologica Espanola*. v.49, p.475-481, 1998.
- RODRIGUES, S.; BEHLAU, M.; PONTES, P. Proporção harmônico-ruído: valores paa indivíduos adultos brasileiros., *Acta Who*, v.13, n.3, p.112-116, 1994.
- RAZERA, D. E. **Determinação de Pitch**. 2004. 94p.. Dissertação (Mestrado em Engenharia Elétrica) – Escola de Engenharia de São Carlos, Universidade de São Paulo.
- RABINER, L.R.; SCHAFER, R.W. Digital Processign of Speech Signals. *Prentice Hall*, New Jersey, 1978.

ROSENBERG, A. E. Effect of Glottal Pulse on the Quality of Natural Vowels. *Journal of Acoustic Society of America*, v. AU-16, pp 81-94, 1971.

SHOJI, K.; REGENBOGEN, E.; YU, J. D.; BLAUGRUND, S. M. High-frequency power ratio of breathy voice. *Laryngoscope*, v.102, p.267-71, 1992a.

SHOJI, K.; REGENBOGEN, E.; YU, J. D.; Stanley M.; BLAUGRUND, S. M. H-Index: A New Measure of Glottal Efficiency for The Pathologic Voice. *Laryngoscope*, v.102, p.1113-7, 1992b.

SORENSEN, D.; HORII, Y. Direction Perturbation Factors for Jitter and Shimmer. *Journal of Communication Disorders*, v.17, p.143-51, 1984.

TAKAHASHII, H.; KOIKE, Y. Some perceptual dimensions and acoustical correlates of pathologic voices. *Acta Oto Laryngologica*, suppl. 338, 1-24, 1975.

WOLFE, V.I.; MARTIN D. P.; PALMER C.I. Perception of dysphonic voice quality by naive listeners, *Journal of Speech, Language, and Hearing Research*, V.43, p.697-705, 2000.

YUMOTO, E.; GOULD, W. J.; BAER, T. Harmonics-to-noise ratio as an index of the degree of hoarseness. *Journal of the Acoustical Society of America*, v.71, p.1544-49, 1982.

ZHANG, Y.; JIANG, J.J.; BIAZO, L.; JORGENSEN, M. Perturbation and Nonlinear Dynamic Analyses of Voices from Patients with Unilateral Laryngeal Paralysis, *Journal of Voice*, no prelo.

Capítulo 9

Bibliografia consultada

- BAKEN, R. J. Clinical Measurements of Speech and Voice, *Singular Publishing Group, Inc.*, San Diego, USA, 1996.
- HILLENBRAND, J.; CLEVELAND, R. A.; ERICKSON, R. L. Acoustic correlates of breathy vocal quality. *Journal of Speech and Hearing Research*, v.37, p.769-778, 1994.
- HIRAOKA, N.; KITAZOE, Y.; UETA, H, TANAKA, S. and TANABE, M. "Harmonic-intensity analysis of normal and hoarse voices". *Journal of the Acoustical Society of America*, n.76, p.1648-51, 1984.
- KOIKE, Y. Application of some acoustic measures for the evaluation of dysfunction. *Studia Phonologic*, 7, 17-23, 1973.
- KOIKE, Y.; MARKEL, J. Application of inverse filtering for detecting laryngeal pathology. *Annals of Otolaryngology*, v.84, p.117-24, 1975.
- KLINGHOLZ, F.; MARTIN F. Quantitative Spectral Evaluation of Shimmer and Jitter. *Journal of Speech and Hearing Research*, v.28, p.169-74, 1985.
- KROM, G. DE. A cepstrum-based technique for determining a harmonics-to-noise ratio in speech signals". *Journal of Speech and Hearing Research*, v.36, p.254-66, 1993.
- LIVELY, M. A.; EMANUEL, F. W. Spectral noise levels and roughness severity ratings for normal and simulated rough vowels produced by adult females. *Journal of Speech and Hearing Research*, p.13, v.498-502, 1970.
- PRESS, W. H.; TEUKOLSKY S. A.; VETTERLING, W. T.; FLANNERY B.P.; Numerical Recipes in C, *Cambridge Press Syndicate of the University of Cambridge*, cap.7 p . 545-7. , 1998
- MUTA, H., BAER, T.; WAGATSUMA, K.; MURAOKA, T.; FUKUDA, H. A pitch-synchronous analysis of hoarseness in running speech. *Journal of the Acoustical Society of America*, v.84, p.1292-1301, 1988.
- RODRIGES S.; BEHLAU M.; PONTES P. Proporção harmônico-ruído: valores para indivíduos adultos brasileiros. *ACTA AWHO*, v.8, n.3, p.112-16, 1994.
- ROSA, M. O. Análise acústica da voz para pré-diagnóstico da patologias da laringe, 1998, 219 p. Dissertação (Mestrado em Engenharia Elétrica), São Carlos, *Escola de Engenharia de São Carlos - Universidade de São Paulo*, 1998
- SANSONE, F. E; EMANUEL, F. W. Spectral noise levels and roughness severity ratings for normal and simulated rough vowels produced by adult males. *Journal of Speech and Hearing Research*, v.13, p.472-88, 1970.

SCHOENTGEN, J.; GUCHTENEERE, R. Time series analysis of jitter. *Journal of Phonetics*, v.23, p.189-201, 1995.

TITZE, I. R.; HORII, Y.; SCHERER R.C. Some technical considerations in voice perturbation measurements. *Journal of Speech and Hearing Research*, v.30, p.252-60, 1987.

VEENEMAN, D. E.; BEMENT, S. L. Automatic glottal Inverse Filtering From Speech and Electroglottographic Signals. *IEEE Transactions on Acoustics, Speech, and signal Processing*, v.33, n.2, 369-77, 1985.