Universidade de São Paulo
Escola de Engenharia de São Carlos

ALAN PETRÔNIO PINHEIRO

Extração e avaliação da dinâmica não-linear da voz utilizando vídeo de alta velocidade da laringe

São Carlos
2012

ALAN PETRÔNIO PINHEIRO

*Extração e avaliação da dinâmica não-linear da voz
utilizando vídeo de alta velocidade da laringe*

Tese apresentada ao Departamento de
Engenharia Elétrica da Universidade de
São Paulo para obtenção do título de
Doutor em Ciências.

Área de Concentração: Processamento de
Sinais e Instrumentação

Orientador: Prof. Dr. José Carlos Pereira

São Carlos
2012

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica preparada pela Seção de Tratamento
da Informação do Serviço de Biblioteca – EESC/USP

P654e Pinheiro, Alan Petrônio
Extração e avaliação dinâmica não-linear da voz
utilizando vídeo de alta velocidade da laringe. / Alan
Pinheiro Petrônio ; orientador José Carlos Pereira. São
Carlos, 2012.

Tese (Doutorado - Programa de Pós-Graduação em
Engenharia Elétrica e Área de Concentração em
Processamento de Sinais e Instrumentação)-- Escola de
Engenharia de São Carlos da Universidade de São Paulo,
2012.

1. Pregas vocais. 2. Dinâmica não-linear. 3. Estimação
de parâmetros. 4. Modos não-lineares. I. Título.

FOLHA DE JULGAMENTO

Candidato: Engenheiro **ALAN PETRÔNIO PINHEIRO**.

Título da tese: "Extração e avaliação da dinâmica não-linear da voz utilizando vídeo de alta velocidade da laringe".

Data da defesa: 03/04/2012

Comissão Julgadora:

Prof. Titular **José Carlos Pereira (Orientador)**
(Escola de Engenharia de São Carlos/EESC)

Prof. Dr. **Adriano de Oliveira Andrade**
(Universidade Federal de Uberlândia/UFU)

Prof. Dr. **Arlindo Neto Montagnoli**
(Universidade Federal de São Carlos/UFSCar)

Prof. Dr. **Marcelo Basílio Joaquim**
(Escola de Engenharia de São Carlos/EESC)

Dr^a. **Maria Eugenia Dájer**
(Pós-doutorado na Faculdade de Medicina/USP)

Resultado:

Aprovado

Aprovado

Aprovado

Aprovado

Aprovado

Coordenador do Programa de Pós-Graduação em Engenharia Elétrica:
Prof. Titular **Denis Vinicius Coury**

Presidente da Comissão de Pós-Graduação:
Prof. Associado **Paulo Cesar Lima Segantine**

Agradecimentos

Agradeço aos meus pais por todo sacrifício, comprometimento, amor e dedicação que tiveram por mim todo este tempo sem jamais se eximirem de seus papéis. Também pelo exemplo de caráter, fé e bondade que constantemente me nortearam. Ainda, à memória de meu irmão Plínio Álvaro que foi meu companheiro por grande parte de minha vida e deixa muitas saudades. À Elisangela Nardini por toda sua humildade e companheirismo.

Ao professor José Carlos Pereira, por ter acreditado no meu trabalho e ter se disposto a me orientar concedendo-me ampla autonomia, suporte e incentivo nos momentos de queda, mas principalmente por seu exemplo de seriedade e retidão em seus atos. Aos professores Arlindo Neto (UFSCar), Domingos Tsuji (HC/USP) e Marcelo Joaquim (EESC/USP) pelos diversos apoios prestados que foram de grande importância para o andamento deste trabalho. Aos professores David Stewart (Universidade de Iowa, EUA), Gaetan Kerschen (Universidade de Liège, Bélgica) e Paavo Alku (Universidade de Aalto, Finlândia) por suas contribuições nos resultados deste trabalho. Aos que se dedicaram a árdua tarefa de coleta de dados, em especial a Maria Eugênia, Regina Pimenta, Marystella Takahashi, Camila Ishikawa, Adriana Hachiya e Gislaine Cordeiro. Por fim, e não menos importante, aos colegas de laboratório mais próximos que sempre se mostraram pessoas de caráter e boa fé e que constantemente dividimos momentos de alegria, angústia e conhecimento. Agradeço também aos colegas da Universidade Federal de São João del Rei pelo apoio prestado assim como o suporte financeiro concedido a esta pesquisa pelo CNPq e FAPESP.

“O conhecimento é como o raio de uma esfera: quanto mais aumenta, mais aumenta a superfície em contato com o desconhecido.”

Sir Isaac Newton (1643 – 1727)

Resumo

PINHEIRO, A. P. **Extração e avaliação da dinâmica não-linear da voz utilizando vídeo de alta velocidade da laringe.** 2012. Tese (Doutorado) - Departamento de Engenharia Elétrica, Universidade de São Paulo, São Carlos, 2012.

No contexto da fonação, as pregas vocais são responsáveis pela produção primária do sinal de voz na laringe. Tão importantes quanto complexas, elas exibem uma dinâmica rica e uma série de comportamentos que determinam as características fundamentais do sinal emitido pela boca. Com o intuito de avaliar suas propriedades sob a luz dos métodos da dinâmica não-linear de sistemas, este trabalho tem por objetivo principal desenvolver uma técnica de extração e simulação de sua dinâmica. Para isto são empregados modelos determinísticos das pregas vocais e imagens de videolaringoscopia de alta velocidade. Estes elementos são usados para reproduzir a dinâmica e cinemática das pregas de indivíduos saudáveis através de um procedimento de otimização dos parâmetros mecânicos destes modelos de pregas. Feito isto, os modelos são usados para avaliar o comportamento não-linear das pregas vocais através de uma análise modal não-linear cujo propósito foi descrever como as não-linearidades estruturais das pregas podem influenciar na voz. Os resultados mostram a eficiência de um método capaz de reproduzir a maior parte dos padrões de vibração das pregas e uma relação de dependência entre sua energia mecânica e sua frequência fundamental de vibração, que é fruto das não-linearidades estruturais das pregas vocais.

Palavras-chave: pregas vocais; dinâmica não-linear; estimação de parâmetros; modos não-lineares.

Abstract

PINHEIRO, A. P. **Extraction and evaluation of vocal folds dynamics using high-speed laryngeal imaging.** 2012. Thesis (Doctoral) - Departamento de Engenharia Elétrica, Universidade de São Paulo, São Carlos, 2012.

As far as phonation is concern, the vocal folds produce the primary voice signal within larynx. As important as complex, they exhibit a rich dynamics capable to influence the fundamental voice features. In light of nonlinear dynamical methods, this work has as purpose to develop a method to extract and simulate the vocal folds dynamics. Deterministic vocal folds models and high-speed videolaryngoscopy are used to reproduce the dynamics of healthy subjects and evaluate their nonlinear behavior and some phenomena raised by the nonlinearities. Nonlinear normal modes are employed to understanding how the nonlinearities may influence the primary phonation. The results show that the method could reflect the most important vibratory patterns. Additionally, the nonlinear analysis has shown energy-frequency dependence and a complex dynamics with structural nonlinearities.

Keywords: vocal folds; nonlinear dynamics; parameter estimation; nonlinear modal analysis.

Lista de ilustrações

Figura 2.1 - Esquema ilustrando a emissão vocal na laringe, trato vocal e cavidades nasal e oral. No esquema são indicadas as representações idealizadas no domínio do tempo e frequência do sinal primário de voz produzido na laringe pelas pregas. ...	8
Figura 2.2 - Diagrama ilustrando corte frontal da laringe. Nele são exibidas algumas das principais estruturas que têm participação durante a produção da voz.	10
Figura 2.3 - Vista de uma seção de uma prega vocal no plano frontal. Nela são ilustradas suas três principais camadas de tecidos que possuem diferentes propriedades e funções.	12
Figura 2.4 - Estágios de um ciclo de vibração das pregas vocais. Os estágios (c) a (e) pertencem ao ciclo de abertura das pregas enquanto que os estágios (f) e (g) fazem parte do ciclo de fechamento.	13
Figura 2.5 - Sinal idealizado de pulso glotal ilustrando suas diferentes fases e sua relação como o movimento das pregas.	13
Figura 2.6 - Exame de videolaringoscopia da laringe. Esquema ilustrando um endoscópio rígido, conectado a uma câmera e inserido até a orofaringe para registrar os movimentos das pregas vocais.	17
Figura 2.7 - (a) Modelo mecânico de pregas vocais com duas e (b) três massas. Nestes modelos, as massas são acopladas por uma mola k_c e podem deslocar-se apenas no plano horizontal. Vista do plano coronal do modelo.	22
Figura 2.8 - Exemplo de sinais gerados por um modelo de duas massas. Figura à esquerda: deslocamento das massas m_1 (linha pontilhada) e m_2 (linha contínua).	24
Figura 2.9 - Padrão assumido pelo modelo de fluxo de ar atravessando a glote. (a) Região de Bernoulli na porção inferior e região de jato na porção superior.	29
Figura 2.10 - (a) Série temporal utilizada como exemplo. (b) Imersão em espaço de fase da série temporal em um espaço com três dimensões ($m=3$).	34
Figura 2.11 - Bifurcação de uma função logística. A partir do ponto a 3, (eixo horizontal) o sistema apresenta uma série de pontos de bifurcação.	35
Figura 2.12 - (a) Série temporal gerada pelas equações de Lorenz e (b) seu correspondente atrator caótico cujo formato é semelhante às asas de uma borboleta e possui duas regiões de “acumulação”.	35
Figura 2.12 - Gráfico de dimensão de correlação. A variável r (eixo horizontal) indica o raio de uma esfera no espaço de fase utilizada para calcular a dimensão de correlação (eixo vertical). A seta na figura indica a região de convergência.	37
Figura 2. 13 - Oscilador mecânico linear e não-amortecido com dois graus de liberdade utilizado para ilustrar os modos de um sistema mecânico em vibração.	38

Figura 2.14 – Ilustração dos dois possíveis modos de um sistema com dois graus de liberdade. Em (a) temos o “modo em fase” onde ambas as massas oscilam em fase..	40
Figura 2.15 – Resposta em frequência do IF72. Em (a) são desconsiderados os efeitos dos amortecedores e são vistos dois picos ressonantes localizados em 120 e 201 Hz...	42
Figura 3.1 – Etapas do método desenvolvido para extrair a dinâmica das pregas vocais. Em (a) é feita a aquisição de dados por videolaringoscopia; em (b) os sinais coletados são processados e a área da glote é segmentada para gerar um “sinal de área glotal”..	47
Figura 3.2 – Ilustração do sistema de videolaringoscopia de alta velocidade empregado nesta pesquisa. Na imagem é possível ver o sistema em funcionamento durante uma coleta de dados.....	49
Figura 3.3 – Na imagem superior é ilustrado um sinal real de área glotal extraída (i.e., SGE) com a correspondente imagem de alguns pontos que o compõem. Na imagem inferior é ilustrada a derivada deste sinal.....	50
Figura 3.4 – Diagrama descrevendo as etapas de processamento utilizadas para segmentar a região da glote nas imagens de videolaringoscopia de alta velocidade.....	51
Figura 3.5 – Representação genérica de um modelo de pregas vocais com suas entradas e saídas mais comuns.....	55
Figura 3.6 – O eixo horizontal dos gráficos representa o deslocamento lateral das pregas e o vertical a força de impacto entre as massas inferiores do modelo IF72. (a) Representação gráfica da equação de impacto (3.7) original do modelo.....	63
Figura 3.7 – (a) Área glotal gerada pelas massas inferiores (linha pontilhada em cinza) e superiores (linha tracejada em cinza) do modelo SH95 com parâmetros padrões. O traço em preto ilustra a área glotal total calculada para o modelo.....	65
Figura 3.8 – Algoritmo genético empregado no primeiro passo do procedimento de otimização.	67
Figura 3.9 – Série temporal correspondendo ao modo “em fase” (a) e “fora de fase” (b) do modo de um sistema linear descrito pelas equações $\ddot{x}_1 + (2x_1 - x_2) = 0$ e $\ddot{x}_2 + (2x_2 - x_1) = 0$ (x_1 : linha contínua; x_2 : linha pontilhada)...	74
Figura 3.10 – Ilustração do método pseudo-arclength usado na predição e correção do cálculo de novos MNNs. A linha contínua representa um ramo de MNNs. Basicamente, este ramo é definido como uma família de MNNs possuindo características similares (e.g., movimentos em fase ou fora de fase).....	78
Figura 4.1 – Amostras representativas de quadros extraídos de vídeos de videolaringoscopia de alta velocidade e seus respectivos sinais de área glotal gerados a partir da segmentação (i.e., SGE).....	85

Figura 4.2 – Parâmetros sintéticos (eixo horizontal) versus parâmetros otimizados (eixo vertical). Cada gráfico representa um dos oito parâmetros do modelo SH95 que foi utilizado para validar o procedimento.....	87
Figura 4.3 – Gráfico superior: erro médio relativo percentual entre os parâmetros sintéticos e simulados. Valores calculados usando a equação (3.19).	88
Figura 4.4 – Sinais extraídos das imagens (curvas em preto) do indivíduo 2 e seus correspondentes valores simulados (curvas em azul) usando o SH95 (gráficos à esquerda) e o IF72 (gráficos à direita).	90
Figura 4.5 – Sinais extraídos das imagens (curvas em preto) do indivíduo 7 e seus correspondentes valores simulados (curvas em azul) usando o SH95 (gráficos à esquerda) e o IF72 (gráficos à direita).	91
Figura 4.6 – Sinais extraídos das imagens (curvas em preto) do indivíduo 16 e seus correspondentes valores simulados (curvas em azul) usando o SH95 (gráficos à esquerda) e o IF72 (gráficos à direita).	92
Figura 4.7 – Figura esquerda: erro médio calculado entre as curvas simuladas pelo modelos mecânicos (SH95: curva em azul; IF72: curva em vermelho) e o sinal de área glotal extraído. Erro calculado pela Equação (3.20).....	95
Figura 4.8 – Evolução do algoritmo genético para um processamento de 50 gerações para cinco diferentes otimizações ilustradas como curvas na cor cinza..	96
Figura 4.9 – Modos não-lineares normais do indivíduo 18. Em (a) são exibidos a família de MNNs em fase (curva inferior) e fora de fase (curva superior)..	103
Figura 4.10 – Modos não-lineares normais do indivíduo 14. Em (a) são exibidos a família de MNNs em fase (curva inferior) e fora de fase (curva superior). A curva em cinza indica a mesma família de MNNs fora de fase mas com frequência duas vezes menor. Estes ramos de MNNs são mostrados individualmente em (b) e (c)..	108
Figura 4.11 – Modos não-lineares normais do indivíduo 19. Em (a) são exibidos a família de MNNs em fase (curva inferior) e fora de fase (curva superior). Estas curvas são mostradas individualmente em (b) e (c)..	111
Figura 4.12 – Diagramas de bifurcação de três indivíduos. O parâmetro Q_d representa o índice de assimetria entre as pregas definido pela expressão (3.33).....	114
Figura 4.13 – Retratos de fase dos indivíduos 1 (gráficos (a) e (b)), 2 (gráficos (c) e (d)) e 3 (gráficos (e) e (f)) usando o modelo SH95. Os eixos horizontal e vertical representam os deslocamentos da prega direita (x_{1d}) e esquerda (x_{1e}) em centímetros.....	115
Figura 4.14 – O gráfico superior ilustra uma forma de onda quasi-periódica do sinal de fluxo glotal U_g simulado pelo modelo mecânico usando os parâmetros do indivíduo 3 com coeficiente de assimetria $Q_d = 0,61$	116

Lista de tabelas

Tabela 2. 1 – Principais constantes mecânicas do modelo de duas massas.	22
Tabela 3. 1 – Constantes do modelo IF72.	58
Tabela 3. 2 – Tabela de distribuição de forças nas pregas vocais.	60
Tabela 4.1 – Parâmetros mecânicos calculados pelo procedimento de otimização.	94
Tabela 4.2 – Modos não-lineares dos indivíduos que compõem a amostra.	105

Lista de abreviaturas e siglas

EEG	–	Eletroglotografia
GFE	–	Gráfico frequência-energia
IF72	–	Modelo mecânico de Ishizaka e Flanagan
MLN	–	Modo linear normal
MNN	–	Modo não-linear normal
PV	–	Pregas vocais
SGE	–	Sinal glotal extraído
SGS	–	Sinal glotal simulado
SH95	–	Modelo mecânico de Steinecke e Herzel
ST95	–	Modelo corpo-cobertura de Story e Titze

Sumário

Agradecimentos	v
Resumo.....	vii
Abstract	viii
Lista de ilustrações.....	ix
Lista de tabelas	xii
Lista de abreviaturas e siglas	xiii
Capítulo 1 – Introdução	1
1.1 – Objetivos da pesquisa.....	3
1.2 – Justificativa	4
1.3 – Organização do texto.....	5
Capítulo 2 - A voz e sua análise dinâmica na laringe	7
2.1 – O sistema fonador.....	7
2.2 – As pregas vocais no contexto da fonação	9
2.3 – Biomecânica básica das pregas vocais	14
2.4 – Métodos empregados na análise das pregas vocais.....	16
2.5 – Modelos de pregas vocais.....	18
2.5.1 - Modelamento mecânico com duas e três massas	22
2.5.2 - Modelo de Ishizaka e Flanagan	25
2.5.3 - Modelo de Steinecke e Herzel	27
2.5.4 - Modelo “corpo-cobertura”	28
2.6 – Análise da voz utilizando métodos da dinâmica não-linear.....	29
2.6.1 – A dinâmica não-linear de sistemas.....	31
2.6.2 – Pesquisas envolvendo métodos da dinâmica não-linear.....	42
2.7 - Resumo.....	44

Capítulo 3 - Métodos empregados na extração e avaliação da dinâmica das pregas vocais.....	46
3.1 – Extração da dinâmica das pregas vocais	47
3.1.1 – Aquisição de dados: sujeitos e métodos.....	48
3.1.2 – Processamento das imagens da laringe	49
3.1.3 – Implementação dos modelos mecânicos das pregas vocais	55
3.1.4 – Otimização dos parâmetros mecânicos dos modelos de pregas vocais	64
3.1.5 – Validação dos procedimentos de extração da dinâmica	69
3.2 – Avaliação da dinâmica das pregas vocais.....	70
3.2.1 – Análise modal não-linear.....	70
3.2.2 – Análises gráficas por bifurcações e retratos de fase	79
3.3 – Resumo	82
Capítulo 4 - Resultados e discussões.....	83
4.1 – Extração da dinâmica das pregas vocais	83
4.1.1 – Processamento das imagens na laringe	83
4.1.2 – Validação do procedimento de extração de dinâmica	86
4.1.3 – Otimização dos parâmetros mecânicos.....	89
4.1.4 – Discussões gerais sobre o método de extração de dinâmica.....	97
4.2 – Avaliação da dinâmica das pregas vocais.....	102
4.2.1 – Análise dos modos não-lineares.....	102
4.2.2 – Avaliação da dinâmica usando diagramas de bifurcação e fase.....	113
4.2.3 – Discussões gerais sobre as avaliações da dinâmica das pregas vocais	117
4.3 – Resumo	119
Capítulo 5 - Conclusões e considerações finais	121
Apêndice A - Gráficos das amostras processadas.....	124
Referências bibliográficas	145

Capítulo 1

Introdução

Em uma economia moderna, a capacidade de comunicação oral assume importância destacada. Segundo Lohscheller e Eysholdt (2008), 60% da população economicamente ativa dos países desenvolvidos necessitam diretamente de suas habilidades de comunicação oral para exercerem sua profissão. Além de importante, a fonação é intrigante a ponto de ser um dos fatores determinantes que nos ajudaram a evoluir e diferenciar-nos na natureza enquanto seres civilizados. Tão intrigante quanto complexa, a produção da fala envolve uma série de órgãos que vão desde o pulmão até a boca gerando uma sofisticada interação entre estes e produzindo a voz como a conhecemos. Desta maneira, existem diversos sistemas do corpo trabalhando de forma harmônica para produzi-la e não um só responsável por tal atividade.

Em especial, as pregas vocais (PV) têm um papel de destaque na fonação. Estas estruturas estão situadas na laringe e através de movimentos de vibração auto-sustentados, geram um sinal primário que é usado como fonte de excitação do trato vocal para formar a voz. Sua importância é tamanha que algumas vezes o sinal gerado pelas pregas é referenciado como “sinal primário de voz”² (Story, 2002). Devido a sua complexa estrutura biológica, sua intensa atividade e intrigante ação aerodinâmica com o fluxo de ar oriundo dos pulmões, as PV despertaram grande atenção dos pesquisadores nos últimos anos e têm sido objeto de estudos em inúmeras pesquisas. Os

² Destaca-se que nem todos autores entendem como sinal primário de voz aquele produzido pelas pregas. Alguns defendem que se trata apenas de um sinal de excitação que gerará o sinal de voz no trato vocal. Independente disto, sua importância não é diminuída.

resultados destes trabalhos têm confirmado seu complexo comportamento na laringe e sua influência determinante na voz emitida. Isto porque inúmeras características da voz, como a sua própria frequência fundamental, são determinadas pelo comportamento dinâmico das pregas na laringe (Childers & Lee, 1991).

Diante deste cenário, pesquisadores vêm lançando mão de variados métodos científicos para avaliar as características das PV e entender algumas de suas propriedades básicas relativas à fonação. Os resultados mais recentes divulgados na literatura têm mostrado uma dinâmica rica e uma série de comportamentos inesperados que vão desde a simples falta de sincronismo de vibração entre a prega esquerda e a direita, até mesmo a comportamentos caóticos (Tokuda *et al.*, 2008; Zhang *et al.*, 2011) onde o sinal produzido tem uma aparência pseudo-aleatória. Ainda, a crescente exposição das pregas a certos fatores tais como, por exemplo, o tabagismo, o abuso de álcool, poluição do ar e outros diversos, tem gerado distúrbios na laringe e criado a necessidade de se desenvolver meios mais eficazes e precisos de avaliação das patologias que acometem as PV.

Respondendo a uma demanda crescente, várias tecnologias e estratégias de processamento de sinais de voz têm sido desenvolvidas para avaliar clinicamente as pregas e/ou buscar achados científicos que justificam alguns de seus comportamentos. Ainda, a necessidade de caracterizar o seu estado funcional é uma necessidade também evidente. Descobertas recentes têm mostrado que a dinâmica não-linear (Jiang *et al.*, 2003; Jiang *et al.*, 2006; Zhang & Jiang, 2008; Vaziri *et al.*, 2010) dispõe de uma série de métodos próprios que superam significativamente em desempenho e qualidade os métodos tradicionais de avaliação (baseados principalmente na análise das perturbações do sinal de voz) e que trazem novas perspectivas na medida em que fornecem meios inéditos para avaliar o comportamento das pregas sob um enfoque dinâmico e não-linear (Kaplan & Glass, 1995). Esta nova abordagem é pertinente já que possibilita a exploração de vários parâmetros da laringe e sistemas adjacentes que têm relação direta

na produção da voz. Adicionalmente, ela traz a possibilidade de análises mais rigorosas das principais fontes conhecidas de não-linearidades das pregas que estão vinculadas principalmente à (i) relação não-linear da pressão exercida pelo fluxo de ar nas paredes (tecidos) das pregas; (ii) as propriedades estruturais de seus tecidos cuja relação tensão mecânica e deformação não corresponde linearmente e (iii) ao impacto entre as pregas durante a colisão (Berry *et al.*, 1994). Esta série de novas potencialidades tem despertado o interesse da comunidade científica na área e é um dos objetivos desta pesquisa.

1.1 – Objetivos da pesquisa

Diante do exposto, esta pesquisa visa contribuir rumo à consolidação da análise dinâmica determinística das PV propondo um método de extração e simulação de sua dinâmica através de modelos teóricos. Em termos práticos, o objetivo principal é desenvolver uma maneira de avaliar *in vivo* o movimento de vibração das pregas de um determinado indivíduo e, a partir destes dados, reconstituir de maneira puramente determinística a sua dinâmica por meio de um modelo mecânico³ de PV capaz de reproduzir (simular) seu comportamento dinâmico não-linear. Ao mesmo tempo, estes mesmos elementos são usados para caracterizar as não-linearidades estruturais dos tecidos das pregas e seus efeitos no sinal produzido na laringe, contribuindo para a elucidação de alguns fenômenos atribuídos a estas não-linearidades. Em função deste objetivo primário, seguem outros relacionados à avaliação das PV à luz dos métodos da dinâmica não-linear de sistemas e dos modelos determinísticos reconstituídos pelo método aqui proposto com o intuito de avaliar as características básicas da dinâmica das pregas e ilustrar o emprego destes métodos.

³ Embora os modelos mecânicos aqui empregados remetam a representação de um sistema biológico, eles serão denominados durante todo o texto simplesmente como modelo “mecânico” e não como “biomecânico”, embora este último termo também se adéque à questão posta.

Para atender a tais propósitos, esta pesquisa utiliza um grupo de indivíduos com vozes saudáveis que tiveram os movimentos de suas PV registrados por um sistema de videolaringoscopia de alta velocidade. A aparente limitação do emprego exclusivo de indivíduos normais não impede teoricamente a aplicação da técnica desenvolvida a casos patológicos. Este enfoque dado a grupos de indivíduos saudáveis justifica-se pelo fato de que esta proposta busca reproduzir e esclarecer os mecanismos básicos da fonação relativos às oscilações das PV em circunstâncias normais. Contudo, em algumas seções são simuladas patologias nos modelos reconstruídos a partir deste grupo de indivíduos saudáveis com o simples intuito de avaliar, e ao mesmo tempo ilustrar, como os métodos da dinâmica não-linear representam estas condições.

1.2 – Justificativa

Embora metodologias consagradas como *jitter* e *shimmer*⁴ tenham suas vantagens, elas são falhas em precisão e exatidão na medida em que são sensíveis a uma série de parâmetros (*e.g.*, tipo de microfone, taxa de amostragem, comprimento da amostra, ruído) e sua aplicação é limitada dentro de contextos mais rigorosos como, por exemplo, o científico ou até mesmo em casos clínicos de disfunções. Diferentemente, métodos da dinâmica não-linear têm se mostrado mais robustos e precisos nestes cenários e têm sido capazes de produzir informações complementares elucidativas (Jiang *et al.*, 2006). Por esta questão, podem servir como valiosos meios para descrever atividades irregulares na laringe, auxiliarem em diagnósticos mais precisos e avaliações de efeitos de tratamento e produzir novos achados clínicos e/ou científicos relacionados à produção primária da voz.

O otimismo proporcionado pelas potencialidades da dinâmica não-linear de sistemas tem levado à generalização do emprego de seus métodos sem que algumas de suas suposições sejam atendidas. Por exemplo, algumas técnicas exigem que seja

⁴ Métodos designados para medir, respectivamente, as variações em frequência e amplitude que causam perturbação em um sinal.

assegurado que o sinal em análise tenha um comportamento totalmente determinístico ou que provenha de um sistema com estas características (Kantz & Schreiber, 2004), o que nem sempre é observado. Neste sentido, o desenvolvimento de um método de extração e reprodução da dinâmica por meio de modelamento mecânico atende à questão de se desenvolver um modelo determinístico da voz e viabiliza o uso de determinados métodos que exigem tal característica. Além disto, possibilita estabelecer uma relação importante entre causa e consequência da ação das pregas no sinal de voz.

O próprio uso de modelos mecânicos padrões de PV é outra limitação frequentemente vista em estudos científicos. Estes trabalhos utilizam constantes padrões que não refletem as características individuais de um sujeito ou grupo e por isto limitam o alcance de seus resultados. Neste sentido, a extração e estimação de parâmetros para diferentes indivíduos pode trazer novas descobertas de padrões dinâmicos em grupos ou mesmo ajudar a elucidar a presença de comportamentos irregulares registrados em experimentos laboratoriais envolvendo PV excisadas (Zhang & Jiang, 2005; Tokuda *et al.*, 2008). Ainda, o uso de um modelo determinístico de PV permite explorar uma série de fatores e assim avaliar a contribuição das não-linearidades estruturais das pregas e como estas comportam-se em diferentes cenários. Em especial, é de interesse desta pesquisa avaliar a relação entre a energia do sistema e como esta pode afetar as vibrações das pregas naquilo que talvez seja sua característica mais básica: sua frequência de vibração.

1.3 – Organização do texto

O texto é dividido em outros quatro capítulos. O capítulo 2 descreve alguns aspectos básicos relativos à produção da voz e ao comportamento das pregas durante a emissão vocal. São também abordados alguns dos trabalhos mais relevantes e atuais que fazem referência à dinâmica não-linear aplicada à voz e seus métodos clássicos que vêm sendo empregados na área. No capítulo 3 são apresentadas as técnicas

desenvolvidas para extrair a dinâmica de PV de um grupo de indivíduos através de ajustes nos parâmetros dos modelos mecânicos empregados nesta pesquisa. São também ilustradas algumas análises e métodos da dinâmica não-linear de sistemas que serão aplicados aos modelos otimizados para estudar os efeitos das não-linearidades estruturais das pregas no sinal produzido na laringe.

No capítulo 4 são ilustrados e discutidos os resultados alcançados por este trabalho dando enfoque ao método da extração da dinâmica das PV assim como as avaliações dinâmicas feitas usando os modelos mecânicos otimizados. Por fim, no capítulo 5, são feitas as conclusões e considerações finais desta pesquisa.

Capítulo 2

A voz e sua análise dinâmica na laringe

Neste capítulo são abordados alguns dos temas e pesquisas mais relevantes no que diz respeito ao estudo da produção da voz na laringe.

Inicialmente é feita uma breve introdução ao sistema fonador dando especial destaque às PV que são o foco primário desta tese. As características básicas destas estruturas também são brevemente descritas. Na sequência, são apresentadas as tecnologias utilizadas para avaliar os sinais produzidos na laringe e alguns modelos mecânicos empregados para simular o comportamento dinâmico das PV. Como mostra a revisão bibliográfica feita aqui, estes recursos têm sido utilizados em associação com métodos da dinâmica não-linear para avaliar sinais de vozes e a dinâmica das PV com sucesso.

O texto também descreve alguns dos principais métodos da dinâmica não-linear empregados em pesquisas de análise de voz e alguns trabalhos relevantes que estão em um contexto comum ao desta tese. Em especial é dada ênfase à análise modal como ferramenta de avaliação dinâmica de osciladores mecânicos que têm comportamento similar ao apresentado pelas pregas vocais.

2.1 – O sistema fonador

O processo de emissão vocal tem início no pulmão onde é formada uma corrente de ar que é expelida até as PV que, através de movimentos de vibração coordenados, modulam a passagem deste ar criando um trem de pulsos conforme se observa na Figura 2.1. Este trem de pulsos, também conhecido como sinal de fluxo glotal, é usado como fonte de excitação para o trato vocal onde o sinal de voz toma sua forma básica através de uma filtragem acústica. Esta filtragem redistribui as amplitudes das componentes em frequências do sinal de fluxo glotal gerando uma série de transformações em sua forma de onda. Em específico, é no trato vocal onde acontece a formação de uma série de picos ressonantes⁵ em frequências específicas do sinal que caracterizam o som produzido e que nos possibilitam distinguir, por exemplo, algumas vogais. A Figura 2.1 também ilustra esquematicamente a característica de resposta em frequência do trato vocal para um determinado som. Por fim, as cavidades nasais e oral realizam um processo de diferenciação no sinal vindo do trato vocal. O resultado final é a voz como a entendemos.

O fluxo glotal propaga-se através do trato vocal e é, sob o ponto de vista

Figura 2.1 - Esquema ilustrando a emissão vocal na laringe, trato vocal e cavidades nasal e oral. No esquema são indicadas as representações idealizadas no domínio do tempo e frequência do sinal primário de voz produzido na laringe pelas pregas (*i.e.*, fluxo glotal) e a saída final (*i.e.*, voz) que é o sinal irradiado pela boca e entendido como som. Ainda, a figura ilustra a resposta em frequência do trato vocal para a produção de um determinado som. Figura adaptada de (Story, 2002).

⁵ Também conhecidos como formantes.

matemático, convoluído com sua resposta impulsiva. Esta resposta, por sua vez, muda segundo a variação da área do trato vocal. Consequentemente, para cada som formado no trato existe uma correspondente configuração física deste e uma resposta impulsiva que modela o sinal. É esta característica que permite a produção da diversidade de sons que conhecemos. Conforme esboçam os gráficos da Figura 2.1, a representação espectral da voz emitida é o produto do espectro do fluxo glotal pela resposta em frequência do trato vocal.

Devido a sua importância primária, o movimento de vibração das PV provê ao trato vocal um sinal relativamente rico em harmônicos. Este sinal será moldado no trato principalmente nas bandas ressonantes que irão caracterizar o som produzido. Mesmo embora o sinal de fluxo glotal possa parecer relativamente simples, os mecanismos responsáveis por induzir uma vibração auto-sustentada nas pregas envolvem uma série de estruturas anatômicas sofisticadas e uma complexa interação entre o fluxo de ar e as pregas. Esta série de fatores são integrados para manter o movimento auto-sustentando das pregas, tornando o sistema mais complexo do que aquele revelado por uma primeira impressão.

2.2 – As pregas vocais no contexto da fonação

Na teoria básica da fonação, as PV são vistas como um oscilador não-linear que envolve uma série de aspectos aerodinâmicos, biomecânicos, fisiológicos e acústicos. Quando o ar é expelido dos pulmões a uma determinada velocidade dirigindo-se ao orifício glotal⁶, as pregas vibram de forma auto-sustentada agindo como uma válvula osciladora que interrompe o fluxo de ar modelando-o em uma série de pulsos e produzindo o sinal de fluxo glotal (U_g). Em termos mais rigorosos, este sinal pulsante representa a velocidade volumétrica média de ar que atravessa a área glotal e induz o movimento auto-sustentado de vibração das pregas e é justamente este movimento que

⁶ Constrição formada pela abertura das pregas por onde passa o ar que se dirige ao trato vocal. Também conhecida como glote.

dá ao sinal esta característica pulsante após passar pelas pregas. Consequentemente, a qualidade de uma voz e suas características fundamentais são fortemente dependentes das propriedades de vibração das pregas e seu comportamento dinâmico, pois estas têm grande influência na forma de onda de U_g . A própria frequência fundamental da voz, por exemplo, é diretamente determinada pela frequência de vibração das pregas.

Na Figura 2.2 são esboçadas as PV e algumas das estruturas periféricas presentes na laringe. A região abaixo da glote é conhecida como área sub-glotal e consiste da superfície inferior das PV e suas adjacências que são ligadas à traquéia e ao sistema brônquio-pulmonar. Acima das PV, é definida a área supra-glotal que está ligada a parte inferior do trato vocal onde estão as pregas ventriculares. Estas últimas também são frequentemente chamadas de pregas vocais falsas, pois, sob certas condições, podem ser responsáveis pela geração de alguns sons.

Apesar de sua grande importância na fonação, a principal função atribuída às PV não está relacionada à emissão vocal e sim a proteção mecânica do sistema pulmonar contra a penetração de substâncias líquidas ou sólidas. Desta maneira, as PV possuem dois modos funcionais relacionados à sua atividade primária e secundária que são, respectivamente, a respiração e fonação. Na respiração, as PV estão

Figura 2.2 - Diagrama ilustrando corte frontal da laringe. Nele são exibidas algumas das principais estruturas que têm participação durante a produção da voz.

constantemente abertas permitindo a entrada de ar nos pulmões e este modo é comumente denotado por *abdução*. Já durante a fonação, elas são aproximadas umas das outras voluntariamente fazendo com que a área glotal seja pequena ou inexistente. Neste caso, diz-se que as pregas foram postas em um movimento de *adução*.

As propriedades das PV relativas à abertura/fechamento e as tensões aplicadas a elas são determinadas por músculos presentes nestas e algumas cartilagens contidas em suas proximidades. Consequentemente, este processo passa por um controle neural. Contudo, a força que age durante a oscilação das PV não é causada pelos músculos na laringe, mas sim pela interação entre o fluxo de ar expelido dos pulmões que passa pela constrição glotal aduzida gerando forças que estarão em contato direto com os tecidos das pregas gerando seu movimento oscilatório auto-sustentando. Este movimento é fruto de uma troca de energia entre as camadas de tecidos das pregas e o escoamento de ar oriundo dos pulmões. Desta maneira, sob o ponto de vista da fonação, os músculos e o sistema neurológico da laringe têm por função básica reduzir o espaço glotal, criando uma área mínima necessária para excitar a vibração das pregas. A laringe por si é incapaz de efetuar o movimento de abrir e fechar das pregas na taxa exigida para a produção da fala que varia, em média, de 100 a 450 Hz.

No que diz respeito à constituição das PV, sabe-se que estas são compostas por um conjunto de tecidos que estão dispostos em camadas (Hirano, 1968; Hirano, 1974) conforme ilustra a Figura 2.3. A camada mais interna é predominantemente constituída por fibras musculares e por isto é conhecida como músculo *vocalis*. Já a camada intermediária é denominada de ligamento (*ligamentum vocale*) enquanto que a camada mais externa é formada por uma membrana epitelial de tecido conhecida como *mucosa*. Esta estrutura em camadas determina uma série de propriedades nos movimentos de vibração das pregas que são refletidos no sinal gerado por esta tendo efeito significativo no sinal de voz emitido.

Figura 2.3 - Vista de uma seção de uma prega vocal no plano frontal. Nela são ilustradas suas três principais camadas de tecidos que possuem diferentes propriedades e funções.

Com a adução e a presença de um fluxo de ar, as pregas são postas a vibrar. A análise de um ciclo de movimento revela a complexidade do movimento destas estruturas. A Figura 2.4 resume em nove diferentes estágios um ciclo de vibração e na Figura 2.5 é ilustrado um pulso glotal (*i.e.*, fluxo glotal U_g) gerado por este ciclo onde são caracterizadas três diferentes fases que estão vinculadas com a abertura ou fechamento das pregas.

Nas Figuras 2.4a e 2.4b, o ciclo começa com a abertura das PV devido à ação aerodinâmica da pressão sub-glotal nas paredes das pregas. Neste mesmo estágio, existe uma ação adicional em sua porção superior dando origem à onda mucosal. Esta onda - revelada em registros de alta velocidade - gera uma diferença de fase na amplitude do movimento entre as porções inferiores e superiores das pregas. Esta diferença de fase é essencial para a oscilação auto-sustentada, pois garante uma diferença na aerodinâmica entre a fase *convergente* do ciclo (Figura 2.4b a 2.4d) e a *divergente* (Figura 2.4e a 2.4g). Ainda, segundo Jiang *et al.* (2001), esta diferença de fase induzida pelas camadas de tecidos e pela porção de massa inferior e superior viabiliza a transferência de energia entre o fluxo de ar e as pregas gerando a oscilação auto-sustentada.

Sob o ponto de vista dinâmico, o movimento oscilatório das pregas é guiado pela pressão de ar dos pulmões e pode ser visto como uma troca de energia cinética e potencial causada pelo equilíbrio de forças em seus tecidos. Como bem destaca Berg (1958), estas forças estão relacionadas com fluxo de ar interagindo com as pregas, as tensões geradas pelos seus músculos, as de amortecimento e outras forças menores

Figura 2.4 - Estágios de um ciclo de vibração das pregas vocais. Os estágios (c) a (e) pertencem ao ciclo de abertura das pregas enquanto que os estágios (f) e (g) fazem parte do ciclo de fechamento. Os demais descrevem a fase fechada em diferentes circunstâncias do ciclo. As setas indicam o sentido de movimento de algumas partes das pregas. Figura adaptada de (Hirano, 1968).

Figura 2.5 - Sinal idealizado de pulso glotal ilustrando suas diferentes fases e sua relação como o movimento das pregas.

oriundas do ciclo de oscilação. No caso específico das interações entre o fluxo de ar e as PV, alguns autores pioneiros na área (Berg *et al.*, 1957; Ishizaka & Flanagan, 1972; Titze, 1973) originalmente utilizaram as leis de Bernoulli para descrever esta relação aerodinâmica na laringe. Contudo, alguns pesquisadores (Xue *et al.*, 2010) têm chamado a atenção para o emprego de Bernoulli nos modelos mecânicos justificando que suas equações modelam com baixo nível de fidelidade, ou até mesmo nenhum, os efeitos de viscosidade do ar, a dinâmica de vórtices na laringe, a variação temporal do fluxo de ar e seu gradiente na direção lateral. Em particular, este mesmo fluxo de ar pode exibir uma assimetria bastante significativa em sua distribuição devido às deflexões de seus jatos⁷ que podem refletir em forças aerodinâmicas assimétricas atuando nas PV. Atento a estas questões, Decker e Thomson (2007) tentaram utilizar abordagens alternativas baseadas nas equações de Navier-Stokkes para descrever com maiores detalhes esta ação aerodinâmica.

Finalmente, a aplicação dos princípios de Bernoulli⁸ gera uma redução da pressão aplicada na porção inferior das pregas e um consequente aumento na velocidade das partículas que passam por esta constrição. Esta queda de pressão resulta em um decréscimo nas forças que agem nas PV fazendo com que estas possam retornar gradativamente ao seu estado inicial produzindo um novo ciclo. A própria elasticidade dos tecidos das pregas também contribui para que estas voltem à posição inicial do ciclo.

2.3 – Biomecânica básica das pregas vocais

Segundo a teoria aerodinâmica-mioelástica das PV (Berg, 1958), a frequência fundamental de vibração das pregas é determinada por um número de fatores interdependentes que incluem (i) a massa, a (ii) viscoelasticidade das pregas e (iii) a pressão sub-glotal. De acordo com esta teoria, as pregas são aduzidas, contraídas e

⁷ Este efeito aerodinâmico é conhecido como “Coanda”.

⁸ No caso das pregas vocais, é considerando um fluxo de ar constante através da constrição glotal.

tensionadas durante a fonação para regular sua elasticidade. A coordenação entre a pressão sub-glotal e a viscoelasticidade das pregas é fundamental para regular o sinal produzido. Além da viscoelasticidade e da tensão das pregas, o processo de controle neuromuscular pode ajustar a configuração da constrição glotal que é também um elemento chave para determinar as forças aerodinâmicas que irão agir durante o ciclo glotal.

Neste sentido, a massa e tensão *efetivas* são fatores importantíssimos na modulação do sinal primário da voz. Sendo assim, as propriedades físicas fundamentais das PV mais relevantes estão ligadas a sua (i) massa efetiva, (ii) rigidez⁹ e (iii) viscosidade.

No que diz respeito à contribuição da(s) massa(s), deve-se destacar que considera-se apenas seu valor efetivo que participa diretamente do processo de modulação. Através do posicionamento das cartilagens e músculos na laringe é possível remodelar a geometria das pregas e a área com que estas entrarão em contato direto com o fluxo de ar e assim modificar a massa efetiva que fará parte deste oscilador. Como exemplificam Jiang *et al.* (2000), um estiramento longitudinal das pregas decrementa a massa efetiva e pode com isto incrementar a frequência fundamental de oscilação. Por outro lado, a contração de fibras musculares pode incrementar a quantidade de massa efetiva e contribuir para um decremento nesta frequência fundamental. Deste modo, as contrações musculares e outros fatores podem afetar a concentração relativa de massa. Estes complexos mecanismos sugerem que pode haver diversos padrões naturais de vibração determinados pela predominância relativa da massa, elasticidade ou pressão assim bem como a variação conjunta destas.

A segunda propriedade física fundamental das PV, *i.e.*, a rigidez, está intimamente ligada à tensão mecânica das pregas. A habilidade dos músculos laríngeos

⁹ Neste texto o termo rigidez será empregado para determinar a medida de resistência oferecida por um corpo elástico a uma deformação ou força.

em alterar as dimensões físicas das pregas (como o seu comprimento, por exemplo) através de contrações ou alongamentos, pode contribuir para a mudança da sua tensão assim como sua massa efetiva, conforme descrito no parágrafo anterior. Esta rigidez está vinculada à resistência que os tecidos das pregas terão quando as forças aerodinâmicas e de impacto atuarem sobre elas. Em modelos mecânicos, geralmente emprega-se uma mola para representar este efeito de resistência a uma determinada força. Já a viscosidade dos tecidos das PV está relacionada com sua propriedade de se opor a quantidade de movimento do fluxo de ar que a excita e também está ligada a própria viscosidade de suas secreções e muco além de fluídos internos aos seus tecidos. Sabe-se que quanto maior a viscosidade, maior deverá ser a força aerodinâmica gerada pela pressão sub-glotal para manter as mesmas características de vibração. Sustenta-se que a própria hidratação das pregas tenha efeito considerável nesta variável (Finkelhor *et al.*, 1998) que também está intimamente relacionada à massa e tensão das pregas.

2.4 – Métodos empregados na análise das pregas vocais

Como mencionado anteriormente, a produção primária da voz tem origem na laringe através das PV. Devido à importância das pregas na emissão vocal, foram desenvolvidas diferentes tecnologias para avaliar sua atividade e registrar seu comportamento para análises clínicas e científicas diversas. Dentre estas tecnologias, destacam-se a análise acústica associada à filtragem inversa (Gómez-Vilda *et al.*, 2009; Pinheiro *et al.*, 2010), a eletroglotografia (Roubeau *et al.*, 2009) e a videolaringoscopia (Citardi *et al.*, 1996).

A análise acústica utiliza registros de vozes captadas por um microfone que são avaliados por métodos computacionais e profissionais treinados. No caso da avaliação dos pulsos glotais produzidos na laringe, tradicionalmente emprega-se algum método de filtragem inversa (Alku *et al.*, 1998) que desfaz os efeitos do trato vocal, reproduzindo o sinal primário de voz originalmente encontrado na laringe. Como o trato vocal executa

um processo de filtragem linear, este método de filtragem inversa é direto. Ao final, é reproduzido o sinal de fluxo glotal U_g gerado pelo indivíduo que teve seus registros acústicos captados por um microfone. A análise da forma de onda de U_g reconstruído revela algumas características do comportamento das PV.

A eletroglotografia (EGG) utiliza dois eletrodos de superfície fixados em lados opostos do pescoço próximo à traquéia para medir a impedância elétrica que muda com a área de contato das PV. Mesmo embora seja um dos métodos menos invasivos, a EGG tem como desvantagem o fato de apresentar distorções naturais no sinal captado pelos eletrodos. Isto acontece porque a pele do pescoço, onde estão fixados os eletrodos, tem o efeito de um filtro passa-baixas que distorce o sinal registrado pelos eletrodos.

No estudo científico da voz, os sinais reconstruídos pela filtragem inversa ou pela EGG não podem registrar diretamente os movimentos das PV ou prover informação visual direta sobre seus padrões de vibração. Neste sentido, as imagens médicas da laringe realizadas por um exame de videolaringoscopia têm sido uma importante ferramenta na medida em que permite a análise visual e quantitativa do padrão de vibração das PV. Basicamente, o exame de videolaringoscopia é feito por meio de um endoscópio conectado a uma câmera e inserido até a região da orofaringe através da via oral ou nasal. A Figura 2.6 ilustra este procedimento.

Contudo, registrar este tipo de movimento exige um sistema de câmeras

Figura 2.6 - Exame de videolaringoscopia da laringe. Esquema ilustrando um endoscópio rígido, conectado a uma câmera e inserido até a orofaringe para registrar os movimentos das pregas vocais. No esquema também é ilustrado uma imagem onde são indicadas as pregas e a área glotal formada.

sofisticado. Usualmente, os sistemas de videolaringoscopia utilizam câmeras e um sistema estroboscópico que captam o movimento a uma taxa de 24 a 30 imagens por segundo. Este tipo de câmera não registra o real movimento das pregas e por isto é aplicável somente a vibrações periódicas, impedindo frequentemente a caracterização e descrição de padrões vibratórios aperiódicos típicos em vozes patológicas. Já os sistemas que utilizam câmeras de alta velocidade são capazes de amostrar o movimento a uma taxa que varia de 2.000 a 4.000 imagens por segundo, sendo por isto capazes de registrar o ciclo completo das PV com maior riqueza de detalhes.

Independentemente da taxa de amostragem, o registro visual do movimento das PV através de imagens é capaz de prover importantes indícios do comportamento dinâmico das PV. Por este motivo, a quantização do movimento a partir das imagens de videolaringoscopia pode ser usada como importante fonte de subsídio de informações para o modelamento das pregas. Além disto, outras informações extraídas das imagens da laringe também têm se mostrado importantes no estudo da produção da voz. A análise da área da glote, por exemplo, tem frequentemente sido usada para caracterizar as vibrações das PV para diferentes condições de vozes (Ahmad *et al.*, 2005; Yan *et al.*, 2007). Outra derivação da videolaringoscopia, conhecida como quimografia (Mul *et al.*, 2009), também utiliza câmeras específicas para registrar o movimento de uma pequena porção das PV.

2.5 – Modelos de pregas vocais

No decorrer das últimas décadas, inúmeros autores modelaram matematicamente o comportamento vibracional das PV. Embora fossem variados os propósitos, estes modelos foram frequentemente utilizados na síntese de voz (Ishizaka & Flanagan, 1972), simulação de fenômenos físicos na laringe (Steinecke & Herzel, 1995) e, principalmente, aplicações biomédicas voltadas ao estudo clínico e científico da voz seja ela saudável ou patológica (Story & Titze, 1995).

Diante da complexidade das interações e fenômenos envolvidos na produção da voz, os modelos foram ao decorrer dos anos sofrendo novas adaptações para incrementar seu realismo. Uma das primeiras iniciativas de modelar o comportamento vibracional das PV foi feita em (Flanagan & Cherry, 1969) por meio da aplicação das leis Newtonianas a um conjunto de massas, molas e amortecedores que quando conectados entre si, formavam um oscilador mecânico que simulava as vibrações das PV com um único grau de liberdade. Anteriormente, Flanagan e Landgraf (1968) sugeriram um modelo mecânico utilizando os mesmos elementos mecânicos e as leis de Bernoulli para descrever a ação aerodinâmica nas pregas.

O modelamento mecânico das pregas deu um grande salto de maturidade no trabalho clássico de Ishizaka e Flanagan (1972) que utilizaram os modelos descritos anteriormente e incorporaram um grau de liberdade a mais com a adição de mais duas massas ao modelo. Com esta adição, este modelo passou a ter duas massas (*i.e.*, dois graus de liberdade) representando cada prega e por isto ficou conhecido na literatura como “modelo de duas massas”. O propósito que levou os autores a incluírem esta massa foi simular a produção da onda mucosal que gera a diferença de fase entre a porção inferior e superior das pregas e é fundamental para a oscilação auto-sustentada. Além disto, utilizaram os resultados de experimentos práticos envolvendo laringes excisadas para adaptar as equações do modelo e definir com maior precisão suas constantes. Por estas e outras características, este modelo tem servido até o presente momento como forte referência para outros modelos e adaptações e por isto é ainda frequentemente usado ou referenciado na literatura científica. Dada sua importância, este será referenciado no texto como **IF72**.

Childers *et al.* (1986) propõem um modelo, baseado no IF72, cujo objetivo fundamental é produzir uma “borda” suave entre a porção superior e inferior das PV de modo que não haja nenhuma descontinuidade na distribuição de pressão na glote, o que não acontece nos modelos precursores. Esta abordagem permite que as forças externas

que agem sobre as pregas possam ser calculadas analiticamente e sem descontinuidades. Esta ideia de se fazer um modelo “suave” sem transições abruptas foi também posteriormente implementada em (Lous *et al.*, 1998) que também se basearam no IF72.

Em meados da década de 90, com o florescimento dos métodos da dinâmica não-linear de sistemas aplicados à análise do comportamento das PV, surgiu a necessidade de se ter um modelo mais simples que tivesse um custo computacional mais baixo já que as equações do IF72 eram computacionalmente onerosas. Isto aconteceu porque estes novos métodos de análise requeriam com frequência simulações extensivas destes modelos mecânicos. Neste contexto, Steinecke e Herzel (1995) desenvolveram uma simplificação do IF72 que ficou bastante conhecida na literatura e continua sendo frequentemente empregado nos dias de hoje. Em função de seu largo emprego, este modelo será referenciado neste texto como **SH95**.

Outros modelos propuseram adaptações bem específicas como, por exemplo, o empregado em (Zhang & Jiang, 2004) para estudar o efeito de pólipos nas pregas. Neste modelo de duas massas, baseado no SH95, foi inserido uma terceira massa em apenas uma das pregas com o propósito de simular o comportamento de um pólipo durante a fonação.

Novas extensões empregando um maior grau de liberdade também foram propostas. Em particular, o modelo proposto em (Titze, 1973; Titze, 1974) descreve uma extensão do IF72 com oito massas. Outro trabalho (Yang *et al.*, 2010) utiliza uma versão com cinco massas baseando-se nas equações simplificadas do SH95. Uma abordagem similar, empregando um modelo multi-massas, foi também proposta em (Schwarz *et al.*, 2008). Contudo, dentre os modelos quem empregaram três ou mais massas, o de Story e Titze (1995) obteve grande expressão na literatura científica produzindo uma versão baseada em três massas. A grosso modo, cada uma das massas representava as três camadas de tecidos que compõem a estrutura fisiológica das PV.

Devido a esta característica, este modelo é algumas vezes conhecido como “modelo corpo-cobertura” e será referenciado neste texto como **ST95**.

Por fim, outros trabalhos (Alipour *et al.*, 2000; Oliveira & Pereira, 2008) vêm utilizando novas técnicas baseando-se principalmente em métodos de elementos finitos, sendo esta uma abordagem consideravelmente diferente da clássica que se baseia em massas, molas e amortecedores interligados. Este novo paradigma tem sido utilizado para analisar os princípios acústicos da voz na laringe, fenômenos no fluxo de ar e sua interação com as pregas e outras estruturas presentes na laringe. Outros modelos tridimensionais baseados em elementos finitos foram empregados para simular deformações físicas nas PV (Gunter, 2003; Tao & Jiang, 2008), análise de colisão e elasticidade durante o ciclo de abertura e fechamento.

A vantagem dos modelos que empregam métodos de elementos finitos reside na sua natureza tridimensional e elementar da representação da estrutura como pequenas unidades interligadas, oferecendo uma riqueza de detalhes não vista em outros métodos. Além disto, o modelo tem vários graus de liberdade e permite incorporar melhor as propriedades geométricas das PV. Já suas desvantagens estão associadas à sua complexidade e custo computacional além de exigir um maior número de parâmetros que pode dificultar sua análise dinâmica ou aplicação em métodos de otimização de parâmetros mecânicos.

Na sequência são estabelecidos os princípios fundamentais do modelamento mecânico das PV empregando duas e três massas. Por essencialmente serem a base para outros modelos com maior número de elementos, pode-se estabelecer analogia a partir destes casos. Ainda, são brevemente descritos os modelos IF72 e SH95, que são as duas principais abordagens do modelamento com duas massas. Por fim, o modelo corpo-cobertura (ou ST95) também é abordado com o propósito de representar o modelamento com três massas.

2.5.1 - Modelamento mecânico com duas e três massas

Considerando que as PV têm um comportamento similar ao de um oscilador mecânico, as mesmas podem ser representadas por elementos mecânicos baseados em um conjunto de massas, molas e amortecedores que irão reproduzir suas oscilações e ao mesmo tempo incorporar as características (bio)mecânicas básicas de seus tecidos através de suas constantes. Tendo isto em vista, os primeiros modelos utilizaram-se destes elementos mecânicos pra descrever o comportamento das pregas. A Figura 2.7 ilustra dois diferentes modelos de PV usando duas e três massas para representar o comportamento de cada prega.

Figura 2.7 - (a) Modelo mecânico de pregas vocais com duas e (b) três massas. Nestes modelos, as massas são acopladas por uma mola k_c e podem deslocar-se apenas no plano horizontal. Vista do plano coronal do modelo.

Tabela 2. 1 – Principais constantes mecânicas do modelo de duas massas.

Constante	Descrição
P_s	Pressão sub-glotal que age nas pregas transferindo energia para as mesmas.
m_1	Massa da porção inferior das pregas.
m_2	Massa da porção superior das pregas.
k_1	Mola representando a tensão dos tecidos das pregas em sua região inferior.
k_2	Mola representando a tensão dos tecidos das pregas em sua região superior.
k_c	Mola de acoplamento entre massas.
r_1	Constante de amortecimento da região inferior.
r_2	Constante de amortecimento da região superior.
L_g	Comprimento da abertura (fenda) da prega vocal.
d_1	Tamanho de m_1 .
d_2	Tamanho de m_2 .

Nos modelos de duas e três massas, cada massa de uma prega indica um grau de liberdade. No caso da Figura 2.7(a), é assumido que as PV sejam representas por um par de osciladores acoplados onde cada um deles é composto por duas massas, três molas e dois amortecedores. A inclusão da massa superior foi feita com o propósito de se reproduzir a onda mucosal e reproduzir uma diferença de fase entre o movimento da porção superior e inferior das pregas. Observe que os movimentos destas massas estão interligados por uma mola k_c . As molas k_1 e k_2 que ligam as massas à parede da laringe representam a tensão nos músculos e os amortecedores estão intimamente ligados as suas perdas viscosas. A Tabela 2.1 fornece as principais constantes do modelo assim como seu significado básico. Analogia pode ser feita para o modelo de três massas visto na Figura 2.7(b).

Utilizando as leis de Newton, estabelecem-se as equações gerais do modelo de duas massas como:

$$\begin{aligned} m_1\ddot{x}_1 + r_1\dot{x}_1 + k_c(x_1 - x_2) + k_1x_1 + I_1(x_1) + N_1(x_1) &= F_1(x_1, x_2); \\ m_2\ddot{x}_2 + r_2\dot{x}_2 + k_c(x_2 - x_1) + k_2x_2 + I_2(x_2) + N_2(x_2) &= F_2(x_1, x_2). \end{aligned} \quad (2.1)$$

Considerando que x indica o movimento das massas e o índice i a porção inferior ($i=1$) ou superior ($i=2$) das pregas, temos três forças que aparecem sob a forma de funções: $F_i(x_i)$, $I_i(x_i)$ e $N_i(x_i)$. A primeira delas refere-se à interação aerodinâmica do fluxo de ar com as PV e é responsável por introduzir energia no sistema. Já $I_i(x_i)$ representa o impacto entre as massas (esquerda e direita) e, por fim, $N_i(x_i)$ inclui as possíveis não-linearidades adicionais do modelo¹⁰. Todos estes termos vistos em (2.1) moldam as leis básicas de um modelo de duas massas. Modelos de duas massas específicos, como o IF72 ou SH95, definem de maneira própria e particular as equações de $F_i(x_1, x_2)$, $I_i(x_i)$ e $N_i(x_i)$ e por isto apresentam diferenças significativas entre seus métodos mesmo embora utilizem exatamente os mesmos elementos mecânicos representados na Figura 2.7(a) e também sejam regidos por (2.1). A Figura 2.8 exemplifica um movimento

¹⁰ O próprio impacto entre as pregas geralmente apresenta caráter não-linear e não tem relação direta com a função N_i . Esta última deve incorporar outras funções não-lineares extras do modelo.

Figura 2.8 - Exemplo de sinais gerados por um modelo de duas massas. Figura à esquerda: deslocamento das massas m_1 (linha pontilhada) e m_2 (linha contínua). Nesta figura é possível ver a diferença de fase entre o movimento destas massas. Figura à direita: fluxo glotal U_g gerado. Sinais gerados pelo IF72 usando seus parâmetros padrões.

padrão de deslocamento das pregas em sua porção superior e inferior assim como o sinal de fluxo glotal U_g produzido por este deslocamento.

Apesar de ser uma aproximação capaz de reproduzir aproximadamente 95% dos movimentos das PV (Berry *et al.*, 1994), sua limitação está relacionada à sua incapacidade de definir de maneira individual os efeitos das três camadas de tecidos das pregas. Para atender a esta questão, foi elaborado outro modelo envolvendo três massas (ver Figura 2.7(b)) onde cada uma destas deve representar de forma aproximada os três diferentes tecidos que compõem a estrutura fisiológica das pregas. Em consequência desta estrutura, o modelo tem uma variedade de parâmetros que podem ser alterados com o intuito de simular diferentes condições produzidas por vários ajustes laringais ou mesmo variações patológicas.

A exemplo do modelo anterior, o de três massas utiliza as leis de Newton para descrever a dinâmica e cinemática do sistema. A equação (2.2) apresenta de maneira resumida estas equações do modelo.

$$\begin{aligned}
 m_u \ddot{x}_u + r_u(\dot{x}_u - \dot{x}_b) + k_c(x_u - x_L) + k_u(x_u - x_b) + I_u(x_u) + N_u(x_u) &= F_u(x_u, x_L, x_b); \\
 m_L \ddot{x}_L + r_L(\dot{x}_L - \dot{x}_b) + k_c(x_L - x_u) + k_L(x_L - x_b) + I_L(x_L) + N_L(x_L) &= F_L(x_u, x_L, x_b); \\
 m_b \ddot{x}_b &= k_b x_b + r_b \dot{x}_b + N_b(x_b) - (k_u x_u + N_u(I_u) + r_u \dot{x}_u + k_L x_L + N_L(x_L) + r_L \dot{x}_L).
 \end{aligned} \tag{2.2}$$

Os índices ‘ u ’ e ‘ L ’ referem-se, respectivamente, a porção superior e inferior do modelo enquanto que ‘ b ’ indica a massa que representa a sua camada da base. A variável j será empregada para designar estes índices. Novamente $N_j(x_j)$, $I_j(x_j)$ e $F_j(x_j)$

indicam, respectivamente, as não-linearidades a serem incorporadas ao modelo, o impacto entre as pregas e as forças aerodinâmicas. No caso da camada base, esta não sofre ação direta do impacto e nem de interação aerodinâmica.

Nas seções subsequentes são apresentadas as principais características dos modelos IF72, SH95 e ST95.

2.5.2 - Modelo de Ishizaka e Flanagan

Baseando-se nas propriedades mioelásticas e aerodinâmicas das PV, Ishizaka e Flanagan (1972) desenvolveram um modelo mecânico de duas massas capaz de simular seus principais movimentos vibracionais e propriedades dinâmicas. Embora tenha sido originalmente desenvolvido para aplicações relacionadas à síntese de voz, o IF72 tem sido amplamente utilizado para outros propósitos e ajudado a entender os princípios acústicos fundamentais da geração da voz e fatores que têm influência neste processo. Segundo os autores do modelo, ele apresenta bom realismo na simulação do sinal produzido na laringe.

Em relação aos outros modelos de duas massas, o IF72 apresenta como diferencial um equacionamento de $F_i(x_1, x_2)$, $I_i(x_i)$ e $N_i(x_i)$ mais detalhado e leva em consideração várias não-linearidades associadas aos tecidos das pregas, impacto e forças aerodinâmicas. Em particular, parte destas não-linearidades foram modeladas para adequar a relação entre a deflexão (a partir do ponto de equilíbrio) e a força necessária para produzir esta deflexão. Para isto, são criados dois coeficientes de não-linearidades η_k e η_h associados, respectivamente, aos tecidos das pregas e ao impacto. Estes estão vinculados nas funções $I_i(x_i)$ e $N_i(x_i)$ como,

$$\begin{aligned} N_i(x_i) &= k_i \eta_{ki} \cdot x_i^3 ; \\ I_i(x_i) &= 3k_i \left(x_i + \frac{A_{g0i}}{2L_g} \right) \left\{ 1 + \eta_{hi} \left(x_i + \frac{A_{g0i}}{2L_g} \right)^2 \right\} \end{aligned} \quad (2.3)$$

Na equação anterior, o índice i novamente é usado para associar à porção superior ($i=2$) e inferior ($i=1$) das PV e a variável x_i denota o deslocamento destas

massas. Ainda, a constante A_{g0i} designa a área de descanso¹¹ da prega. Outra característica incorporada ao modelo diz respeito à colisão entre a superfície das PV. Esta colisão ocorre durante a fase de fechamento e causa uma deformação nas massas que é de grande importância na dinâmica das pregas. Esta força de deformação (ou colisão) é representada por uma mola “invisível”. Nos momentos do ciclo do pulso glotal onde não há impacto (*i.e.*, $x_i > A_{g0i}/(2L_g)$), a função $I_i(x_i)$ é nula.

Quanto ao amortecimento r_i do modelo, este é definido pela Equação (2.4) e por sua constante de razão de amortecimento ζ_i . Vale destacar que durante o ciclo de abertura e fechamento, a razão de amortecimento assume diferentes valores.

$$r_i = 2\zeta_i\sqrt{m_i k_i} \quad (2.4)$$

Já no que diz respeito às forças de interação aerodinâmica $F_i(x_1, x_2)$, estas são distribuídas nas paredes das pregas e por isto deve-se levar em consideração o tamanho da abertura da fenda glótica (constante L_g) e o comprimento das massas (constantes d_i). Deste modo, a área exposta à ação aerodinâmica do sistema com o fluxo de ar é dada pela razão $L_g d_i$. Tendo isto em mente, as forças F_1 e F_2 agindo, respectivamente, em m_1 e m_2 são dadas por:

$$\begin{aligned} F_1(x_1, x_2) &= L_g d_1 \left[P_s - 1.37 \frac{\rho}{2} \left(\frac{U_g}{A_{g1}} \right)^2 - 0.5(R_{v1}U_g + L_{g1}\dot{U}_g) \right]; \\ F_2(x_1, x_2) &= L_g d_2 \left[F_1 - 0.5\{(R_{v1} + R_{v2})U_g + (L_{g1} + L_{g2})\dot{U}_g\} - \frac{\rho}{2} U_g^2 \left(\frac{1}{A_{g2}^2} - \frac{1}{A_{g1}^2} \right) \right] \end{aligned} \quad (2.5)$$

Nestas equações são vistas uma série de variáveis que estão vinculadas com os princípios de Bernoulli, resistências e perdas acústicas. Mais informações sobre as mesmas podem ser consultadas em (Ishizaka & Flanagan, 1972).

¹¹ Considera-se que sob a ausência de nenhuma força aerodinâmica, a prega vocal terá uma área de abertura (fenda glótica) que é chamada de “área de descanso”.

2.5.3 - Modelo de Steinecke e Herzel

Após o surgimento do modelo de duas massas de Ishizaka e Flanagan, outros baseados no IF72 foram propostos. Dentre eles, o sugerido por Steinecke e Herzel (1995) merece especial destaque por ter sido amplamente utilizado em outras pesquisas.

O SH95 teve como contribuição principal a simplificação das equações aerodinâmicas, de impacto e de cálculo de fluxo glotal do IF72. Em consequência disto, ele é assumidamente uma simplificação do IF72, tido como complexo e computacionalmente dispendioso. Basicamente, os autores deste modelo fizeram três considerações que conduziram a uma significativa simplificação¹². São elas: (i) as não-linearidades cúbicas podem ser negligenciadas para baixas amplitudes de vibração; (ii) as interações do pulso glotal com ressonâncias sub e supra-glotaes também podem ser negligenciadas¹³ e (iii) o fluxo de Bernoulli age principalmente nas porções inferiores das pregas sendo desconsiderado sua ação na região superior. Por estas razões, neste modelo $F_2(x_1, x_2)$ e $N_i(x_i)$ são nulos.

As simplificações propostas deram a muitos pesquisadores a possibilidade de simulação extensiva da dinâmica das pregas principalmente em situações onde se deseja avaliar a estabilidade do sistema utilizando, por exemplo, diagramas de bifurcações. Em especial, o próprio autor do modelo adaptado o utilizou com sucesso para levantar os seus diagramas de bifurcação e regiões de estabilidade utilizando as constantes padrões do modelo. As equações gerais do modelo são apresentadas como

$$\begin{aligned} I_i(x_i) &= 3k_i \frac{a_i}{2L_g}; \\ F_1(x_i) &= \frac{L_g d_1 P_s}{m_1} \left[1 - \Omega(a_{min}) \left(\frac{a_{min}}{a_1} \right) \right] \Omega(a_1); \end{aligned} \tag{2.6}$$

onde, para um modelo simétrico, são válidas as relações de (2.7).

¹² Estas simplificações foram baseadas nas descobertas discutidas em (Pelorson, 1994).

¹³ Isto acontece porque considera-se que a pressão subglotal é constante e a supra-glotal é zero.

$$\begin{aligned}
\Omega(x) &= \begin{cases} \tanh[50(x/A_{g01})], & x > 0 \\ 0, & x \leq 0; \end{cases} \\
a_i &= A_{g0i} + 2L_g(x_i); \\
a_{min} &= \min(a_1, a_2).
\end{aligned} \tag{2.7}$$

O termo $I_i(x_i)$ só assume a forma de (2.6) quando há impacto entre as pregas. Em outros casos, $I_i(x_i)$ é nulo.

2.5.4 - Modelo “corpo-cobertura”

Buscando atender a propósitos de realismo fisiológico maiores, Story e Titze (1995) produziram um modelo baseado em três massas que foi projetado para ser utilizado em pesquisas que visam simular condições normais e patológicas de voz mantendo o realismo fisiológico. As constantes empregadas no modelo foram baseadas em valores medidos em laboratório a partir de PV humanas excisadas.

Considerando x_j como a variável que designa os deslocamentos das massas e j para representar a porção superior ($j=u$) e inferior ($j=L$) das PV (Figura 2.7(b)), temos que:

$$\begin{aligned}
N_j(x_j) &= k_j \eta_{kj} (x_j - x_b)^3; \\
I_j(x_j) &= 3k_j (x_j + \eta_{hj} x_j^3).
\end{aligned} \tag{2.8}$$

A exemplo do IF72, η_{kj} e η_{hj} referem-se a coeficientes de não-linearidade ligados, respectivamente, às molas e ao impacto entre as massas. Destaca-se que para o caso onde $j = b$, vale a relação $N_b(x_b) = I_b(x_b) = 0$.

De forma similar ao IF72, as forças geradas através da interação entre o fluxo de ar e as paredes das pregas na sua porção superior (F_u) e inferior (F_L) são calculadas baseadas em dois diferentes cenários que são descritos na Figura 2.9. No ST95, existem somente dois pontos na glote onde a pressão deve ser calculada e estes estão posicionados nas massas superiores e inferiores e são designados, respectivamente, por P_u e P_L .

Para a configuração ilustrada na Figura 2.9(a), a pressão na região inferior da glote é dada por

$$P_L = P_S - (P_S - P_i)(a_u/a_L)^2. \quad (2.9)$$

Os termos a_u e a_L referem-se à área formada entre as pregas na sua porção superior e inferior, respectivamente. A constante P_S indica a pressão supra-glotal presente na entrada do trato vocal. Para o segundo cenário exibido na Figura 2.9(b), a pressão superior e inferior são ambas dadas pela relação (2.10) onde P_i é a pressão presente no trato vocal.

$$P_u = P_L = P_i. \quad (2.10)$$

As forças exercidas em cada uma das massas são enunciadas como $F_u = L_g d_2 P_u$ e $F_L = L_g d_1 P_L$. Aqui, mais uma vez, L_g indica o comprimento da abertura da glote e as constantes d_1 e d_2 remetem às dimensões das massas que estão em contato com o fluxo de ar. Em Titze (2002) são propostas algumas atualizações no modelo.

2.6 – Análise da voz utilizando métodos da dinâmica não-linear

A dinâmica de sistemas trata da modelagem matemática de um determinado sistema e da análise de sua resposta a partir de um conjunto de variáveis de entrada. Por ser uma área ampla, é frequentemente usada para modelar fenômenos biológicos, logísticos, químicos, climáticos, econômicos, sociais e mecânicos na forma de um sistema regido por equações que tentam traduzir seu comportamento. No caso de sistemas mecânicos, as leis da mecânica clássica são usualmente empregadas para

Figura 2.9 - Padrão assumido pelo modelo de fluxo de ar atravessando a glote. (a) Região de Bernoulli na porção inferior e região de jato na porção superior. (b) Região de jato em ambas as porções das pregas vocais.

avaliar as forças envolvidas no modelo (*e.g.*, sistema) e os movimentos produzidos por este. Ainda, técnicas de análise de sinais são frequentemente empregadas para desvendar as características destes sistemas e inferir conhecimento sobre os mesmos considerando sua evolução temporal e as relações entre suas entradas e saídas.

No estudo da voz, tradicionalmente sinais de vozes são frequentemente avaliados em termos de *jitter*, *shimmer* e avaliação espectral das componentes do sinal. Em particular, estes dois primeiros parâmetros buscam quantizar as perturbações de um sinal periódico a partir da sua variação em frequência e amplitude ciclo a ciclo. Mesmo embora haja controvérsia no seu emprego devido a sua “superficialidade” na análise, são importantes indicadores dado seu uso extensivo na prática clínica. Contudo, eles são reconhecidamente falhos na análise de determinados padrões de voz como, por exemplo, aquelas que são patológicas ou apresentam consideráveis irregularidades (Zhang & Jiang, 2003) ou até mesmo em casos normais onde, sob um ponto de vista mais criterioso, não são capazes de definir as características básicas do sistema que produz o sinal e sim apenas suas perturbações que podem ser fruto de fatores diversificados.

Neste contexto, técnicas de análise dinâmica não-linear têm sido empregadas no estudo da voz para descrever e caracterizar por meio de parâmetros próprios desta área, diferentes sinais de vozes e o próprio comportamento dinâmico da voz na laringe. Esta metodologia traz uma visão dinâmica voltada mais ao sistema do que ao sinal produzido por ele, sua interação com fatores internos e externos e a relação de entrada (excitação) e saída produzida. Adicionalmente, o fato de se usar técnicas não-lineares possibilita incorporar este tipo de característica às análises trazendo uma visão mais profunda e completa do sistema que produz a voz.

Independentemente de saudáveis ou patológicas, as PV são um sistema essencialmente não-linear (Giovanni *et al.*, 1999) e quando estudadas sob a ótica puramente linear, características importantes de seu comportamento são suprimidas.

Por exemplo, a própria transformada de Fourier deve, em determinadas circunstâncias, ter sua aplicação limitada, pois utiliza como base uma superposição linear de funções trigonométricas. Em consequência disto, a transformada emprega componentes em frequências inexistentes no sistema para tentar representar o sinal produzido por este (Worden & Tomlinson, 2001). Tendo em vista as potenciais vantagens trazidas pela dinâmica, em especial a não-linear, inúmeros pesquisadores vêm explorando a aplicação desta técnica ao estudo da voz e com isto produzindo achados relevantes, consolidando-a como uma importante ferramenta na pesquisa científica (Jiang *et al.*, 2006).

Pesquisas utilizando PV excisadas e análises de sinais de voz têm confirmado a existência de comportamento não-linear das PV e até mesmo, em alguns casos, comportamento caótico (Berry *et al.*, 1996; Giovanni *et al.*, 1999; Jiang & Zhang, 2002; Zhang & Jiang, 2004; Tokuda & Herzel, 2005; Zhang & Jiang, 2005; Jiang *et al.*, 2006; Tokuda *et al.*, 2008). Tais observações são importantes, pois, justificam (em parte) o uso de métodos da dinâmica não-linear para tentar elucidar os complexos padrões vibracionais associados às PV e desvendar fenômenos envolvidos na produção da voz em termos de modelos dinâmicos não-lineares. Tal aplicação pode, idealmente, ser usada tanto na análise de sujeitos normais quanto patológicos. Na sequência são brevemente descritos alguns métodos da dinâmica não-linear de sistemas que podem ser usados para extrair valiosas informações sobre a dinâmica das PV.

2.6.1 – A dinâmica não-linear de sistemas

Na tentativa de estudar fenômenos climáticos através de um modelo dinâmico, Lorenz (1963) observou que alguns sinais complexos e imprevisíveis¹⁴ podem ser gerados por sistemas puramente determinísticos e de baixa dimensão¹⁵ (May, 1976). Até então,

¹⁴ Neste sentido, imprevisível refere-se a sinais que não podem ser preditos.

¹⁵ Segundo Kantz e Schreiber (2004), sistemas com duas ou mais dimensões estão sujeitos a comportamentos imprevisíveis.

acreditava-se que sistemas puramente determinísticos (como aqueles utilizados para modelar mecanicamente as PV) poderiam ser completamente especificados por suas condições iniciais e leis da dinâmica. Contudo, estes sistemas não-lineares determinísticos, quando submetidos a pequenos erros ou pequenas variações em suas condições iniciais, podem apresentar grandes desvios de comportamento. Isto porque as não-linearidades presentes no sistema poderiam induzir estas pequenas variações a grandes crescimentos exponenciais no tempo. Nestas circunstâncias, estas incertezas nas medidas das condições iniciais podem levar o sistema a outros estados completamente diferentes daqueles esperados, dificultando sua previsibilidade. Tais fenômenos não são observados em sistemas lineares onde o princípio da superposição pode ser aplicado. Devido a esta questão, a dinâmica de sistemas é tradicionalmente dividida em sistemas lineares, baseados no princípio da superposição, e sistemas dinâmicos não-lineares (Kaplan & Glass, 1995).

Neste cenário, sistemas que são determinísticos e podem apresentar comportamentos imprevisíveis e pseudo-randômicos são chamados de caóticos. Naturalmente, apenas sistemas não-lineares podem ser caóticos, pois esta caoticidade é fruto da não-linearidade (Abarbanel, 1995). As próximas subseções ilustram alguns dos métodos da dinâmica não-linear de sistemas que têm tido grande visibilidade e/ou potencialidade para desenvolvimento na área a que se dedica esta tese. Parte deles são empregados nos próximos capítulos deste trabalho com o objetivo de estudar as características dinâmicas das PV usando modelos determinísticos.

Análise em espaço de fase

Tradicionalmente, uma série temporal pode ser representada no domínio do tempo ou da frequência. Diferentemente desta visão convencional, uma série temporal pode ser representada na dinâmica não-linear através de uma relação matemática de imersão apresentada na equação (2.11) que representa este sinal em um espaço alternativo de dimensão m . Esta equação definirá as coordenadas da série temporal

neste espaço m -dimensional. Ainda, o valor de m deve ser calculado por um método específico, como o apresentando em (Kennel *et al.*, 1992).

$$x(t) = \{x(t); x(t + \tau); x(t + 2\tau); \dots; x(t + [m - 1]\tau)\}. \quad (2.11)$$

Esta representação espacial, conhecida como espaço de fase (*phase space*) (Sauer *et al.*, 1991), também utiliza uma constante de atraso de tempo τ calculada por um método específico¹⁶. Um baixo valor de τ faz com que as variáveis sejam fortemente correlacionadas formando uma figura esticada próxima à linha diagonal do espaço de fase. Pouca informação pode ser extraída neste caso. De forma semelhante, um alto valor de τ gera uma trajetória caracterizada por muitas auto-intersecções. Uma escolha ótima de atraso produz uma reconstrução otimizada no espaço de fase (Kantz & Schreiber, 2004).

A Figura 2.10 ilustra uma série temporal arbitrária sendo representada no espaço de fase. Esta representação geométrica pode revelar várias características da dinâmica do sistema. Mesmo embora o cálculo exato da dimensão m e da constante de atraso τ seja importante, algumas características do sistema são independentes destes valores (Herzel *et al.*, 1994). Ocasionalmente, o sinal representado no espaço de fase é também chamado de retrato de fase (*phase portrait*).

A trajetória no espaço de fase descreve qualitativamente a dinâmica do sistema e suas características. No caso de sistemas oscilantes, como as PV, vibrações periódicas produzem uma trajetória fechada com alguma regularidade, ao contrário das vibrações aperiódicas que produzem uma trajetória irregular neste espaço.

Atratores do sistema

Conforme visto na Figura 2.10, a série temporal quando representada em espaço de fase forma uma figura geométrica. Após seu regime transiente, representando como uma espiral no gráfico da Figura 2.10(b), a série temporal entra em regime permanente

¹⁶ Para o cálculo de um valor de τ otimizado, consultar (Fraser & Swinney, 1986).

e pode apresentar um comportamento assintótico gerando uma órbita. O formato geométrico formado após o regime permanente é conhecido como atrator¹⁷. No caso da Figura 2.10(b), o atrator formado é conhecido como ciclo limite (*limit cycle*) devido suas características geométricas e seu formato orbital bem definido. Este tipo de atrator é frequentemente encontrado em sistemas dinâmicos.

Na maior parte dos casos, os atratores são classificados em quatro categorias. A primeira delas, chamada de “estados estacionários estáveis” são representados por figuras geométricas simples e bem definidas. A segunda, chamada de ciclo limite, refere-se às oscilações periódicas similares as mostradas na Figura 2.10b. O terceiro tipo, conhecido como *torus*, consiste na superposição de duas ou mais oscilações sem nenhuma relação de dependência racional entre suas frequências. Ele leva este nome devido ao fato de que a figura geométrica gerada por sinais desta natureza tem o formato similar de um *torus*. Sistemas *quasi*-periódicos apresentam frequentemente este tipo de atrator. Por fim, o quarto e último tipo é o atrator caótico que exibe um comportamento não periódico entre as trajetórias do sistema representadas no espaço de fase ou uma estrutura fractal. A Figura 2.11 ilustra uma série temporal caótica e seu correspondente atrator.

Figura 2.10 - (a) Série temporal utilizada como exemplo. (b) Imersão em espaço de fase da série temporal em um espaço com três dimensões ($m=3$). Figura adaptada de (Herzel *et al.*, 1994).

¹⁷ Este termo é também as vezes empregado para indicar a figura geométrica que descreve as trajetórias de um sistema dinâmico em um espaço de fase. As trajetórias podem representar a(s) saída(s) do sistema em um espaço onde os eixos representam seus estados.

Bifurcações

Se os parâmetros de um sistema dinâmico são alterados continuamente dentro de uma faixa de valores com pequenos passos, possíveis alterações significativas no atrator podem aparecer gerando transições conhecidas como bifurcações. Estas transições representam mudanças bruscas na dinâmica do sistema quando pequenas alterações em um determinado parâmetro são feitas. A Figura 2.11 ilustra um exemplo clássico de bifurcação de uma função.

No gráfico da Figura 2.11, o eixo horizontal representa uma constante a do sistema que teve seu valor variado dentro do intervalo $[0, 4]$. O eixo vertical representa um estado y do sistema que pode ser, por exemplo, a máxima amplitude de saída produzida para um dado valor fixo de a . Assim, para cada valor de a é calculado várias vezes (utilizando condições iniciais diferentes) o máximo valor¹⁸ de sua saída e este é associado ao eixo vertical. Assim, quando $a = 3$, o sistema pode apresentar dois estados diferentes para um mesmo valor de a . Isto quer dizer que qualquer uma das duas possibilidades de saída do sistema pode acontecer e vão depender das condições iniciais deste sistema ou de outros fatores. Este ponto é conhecido como ponto de bifurcação. Antes, quando $a < 3$, o sistema tinha um comportamento bem definido (*i.e.*, um ponto

Figura 2.11 – Bifurcação de uma função logística. A partir do ponto $a \geq 3$, (eixo horizontal) o sistema apresenta uma série de pontos de bifurcação.

¹⁸ Consideram-se apenas os valores produzidos após o regime transiente do sistema.

para cada valor de a) e as condições iniciais não eram capazes de produzir saídas distintas para uma mesma contate a . Em $a = 3.2$, por exemplo, o sistema tem dois estados distintos ($y = 0.5$ e $y = 0.8$). Para $a = 3.9$, o sistema apresenta uma infinidade de estados diferentes para um mesmo a podendo gerar um comportamento caracterizado por muitas irregularidades e grande sensibilidade às condições iniciais. Através deste gráfico percebe-se que a constante a pode introduzir instabilidades no sistema quando tem seu valor igual ou maior que 3. Neste exemplo, é possível perceber que o diagrama pode ajudar a avaliar a estabilidade do sistema.

A teoria de sistemas dinâmicos salienta que a existência de bifurcações e a coexistência com atratores com vários ciclos-limite ou *torus* são precursores para caos determinístico (Kantz & Schreiber, 2004). Consequentemente, a detecção de bifurcações e atratores estranhos em sinais de vozes dão base à ideia que algumas irregularidades na voz podem ser manifestações de caos ou fortes não-linearidades.

Dimensão de correlação

A dimensão de correlação (Grassberger & Procaccia, 1983) foi originalmente desenvolvida para avaliar o quanto complexo é um atrator. Trata-se de uma medida geométrica no espaço de fase que descreve a distância entre os pontos de uma trajetória e sua correlação. Desta forma, é possível descrever quantitativamente a irregularidade de uma trajetória. Sistemas que possuem trajetórias caóticas com oscilações aperiódicas possuem dimensão fractal¹⁹. Graças a esta característica, é possível distinguir se uma série temporal tem comportamento randômico ou caótico. Isto é feito analisando o valor de dimensão de correlação, pois este valor não converge para comportamentos randômicos quando a dimensão de espaço de fase m aumenta. De forma contrária, em caso de comportamento caótico, esta convergência da dimensão de correlação acontece quando se aumenta a dimensão de imersão m do espaço de fase. A Figura 2.12 ilustra

¹⁹ Resumidamente, uma dimensão fractal é caracterizada por ter um valor não-inteiro como grandeza de dimensão.

Figura 2.13 – Gráfico de dimensão de correlação. A variável r (eixo horizontal) indica o raio de uma esfera no espaço de fase utilizada para calcular a dimensão de correlação (eixo vertical). A seta na figura indica a região de convergência. As curvas indicam, respectivamente, diferentes dimensões no espaço da fase ($m = 1, 2, 3, \dots, 10$) da curva inferior a superior. Figura adaptada de (Jiang *et al.*, 2006).

um gráfico de dimensão de correlação para um sistema caótico cuja dimensão de correlação converge para o valor aproximado de 3,5 dentro da região $13,6 < r < 14$ no eixo horizontal do gráfico (ver indicação da seta).

Além de determinar a característica caótica do sinal (e não aleatória), o gráfico da Figura 2.12 indica que a dimensão adequada para se representar este sistema no espaço de fase é próxima ao valor de convergência. Note que as curvas inferiores para $m = 1, 2$ e 3 não convergem na região indicada pela seta pois não possuem uma dimensão suficiente para alcançar a convergência que só acontece para dimensões $m \geq 3,5$.

Análise modal de osciladores

A análise modal (Worden & Tomlinson, 2001) é uma técnica utilizada para estudar osciladores com n graus de liberdade. Basicamente, ela utiliza de recursos matemáticos para desacoplar as equações do sistema e remodelá-las em partes menores e independentes onde cada uma destas partes representa um oscilador com um grau de liberdade e características de frequência próprias. Estes “osciladores menores” são

chamados de *modos*²⁰ do sistema. Para ilustrar seu princípio, considere o sistema ilustrado na Figura 2.13.

Usando a teoria mecânica clássica, sabe-se que as equações do movimento estão acopladas pela mola k_{12} por onde acontece uma transmissão de movimento e que ambas equações dos modelos estão “interligadas” por esta mola. Considerando as equações do sistema na sua forma matricial, temos que

$$\begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix} \begin{Bmatrix} \ddot{y}_1 \\ \ddot{y}_2 \end{Bmatrix} + \begin{bmatrix} k_1 & k_{12} \\ k_{12} & k_2 \end{bmatrix} \begin{Bmatrix} y_1 \\ y_2 \end{Bmatrix} = \begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix}. \quad (2.12)$$

Para encontrar as frequências naturais do modelo (ou frequências de ressonância), consideremos a resposta não-forçada do sistema na forma

$$[m]\{\ddot{y}\} + [k]\{y\} = \{0\} \quad (2.13)$$

onde $[m]$ e $[k]$ representam, respectivamente, a matriz de massas e molas e $\{\bullet\}$ designa um vetor de estados. Chaves serão empregadas para representar vetores. Para se descobrir os modos deste sistema é necessário que se faça uma transformação algébrica de mudança de base de tal forma que se diagonalize de forma simultânea as matrizes $[m]$ e $[k]$. Considere que esta matriz de diagonalização seja representada por $[\Psi]$ e que a mesma tenha que atender à relação $[\Psi]^T = [\Psi]^{-1}$ onde T indica a transposição da matriz e $^{-1}$ a sua inversão (*e.g.*, a matriz $[\Psi]$ de ser ortogonal). Sendo assim, temos a

Figura 2. 14 – Oscilador mecânico linear e não-amortecido com dois graus de liberdade utilizado para ilustrar os modos de um sistema mecânico em vibração. Considere x_1 e x_2 como as forças que impulsionam, respectivamente, os deslocamentos y_1 e y_2 das massas conforme a ilustração.

²⁰ Este termo também é conhecido como *eigenmode*.

diagonalização de $[m]$ e $[k]$ como

$$\begin{aligned} [\Psi]^T [m] [\Psi] &= [\tilde{m}] = \begin{bmatrix} \tilde{m}_1 & 0 \\ 0 & \tilde{m}_2 \end{bmatrix}, \\ [\Psi]^T [k] [\Psi] &= [\tilde{k}] = \begin{bmatrix} \tilde{k}_1 & 0 \\ 0 & \tilde{k}_2 \end{bmatrix}. \end{aligned} \quad (2.14)$$

Repare que o último termo de ambas as equações de (2.14) são matrizes diagonalizadas. O símbolo $(\sim)$ foi empregado para indicar que as constantes de massa e mola passaram por uma transformação linear e estão representadas em uma outra base. Mudadas as constantes do sistema, é também necessário mudar suas coordenadas $\{y\}$ para outra base $\{z\}$ onde $\{y\} = [\Psi]\{z\}$, *i.e.*,

$$\begin{aligned} y_1 &= \Psi_{11}z_1 + \Psi_{12}z_2 \\ y_2 &= \Psi_{21}z_1 + \Psi_{22}z_2 \end{aligned} \quad (2.15)$$

Logo, a Equação (2.13) é reescrita como

$$[m][\Psi]\{\ddot{z}\} + [k][\Psi]\{z\} = \{0\}. \quad (2.16)$$

Pré-multiplicando (2.16) por $[\Psi]^T$, temos

$$[\tilde{m}]\{\ddot{z}\} + [\tilde{k}]\{z\} = \{0\} \quad (2.17)$$

que na forma escalar é dada por

$$\begin{aligned} \tilde{m}_1 \ddot{z}_1 + \tilde{k}_1 z_1 &= 0, \\ \tilde{m}_2 \ddot{z}_2 + \tilde{k}_2 z_2 &= 0. \end{aligned} \quad (2.18)$$

Em (2.18) temos dois osciladores desacoplados com um grau de liberdade cada em um espaço definido pelas coordenadas z . Estas coordenadas estão linearmente associadas às coordenadas y por (2.15). Note que a mola que unia as massas e era comum a ambas equações não existe mais e que agora temos dois osciladores independentes representados em um sistema de coordenadas z diferente de y . Cada um dos dois osciladores produzidos em (2.18) é um “modo do sistema” e está associado a uma frequência natural w dada por:

$$w_1 = (2\pi)^{-1} \sqrt{\tilde{k}_1/\tilde{m}_1} \quad \text{e} \quad w_2 = (2\pi)^{-1} \sqrt{\tilde{k}_2/\tilde{m}_2} \quad (2.19)$$

Cada eixo no sistema de coordenadas z está ligado a uma frequência e um modo de vibração e por esta razão o sistema de coordenadas z é designado como “coordenada ou espaço modal”. Repare que ao voltarmos ao sistema de coordenadas original usando

a relação (2.15), estes dois modos são linearmente ponderados e somados para produzir uma coordenada do sistema original.

A dissociação em modos permite analisar o sistema de uma forma mais simples e decompor sua saída como uma superposição de seus modos. No caso do modelo da Figura 2.14, similar ao de duas massas das PV, os dois possíveis modos deste sistema podem ser exemplificados através da Figura 2.14. Em seu item (a), a Figura 2.14 mostra um caso onde as massas estão se deslocando em fase sem nenhuma defasem de movimento entre as mesmas e produzindo o modo conhecido como “*modo em fase*”. No caso das pregas vocais, este movimento está associado ao movimento lateral das pregas que determina o ciclo de abertura e fechamento.

No outro caso desta mesma figura, as massas têm uma diferença de fase de exatamente 180° produzindo o “*modo fora de fase*”. Este movimento está associado à produção de uma constrição glotal convergente ou divergente. Alguns trabalhos (Titze, 1976; Titze, 1988; Zhang *et al.*, 2006) empregando a teoria modal revelaram que a frequência de oscilação das pregas é geralmente próxima à oscilação do modo que tem menor frequência natural.

Neste ponto é importante observar que este tipo de transformação ou análise modal só é válida pra sistemas lineares. Apesar desta grande limitação, alguns

Figura 2.15 – Ilustração dos dois possíveis modos de um sistema com dois graus de liberdade. Em (a) temos o “modo em fase” onde ambas as massas oscilam em fase. Em (b) temos um movimento com uma diferença de fase de 180° produzindo o “modo fora de fase”. Qualquer outro valor de fase entre os deslocamentos das massas é uma combinação destes dois principais modos.

trabalhos (Titze, 1976; Berry *et al.*, 1994; Berry & Titze, 1996; Berry, 2001; Zhang *et al.*, 2006) utilizaram esta técnica para estudar o comportamento dinâmico das PV durante sua vibração através de modelos linearizados.

Em (Berry *et al.*, 1994) foi concluído que dois modos eram capazes de reproduzir a maior parte dos padrões de vibrações de PV normais e patológicas. Em outro trabalho (Berry & Titze, 1996), os autores analisaram a correlação entre a frequência fundamental da voz e o comportamento modal de um modelo linearizado de PV. Utilizando metodologia similar, Zhang *et al.* (2006) estudaram alguns padrões de vibrações em que o modo em fase é predominante e em outros em que o modo fora de fase que assume esta predominância.

Utilizando uma versão linearizada do IF72, Berry (2001) analisaram a resposta em frequência do modelo considerando um caso onde o sistema era conservativo ($r_1=r_2=0$) e outro onde era amortecido. A Figura 2.15 ilustra os espectros de ressonância encontrados pelos autores.

Utilizando a análise modal, os autores deste último trabalho perceberam que os dois picos de ressonância encontrados correspondem às frequências de oscilações dos modos em fase e fora de fase do modelo. O modo em fase tinha um pico próximo de 120 Hz e o fora de fase próximo de 201 Hz. Foi ainda notado que os picos ressonantes tinham frequências próximas em ambos os casos que se considerou o amortecimento do sistema. Com isto confirmaram que a análise modal pra sistemas conservativos²¹ tem correspondência com sua contrapartida amortecida.

Mesmo sendo uma técnica poderosa, a análise modal aplicada ao estudo dinâmico das PV tem seu alcance limitado devido à questão da linearidade como requerimento básico. Todavia, tem-se desenvolvido uma nova abordagem da análise modal voltada a sistemas não-lineares. Esta técnica nova, conhecida como *nonlinear*

²¹ É comum em análise modal considerar um sistema conservativo, negligenciado seus efeitos de amortecimento, com o propósito de facilitar ou tornar viável o cálculo de uma matriz de diagonalização capaz de diagonalizar simultaneamente $[m]$, $[k]$ e $[r]$.

normal modes (Kerschen *et al.*, 2009), vem mostrando ser capaz de reproduzir alguns efeitos das não-linearidades nos modos do sistema e mostra-se como promissora. Na tentativa de explorar estas potencialidades, esta tese emprega em seus próximos capítulos este método pioneiro para descrever o comportamento estrutural das PV e avaliar seus modos levando em consideração suas não-linearidades.

2.6.2 – Pesquisas envolvendo métodos da dinâmica não-linear

No decorrer das últimas décadas, inúmeros pesquisadores vêm desenvolvendo esforços para avaliar a voz e sua dinâmica na laringe. Como citado anteriormente, os métodos da análise dinâmica não-linear têm mostrado resultados promissores e por isto vêm atraindo considerável atenção.

Estes métodos têm sido empregados para avaliar qualitativamente e quantitativamente variados tipos de vozes (Zhang & Jiang, 2003) e estabelecer uma diferença entre seus padrões. Em (Jiang & Zhang, 2002), eles foram usados para diferenciar vozes normais e patológicas de pacientes com pólipos nas pregas através de índices de dimensão de correlação. Utilizando este mesmo índice, Titze *et al.* (1993) já haviam afirmado que sujeitos com disfonia apresentavam uma maior dimensão de correlação. Kakita e Okamoto (1995) também descobriram que este índice é maior para

Figura 2.16 – Resposta em frequência do IF72. Em (a) são desconsiderados os efeitos dos amortecedores e são vistos dois picos ressonantes localizados em 120 e 201 Hz. No gráfico (b) o amortecimento do IF72 é levado em conta e novamente são destacados dois pontos ressonantes localizados em 121 e 198 Hz. Figura extraída e adaptada de (Berry, 2001).

sujeitos com rouquidão, indicando que esta condição de voz remete a um estado dinâmico da laringe mais complexo. Analisando indivíduos com Parkinson a partir de sinais de eletroglotografia, foi confirmado que estes possuíam uma maior dimensão fractal em seus sinais de vozes (Hertrich *et al.*, 1997). Em (Giovanni *et al.*, 1998) foi utilizado o maior expoente de Lyapunov para diferenciar sujeitos com diferentes graus de paralisia laringeal. Jiang *et al.* (2006) sugeriram que alterações nas medidas dinâmicas na laringe podem indicar estados de disfunção e que por isto estes parâmetros podem ser potencialmente aplicados para diagnosticar distúrbios fisiológicos e avaliar efeitos de tratamentos clínicos.

Outros trabalhos dedicaram-se a analisar a estabilidade das PV saudáveis e patológicas utilizando diagramas de bifurcação (Mende *et al.*, 1990; Zhang & Jiang, 2004; Tokuda *et al.*, 2008; Pinheiro *et al.*, 2012) na tentativa de avaliar em quais condições poderiam ser observados comportamentos estranhos. Estes mesmos comportamentos de instabilidades foram estudados em (Giovanni *et al.*, 1999) utilizando diagramas de espaço de fase a fim de se verificar um comportamento não-linear em um conjunto de pregas caninas excisadas.

Outros métodos análogos envolvendo cálculos nas trajetórias e avaliações da topologia do atrator tais como, por exemplo, mapas de Poincaré, entropia, expoente de Liapunov e outros (Eckmann *et al.*, 1986; Pincus, 1991) também têm sido aplicados para avaliar sinais de vozes. Mais detalhes sobre a aplicação destes outros métodos ao estudo da voz podem ser consultados em (Herzel, 1993; Ott *et al.*, 1994; Herzel *et al.*, 1995; Kumar & Mullick, 1996; Behrman & Baken, 1997; Boek *et al.*, 1997; Jiang *et al.*, 2001; Jiang *et al.*, 2003).

No que diz respeito à extração e/ou avaliação da dinâmica das pregas, uma das primeiras iniciativas de estimação de seus parâmetros mecânicos a partir de vídeos de alta velocidade aconteceu em (Mergell *et al.*, 2000). Neste trabalho, o sinal de deslocamento lateral das pregas registrado pela câmera foi utilizado para caracterizar o

comportamento dinâmico das PV. Variando de maneira manual alguns coeficientes do SH95, os autores tentaram fazer com que este fosse capaz de reproduzir os sinais de deslocamentos extraídos. De maneira similar, Dollinger *et al.* (2002) utilizaram vídeos de alta velocidade de PV que foram utilizadas para quantizar o grau de assimetria entre as pregas direita e esquerda de vozes patológicas. Isto foi feito a partir da análise espectral por Fourier dos sinais de deslocamento da prega esquerda e direita individualmente. A partir destes sinais os autores determinaram um índice para medir a assimetria de movimento entre as pregas. Em (Qin *et al.*, 2009) foram utilizadas câmeras de alta velocidade e sinais de EGG para mostrar que a combinação de ambos pode incrementar a qualidade da estimação dos parâmetros das PV. Outros autores (Zhang *et al.*, 2006) preferiram utilizar uma nova abordagem baseada na “sincronização de caos” das pregas com o propósito de encontrar parâmetros biomecânicos capazes de representar potenciais assimetrias entre pregas. Há também autores (Yang *et al.*, 2011) que investiram esforços em modelos complexos de PV com vários graus de liberdade com o objetivo de fazer com que o modelo fosse capaz de simular os padrões de movimento das pregas e extrair parâmetros de vibração fisiológicos.

2.7 - Resumo

A importância das PV na fonação é evidente e pouco questionável. O fato de apresentarem um comportamento não-linear introduz uma série de complexidades não percebidas nas outras estruturas que fazem parte do sistema fonador. Diante disto, a videolaringoscopia de alta velocidade trouxe um grande impulso ao estudo científico e clínico da voz na medida em que possibilitou um real registro dos movimentos das pregas *in vivo*.

No que diz respeito ao modelamento mecânico das PV, percebe-se que os modelos mais utilizados na literatura têm todos eles, pelo menos, uma década. Mesmo embora seja um dos pioneiros, o proposto por Ishizaka e Flanagan (IF72) ainda pode

ser tido como aplicável, relevante e complexo e por isto merece destaque. Sua grande desvantagem está relacionada à complexidade computacional de suas equações e também ao fato de que alguns de seus parâmetros mecânicos não apresentam grande realismo fisiológico direto como destacam Story e Titze (1995). Apesar disto, é um dos modelos (com dois graus de liberdade) mais completo e rico em detalhes. Suas equações revelam estas riquezas quando comparadas aos modelos mais atuais com mesmo número de elementos. Ainda, há muito que se produzir em termos de modelamento das PV e pesquisas ainda continuam sendo desenvolvidas na área. Os trabalhos recentes citados anteriormente comprovam esta afirmação. Neste sentido, os achados avaliados por registros de câmeras de alta velocidade podem potencialmente ajudar a melhorar o realismo dos modelos e suas constantes além de conduzir a novos aperfeiçoamentos e descobertas até então não contempladas em modelos padrões.

A aplicação de métodos da dinâmica não-linear de sistemas para o estudo de complexos fenômenos da natureza, especialmente aqueles envolvendo sistemas biológicos, tem produzido achados preciosos e permitido um melhor entendimento de tais sistemas. Em particular, estes métodos têm sido aplicados com sucesso na eletroencefalografia (Stam, 2005) e eletrocardiografia (Wang *et al.*, 2000). Como descrito anteriormente, o método foi empregado com mesmo sucesso no estudo de vozes e avaliação dinâmica da produção de voz na laringe e seu uso têm sido uma tendência nesta área. Devido a este sucesso, alguns pesquisadores (Jiang *et al.*, 2006; Tokuda *et al.*, 2008) têm apontado a dinâmica não-linear como uma promissora ferramenta para caracterização e análise de sinais de vozes e avaliação das PV.

No próximo capítulo são descritos os métodos utilizados para reconstruir a dinâmica das PV de um grupo de indivíduos que tiveram os movimentos de suas pregas estimados por um sistema de videolaringoscopia de alta velocidade. São também descritos alguns métodos utilizados para avaliar a dinâmica destes modelos reconstituídos utilizando análise modal não-linear, retratos de fase e bifurcações.

Capítulo 3

Métodos empregados na extração e avaliação da dinâmica das pregas vocais

A avaliação dinâmica de sistemas tem trazido inúmeras descobertas valiosas a várias áreas do conhecimento. Em casos de sistemas mais complexos - como o aqui estudado - é importante desenvolver métodos para reproduzir e simular a sua dinâmica de maneira determinística tornando possível analisar seu comportamento utilizando os recursos disponíveis na área. Ademais, o conhecimento da interação de forças entre os componentes do sistema e seu ambiente traz uma nova abordagem complementar às existentes no estudo da voz. Para que tal propósito possa ser alcançado com sucesso, é importante que estes modelos dinâmicos estejam respaldados em informações realísticas extraídas de sistemas reais. Neste sentido, este capítulo descreve em sua primeira parte um método que usa sinais glotais reais extraídos *in vivo* de um grupo de indivíduos e os emprega em uma técnica de otimização numérica. Esta técnica utiliza estes sinais para estimar os parâmetros mecânicos de dois diferentes modelos de PV com o propósito que estes possam reproduzir os sinais extraídos e a dinâmica básica das pregas de cada indivíduo analisado.

Dispondo destes modelos capazes de representar as características dinâmicas do grupo avaliado, a segunda parte do capítulo dedica-se a analisar a não-linearidade estrutural das pregas estudando o comportamento de seus modos e os relacionando a

energia mecânica do sistema. Diagramas de bifurcação e gráficos de espaço de fase também são usados de maneira complementar para avaliar o comportamento das pregas em condições simuladas de assimetrias.

3.1 – Extração da dinâmica das pregas vocais

Alguns métodos de análise mais elaborados como, por exemplo, a modal, necessitam que se tenha um modelo determinístico do sistema em estudo. Logo, antes de se aventurar em qualquer tipo de análise mais profunda, é importante produzir meios para que esta dinâmica determinística possa ser modelada a partir de dados reais. Para atingir tal propósito e permitir um estudo mais fundamentado da dinâmica vocal neste trabalho (e em outros futuros), esta seção descreve uma técnica desenvolvida para extrair alguns parâmetros de vibração de indivíduos e utilizá-los em modelos mecânicos de PV. A Figura 3.1 resume alguns passos básicos do método proposto para esta extração de dinâmica. O desenvolvimento deste método, sua validação e aplicação a um grupo de indivíduos é um dos objetivos principais desta pesquisa.

Na primeira etapa (Figura 3.1(a)) do método é feita a aquisição de dados

Figura 3.1 – Etapas do método desenvolvido para extrair a dinâmica das pregas vocais. Em (a) é feita a aquisição de dados por videolaringoscopia; em (b) os sinais coletados são processados e a área da glote é segmentada para gerar um “sinal de área glotal”. Na etapa (c) são implementados modelos mecânicos de pregas vocais. Os dados extraídos em (b) são utilizados por um método de otimização em (d) para alterar os parâmetros mecânicos do modelo produzindo um sistema capaz de reproduzir a dinâmica das pregas em análise

através da videolaringoscopia de alta velocidade. Esta técnica gera um conjunto de imagens que serão analisadas por um algoritmo de processamento de imagens na segunda etapa (Figura 3.1(b)), que as segmentará para estimar o sinal de área glotal. Este sinal extraído será usado pelas etapas subsequentes.

Dando sequência ao procedimento da Figura 3.1, um método de otimização (Figura 3.1(d)) utiliza o sinal produzido pela etapa (b) para otimizar os parâmetros de um modelo mecânico das PV (Figura 3.1(c)). Este modelo otimizado reproduz as características dinâmicas das PV registradas pelo exame de videolaringoscopia de um determinado indivíduo como ilustra a Figura 3.1(e). Para fins elucidativos, estas etapas são descritas individualmente nas próximas subseções.

3.1.1 – Aquisição de dados: sujeitos e métodos

Como mencionado anteriormente, esta pesquisa utiliza a videolaringoscopia de alta velocidade para estimar o formato do pulso glotal e avaliar os movimentos de vibração das PV com o propósito de utilizar estes dados para reconstrução de sua dinâmica.

Foram selecionados uma amostra²² de 20 indivíduos (11 mulheres e 9 homens com média de idade de 34 anos) com vozes normais que foram submetidos ao exame de videolaringoscopia de alta velocidade utilizando o sistema Endocam 5262 (Richard Wolf, Knittlingen, Alemanha) que é ilustrado na Figura 3.2. Este sistema é capaz de registrar 4.000 imagens por segundo com resolução de 256x256 *pixels*. Foram selecionados para amostra apenas indivíduos com frequência fundamental de voz inferior a 300 Hz. Considerando uma frequência de amostragem de 4.000 imagens por segundo, frequências maiores que 300Hz produziriam poucas imagens por ciclo de vibração e resultariam em um sinal de área glotal com baixa resolução. Foram processados, em média, cerca 70 ms de vídeo para cada indivíduo que emitiu de forma

²² Este projeto foi aprovado pela Comissão de Ética para Análise de Projetos de Pesquisa (CAPPesq) da Diretoria Clínica do Hospital de Clínicas e da Faculdade de Medicina da Universidade de São Paulo sob o protocolo 0767/09.

sustentada a vogal /a/ ou /e/ durante o exame. Os procedimentos foram executados pelo grupo de otorrinolaringologia do Hospital de Clínicas da Universidade de São Paulo, São Paulo.

3.1.2 – Processamento das imagens da laringe

A correta aplicação de técnicas de processamento e segmentação de imagens é um processo importante para o cálculo da área glotal já que de forma manual esta tarefa é inviável em muitas circunstâncias e de precisão limitada. Este processamento deve ser aplicado em cada quadro de imagem (*frame*) que compõe o vídeo e ao final do processo é gerada uma série temporal que representa a variação da área glotal e os movimentos de vibração das PV. Esta série temporal será denominada de *sinai glotal extraído* (SGE) e no caso da videolaringoscopia representa a área glotal de um determinado indivíduo. Neste caso, cada amostra do SGE representa uma imagem processada e seu respectivo valor de área glotal. Este sinal é um importante reflexo do comportamento dinâmico das PV e pode, de maneira resumida, ser considerado como proporcional ao pulso glotal U_g (Lucero & Koenig, 2005) em condições de normalidade de voz. A Figura 3.3 ilustra um SGE e alguns de seus pontos.

Figura 3.2 – Ilustração do sistema de videolaringoscopia de alta velocidade empregado nesta pesquisa. Na imagem é possível ver o sistema em funcionamento durante uma coleta de dados.

Vários métodos têm sido empregados para segmentar a glote a partir de imagens da laringe, incluindo métodos de binarização simples (Hong *et al.*, 2002), contornos ativos (Lohscheller *et al.*, 2004), detecção de bordas por Canny (Zhang *et al.*) dentre outros métodos tradicionais (Yan *et al.*, 2006). Contudo, estas técnicas frequentemente apresentam uma série de restrições que dificultam seu uso ou limitam sua precisão. Métodos de binarização simples assumem que existe uma diferença significativa entre a intensidade da cor dos *pixels* da glote e os das pregas, regra esta nem sempre válida principalmente nos casos onde a área da glote é pequena fazendo com que esta diferença seja decrementada e assim dificultando o processo de binarização. Contornos ativos, mesmo embora sejam bastante populares, apresentam um alto custo computacional e podem facilmente conduzir a falsas detecções durante a fase onde as pregas estão fechadas. De um modo em geral, métodos tradicionais de processamento de imagens apresentam pouca eficácia, pois as imagens da laringe geralmente apresentam baixa resolução espacial, iluminação pobre e/ou não-uniforme, movimentos relativos e contraste precário até mesmo em sistemas mais atuais como o empregado nesta pesquisa. Diante deste cenário, foi desenvolvido pelo autor deste trabalho um algoritmo para extrair a região da glote de forma semi-automática conjugando

Figura 3.3 – Na imagem superior é ilustrado um sinal real de área glotal extraída (*i.e.*, SGE) com a correspondente imagem de alguns pontos que o compõem. Na imagem inferior é ilustrada a derivada deste sinal.

diversificadas técnicas de processamento de imagens.

O algoritmo desenvolvido para extrair (*i.e.*, segmentar) a região glotal nas imagens é composto de uma série de etapas sequenciais que são aplicadas a cada um dos *frames* de vídeo. O diagrama da Figura 3.4 ilustra estas etapas.

A primeira etapa do algoritmo, rotulada como “Inicialização” no diagrama da Figura 3.4(a), começa com a exibição da imagem de um quadro do vídeo selecionado pelo usuário. Este primeiro quadro marca o início do SGE. Nesta etapa, o usuário definirá com o *mouse*, um ou mais pontos contidos dentro da região da glote. Estes pontos serão utilizados pelo algoritmo para estimar a intensidade média da cor da região da glote e assim definir um limiar L que será utilizado posteriormente na etapa de “segmentação”. Eles também serão utilizados em etapas futuras para definir a posição inicial da glote na imagem. Ressalta-se que esta primeira etapa é a única onde há necessidade de intervenção manual do usuário e que ela é executada somente uma vez para cada vídeo processado.

A segunda etapa da Figura 3.4(b), rotulada como “Binarização”, é aplicada a cada quadro do vídeo. Sua principal função é separar a região da glote das demais regiões presentes na imagem a partir do limiar L definido na etapa anterior. Até esta

Figura 3.4 – Diagrama descrevendo as etapas de processamento utilizadas para segmentar a região da glote nas imagens de videolaringoscopia de alta velocidade.

etapa, cada *pixel* da imagem tem um valor de intensidade que pode variar de 0 a 255. Os *pixels* com valor inferior ao limiar L terão sua intensidade alterada para o valor 1 enquanto que os demais assumirão o valor 0. O resultado deste processo é ilustrado na Figura 3.4(b). As regiões onde a intensidade do *pixel* é 1 são representadas pela cor branca enquanto que as demais são definidas pela cor preta. Como as imagens têm uma iluminação precária e baixo contraste, é comum a presença de artefatos. Apesar disto, o principal objetivo desta etapa é esboçar grosseiramente a região da glote ou parte dela.

Conforme mencionado, o processo de binarização gera muitos artefatos na imagem. Além disto, a região correspondente à glote nem sempre é segmentada por este processo devido à má qualidade das imagens. Como os processos diretos de segmentação são muitas vezes insuficientes para extrair a glote, foi utilizado um algoritmo de “crescimento de região” (*region growing*) (Wood *et al.*, 1995) para encontrar na imagem a área correspondente à glote. Basicamente, este algoritmo parte de uma região ou ponto inicial contido dentro da área que se deseja segmentar e a partir de um critério de homogeneidade, aumenta iterativamente esta área inicial até que este critério não possa mais ser satisfeito pelos *pixels* adjacentes à área segmentada. Esta região, ou ponto inicial, recebe o nome de “semente”. O grande desafio deste método de segmentação é definir corretamente a posição da semente e o critério de homogeneidade. Neste contexto, a etapa de “filtragem e estimação da semente”, ilustrada na Figura 3.4(c), tem como principal objetivo definir uma semente que será utilizada pela etapa de “crescimento de região” ilustrada na Figura 3.4(d).

A primeira etapa do processo de “filtragem e estimação da semente” (Figura 3.4(c)) é amenizar os efeitos dos artefatos na imagem. Para isto é aplicado à imagem sucessivas operações de erosão (Gonzalez & Woods, 2002). Esta operação morfológica, muito comum em imagens binárias, suprime os *pixels* da borda que possuem uma conexão “fraca” ou *pixels* isolados. Em termos práticos, esta operação elimina os detalhes irrelevantes da imagem e os *pixels* isolados. Depois de erodida, a imagem passa

pelo algoritmo *Blob* (Lindeberg, 1993). Este identifica na imagem binária os elementos na cor branca e permite que sejam analisados individualmente e classificados. Aquele elemento que apresentar o maior tamanho e altura será selecionado enquanto que os demais serão suprimidos.

Na sequência, o elemento selecionado tem seu centroide estimado. As coordenadas do centróide serão usadas como semente na etapa de crescimento de região. Em casos especiais onde a área da glote é pequena ou inexistente, o processo de erosão elimina todos os elementos da imagem e nenhuma semente é definida. Ainda, para evitar falsos positivos nesta etapa é avaliada a variação da posição das sementes durante os últimos quadros. Quando há uma grande variação na direção horizontal (indicando um possível falso positivo) a semente é descartada e nenhuma área é segmentada.

Realizada a estimação da semente, dá-se início ao processo de crescimento de região. Este é um método que vem ganhando grande destaque na área de segmentação de imagens, em especial, as médicas. Mesmo embora sua ideia inicial de crescer uma região “semente” até certo limite de tolerância mantenha-se conservada, esta técnica ganhou inúmeras variações. Nesta pesquisa, o algoritmo de crescimento de região utilizado (Wood *et al.*, 1995) faz com que a região inicial (semente) cresça iterativamente através da análise de todos os *pixels* que fazem vizinhança com a fronteira desta região e segundo um critério de homogeneidade que avalia a diferença entre o valor da intensidade dos *pixels* vizinhos e a intensidade média da área já segmentada. Se esta diferença for menor que um certo valor percentual definido inicialmente pelo usuário, o *pixel* com menor diferença é anexado à área segmentada. Este processo é novamente repetido até que a diferença de intensidade entre a média da região segmentada e seus *pixels* vizinhos torne-se maior que o patamar definido pelo usuário como critério de homogeneidade. A prática tem mostrado que uma tolerância média de 15% para este tipo de imagem tem sido suficiente em muitos casos. Ao final

do processo de crescimento de região, a área definida por esta etapa deve ser correspondente à área da glote.

Depois de segmentada, a região da glote pode apresentar algumas irregularidades indesejadas em sua forma devido aos artefatos da imagem ou imperfeições presentes no processo de crescimento da região (ver Figura 3.4(d)). Para amenizar estas “imperfeições” é aplicada à área segmentada uma operação morfológica conhecida como “fechamento”. Basicamente, o processo de fechamento é uma combinação das operações de erosão e dilatação (Gonzalez & Woods, 2002). O processo inicia com a aplicação sucessiva de duas ou três operações de dilatação e ao resultado produzido, aplica-se o mesmo número de vezes a operação de erosão. O resultado final é uma imagem mais suavizada como a ilustrada na Figura 3.4(e).

Realizadas todas as etapas descritas nos parágrafos anteriores, o algoritmo seleciona o próximo quadro do vídeo que seguirá a sequência descrita pelo diagrama da Figura 3.4. Este processo se repete até que todos os quadros do vídeo tenham sido processados e o SGE construído. Todas estas etapas foram desenvolvidas em MatLab (MathWorks, Massachusetts, EUA).

Ainda, vale destacar que a área segmentada é medida em *pixels*. Para que se possa estabelecer a correspondência entre *pixels* e a real dimensão da área glotal, é necessário que se tenha marcas de referência na imagem. Wurzbacher *et al.*, (2008) propuseram uma adaptação em um endoscópio rígido de modo que este passasse a projetar nas PV dois pontos luminosos cuja distância entre eles é conhecida e fixa. Com esta informação métrica, foi possível fazer a conversão de *pixels* em centímetros. Contudo, nem sempre esta adaptação é possível. Em outros trabalhos (Mergell *et al.*, 2000; Tao *et al.*, 2007) foi assumido que a distância entre o ponto distal e ventral da glote em sua abertura máxima é de 1.0 a 1.4 cm. Através desta suposição, é possível a conversão da área medida em *pixels* para centímetros. Esta também foi a estratégia adotada nesta pesquisa.

3.1.3 – Implementação dos modelos mecânicos das pregas vocais

Os modelos mecânicos de PV estabelecem uma relação matemática entre um conjunto de dados de entrada e sua correspondente saída a partir das leis dinâmicas. Esta relação de entrada e saída é ilustrada esquematicamente na Figura 3.5.

A entrada do modelo está associada a um conjunto de parâmetros que refletem as propriedades das pregas e as suas condições de trabalho relativas, principalmente, à pressão aerodinâmica com que vibram. Devem também estar presentes as condições de contorno necessárias para a solução das equações diferenciais do modelo. No caso específico dos parâmetros viscoelásticos, estes definem de maneira direta a resposta estática e dinâmica dos músculos das pregas a um esforço externo aplicado sobre sua superfície, como é o caso da pressão sub-glotal. Seus valores afetam diretamente os modos e as frequências de vibração das PV.

Nesta pesquisa são empregados modelos de duas massas para representarem a dinâmica das PV. Embora existam opções mais sofisticadas, acredita-se que um modelo com poucos graus de liberdade se adeque melhor à proposta na medida em que torna mais simples as interpretações e utiliza uma menor quantidade de parâmetros contribuindo para a eficiência da otimização. Especificamente, foram utilizados o IF72 e SH95.

O uso de dois diferentes modelos visa atender a diferentes expectativas de análises. Por exemplo, o IF72 é usado aqui para avaliar as não-linearidades estruturais das pregas já que é um modelo que oferece uma abordagem mais completa e detalhada

Figura 3.5 – Representação genérica de um modelo de pregas vocais com suas entradas e saídas mais comuns.

das não-linearidades identificadas em PV reais durante experimentos laboratoriais. Por outro lado, o SH95, apesar de não apresentar este tipo de detalhamento, tem a vantagem de ter um custo computacional mais eficiente e se atém mais aos mecanismos básicos de vibração das pregas desconsiderando alguns aspectos tidos como irrelevantes para análises mais elementares. Isto permite a sua aplicação em métodos da dinâmica não-linear que empregam cálculos extensivos e mesmo sendo uma simplificação, preserva as características dinâmicas básicas das PV segundo seus autores. Ambos modelos têm sido amplamente empregados na literatura científica e o uso de ambos visa dar um caráter mais flexível às análises na medida em que associa a uma determinada técnica de avaliação um modelo mais adequado.

Cabe destacar que neste trabalho é negligenciado o acoplamento das PV ao trato vocal a exemplo de outros importantes estudos (Mergell *et al.*, 2000; Dollinger *et al.*, 2002; Tao *et al.*, 2007; Qin *et al.*, 2009) que assumiram que as vibrações das PV são independentes do trato vocal. Esta suposição nem sempre é válida como ressaltam os trabalhos (Titze & Story, 1997; Jiang *et al.*, 2000) que descrevem algumas interações acústicas entre as pregas e a porção inferior do trato vocal. Estas interações podem interferir na pressão desta região e assim potencialmente contribuir para o efeito Bernoulli que mantém a vibração das pregas. A simplificação adotada geralmente é aceita em casos onde são analisados indivíduos emitindo alguns tipos determinados de vogais sustentadas. Nos próximos parágrafos são descritas as implementações do IF72 e SH95.

Modelo de Ishizaka e Flanagan (IF72)

A implementação do IF72 pode ser feita de dois modos: (i) de maneira contínua, através da solução das equações diferenciais do modelo, ou (ii) por aproximação das derivadas usando equações de diferença. Neste último caso, as derivadas são substituídas pela relação (3.1)

$$\dot{f}(t_n) \cong \frac{f(t_n) - f(t_{n-1})}{t_n - t_{n-1}} = \frac{f_n - f_{n-1}}{T} \quad (3.1)$$

onde T é o tempo de amostragem dos sinais produzidos pelo modelo. Neste trabalho utilizou-se o segundo modo por considerar que sua implementação é mais direta e rápida haja visto que programas de simulação numérica utilizam funções com um alto custo computacional para resolver analiticamente as equações diferenciais deste modelo. Ainda, a taxa de amostragem escolhida²³ foi de 4 kHz para acompanhar as amostras de videolaringoscopia.

A primeira etapa da implementação do modelo começa com o cálculo do fluxo glotal (U_g) através de (3.2).

$$\begin{aligned} & (R_{k1(n)} + R_{k2(n)})|U_{g(n)}|U_{g(n)} + (R_{v1(n)} + R_{v2(n)})U_{g(n)} + (L_{g1(n)} + L_{g2(n)})\frac{U_{g(n)} - U_{g(n-1)}}{T} \\ & + L_1\frac{U_{g(n)} - U_{g(n-1)}}{T} + R_1U_{g(n)} + \frac{T}{C_1}\sum_{n=0}^{n-1}U_{g(n)} - P_s = 0. \end{aligned} \quad (3.2)$$

Os termos desta equação são dados por (3.3). As constantes ρ , c e μ indicam, respectivamente, a densidade do ar, velocidade do som e viscosidade do ar.

$$\begin{aligned} A_{gi} &= (A_{g0i} + 2L_g x_i); & L_{gi} &= \frac{\rho d_i}{A_{gi}}, \\ R_{k1} &= \frac{0.19\rho}{A_{g1}}; & L_1 &= \frac{\rho L}{2A_1}; \\ R_{k2} &= \frac{\rho \left[0.5 - \frac{A_{g2}}{A_1} \left(1 - \frac{A_{g2}}{A_1} \right) \right]}{A_{g2}^2}; & C_1 &= \frac{LA_1}{\rho c^2}; \\ R_{vi} &= \frac{12\mu l_g^2 d_i}{A_{gi}^3}; & R_1 &= \frac{S}{2A_1} \sqrt{\rho \mu \frac{\sqrt{k_1/m_1}}{\pi}}. \end{aligned} \quad (3.3)$$

²³ Os autores do modelo IF72 recomendam uma frequência superior a 10 kHz com o propósito de se produzir um pequeno valor de T (ver (3.1)) e tornar a equação mais próxima possível do ideal de derivada evitando pequenos artefatos na solução das equações.

Outras importantes constantes associadas ao modelo são A_{g01} e A_{g02} . Estas indicam a área inicial formada entre as PV quando as mesmas estão em repouso. O índice i designa a porção inferior ($i=1$) ou superior ($i=2$) das PV (Figura 2.7(a)) e n as iterações da equação realizadas durante o período de simulação. Os valores iniciais de U_g são considerados nulos e a área glotal A_{gi} é inicializada como $A_{gi} = A_{g0i} = 5.10^{-6} \text{m}^2$. Em consequência disto, as PV estão consideradas em repouso na primeira iteração e seus deslocamentos são inicializados como $x_i = x_1 = x_2 = 0,0 \text{m}$. Estas e outras constantes do modelo são apresentadas na Tabela 3.1.

A primeira amostra de U_g é calculada por (3.2) e usada para estimar as amostras

Tabela 3. 1 – Constantes do modelo IF72 e seus valores no sistema internacional de medidas.

Constante	Valor padrão	Constante	Valor padrão
P_s	800 Pa	d_2	$0,5.10^{-3} \text{ m}$
m_1	$1,25.10^{-4} \text{ Kg}$	A_1	5.10^{-4} m^2
m_2	$0,25.10^{-4} \text{ Kg}$	A_{g01}	5.10^{-6} m^2
k_1	80 N/m	A_{g02}	5.10^{-6} m^2
k_2	8 N/m	L	16.10^{-2} m
k_c	25 N/m	S	$7,85.10^{-2} \text{ m}^2$
h_1	$3k_1$	A_{TT}	25
h_2	$3k_2$	$\eta_{h1} \text{ e } \eta_{h2}$	500
r_1	$0,2(m_1 k_1)^{\frac{1}{2}}$	$\eta_{k1} \text{ e } \eta_{k2}$	100
r_2	$1,2(m_2 k_2)^{\frac{1}{2}}$	ρ	$1,29 \text{ kg/m}^3$
L_g	$1,4.10^{-2} \text{ m}$	μ	$1,86.10^{-5} \text{ N.s/m}^2$
d_1	$2,5.10^{-3} \text{ m}$	c	350 m/s

iniciais de forças que agem sobre as massas das pregas e seus respectivos deslocamentos (x_1 e x_2) produzidos. Em (3.4) são ilustradas as equações gerais do movimento do modelo de duas massas na sua forma discretizada.

$$\begin{aligned}
\frac{m_1}{T^2}(x_{1(n)} - 2x_{1(n-1)} + x_{1(n-2)}) + \frac{r_1}{T}(x_{1(n)} - x_{1(n-1)}) + k_c(x_{1(n-1)} - x_{2(n-1)}) + k_1 x_{1(n-1)} + I_1 + N_1 &= F_1 \\
\frac{m_2}{T^2}(x_{2(n)} - 2x_{2(n-1)} + x_{2(n-2)}) + \frac{r_2}{T}(x_{2(n)} - x_{2(n-1)}) + k_c(x_{2(n-1)} - x_{1(n-1)}) + k_2 x_{2(n-1)} + I_2 + N_2 &= F_2
\end{aligned} \tag{3.4}$$

As amostras $x_{i(n)}$, para $n < 0$, são consideradas nulas (*i.e.*, pregas em repouso). Em outros casos, os valores dos deslocamentos $x_{i(n)}$ realizados pelas massas estão ligados diretamente à ação das forças F_i que o fluxo de ar U_g exerce nas paredes das PV. O cálculo destas forças está ligado às diferentes configurações físicas que as PV assumem durante o ciclo glotal e seus valores estão indicados na Tabela 3.2. As constantes $x_{i \min}$, para $i = 1, 2$, são matematicamente equivalentes a $x_{i \min} = -A_{g0i}/(2L_g)$ e indicam as coordenadas de impacto entre as massas na sua porção inferior e superior.

Em algumas situações descritas na Tabela 3.2, é necessário calcular as pressões P_{m1} e P_{m2} que o fluxo de ar exerce sobre a área $L_g d_i$ das pregas. Estes valores são dados por (3.5).

$$\begin{aligned} P_{m1(n)} &= P_s - 1.37 \frac{\rho}{2} \left(\frac{U_{g(n)}}{A_{g1(n-1)}} \right)^2 - \frac{1}{2} \left\{ R_{v1(n)} U_{g(n)} + \frac{L_{g1(n)}}{T} (U_{g(n)} - U_{g(n-1)}) \right\}; \\ P_{m2(n)} &= P_{1(n)} - \frac{1}{2} \left\{ (R_{v1(n)} + R_{v2(n)}) U_{g(n)} + (L_{g1(n)} + L_{g2(n)}) \left(\frac{U_{g(n)} - U_{g(n-1)}}{T} \right) \right\} \\ &\quad - \frac{\rho}{2} U_{g(n)}^2 \left(\frac{1}{A_{g2(n-1)}^2} - \frac{1}{A_{g1(n-1)}^2} \right). \end{aligned} \quad (3.5)$$

Além de calcular F_i para a solução de (3.4), faz-se necessário que também sejam avaliadas as funções que indicam as não-linearidades adicionais do modelo (N_i) e o impacto (I_i). No caso de N_i , esta é dada pela relação (3.6) e incorpora não-linearidades às molas k_i do modelo. A constante η_{ki} de (3.6) indica o coeficiente de não-linearidade da mola k_i .

$$N_{i(n)} = k_i \eta_{ki} \cdot x_{i(n-1)}^3 \quad (3.6)$$

Já as equações de impacto I_i são originalmente definidas como em (3.7).

$$I_{i(n)}(x_{i(n-1)}) = \begin{cases} 0, & x_{i(n)} > -\frac{A_{g0i}}{2L_g} \\ \left(x_{i(n-1)} + \frac{A_{g0i}}{2L_g} \right) + \eta_{hi} \left(x_{i(n-1)} + \frac{A_{g0i}}{2L_g} \right)^3, & x_{i(n)} \leq -\frac{A_{g0i}}{2L_g} \end{cases} \quad (3.7)$$

Novamente, a constante η_{hi} indica o coeficiente de não-linearidade utilizado para descrever a força de contato entre as pregas durante a colisão. Como mencionado previamente, a constante $-A_{g0i}/(2L_g)$ indica a coordenada em x onde acontece a colisão.

O modelo IF72 é usado mais adiante neste capítulo para estudar o comportamento modal das PV utilizando uma nova abordagem numérica que leva em consideração as não-linearidades do modelo. Esta abordagem considera que para o cálculo dos modos não-lineares, o modelo tenha equações suaves e não apresente descontinuidades ou mudanças abruptas nas funções do sistema. Contudo, percebe-se em (3.7) que o IF72 não atende a este requerimento. Para contornar esta questão, as equações de impacto foram remodeladas nesta pesquisa de modo a introduzir uma suavidade no ponto de impacto e tentar preservar ao máximo sua dinâmica original. Inicialmente, considerou-se utilizar modelos de suavização de impactos baseados na técnica de Zuralev (Haiwu *et al.*, 2009), funções trigonométricas hiperbólicas (Manevitch & Gendelman, 2008) ou métodos de interpolação polinomial com propósitos de suavização. Testes preliminares indicaram que a última estratégia apresentou resultados mais eficientes e por isto foi escolhida.

A suavização da função de impacto foi desenvolvida utilizando o polinômio de Hermite entre a transição das duas equações de (3.7). Para isto, é considerada uma região de suavização definida entre os intervalos $[a-\Delta, a+\Delta]$, onde a indica o ponto de transição e no caso do IF72 coincide com a coordenada de impacto (*i.e.*, $a=-A_{g0i}/(2L_g)$) e Δ a largura do intervalo de suavização. A equação (3.8) ilustra a nova abordagem de

Tabela 3. 2 – Tabela de distribuição de forças nas pregas vocais.

Deslocamentos	F_1	F_2
$x_1 > x_{1min}$ e $x_2 > x_{2min}$	$P_{m1}L_g d_1$	$P_{m2}L_g d_2$
$x_1 \leq x_{1min}$ e $x_2 > x_{2min}$	$P_s L_g d_1$	0
$x_1 > x_{1min}$ e $x_2 \leq x_{2min}$	$P_s L_g d_1$	$P_s L_g d_2$
$x_1 \leq x_{1min}$ e $x_2 \leq x_{2min}$	$P_s L_g d_1$	0

impacto empregando a técnica de suavização utilizada.

$$I_{i(n)}(x_{i(n-1)}) = \begin{cases} 0, & x_{i(n)} \geq a + \Delta \\ P(t(x_{i(n-1)})) & a - \Delta < x_{i(n)} < a + \Delta \\ \left(x_{i(n-1)} + \frac{A_{g0i}}{2L_g}\right) + \eta_{hi} \left(x_{i(n-1)} + \frac{A_{g0i}}{2L_g}\right)^3 & x_{i(n)} \leq a - \Delta \end{cases} \quad (3.8)$$

A função $t(x_i)$ é representada em (3.9) e usada na definição do polinômio de interpolação $P(t(x_i))$ definido em (3.10).

$$t(x) = \frac{x - x_k}{x_{k+1} - x_k} \quad (3.9)$$

$$P(t) = h_{00}(t)p_k + h_{10}(t)(x_{k+1} - x_k)s_k + h_{01}(t)p_{k+1} + h_{11}(t)(x_{k+1} - x_k)s_{k+1} \quad (3.10)$$

Os termos p_k e p_{k+1} representam, em (3.10), os valores do impacto nos pontos $x_k=a-\Delta$ e $x_{k+1}=a+\Delta$, respectivamente. Nesta mesma equação, s_k e s_{k+1} são os valores das derivadas nos mesmos pontos x_k e x_{k+1} e as funções $h(t)$ são determinadas em (3.11).

$$\begin{aligned} h_{00}(t) &= 2t^3 - 3t^2 + 1 \\ h_{10}(t) &= t^3 - 2t^2 + t \\ h_{01}(t) &= -2t^3 + 3t^2 \\ h_{11}(t) &= t^3 - t^2 \end{aligned} \quad (3.11)$$

Embora (3.8) mantenha um formato *step-wise*, é alcançado um efeito de suavidade, pois o polinômio $P(t)$ define uma região de transição entre os pontos $a-\Delta < x < a+\Delta$ sem alterar significativamente os valores de força de impacto conforme ilustra a Figura 3.6. Note que nesta figura o comportamento não-linear é visível no modelo original. Contudo, como deslocamentos superiores a 2 mm são geralmente irreais, a Figura 3.6(b) ilustra parte da região de deslocamento de interesse e exibe graficamente a expressão original (3.7) e a modificada em (3.8) para fins de comparação. Durante as simulações foi empregado um $\Delta=1.10^{-5}\text{m}$.

Depois de calculados os valores de F_i , N_i e I_i , os deslocamentos x_i são estimados por (3.4) e novos valores de área glotal A_{gi} são gerados e novamente reintroduzidos em (3.2) para o cálculo da próxima amostra de U_g . O processo é repetido até que o tempo de simulação desejado seja alcançado.

Modelo de Steinecke e Herzel (SH95)

Embora seja uma simplificação do IF72, a extensão das simplificações foram abrangentes o suficiente para que o SH95 pudesse ser empregado em várias pesquisas que buscavam equações mais simples para cálculos mais extensivos. Seu amplo uso deu um ‘caráter próprio’ ao SH95 na medida em que este deixa de ser abordado como uma simplificação e passa a ser descrito como um modelo próprio com características dinâmicas específicas.

Para implementar o modelo neste trabalho, será considerada a representação em espaço de estados (Chen, 1999) de suas equações dinâmicas. Esta forma torna mais evidente na formulação matemática do modelo os estados que são representados em uma forma vetorial. Esta representação é preferível em situações onde se deseja evidenciar a importância das condições iniciais dos estados e analisar seus efeitos. A relação (3.12) exhibe a forma das equações do modelo no espaço de estados e é uma representação alternativa de (2.1). As constantes do modelo são as mesmas ilustradas na Tabela 3.1.

$$\begin{bmatrix} \dot{x}_1 \\ \dot{v}_1 \\ \dot{x}_2 \\ \dot{v}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\left(\frac{k_1 + k_c}{m_1}\right) & -\frac{r_1}{m_1} & \frac{k_c}{m_1} & 0 \\ 0 & 0 & 0 & 1 \\ \frac{k_c}{m_2} & 0 & -\left(\frac{k_2 + k_c}{m_2}\right) & -\frac{r_2}{m_2} \end{bmatrix} \begin{bmatrix} x_1 \\ v_1 \\ x_2 \\ v_2 \end{bmatrix} + \begin{bmatrix} 0 \\ I_1(x_1) + N_1(x_1) + F_1(x_1) \\ 0 \\ I_2(x_2) + N_2(x_2) \end{bmatrix}. \quad (3.12)$$

As variáveis x_i e v_i representam os estados do sistema e correspondem às amplitudes de oscilação das massas e suas velocidades, respectivamente. Novamente, os termos $I_i(x_i)$, $N_i(x_i)$ e $F_i(x_i)$ indicam, respectivamente, as equações de impacto entre as massas, as não-linearidades adicionais ao modelo e a força de excitação aerodinâmica. Como o SH95 desconsidera não-linearidades adicionais²⁴, temos que $N_i(x_i)=0$. Já para os demais termos, são válidas as relações (3.13) e (3.14).

$$I_i(x_i) = -\theta(-a_i)3k_i \frac{A_{gi}}{2L_g}. \quad (3.13)$$

²⁴ As principais fontes de não-linearidade do SH95 estão ligadas à excitação e ao impacto (função afim).

$$F_1(x_i) = \frac{L_g d_1 P_s}{m_1} \left[1 - \Omega(a_{min}) \left(\frac{A_{gmin}}{A_{g1}} \right) \right] \Omega(A_{g1}) \quad (3.14)$$

Com,

$$\begin{aligned} \theta(x) &= \begin{cases} 1, & x > 0, \\ 0, & x \leq 0. \end{cases} \\ \Omega(x) &= \begin{cases} \tanh[50(x/A_{g01})], & x > 0, \\ 0, & x \leq 0. \end{cases} \\ A_{gmin} &= \min(A_{g1}, A_{g2}). \end{aligned} \quad (3.15)$$

As variáveis A_{g1} e A_{g2} indicam a área formada entre as pregas na sua porção inferior e superior, respectivamente, e são dadas em (3.3). O cálculo do fluxo glotal U_g é indicado por (3.16).

$$U_g = \sqrt{\frac{2P_s}{0.00113}} \cdot A_{gmin} \cdot \Omega(A_{gmin}) \quad (3.16)$$

Para a resolução de (3.12), é necessário que se estipule um vetor contendo os valores iniciais dos estados para definir as condições de contorno necessárias para resolução das equações diferenciais do modelo. Os valores padrões adotados para este

Figura 3.6 – O eixo horizontal dos gráficos representa o deslocamento lateral das pregas e o vertical a força de impacto entre as massas inferiores do modelo IF72. (a) Representação gráfica da equação de impacto (3.7) original do modelo. Em (b) são comparadas as curvas da equação de impacto original (linha preta contínua) e a versão suavizada (linha pontilhada cinza) dada por (3.8) dentro de uma faixa de valores de deslocamento fisiologicamente compatíveis com a realidade. Para destacar a região suavizada, foi considerado um $\Delta=6.10^{-5}$ m para este gráfico. O ponto indicado pela seta mostra a inflexão indesejada da equação original na coordenada de colisão. Unidades de força dadas em Newton.

vetor foram: $[x_1(0) = 1.10^{-4} \text{ m}, v_1(0) = 0.0, x_2(0) = 1.10^{-4} \text{ m}, v_2(0) = 0.0]$. A resolução destas equações foi feita através da função *ode45* do MatLab que emprega um método de Runge-Kutta de quarta ordem com tamanho de passo de integração variável.

Outra importante questão relativa tanto ao modelo SH95 quanto ao IF72 diz respeito a como a área glotal é representada pelos modelos. Conforme destaca a Figura 3.7(a), são produzidos dois sinais de área glotal: um para cada massa. Como revela a relação (3.15), a área glotal do modelo é a menor das duas áreas produzidas conforme esboça o traço negro neste gráfico. Sendo assim, este será o sinal usado para comparar a saída do modelo com o SGE das imagens. Vale ainda observar que ambos modelos geram áreas negativas, o que em termos práticos parece ser incoerente. A solução encontrada é considerar nestes casos que a prega está fechada e, portanto, não existe área glotal (*i.e.*, $A_{gi}=0$). As Figuras 3.7(b) e 3.7(c) ilustram esquematicamente que a área registrada por uma câmera de videolaringoscopia faz um processo similar ao adotado pelo modelo para definir a área glotal entre as pregas. Isto porque quando as imagens são segmentadas, a área tomada como referência é a de cor próxima ao negro (área pouca iluminada) que circunda as bordas das pregas. Devido a pouca luminosidade desta área que corresponde à glote, é possível ver com melhor nível de detalhamento as bordas das pregas e estas bordas remetem à menor distância entre as porções inferiores ou superiores das pregas como esboçam as Figuras 3.7(b) e 3.7(c).

3.1.4 – Otimização dos parâmetros mecânicos dos modelos de pregas vocais

Cada indivíduo tem suas próprias características de voz que refletem, dentre outras coisas, a dinâmica de suas PV. Na pesquisa científica da voz é importante encontrar mecanismos para extrair estas características dinâmicas possibilitando uma análise mais criteriosa dos fatores capazes de influenciar diretamente a voz, avaliar o estado das PV e entender de forma mais clara complexos fenômenos produzidos na

larínge e patologias diversas. São inúmeras as aplicações que justificam a elaboração de um método voltado à reconstrução das propriedades dinâmicas das PV para um determinado indivíduo. Até então, boa parte dos trabalhos científicos empregavam modelos mecânicos usando somente parâmetros padrões similares aos ilustrados na Tabela 3.1. Com isto, uma série de restrições e inviabilidades eram impostas aos estudos e determinados grupos específicos de indivíduos não podiam ser avaliados, pois seus padrões de dinâmica eram diferentes e não haviam métodos capazes de capturar de maneira quantitativa esta diferença.

Estimar diretamente as propriedades dinâmicas das PV *in vivo* é uma tarefa desafiadora devido à dificuldade de acesso a estas estruturas e a sua sensibilidade. Sendo assim, é necessário um meio, ainda que indireto, de avaliar estas propriedades. Neste sentido, o uso dos padrões de vibração extraídos das pregas (via exame de videolaringoscopia) e modelos mecânicos podem ser empregados para esta finalidade. A idéia básica é utilizar o SGE de indivíduos e fazer com que um método de otimização altere as principais propriedades mecânicas de um modelo de PV de tal forma que este possa incorporar o padrão apresentado pelo SGE e assim reproduzir suas características

Figura 3.7 – (a) Área glotal gerada pelas massas inferiores (linha pontilhada em cinza) e superiores (linha tracejada em cinza) do modelo SH95 com parâmetros padrões. O traço em preto ilustra a área glotal total calculada para o modelo. As imagens (b) e (c) esboçam esquematicamente o campo de visão de uma câmera para diferentes circunstâncias do ciclo glotal.

dinâmicas. As massas (m_1 e m_2), coeficientes de amortecimento (k_1 e k_2), constante de acoplamento entre massas (k_c) e pressão sub-glotal (P_s) representam alguns dos mais importantes parâmetros do modelo capazes de definir o comportamento básico das pregas (Mergell *et al.*, 2000; Tao *et al.*, 2007). Estas constantes serão matematicamente representadas por um grupo de parâmetros Φ que serão representados na forma de um vetor tal que $\Phi := [m_i, k_i, k_c, r_i, P_s \mid i=1, 2]$. Estes valores determinarão a saída do modelo mecânico e seu comportamento.

O método de otimização deve buscar por um grupo de valores adequados para o vetor Φ de forma que o modelo possa reproduzir e prever, com relativa precisão, o sinal extraído das imagens de um determinado indivíduo. Para este propósito, o procedimento de otimização combina algoritmos genéticos (Goldberg, 1989) e o método simplex de otimização (Bonnans *et al.*, 2006). Basicamente, o método compara o SGE do indivíduo com o *sinal glotal simulado* (SGS) pelo modelo mecânico usando um vetor Φ produzido pelo procedimento de otimização na iteração corrente.

Como as equações do modelo mecânico apresentam descontinuidades e podem apresentar um espaço de busca não convexo, na primeira das duas etapas do procedimento de otimização é aplicado um algoritmo genético para buscar uma solução aproximada. Seu uso na primeira etapa justifica-se pelo fato de que ele executa uma busca global e que por isto pode evitar mínimos locais. Na segunda etapa, esta solução aproximada serve como solução inicial para o simplex que deverá refiná-la. Esta estratégia de utilizar dois diferentes métodos deve ajudar a reduzir o tempo gasto no processo de otimização e evitar o problema de convergência em mínimos indesejados.

O algoritmo genético empregado nesta pesquisa é esquematicamente ilustrado na Figura 3.8. Nele é utilizada a regra da roleta para seleção dos indivíduos da população. Neste caso, cada indivíduo representa uma potencial solução: ou seja, um vetor de parâmetros Φ . Os cromossomos do indivíduo representam os oito parâmetros mecânicos

contidos no vetor e todos os indivíduos de uma mesma iteração formam a dita população. Foi também empregada a regra do elitismo para acelerar a convergência da solução. Com isto, é garantido que em cada geração o melhor indivíduo seja mantido já que estatisticamente a regra da roleta não assegura a permanência deste.

Foi convencionada uma população de dois mil indivíduos e o algoritmo foi programado para processar até 50 gerações com o objetivo de evitar um tempo de processamento excessivo. O espaço de busca compreendeu o intervalo $[0,1 \times 10^{-4}; 0,1 \times 10^{-5}; 10; 1; 2; 0,1 \times 10^{-4}; 0,1 \times 10^{-4}; 500]$ a $[1,0 \times 10^{-3}; 5,0 \times 10^{-4}; 200; 50; 150; 0.4; 0.4; 40000]$ para os oito parâmetros (leia-se cromossomos) de cada vetor Φ (leia-se indivíduo). De acordo com (Mergell *et al.*, 2000), um espaço de busca similar a este deve cobrir a maioria dos casos de vozes normais e patológicas. O processo de otimização empregou ambos os modelos mecânicos implementados nesta pesquisa. Contudo, é importante destacar que os coeficientes de amortecimento r_i do IF72 não foram estimados para o referido modelo já que estes parâmetros são definidos analiticamente pela equação listada na Tabela 3.1. Os parâmetros do modelo foram otimizados considerando-se todas as amostras.

A primeira população do algoritmo genético é randomicamente gerada a partir dos intervalos descritos no parágrafo anterior. Os valores padrões de probabilidades de cruzamento e números de indivíduos selecionados a cada geração são de 16% e 75%,

Figura 3.8 – Algoritmo genético empregado no primeiro passo do procedimento de otimização.

respectivamente. Contudo, estas taxas podem ser alteradas para cada caso de modo a assegurar uma convergência mais eficiente. Vale ainda destacar que na implementação deste algoritmo nesta pesquisa, os indivíduos descartados dão lugar a novos indivíduos completamente novos que são gerados randomicamente e não a indivíduos resultantes do cruzamento entre os que foram selecionados. Ao contrário das implementações clássicas, esta alteração garante uma diversificação genética grande na população. Naturalmente, a aptidão média desta população é comprometida com esta estratégia. Ainda, o processo de mutação é aplicado aos indivíduos selecionados para também garantir alguma diversidade genética à população em uma mesma iteração. Foi considerada uma taxa mutação de 0.5%. Vale destacar que com esta nova abordagem motivada pela necessidade de se ter uma grande variabilidade genética a nível cromossômico, o processo de mutação perde significativamente sua importância. Contudo, por questões de flexibilidade, ele é mantido no algoritmo.

No algoritmo genético, as operações de seleção e cruzamento em cada geração dão origem a um processo evolucionário produzindo indivíduos mais aptos a cada geração. Este processo continua até que todas as gerações sejam processadas ou o indivíduo mais apto assuma um valor de aptidão superior a um patamar estabelecido pelo usuário. Uma alta aptidão indica que o modelo mecânico das PV produz um comportamento similar ao das pregas registradas no vídeo. Nesta pesquisa a função objetiva que avalia a aptidão dos indivíduos é definida em (3.17) como o erro quadrático médio entre o sinal extraído x_e (*i.e.*, SGE) e o simulado x_s (*i.e.*, SGS). A constante n indica o número de amostras extraídas do vídeo (*i.e.*, número de quadros segmentados).

$$\psi_1 = \frac{1}{n} \sum_{k=1}^n (x_e[k] - x_s[k])^2. \quad (3.17)$$

Na segunda etapa do procedimento de otimização, a solução encontrada pelo algoritmo genético é usada como entrada para o simplex para refinar a solução. Este é

um método comum empregado para a solução de problemas de otimização usando técnicas de programação linear. Por ser uma técnica bem conhecida e difundida, sua explicação é omitida²⁵. A função objetiva usada nesta etapa é diferente da empregada na etapa anterior, pois aqui pretende-se aproximar o espectro do sinal simulado $\hat{E}(w)$ com o extraído $E(w)$. A Equação (3.18) avalia a diferença espectral logarítmica entre os sinais. Neste caso, n representa o número de componentes de frequência que os sinais possuem.

$$\psi_2 = \sqrt{\frac{1}{n} \sum_{w=1}^n \left(10 \log_{10} \frac{E(w)}{\hat{E}(w)} \right)^2} \quad (3.18)$$

A utilização de uma função objetiva similar à Equação (3.17) poderia não levar em conta detalhes importantes e específicos do pulso glotal como, por exemplo, a fase de fechamento. Desta maneira, potencialmente poderia ser produzido um bom “encaixe” entre o SGE e o SGS utilizando esta função mesmo que o SGS não produzisse adequadamente o rápido fechamento do ciclo glotal. Este rápido decrescimento (ver gráfico da Figura 3.3) da fase de fechamento tem grande contribuição na produção das altas frequências do pulso glotal e é uma importante característica do sinal produzido na laringe. Para ponderar melhor esta característica do pulso, a comparação entre os sinais extraídos e simulados é feita no domínio da frequência na segunda etapa da otimização e as altas frequências produzidas pelo fechamento são levadas em conta. A função *fminsearch* do MatLab foi utilizada para executar o simplex. O procedimento de otimização foi aplicado para os modelos de SH95 e IF72.

3.1.5 – Validação dos procedimentos de extração da dinâmica

Esta seção tem o propósito de descrever os procedimentos utilizados para descrever o grau de exatidão do método desenvolvido para extrair a dinâmica *in vivo*

²⁵ Informações sobre este método podem ser consultadas em (Bonnans & Gilbert, 2006).

das PV. Para isto é analisada a conformidade (*i.e.*, exatidão) dos resultados encontrados pelo método com seus respectivos valores verdadeiros. São geradas sinteticamente 50 amostras de sinais cujos parâmetros reais foram estipulados pelo autor e são totalmente conhecidos. Tentou-se produzir a faixa mais ampla possível de valores no espaço de busca capazes de gerar oscilações normais e regulares. O grau de conformidade foi avaliado pela relação (3.19) que determina o valor médio do erro relativo percentual entre os n parâmetros $p(k)$ conhecidos, com os parâmetros $p'(k)$ estimados pelo método. Cada parâmetro representa um elemento do vetor Φ .

$$E_p(\Phi) = \frac{1}{n} \sum_{k=1}^n \frac{|p(k) - p'(k)|}{p(k)} 100\% \quad (3.19)$$

De maneira adicional, é avaliada a qualidade dos sinais simulados em relação aos seus correspondentes sintéticos a partir da relação (3.20). Nela, é estimada a diferença absoluta média entre cada ponto do sinal sintético $x(k)$ e seu correspondente simulado $x'(k)$. O termo x_p indica o máximo valor de pico do sinal sintético gerado após seu regime transiente e m a quantidade de pontos avaliados.

$$E_s(x, x') = \frac{1}{m} \sum_{k=1}^m \frac{|x(k) - x'(k)|}{x_p} 100\% \quad (3.20)$$

3.2 – Avaliação da dinâmica das pregas vocais

3.2.1 – Análise modal não-linear

Tradicionalmente, um oscilador de n graus de liberdade com comportamento linear pode ser decomposto em outros n osciladores independentes com um grau de liberdade cada. Vale ressaltar que o princípio de superposição deve ser satisfeito para que este método possa ser empregado corretamente. Como visto, esta decomposição é feita pela análise modal (Worden & Tomlinson, 2001) e cada um dos osciladores resultantes serão chamados de *modos lineares normais* (MLN) do sistema.

Experimentos práticos envolvendo laringes excisadas têm reportado padrões de vibrações não-lineares (Giovanni *et al.*, 1999; Jiang *et al.*, 2003) e uma variedades de

fenômenos não-lineares tais como bifurcações, surgimento de harmônicos e caos (Zhang & Jiang, 2005; Tokuda *et al.*, 2008). Tais fenômenos têm sido atribuídos às propriedades não-lineares das PV (Jiang *et al.*, 2006) que estão vinculadas a três principais fontes: (i) relação não-linear da pressão exercida pelo fluxo de ar que excita as pregas, (ii) propriedades estruturais de seus tecidos cuja relação tensão mecânica e deformação não correspondem linearmente e (iii) ao impacto entre as pregas (Berry *et al.*, 1994). Contudo, tais investigações práticas não têm apontado os fundamentos teóricos destes fenômenos.

Outros resultados teóricos (Steinecke & Herzel, 1995; Jiang *et al.*, 2001) envolvendo modelos simplificados padrões das PV, reportaram fenômenos não-lineares similares. Em (Thomson *et al.*, 2005) foi explorada a transferência aerodinâmica de energia do pulso glotal para as pregas durante a fonação usando dados sintéticos e modelos mecânicos padrões de PV. Os resultados mostraram uma relação não-linear e identificaram potenciais efeitos relacionados à dessincronização do movimento das pregas. Contudo, não foram encontrados na literatura trabalhos que avaliaram de forma isolada as outras duas principais fontes de não-linearidades associadas às propriedades estruturais das pregas (*i.e.*, características elásticas dos tecidos e o impacto entre as paredes das pregas) e as possíveis contribuições que estas promovem nos movimentos das PV, em especial na sua frequência de vibração.

Neste contexto, a análise modal pode revelar interessantes características dinâmicas de um sistema e explorar seu comportamento estrutural para propósitos distintos. Particularmente, os *modos não-lineares normais*²⁶ ou MNNs (Rosenberg, 1966; Kerschen *et al.*, 2009) oferecem um meio efetivo para interpretar uma ampla classe de sistemas dinâmicos não-lineares e fenômenos relacionados às suas não-linearidades. Mesmo embora sua teoria seja considerada recente e ainda em consolidação, os MNNs podem ser grosseiramente entendidos como uma extensão não-

²⁶ Termo originalmente conhecido como *nonlinear normal modes*.

linear dos MLNs²⁷, mesmo embora muitas de suas propriedades sejam essencialmente diferentes de seu correspondente linear. A grande diferença está no fato que o princípio da superposição não é válido para os MNNs e logo estes não podem ser decompostos em osciladores menores com um grau de liberdade cada. Contudo, a vantagem dos modos não-lineares está associada a sua habilidade de analisar um sistema não-linear cujo comportamento desvia significativamente da linearidade e ainda é capaz de lidar com não-linearidades estruturais fortes. Neste sentido, MNNs podem ajudar a entender como curvas modais²⁸ deformam sobre a ação da(s) não-linearidade(s).

A análise modal será empregada para elucidar a ocorrência de vibrações não-lineares e avaliar as não-linearidades estruturais das PV considerando apenas as propriedades mecânicas de seus tecidos e seu impacto. Para isto, será desconsiderado a ação de forças externas. Em particular, é interesse desta pesquisa desvendar a relação entre a frequência de oscilação das PV e o seu nível de energia mecânica e avaliar como possíveis não-linearidades podem influenciar este sistema em diferentes amplitudes de vibração (*i.e.*, níveis de energia). Vale destacar que ao contrário dos sistemas lineares, a função de resposta em frequência para sistemas não-lineares pode variar segundo o tipo de entrada e sua energia mecânica (Worden & Tomlinson, 2001). Ainda, para que os comportamentos não-lineares do sistema possam ser bem caracterizados e baseados em dados reais, serão utilizados nesta parte da pesquisa os vinte indivíduos analisados pelo método de extração de dinâmica descrito anteriormente.

Uma avaliação direta dos modos de vibração das pregas vocais usando pregas excisadas ou métodos *in vivo* pode ser inviável e problemática devido às pequenas dimensões das PV (na ordem de 1 cm). Isto dificulta o posicionamento de acelerômetros pra medir suas trajetórias e seu número seria limitado. O próprio peso dos acelerômetros poderia modificar os modos. Além disto, a fixação destes seria

²⁷ Esta comparação só é válida em ocasiões onde a amplitude do movimento (ou energia) não é suficiente para excitar as não-linearidades do sistema.

²⁸ As curvas modais associam os deslocamentos x_1 e x_2 (para um oscilador com dois graus de liberdade).

invasiva assim como os procedimentos de avaliação e extração da informação necessária para se determinar de maneira prática o comportamento modal das pregas. Neste cenário, o uso de métodos teóricos para investigar os modos do sistema, com a ajuda de um modelo determinístico do sistema, viabiliza a análise modal. O uso de modelos extraídos de casos de indivíduos reais (através do procedimento de otimização anteriormente descrito) pode contribuir para uma maior coerência e embasamento dos resultados.

O modelo de pregas empregado no cálculo dos MNNs é o IF72, pois este permite uma melhor caracterização das não-linearidades já que incorpora uma série de parâmetros que contemplam esta característica. A próxima subseção descreve o método numérico para cálculo dos MNNs.

Cálculo dos modos não-lineares das pregas vocais

Rosenberg (1966) definiu um MNN como “uma oscilação síncrona onde todos os pontos de deslocamento do material que o constituem podem ser expressos em termos de uma referência de deslocamento simples”. Em sistemas osciladores conservativos, o MNN pode ser considerado como um movimento periódico que, sob a ação de não-linearidades, estabelece uma relação não-linear entre suas coordenadas modais durante o movimento. A Figura 3.9 ilustra duas séries temporais geradas por um sistema linear e suas respectivas MLNs. Uma terceira série temporal de um sistema não-linear e seus correspondentes MNNs são também ilustrados. As linhas de modo (ou curvas modais) dos sistemas lineares são vistas como retas no espaço modal, enquanto que no caso de um sistema não-linear, sua curva descreve uma ligação não-linear entre as coordenadas do movimento.

De forma a contornar a restrição de Rosenberg relativa a sistemas conservativos, será considerada nesta pesquisa somente a resposta não amortecida do sistema (*i.e.*, $r_1=r_2=0$). Esta aproximação é justificável devido ao fato que a dinâmica de um sistema amortecido pode ser frequentemente interpretada baseada na estrutura topológica dos

MNNs de seu correspondente não-amortecido, haja visto que ambos possuem MNNs semelhantes (Kerschen *et al.*, 2009; Peeters *et al.*, 2009). Ainda, para enfatizar as características estruturais dos tecidos das pregas e do impacto, será considerada uma resposta livre do sistema mecânico (*i.e.*, $F_1=F_2=0$). Isto é particularmente interessante já que fará com que a análise não-linear seja limitada às estruturas das PV, desconsiderando sua excitação que também tem características não-lineares. Levando em conta estes fatores, as equações do movimento podem ser vistas na forma (3.21).

$$M\ddot{x}(t) + Kx(t) + E\{x(t)\} = 0 \quad (3.21)$$

Na equação anterior, M indica a matriz de massas, K a matriz de molas e E o vetor contendo as não-linearidades do modelo. Para calcular os MNNs nesta pesquisa, foi usada uma abordagem baseada no trabalho (Peeters *et al.*, 2009). Basicamente, este método combina as técnicas de continuação de *shooting* e *pseudo-arclength* (Deuffhard

Figura 3.9 – Série temporal correspondendo ao modo “em fase” (a) e “fora de fase” (b) do modo de um sistema linear descrito pelas equações $\ddot{x}_1 + (2x_1 - x_2) = 0$ e $\ddot{x}_2 + (2x_2 - x_1) = 0$ (x_1 : linha contínua; x_2 : linha pontilhada). Em (d) e (e) são ilustrados suas respectivas curvas modais. Finalmente, (c) ilustra uma série temporal do sistema não-linear cuja equação do movimento é $\ddot{x}_1 + (2x_1 - x_2) + 0.5x_1^3 = 0$ e $\ddot{x}_2 + (2x_2 - x_1) = 0$. A curva modal correspondente é mostrada em (f). As condições iniciais são $[x_1(0)=0.607, x_2(0)=0.607, \dot{x}_1=0, \dot{x}_2=0]$ em (a); $[x_1(0)=-0.807, x_2(0)=0.807, \dot{x}_1=0, \dot{x}_2=0]$ em (b), e $[x_1(0)=5.77, x_2(0)=34.77, \dot{x}_1=0, \dot{x}_2=0]$ em (c).

et al., 1976) para resolver o problema das condições de contorno que definem uma família de MNNs de um sistema não-linear.

O método de cálculo exige que as equações de movimento sejam agora convertidas para o espaço de estados considerando a relação $z=[x,\dot{x}]'$. O símbolo $(')$ indica a operação de transposição do vetor. Logo, o sistema é matematicamente descrito como

$$\dot{z} = g(z) = \begin{bmatrix} \dot{x} \\ -M^{-1}\{Kx + E(x)\} \end{bmatrix}. \quad (3.22)$$

Considerando a condição inicial $z_0=[x_0,\dot{x}_0]'$, a solução do sistema dinâmico pode ser escrita como $z(t)=z(t,z_0)$ e a solução periódica z_p , com período mínimo T , é então definida como $z_p(t,z_{p0})=z(t+T,z_{p0})$. Basicamente, o cálculo dos MNNs é feito a partir de uma busca por soluções periódicas a partir da análise das equações do movimento do sistema. Para este propósito, o método de *shooting* é usado para resolver numericamente as condições de contorno definidas pelo requerimento de periodicidade. Este método encontra iterativamente a condição inicial z_{p0} e o período T que satisfazem a condição de movimento periódico. Matematicamente, isto é definido pela função:

$$H(z_{p0}, T) = z_p(z_{p0}, T) - z_{p0} = 0 \quad (3.23)$$

A Equação (3.23) representa a diferença entre a resposta do sistema no tempo T e as condições iniciais. Na primeira iteração é assumida a condição inicial $z_{p0}^{(k)}$ e o período $T^{(k)}$, onde k denota o índice da iteração do processo de *shooting* (*e.g.*, $k=0$ para a primeira iteração). Usando estes valores e métodos de integração numérica, o movimento $z_p^{(0)}(t, z_{p0}^{(0)})$ é calculado na primeira iteração. Os dois valores iniciais são estimados a partir do MLN da porção linear do sistema e a sua correspondente frequência natural. Para o cálculo destes valores iniciais, o IF72 é linearizado e suas equações diferenciais são reescritas em outro domínio como ilustra (3.24).

$$s^4 + \left(\frac{r_1}{m_1} + \frac{r_2}{m_2}\right)s^3 + \left(\frac{k_1 + k_c}{m_1} + \frac{k_2 + k_c}{m_2} + \frac{r_1 r_2}{m_1 m_2}\right)s^2 + \left(\frac{r_1(k_2 + k_c) + r_2(k_1 + k_c)}{m_1 + m_2}\right)s + \left(\frac{k_1 k_c + k_2(k_1 + k_c)}{m_1 m_2}\right) = 0 \quad (3.24)$$

Considerando $r_1=r_2=0$, a equação anterior apresenta duas soluções com frequências positivas que são usadas para a estimativa do período inicial. A solução com menor frequência é usada como estimativa de T para o modo em fase e a segunda para o outro modo. Foi considerado no primeiro MNN a condição inicial $[x_1=1.10^{-5}\text{m}, v_1=0\text{m/s}, x_2=1.10^{-5}\text{m}, v_2=0\text{ m/s}]$ para os movimentos em fase e $[x_1=1.10^{-5}\text{m}, v_1=0\text{m/s}, x_2=-1.10^{-5}\text{m}, v_2=0\text{ m/s}]$ para os fora de fase. Estes baixos valores de condições iniciais introduzem pouca energia no sistema que geralmente não é suficiente para excitar o comportamento não-linear e assim os MLNs são próximos dos MNNs.

Todavia, os valores iniciais encontrados pelos MLNs da iteração inicial raramente satisfazem a condição de periodicidade (3.23). Para ajustá-los, é executada uma iteração do método de aproximação de Newton-Raphson para alcançar a convergência baseando-se nos parâmetros de correção $\Delta z_{p0}^{(0)}$ e $\Delta T^{(0)}$. Estes valores são encontrados através da expansão da função $H(z_{p0}^{(0)} + \Delta z_{p0}^{(0)}, T^{(0)} + \Delta T^{(0)})$ em uma série de Taylor como ilustra (3.25).

$$H(z_{p0}^{(0)}, T^{(0)}) + \left. \frac{\partial H}{\partial z_{p0}} \right|_{(z_{p0}^{(0)}, T^{(0)})} \Delta z_{p0}^{(0)} + \left. \frac{\partial H}{\partial T} \right|_{(z_{p0}^{(0)}, T^{(0)})} \Delta T^{(0)} + \dots = 0 \quad (3.25)$$

Negligenciando os termos de mais alta ordem e usando o índice k do processo de *shooting*, a equação anterior pode ser reescrita como

$$\left. \frac{\partial H}{\partial z_{p0}} \right|_{(z_{p0}^{(k)}, T^{(k)})} \Delta z_{p0}^{(k)} + \left. \frac{\partial H}{\partial T} \right|_{(z_{p0}^{(k)}, T^{(k)})} \Delta T^{(k)} = -H(z_{p0}^{(k)}, T^{(k)}). \quad (3.26)$$

Os valores finais para a condição inicial z_{p0} e o período T que satisfazem a condição de periodicidade (3.23) são calculados através das iterações

$$\begin{aligned} z_{p0}^{(k+1)} &= z_{p0}^{(k)} + \Delta z_{p0}^{(k)} \\ T^{(k+1)} &= T^{(k)} + \Delta T^{(k)} \end{aligned} \quad (3.27)$$

A convergência é alcançada quando (3.23) é satisfeito (*i.e.*, $(z_{p0}, T) \approx 0$).

Matematicamente, a expressão (3.26) possui $2n$ equações e $2n+1$ parâmetros desconhecidos e por isto necessita de um valor suplementar. Este valor corresponde à

fase da solução periódica. Esta condição adicional, denominada aqui de “condição de fase”, é incorporada a (3.23) tornando-se a condição de contorno ilustrada em (3.28).

$$F(z_{p0}, T) \equiv \begin{cases} H(z_{p0}, T) = 0 \\ h(z_{p0}) = 0 \end{cases} \quad (3.28)$$

Aqui, a função $h(z_{p0})$ refere-se à condição de fase mencionada anteriormente. Para resolver esta questão, é atribuindo o valor zero a pelo menos uma componente do vetor condição inicial (*e.g.*, velocidade inicial). A equação que irá governar as famílias de MNNs é a (3.28).

Para determinar os próximos MNNs, a energia do sistema é gradualmente incrementada. Em termos práticos, isto é feito através do aumento dos valores de condição inicial do sistema. Os MNNs computados anteriormente são utilizados como valores iniciais para calcular os próximos MNNs usando o método de continuação *pseudo-arclength*. Basicamente, este método começa a partir de um ponto inicial $(z_{p0,(j)}, T_{(j)})$, que neste caso é o MNN computado anteriormente pelo método de *shooting*, e o usa para estimar a próxima solução periódica representada pelo ponto $(z_{p0,(j+1)}, T_{(j+1)})$ no ramo (*branch*) de MNNs. Esta computação numérica envolve um passo de predição e outro de correção. Aqui, o termo ramo é empregado para descrever vários MNNs que quando agrupados geram uma curva. Cada ponto da curva (*i.e.*, ramo) representa um MNN calculado em diferentes níveis de energia e todos eles possuem características similares (representam movimentos em fase ou fora de fase). A Figura 3.10 ilustra um ramo com quatro MNNs calculados.

Conforme ilustra a Figura 3.10, o passo de predição usa um vetor tangente ao ramo no ponto corrente $(z_{p0,(j)}, T_{(j)})$ para calcular o ponto de predição $(\tilde{z}_{p0,(j)}, \tilde{T}_{(j)})$, que representa a solução aproximada para o verdadeiro valor do MNN $(z_{p0,(j+1)}, T_{(j+1)})$. Esta predição pode ser matematicamente descrita pelo vetor tangente $p_{(j)} = [p_{z,(j)}, p_{T,(j)}]'$ e um tamanho de predição $s_{(j)}$ como mostra (3.29).

$$\begin{bmatrix} \tilde{z}_{p0,(j+1)} \\ \tilde{T}_{(j+1)} \end{bmatrix} = \begin{bmatrix} z_{p0,(j)} \\ T_{(j)} \end{bmatrix} + s_{(j)} \begin{bmatrix} p_{z,(j)} \\ p_{T,(j)} \end{bmatrix} \quad (3.29)$$

O vetor $p_{(j)}$ pode ser calculado usando (3.30) com a condição adicional que $\|p_{(j)}\|=1$.

$$\begin{bmatrix} \left. \frac{\partial H}{\partial z_{p0}} \right|_{(z_{p0,(j)}, T_{(j)})} & \left. \frac{\partial H}{\partial T} \right|_{(z_{p0,(j)}, T_{(j)})} \\ \left. \frac{\partial h}{\partial z_{p0}} \right|_{(z_{p0,(j)})} & 0 \end{bmatrix} \begin{bmatrix} p_{z,(j)} \\ p_{T,(j)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (3.30)$$

O passo de predição gera o ponto $(\tilde{z}_{p0,(j+1)}, \tilde{T}_{(j+1)})$. O verdadeiro valor é alcançado por meio do método de correção através de um processo iterativo também baseado no método de *shooting*. Este processo usa como estimativa inicial o ponto predito e em cada iteração k , calcula (3.28) para encontrar novos valores de condições iniciais e períodos forçando-os a serem ortogonais ao vetor predictor. A Figura 3.10 ilustra esquematicamente todos estes passos de predição e correção. A falta de suavidade nas equações do modelo pode prejudicar esta etapa e outras e por isto foi considerada a versão do IF72 com interpolação polinomial.

Analiticamente, no $(j+1)$ ésimo MNN, a correção das variações nas condições iniciais $\Delta z_{p0,(j+1)}^{(k)}$ e período $\Delta T_{(j+1)}^{(k)}$, na iteração k , são dadas por

Figura 3.10 – Ilustração do método *pseudo-arclength* usado na predição e correção do cálculo de novos MNNs. A linha contínua representa um ramo de MNNs. Basicamente, este ramo é definido como uma família de MNNs possuindo características similares (*e.g.*, movimentos em fase ou fora de fase). No ramo em questão, os pontos negros indicam diferentes MNNs de movimentos periódicos. A seta no gráfico ilustra o vetor predictor tangente ao ramo. Os pontos em branco, perpendiculares ao vetor predictor, ilustram as iterações da convergência da solução calculada pelo método de *shooting* durante a fase de correção.

$$\begin{aligned} \mathbf{z}_{p0,(j+1)}^{(k+1)} &= \mathbf{z}_{p0,(j+1)}^{(k)} + \Delta \mathbf{z}_{p0,(j+1)}^{(k)}; \\ \mathbf{T}_{(j+1)}^{(k+1)} &= \mathbf{T}_{(j+1)}^{(k)} + \Delta \mathbf{T}_{(j+1)}^{(k)}. \end{aligned} \quad (3.31)$$

Estes fatores de correções são calculados pela equação (3.32) que impõe a condição de ortogonalidade ao processo corretor. Suas derivadas parciais são calculadas numericamente. O ponto predito $(\tilde{\mathbf{z}}_{p0,(j+1)}, \tilde{\mathbf{T}}_{(j+1)})$ é usado como estimativa inicial.

$$\begin{bmatrix} \left. \frac{\partial H}{\partial \mathbf{z}_{p0}} \right|_{(\mathbf{z}_{p0,(j+1)}^{(k)}, \mathbf{T}_{(j+1)}^{(k)})} & \left. \frac{\partial H}{\partial \mathbf{T}} \right|_{(\mathbf{z}_{p0,(j+1)}^{(k)}, \mathbf{T}_{(j+1)}^{(k)})} \\ \left. \frac{\partial h}{\partial \mathbf{z}_{p0}} \right|_{(\mathbf{z}_{p0,(j+1)}^{(k)})} & 0 \\ p_{\mathbf{z},(j)} & p_{\mathbf{T},(j)} \end{bmatrix} \begin{bmatrix} \Delta \mathbf{z}_{p0,(j+1)}^{(k)} \\ \Delta \mathbf{T}_{(j+1)}^{(k)} \end{bmatrix} = \begin{bmatrix} -H(\mathbf{z}_{p0,(j+1)}^{(k)}, \mathbf{T}_{(j+1)}^{(k)}) \\ -h(\mathbf{z}_{p0,(j+1)}^{(k)}) \\ 0 \end{bmatrix}. \quad (3.32)$$

O processo de iteração de índice k continua até que a condição de periodicidade (3.28) seja satisfeita. Em termos práticos, o valor desta função deve ser menor que um determinado erro de tolerância e previamente definido pelo usuário (*i.e.*, $H(\mathbf{z}_{p0}, \mathbf{T}) \leq e$).

3.2.2 – Análises gráficas por bifurcações e retratos de fase

Segundo achados anteriores (Herzel *et al.*, 1994; Steinecke & Herzel, 1995), um dos motivos de instabilidades percebidas em vozes podem ser resultado da dessincronização de alguns modos de vibração das PV causados por uma tensão de desbalanço crítica percebida em uma das pregas. Estas tensões geram assimetrias nos parâmetros de vibração das pregas que começam, individualmente, a vibrar com características diferentes e geram dessincronismo nos movimentos produzindo sinais de vozes com grandes perturbações.

Como o intuito de analisar como a dinâmica determinística do sistema se comporta diante de tais assimetrias, que podem representar até determinado ponto patologias nas pregas, serão analisados os modelos de PV reconstituídas dos indivíduos que passaram pelo exame de videolaringoscopia de alta velocidade nesta pesquisa. Mesmo embora seja um grupo de indivíduos com vozes saudáveis que produziram modelos mecânicos normais (*i.e.*, sem patologias e assimetrias), introduzimos um parâmetro Q_d , dado pela relação (3.33), que muda as principais constantes mecânicas

da prega direita destes indivíduos simulando uma patologia. Serão usados os índices ‘ e ’ e ‘ d ’ para designarem a prega esquerda e direita, respectivamente. Neste trabalho, a prega esquerda permanecerá com seus parâmetros inalterados (*i.e.*, $Q_e=1$) enquanto a direita terá seus valores variados dentro da faixa $0,5 \leq Q_d \leq 1,0$ com passos de 0.001 unidades. Somente a porção inferior ($i=1$) será variada por Q_d conforme destaca (3.33). Contudo, as alterações introduzidas nesta porção são percebidas também na região superior devido à transferência de movimento realizada pela mola de acoplamento k_c .

$$\tilde{m}_{1d} = \frac{m_{1d}}{Q_d}; \quad \tilde{k}_{1d} = k_{1d}Q_d; \quad \tilde{k}_{cd} = k_{cd}Q_d; \quad \tilde{r}_{1d} = r_{1d}Q_d. \quad (3.33)$$

Investigamos como o comportamento e as vibrações do modelo SH95²⁹ mudam com o decremento, em pequenos passos, de Q_d . Em outros termos, é avaliada a estabilidade das vibrações na medida em que a assimetria é incrementada e são vistos os efeitos produzidos pelas assimetrias que simulam patologias. O desbalanço introduzido pela constante Q_d resulta em novos parâmetros para a prega direita fazendo com que o modelo não seja mais simétrico. Sob outro ponto de vista, os parâmetros mecânicos empregados para cada indivíduo serão $\Phi_e = [m_{1e}, m_{2e}, k_{1e}, k_{2e}, k_{ce}, r_{1e}, r_{2e}, P_s]$ para a prega esquerda (que é preservada), e $\Phi_d = [m_{1d} \cdot Q_d, m_{2d}, k_{1d} \cdot Q_d, k_{2d}, k_{cd} \cdot Q_d, r_{1d} \cdot Q_d, r_{2d}, P_s]$ para a direita. Vale lembrar que originalmente $\Phi_e = \Phi_d$ (*i.e.*, pregas balanceadas e simétricas) já que os indivíduos analisados apresentaram vozes saudáveis.

Para avaliar os efeitos da assimetria no sistema e relacioná-la com as mudanças dos estados na saída do modelo, são empregados diagramas de bifurcação. Basicamente, o diagrama de bifurcação irá associar de maneira gráfica uma saída do sistema (representando um estado do mesmo) com uma correspondente assimetria Q_d . O máximo valor de amplitude de movimento x_{1d} da prega direita representará o estado do sistema. Para incrementar o desempenho da simulação, o estado final da última

²⁹ Conforme descrito em seções anteriores, o uso do SH95 nesta parte do trabalho justifica-se pelo fato de que o custo computacional deste modelo é significativamente mais simples. Como os métodos de análise de variação de parâmetros produzem cálculos extensivos, é necessário utilizar um modelo com estas características.

simulação é usado como condição inicial para a próxima simulação. Os diagramas de bifurcação foram computados usando o algoritmo *parameter stepping*. Seu pseudocódigo pode ser consultado em (Parker & Chua, 1989). Foi considerado um tempo transiente de 500 ms e só após este período é avaliado o máximo valor de x_{1d} .

Para o modelo estudado aqui, sua interpretação básica em um diagrama de bifurcação denota que para um dado valor de Q_d , um único ponto associado a este valor no gráfico, gera no espaço de fase um atrator com geometria similar a de um ciclo limite de período simples (*period-one limit cycle*). n pontos indicam um ciclo limite de período n e um aglomerado de pontos podem indicar caos ou comportamento *quasi*-periódico. Neste último caso, é necessária outra técnica para diferenciar estes comportamentos tais como, por exemplo, análise do atrator do sistema (ou sua trajetória no espaço de fase) ou expoentes de Lyapunov.

Conforme mencionado no parágrafo anterior, a simples análise dos diagramas de bifurcação não é suficiente para distinguir certas características dinâmicas do sistema. Um dos recursos gráficos utilizados para avaliar estas propriedades é o retrato de fase. Mesmo sendo uma ferramenta simples e de implementação direta, ela é um recurso poderoso na medida em que permite a descrição gráfica das trajetórias do sistema em um espaço de fase. Ela é usada quando se tem mais do que um estado do sistema e cada eixo do espaço reconstruído constitui um de seus estados. No caso do SH95, ilustrado na equação (3.12) e utilizado por ser computacionalmente mais simples no cálculo de bifurcações, é possível perceber que nas equações do modelo são diretamente calculados quatro estados do sistema para cada prega. Logo, o retrato de fase pode apresentar até oito eixos. Contudo, para tornar simples a sua representação gráfica são utilizados apenas dois estados: x_{1e} e x_{1d} . Consequentemente, o que se terá é uma projeção do atrator em um plano de duas dimensões. A escolha dos estados é fundamental já que ela tem influência na forma do atrator na medida em que determina diferentes planos de projeção.

Em casos onde se possui apenas um sinal produzido pelo sistema e operações de derivação são desencorajadas³⁰, possivelmente uma análise em espaço de fase feita por imersão seja mais eficiente e fidedigna à verdadeira forma do atrator. Geralmente isto acontece em sistemas que possuem componentes estocásticas.

3.3 – Resumo

Os métodos empregados nesta pesquisa podem ser divididos em duas partes. Na primeira delas são mostradas as técnicas usadas para extrair a dinâmica das PV e representá-la em dois diferentes modelos mecânicos: o IF72 e o SH95. Este procedimento de dinâmica inversa consiste basicamente em ajustar os modelos mecânicos de PV usando os sinais coletados de indivíduos através de exames de videolaringoscopia de alta velocidade. Este ajuste é feito por meio do cálculo dos parâmetros do modelo que melhor reproduzem um dado sinal registrado.

Na segunda parte são descritos os métodos empregados para avaliar a dinâmica dos modelos reconstruídos utilizando uma nova abordagem conhecida como ‘análise modal não-linear’ cujo objetivo para este trabalho é desvendar a influência das não-linearidades estruturais e sua relação com a frequência fundamental de oscilação das pregas e sua energia mecânica. Técnicas bem conhecidas da dinâmica não-linear de sistemas como representação em espaço de fase e diagramas de bifurcações são também empregadas para analisar o comportamento dinâmico das pregas em situações onde são simuladas patologias que remetem a uma assimetria entre as pregas direita e esquerda.

No próximo capítulo são apresentados e discutidos os resultados obtidos na mesma ordem com que apareceram aqui.

³⁰ Derivações numéricas podem amplificar os ruídos presentes no sinal gerando distorções maiores.

Capítulo 4

Resultados e discussões

Neste capítulo são apresentados os resultados alcançados nesta pesquisa. A exemplo do capítulo anterior, este também é dividido em duas partes. Em sua primeira parte são descritos e discutidos os resultados referentes ao método de extração da dinâmica das PV usando os modelos mecânicos de SH95 e IF72 bem como os resultados da validação e desempenho do método quando aplicado a sinais sintéticos.

A segunda parte descreve os resultados alcançados relativos à análise modal não-linear dos indivíduos avaliados nesta pesquisa. De forma complementar, foram simuladas assimetrias entre as pregas esquerda e direita nos modelos mecânicos dos indivíduos estudados com o propósito de avaliar como o sistema dinâmico se comporta nestes cenários. Estes resultados são descritos através de diagramas de bifurcação e retratos de fase.

4.1 – Extração da dinâmica das pregas vocais

4.1.1 – Processamento das imagens na laringe

A segmentação de imagens da laringe é um importante meio para entender os principais aspectos envolvidos na produção de vozes saudáveis e patológicas. Sendo assim, sua correta segmentação é tarefa essencial para uma análise mais confiável. Neste sentido, o algoritmo proposto no capítulo anterior conseguiu satisfatoriamente processar as imagens dos vinte indivíduos recrutados e produzir o SGE destes. A Figura 4.1 ilustra alguns resultados da segmentação das imagens. Os quadros ilustrados nesta figura mostram que a área extraída da imagem acompanhou na maior parte as bordas

das pregas. Resultados similares foram obtidos para os demais quadros que não puderam ser representados na figura. Nela, também são mostrados os SGEs correspondentes.

Em particular, o segundo quadro da Figura 4.1(e) apresentou uma falha de segmentação. Isto porque delimitou somente a porção inferior da glote, desconsiderando uma pequena abertura na sua porção superior. Isto ocorre porque o algoritmo considera que existe somente uma área glotal - e não duas como no presente exemplo - e descarta automaticamente aquela que apresenta uma grande variação nas coordenadas de seu centroide segundo a análise de histórico dos centroides calculados anteriormente ao quadro corrente. Todavia, o usuário pode intervir no processo e indicar manualmente com o *mouse* a área não processada pelo algoritmo.

Apesar das considerações feitas no parágrafo anterior, o algoritmo mostrou bom automatismo no processamento dos quadros e exigiu pouca intervenção do usuário. Como esta questão do automatismo está ligada a qualidade dos vídeos, as imagens mais degradadas das amostras selecionadas tiveram seu grau de automatismo consideravelmente menor. Mesmo assim, muitas das limitações encontradas pelos algoritmos tradicionais de segmentação não foram percebidas nos experimentos realizados. Ainda, o algoritmo mostrou-se robusto na medida em que não demonstrou grande sensibilidade às desigualdades de cores presentes dentro da região glotal e pôde trabalhar com quadros com baixo contraste. Até mesmo após processar um quadro onde não havia área glotal, como no caso do primeiro quadro da Figura 4.1(e), as regiões sementes dos quadros subsequentes foram encontradas e as áreas segmentadas.

Em alguns SGEs é possível notar potenciais problemas que podem ser atribuídos aos erros de segmentação. Embora indesejáveis, estes erros são comuns neste tipo de processo que envolve uma grande quantidade de imagens com diferentes características que variam quadro a quadro. Porém a grande dificuldade da segmentação está no formato das bordas das pregas que não é nítido devido sua característica

“arredondada”, o que torna difícil estabelecer uma correta delimitação onde termina a área glotal e onde começa as bordas das pregas na medida em que não existe uma fronteira nítida entre estas e por isto, às vezes, nem mesmo é percebida com clareza para o próprio olho humano. No caso do sujeito 2 (Figuras 4.1(a) e 4.1(b)), esta característica indesejada da ausência de uma borda nítida é um pouco mais evidente e por isto o SGE deste indivíduo está mais sujeito à distorções. Por exemplo, a fase de

Figura 4.1 – Amostras representativas de quadros extraídos de vídeos de videolaringoscopia de alta velocidade e seus respectivos sinais de área glotal gerados a partir da segmentação (*i.e.*, SGE). As linhas brancas delimitando as pregas definem a região da glote segmentada.

fechamento de seu ciclo glotal aparenta ter uma velocidade de fechamento mais lenta do que seu período de abertura quando tomamos por referência o formato ideal de um pulso discutido em (Childers & Lee, 1991) (análises subsequentes confirmarão esta impressão). Apesar destas limitações, o uso de um método de segmentação (semi)automático pode contribuir para a precisão na medida em que este possui um critério fixo e numérico para determinar onde termina a área glotal e onde começa a borda da prega ajudando potencialmente a decrementar a subjetividade da análise.

No que diz respeito ao tempo de processamento das imagens, para um vídeo com 270 quadros (256 x 256 *pixels*), foi gasto em média 10 segundos³¹. Estes valores de tempo também dependem da qualidade da imagem utilizada e não consideram a intervenção manual do operador.

4.1.2 – Validação do procedimento de extração de dinâmica

Os resultados relativos à validação do procedimento de extração da dinâmica a partir de 50 sinais sintéticos são vistos na Figura 4.2. Cada gráfico desta Figura possui 50 pontos que associam individualmente o valor verdadeiro do parâmetro (eixo horizontal) com o valor encontrado pelo procedimento de otimização (eixo vertical). Quanto mais próximo o ponto está da linha diagonal do gráfico, maior a exatidão do método. Os dados sintéticos não puderam ser uniformemente distribuídos ao longo do eixo horizontal, pois nem sempre seus valores produziam oscilações regulares.

Os resultados ilustrados na Figura 4.2 mostram uma grande concentração dos pontos ao longo da linha diagonal. Contudo, alguns destes pontos apresentam um considerável distanciamento desta referência. Para avaliar o efeito deste distanciamento, foi calculado (relação (3.19)) o erro médio relativo entre os parâmetros encontrados e os seus respectivos valores verdadeiros (*i.e.*, valores sintéticos) para expressar em termos percentuais o quanto o vetor de parâmetros estimados Φ_e é

³¹ Valores obtidos usando um computador com 2 núcleos de processamento de 2,9GHz cada e 4 GB de memória RAM

diferente da referência sintética Φ_s . A Figura 4.3 (superior) ilustra os resultados que mostraram erros de até 30,2% entre estes vetores e um valor médio de 9,9% entre as 50 amostras do gráfico.

O resultado anterior leva a crer que nos casos onde o erro entre os parâmetros é maior, este erro produzirá por consequência uma maior diferença entre os sinais simulados e os sintéticos. Por exemplo, a amostra sintética de índice 20 tem um erro médio entre seus parâmetros consideravelmente maior que a amostra 35. A princípio, espera-se como consequência disto a produção de uma curva com erro maior do que a produzida pela amostra 35. Contudo, uma análise mais detalhada do erro médio das curvas produzidas pelos parâmetros sintéticos e as geradas pelos parâmetros extraídos, revela no gráfico inferior da Figura 4.3 que este erro para a amostra 20 é menor do que a 35. Mesmo que esta última tenha encontrado um vetor de parâmetros mais próximo do valor verdadeiro, estes parâmetros produzem um sinal com erro maior do que

Figura 4.2 – Parâmetros sintéticos (eixo horizontal) *versus* parâmetros otimizados (eixo vertical). Cada gráfico representa um dos oito parâmetros do modelo SH95 que foi utilizado para validar o procedimento. Cada ponto nos gráficos tem como coordenada horizontal o valor do parâmetro sintético utilizado para gerar a curva e na coordenada vertical o valor encontrado pelo procedimento de otimização. Um ponto na reta diagonal representa a igualdade entre estes dois valores.

aqueles produzidos pela amostra 20. Uma possível consequência disto é que o método de otimização encontrou uma solução local para a amostra 20 e esta solução foi capaz de produzir um sinal de área glotal muito próximo da curva sintética produzida pela solução de mínimo global. Ainda, ressalta-se que não foi levado em consideração a ponderação entre os parâmetros o que também pode contribuir para casos similares a este exemplificado.

Quando analisado, o gráfico inferior da Figura 4.3 revela um erro médio entre as 50 amostras sintéticas de 1,6%. Isto significa que o erro de amplitude ponto a ponto entre as curvas sintéticas e as simuladas foi, em média, 1,6% do valor de pico da curva sintética. Este resultado mostra que as curvas apresentam um bom encaixe e que o procedimento de otimização consegue produzir um modelo que reproduz com satisfação este movimento. Foi também avaliada a correlação entre as curvas e encontrou-se uma correlação média de 99,2% sendo o pior resultado calculado como 97,8%. Todos estes valores indicam uma correlação fortíssima reforçando a qualidade da simulação produzida.

No que diz respeito aos parâmetros mecânicos, mesmo embora estes tenham

Figura 4.3 – Gráfico superior: erro médio relativo percentual entre os parâmetros sintéticos e simulados. Valores calculados usando a equação (3.19). Gráfico inferior: erro médio percentual entre as curvas sintéticas e as simuladas. Valores calculados usando a equação (3.20)

mostrado um erro médio maior do que o erro entre as curvas, seus resultados são avaliados como aceitáveis na medida em que os parâmetros variaram, em média, 9,9% do valor ideal. Esta questão é reforçada se considerarmos a ampla faixa de valores (*i.e.*, grande espaço de busca) com que o procedimento trabalha e a considerável quantidade de parâmetros que são avaliados pelo processo de otimização. Dadas estas questões, mesmo assim há de se considerar que a otimização possa produzir soluções que possam representar um mínimo local e não um global.

4.1.3 – Otimização dos parâmetros mecânicos

Os sinais produzidos pela segmentação das imagens (*i.e.*, SGE) foram utilizados pelo método de otimização que encontrou novos valores para os parâmetros dos modelos mecânicos das PV permitindo a estes reproduzirem o SGE dos indivíduos analisados. As Figuras 4.4 a 4.6 ilustram graficamente três casos (cada um deles associados a um determinado indivíduo) que representam em parte as principais características das amostras processadas. No anexo A é possível ver os gráficos de todos os indivíduos analisados nesta pesquisa.

Nos gráficos das Figuras 4.4 a 4.6 são visualizados o sinal de área glotal extraído das imagens (cor preta), os simulados pelo modelo SH95 (gráficos à esquerda na cor azul) e IF72 (gráficos à direita na cor azul). Para cada um destes sinais é ilustrada sua correspondente derivada e espectro de potência³² para avaliar as características dos sinais e estabelecer uma comparação entre os mesmos³³. Neste ponto é importante observar alguns pequenos artefatos produzidos pelo IF72. Note, por exemplo, que os picos negativos da derivada na Figura 4.5 não apresentam em todo ciclo a mesma amplitude, o que era de se esperar de um modelo determinístico.

³² Foi considerada uma janela de Hanning e 256 amostras para o cálculo da transformada de Fourier dos sinais.

³³ Vale destacar que a forma de onda do fluxo glotal U_g e o sinal de área glotal A_g têm formato bastante próximos (a principal diferença está na amplitude) e por isto também possuem espectros semelhantes. Sendo assim, o espectro de A_g será usado para avaliar indiretamente as propriedades de U_g .

Figura 4.4 – Sinais extraídos das imagens (curvas em preto) do indivíduo 2 e seus correspondentes valores simulados (curvas em azul) usando o SH95 (gráficos à esquerda) e o IF72 (gráficos à direita).

Os artefatos produzidos são um resultado do processo de diferenciação usado para resolver as equações do modelo (ver (3.1)) já que foi considerado um $T=1/4000$ s para acompanhar a frequência de registro de imagens do sistema de videolaringoscopia empregado nesta pesquisa. A derivação do sinal também amplifica estas diferenças. Um período de amostragem igual ou superior a $1/10000$ s não produz estes efeitos de forma tão visível. Apesar disto, os efeitos causados por estes artefatos são desconsiderados já que o propósito primário do trabalho é avaliar as características essenciais das pregas e por isto prefere-se assegurar uma análise ponto a ponto entre os sinais extraídos e os

simulados, o que conduz a uma simulação mais rápida³⁴ deixando de lado detalhes.

Os resultados das amostras ilustradas (e as demais exibidas no Anexo A) revelam que os modelos SH95 e IF72 foram capazes de reproduzir os mais importantes padrões de vibração presentes no SGE em praticamente todos os casos analisados. Em particular, nos resultados exibidos nas Figuras 4.4 e 4.5, o SGS produzido pelo SH95 é aparentemente melhor que o produzido pelo IF72. O mesmo não acontece no caso da Figuras 4.6 onde o SH95 não conseguiu encontrar valores de parâmetros mecânicos capazes de produzir uma forma de onda que acompanhasse em amplitude o sinal

Figura 4.5 – Sinais extraídos das imagens (curvas em preto) do indivíduo 7 e seus correspondentes valores simulados (curvas em azul) usando o SH95 (gráficos à esquerda) e o IF72 (gráficos à direita).

³⁴ O mesmo problema não foi percebido para o SH95 já que este apresenta equações mais simples que viabilizam a sua solução analiticamente, decrementando a criação de artefatos na solução das equações diferenciais do modelo.

extraído. Neste exemplo, o IF72 obteve melhor êxito produzindo uma curva mais próxima da extraída.

No caso do indivíduo 2 (Figura 4.4), a derivada do sinal de área glotal apresenta um pico positivo maior que o negativo o que contradiz a ilustração da Figura 3.3 e do modelo idealizado de pulso glotal. Apesar desta distorção, os modelos tiveram a habilidade de produzir derivadas mais coerentes que a vista pelo sinal extraído da imagem. O bom encaixe de curvas de área glotal não é visto com mesma qualidade na derivada deste sinal principalmente porque a diferenciação acentua as diferenças entre os sinais amplificando os ruídos e diferenças. No caso do IF72, este produz um sinal com maior decaimento da derivada na fase de fechamento e um maior pico negativo. Segundo Childers e Lee (1991), esta propriedade é fundamental para produzir um sinal

Figura 4.6 – Sinais extraídos das imagens (curvas em preto) do indivíduo 16 e seus correspondentes valores simulados (curvas em azul) usando o SH95 (gráficos à esquerda) e o IF72 (gráficos à direita).

rico em componentes harmônicas para o trato vocal e é uma das principais características do sinal produzido na glote. Estes mesmos aspectos podem ser percebidos nos demais indivíduos revelando uma boa capacidade do IF72 em reproduzir uma fase de fechamento mais coerente, ainda que o sinal de referência (*i.e.*, SGE) não tenha esta propriedade tão proeminente. Isto mostra que os modelos empregados são capazes de reproduzir algumas de suas características dinâmicas mais básicas mesmo passando por uma otimização que alterou seus principais parâmetros.

A análise espectral também revela que o IF72 produz na maior parte dos casos um espectro mais próximo do sinal extraído em comparação com o SH95. De mesma maneira, aponta que as principais componentes harmônicas do sinal são reproduzidas gerando um espectro rico em frequências que servirá de excitação para o trato vocal e nele será modelado. Apesar de não ter conseguido um desempenho melhor no encaixe de curvas de área glotal, os sinais de derivada do IF72 mostram-se na maior parte dos casos mais adequados e têm forte influência no espectro do sinal. A Tabela 4.1 apresenta os parâmetros mecânicos encontrados pelo processo de otimização de todos os indivíduos analisados e o erro entre os SGE e SGS. No caso do IF72, os coeficientes de amortecimento r_1 e r_2 não foram estimados pelo procedimento, visto que estes são dados pela relação matemática da Tabela 3.1. Em contraposição, o SH95 deixa livre a escolha dos valores para coeficientes de amortecimento e por isto foram estimados para este modelo. Na Figura 4.7 é ilustrado graficamente os erros calculados pelas equações (3.18) e (3.20) para todos os indivíduos processados.

A análise dos erros entre a aproximação de SGS e o SGE indica que o SH95 apresenta em média um menor erro e por isto é capaz de produzir com melhor fidelidade os movimentos de vibração das pregas confirmando esta questão levantada anteriormente pela análise dos gráficos produzidos. Contudo, o mesmo não vale para o espectro produzido pelo modelo já que neste caso o IF72 mostrou ser capaz de dar origem a um espectro mais próximo do padrão extraído e com características mais

Tabela 4.1 – Parâmetros mecânicos calculados pelo procedimento de otimização usando o modelo SH95 e IF72.

Ind.	Modelo SH95				Modelo IF72			
	$(m_1, m_2, k_1, k_2, k_c, r_1, r_2, P_s)^1$	Erro			$(m_1, m_2, k_1, k_2, k_c, -, -, P_s)^1$	Erro		
		(%) ²	(dB) ³			(%) ²	(dB) ³	
1	[0.000267, 0.000064, 314.51, 17.42, 13.81, 0.003530, 0.073542, 2659.07]	6.8	0.73		[0.000198, 0.000065, 155.87, 4.32, 66.24, -, -, 1070.51]	6.4	0.68	
2	[0.000248, 0.000097, 238.53, 51.52, 65.21, 0.013790, 0.027146, 2043.46]	3.9	0.8		[0.000172, 0.000062, 136.84, 47.50, 87.36, -, -, 2226.69]	7.1	0.42	
3	[0.000114, 0.000011, 238.17, 32.19, 13.52, 0.03034, 0.02837, 3597.12]	8.1	0.79		[0.000066, 0.000028, 110.76, 3.39, 122.91, -, -, 1336.12]	8.9	0.31	
4	[0.000305, 0.000015, 259.97, 7.83, 50.65, 0.029400, 0.016710, 2987.34]	6.6	0.71		[0.000168, 0.000010, 136.23, 1.69, 6.76, -, -, 3224.18]	6.4	0.44	
5	[0.000257, 0.000120, 253.24, 74.02, 72.87, 0.007169, 0.038355, 2924.95]	3.3	0.56		[0.000083, 0.000147, 21.17, 45.20, 124.92, -, -, 3086.16]	4.2	0.52	
6	[0.000348, 0.000053, 342.40, 44.54, 97.30, 0.017704, 0.023644, 2893.78]	3.6	0.62		[0.000103, 0.000046, 96.12, 39.63, 67.63, -, -, 2827.51]	6.8	0.29	
7	[0.000234, 0.000064, 180.74, 43.65, 173.05, 0.001733, 0.051824, 2974.19]	4.2	0.57		[0.000201, 0.000075, 153.18, 39.95, 70.61, -, -, 4074.44]	8.2	0.32	
8	[0.000221, 0.000077, 202.72, 49.25, 40.14, 0.008928, 0.026743, 2721.32]	3.2	0.48		[0.000113, 0.000084, 69.30, 30.95, 107.47, -, -, 3680.25]	3.9	0.37	
9	[0.000221, 0.000066, 99.14, 18.78, 30.24, 0.015790, 0.013883, 1788.15]	5.9	0.62		[0.000302, 0.000195, 85.54, 16.53, 109.45, -, -, 4055.52]	3.6	0.30	
10	[0.000284, 0.000070, 329.58, 54.01, 123.23, 0.029499, 0.047637, 2932.05]	4.6	0.63		[0.000204, 0.000094, 185.5, 45.77, 141.74, -, -, 3433.43]	4.5	0.46	
11	[0.000300, 0.000075, 473.95, 38.73, 132.85, 0.003172, 0.060080, 2513.83]	4.1	0.67		[0.000108, 0.000104, 94.78, 27.31, 138.28, -, -, 2204.59]	3.3	0.64	
12	[0.000324, 0.000079, 358.06, 36.52, 97.41, 0.018012, 0.074540, 2889.33]	2.8	0.68		[0.000169, 0.000038, 202.08, 32.71, 61.63, -, -, 1642.33]	8.6	0.46	
13	[0.000248, 0.000100, 156.50, 36.74, 69.33, 0.008509, 0.050340, 1803.29]	2.7	0.49		[0.000347, 0.000102, 200.78, 35.11, 63.35, -, -, 3187.45]	7.9	0.55	
14	[0.000096, 0.000035, 102.65, 77.06, 254.92, 0.008503, 0.094103, 3036.34]	6.0	1.00		[0.000120, 0.000014, 182.40, 38.12, 39.38, -, -, 3202.83]	5.0	0.44	
15	[0.000232, 0.000051, 149.20, 37.04, 122.07, 0.003121, 0.016204, 2981.04]	4.2	0.57		[0.000085, 0.000051, 55.20, 41.60, 49.95, -, -, 2942.82]	9.0	0.40	
16	[0.000281, 0.000066, 66.89, 56.65, 23.86, 0.106572, 0.001248, 2407.32]	8.6	0.86		[0.000054, 0.000004, 56.86, 46.93, 62.54, -, -, 2881.81]	4.9	0.37	
17	[0.000116, 0.000035, 134.15, 23.78, 139.35, 0.004810, 0.041548, 2616.06]	4.7	0.89		[0.000145, 0.000085, 122.2, 59.48, 123.38, -, -, 4153.74]	4.7	0.97	
18	[0.000321, 0.000146, 125.25, 37.49, 45.01, 0.022429, 0.050504, 1808.67]	3.0	0.74		[0.000311, 0.000192, 93.04, 47.88, 76.48, -, -, 2483.79]	5.7	0.71	
19	[0.000293, 0.000019, 327.30, 16.43, 56.19, 0.017947, 0.023184, 3131.58]	3.3	0.61		[0.000060, 0.000055, 42.59, 32.75, 101.88, -, -, 2904.03]	6.9	0.37	
20	[0.000206, 0.000010, 168.90, 59.85, 189.86, 0.004330, 0.026891, 2951.68]	3.7	0.63		[0.000044, 0.000004, 45.81, 47.27, 56.62, -, -, 2774.41]	5.1	0.36	

¹Unidades dados no sistemas internacional de medidas (Kg, N/m, Pa). ²Erro calculado por (3.20). ³Erro calculado por (3.18).

próximas do padrão ideal. Tendo em mente esta questão, cada sistema apresenta suas

vantagens e desvantagens no modelamento do comportamento das PV. Em aplicações envolvendo síntese de voz ou avaliações acústicas, o IF72 pode produzir sinais mais ricos e realísticos. Já em casos onde se deseja reproduzir os movimentos básicos de vibração, o SH95 parece ser mais vantajoso na medida em que incorpora melhor os padrões de forma de onda da vibração e apresenta menor custo computacional. Vale destacar que a mudança proposta na modelagem do impacto utilizando uma função polinomial não é usada no procedimento de extração de parâmetros e sim apenas nas avaliações dinâmicas modais que são mostradas nas próximas seções deste capítulo.

Figura 4.7 – Figura esquerda: erro médio calculado entre as curvas simuladas pelo modelos mecânicos (SH95: curva em azul; IF72: curva em vermelho) e o sinal de área glotal extraído. Erro calculado pela Equação (3.20). Figura direita: erro médio entre os espectro de potência dos sinais gerados pelos modelos mecânicos (SH95: curva em azul; IF72: curva em vermelho) e o espectro do sinal de área glotal extraído das imagens. Erro calculado pela Equação (3.18).

Outra questão relevante vista na Tabela 4.1 diz respeito à comparação direta entre os parâmetros do SH95 e IF72 para um mesmo indivíduo. Os resultados revelam que para um mesmo indivíduo os modelos têm parâmetros bastante diversos na maioria dos casos. Mesmo se tratando essencialmente de modelos de duas massas e com muitas características em comum, estes apresentam parâmetros diferentes. Um dos fatores que pode explicar parcialmente esta divergência pode estar ligado à forma como o modelo representa seu amortecimento (que é bastante distinta entre eles) e outras questões

Figura 4.8 – Evolução do algoritmo genético para um processamento de 50 gerações para cinco diferentes otimizações ilustradas como curvas na cor cinza. As curvas em preto indicam a média calculada a partir destas cinco otimizações tomadas como exemplo. À esquerda, as curvas indicam o erro médio da geração e a direita o erro do melhor indivíduo de uma dada geração.

relacionadas, principalmente, a maneira como é modelada a excitação. No caso do SH95, é determinada apenas a excitação das camadas inferiores das pregas (*i.e.*, $F_2=0$) e uma série de simplificações nas leis de Bernoulli são aplicadas de modo a produzir, em termos práticos, outra abordagem distinta da utilizada originalmente no IF72. Sendo assim, ainda que o SH95 seja uma adaptação simplificada do IF72, este grau de adaptação parece ser grande o suficiente para dar ao modelo traços próprios que podem tornar inviável a comparação direta de seus parâmetros. Ainda, outro fator que pode ter apresentado contribuição para a produção de parâmetros divergentes pode estar ligado a uma otimização com muitos mínimos locais. Contudo, a contribuição deste último pode ser limitada tendo em vista os resultados alcançados pela validação do procedimento de otimização.

De maneira geral, os resultados mostraram que o método proposto é eficiente para reproduzir os movimentos de vibração das pregas e extrair parâmetros. As diferenças entre as curvas podem ser justificadas por erros de segmentação, componentes estocásticas no sinal extraído, as propriedades dos modelos mecânicos que não foram capazes de acompanhar o padrão de curvas experimentais ou falhas no procedimento de otimização. Ainda, o algoritmo genético requereu poucas gerações para

produzir boas aproximações. A Figura 4.8 mostra o desempenho médio do algoritmo considerando sua convergência. Repare que no gráfico à direita da Figura 4.8, o erro médio da população cai até por volta da geração 10. A partir deste ponto o erro médio desta população varia e sua convergência é prejudicada devido à grande quantidade de novos cromossomos inseridos. É esta estratégia que garante uma grande variabilidade genética e uma maior exploração do espaço de busca. O outro gráfico da Figura 4.8 mostra que após a geração 20, a otimização alcança uma boa convergência.

4.1.4 – Discussões gerais sobre o método de extração de dinâmica

Os resultados apresentados mostram que o método é capaz de reproduzir os principais padrões de movimento das pregas extraídas usando dois diferentes modelos de PV. Quando comparado a outros trabalhos (Mergell *et al.*, 2000; Dollinger *et al.*, 2002; Gómez-Vilda *et al.*, 2007; Tao *et al.*, 2007; Zhang *et al.*, 2010) com propósitos similares, o método aqui desenvolvido apresenta significativas contribuições que estão relacionadas a (i) uma maior quantidade de parâmetros mecânicos processados e (ii) um espaço de busca mais abrangente; (iii) otimização mais rápida; (iv) emprego de dois diferentes modelos mecânicos de PV (SH95 e IF72); (v) avaliação espectral dos sinais na tentativa de reproduzir com maior detalhamento a fase de fechamento e (vi) o uso de imagens clínicas refletindo casos reais. O emprego de um método misto que utiliza de forma conjunta a “força bruta” dos algoritmos genéticos e a análise refinada do simplex é também uma nova abordagem que mostrou um bom resultado.

Nesta pesquisa foram avaliados oito parâmetros mecânicos para o SH95 e seis para o IF72. Em outros trabalhos, foram vistos a otimização de dois (Mergell *et al.*, 2000), três (Dollinger *et al.*, 2002) e sete (Tao *et al.*, 2007; Qin *et al.*, 2009) parâmetros para o SH95, sendo que neste último caso os autores consideraram apenas experimentos envolvendo uma única prega excisada controlada por um aparato mecânico em laboratório. No caso do IF72, não foram encontradas tentativas de otimização de suas

constantes. Ainda, foi considerado um espaço de busca maior para estas constantes, o que potencialmente permite ao método explorar padrões de vibrações mais abrangentes.

Apesar da maior quantidade de constantes e maior espaço de busca, o algoritmo genético conseguiu encontrar uma solução aceitável processando um valor máximo de 50 gerações. Em outros casos (Tao *et al.*, 2007; Qin *et al.*, 2009) onde empregou-se de maneira única esta técnica, foram necessárias mais que 1.500 gerações para produzir um resultado aceitável considerando uma população similar. Possivelmente os dois principais fatores que possibilitaram o desempenho expressivo nesta pesquisa estão vinculados à (i) alta taxa de variabilidade genética do algoritmo, muito acima dos padrões tradicionais, e (ii) o fato de considerarmos um método de otimização em duas etapas sendo que a primeira delas (*i.e.*, algoritmo genético) não é sobrecarregada já que sua função é apenas produzir uma estimativa grosseira da solução. A segunda etapa, envolvendo a técnica simplex, tem um tempo de processamento geralmente inferior a 1 minuto³⁵.

O uso dos dois principais modelos de duas massas das pregas vocais (IF72 e SH95) introduz uma flexibilidade a mais no método possibilitando a escolha de qual modelo se adéqua melhor a aplicação desejada já que as vantagens e desvantagens destes sistemas estão intimamente ligadas a sua aplicação. Mesmo embora os dois modelos tenham essencialmente a mesma dinâmica de duas massas, o SH95 é preferível em circunstâncias onde são aplicados cálculos extensivos e não são requeridos a reprodução de detalhes. Já o IF72 simula com maior detalhamento os principais sinais. Apesar de sua importância e ser um modelo de duas massas mais ‘minucioso’ que o SH95, não foram encontrados na literatura trabalhos que buscaram otimizar suas constantes. Todos artigos referenciados neste texto que procuraram otimizar os parâmetros mecânicos de um modelo de duas massas, o fizeram utilizando o SH95. Todavia, os resultados revelaram que o IF72 pode ser mais eficiente na representação

³⁵ Considerando um computador de 2 núcleos de processamento de 2,9GHz cada e 4 GB de memória RAM

de um sinal principalmente no que diz respeito a sua composição espectral e apesar de sua complexidade e custo computacional, é uma abordagem importante em determinados casos.

Mesmo apresentando significativas vantagens, o emprego do modelo de três massas (ST95) não foi considerado neste trabalho, pois dificultaria significativamente os procedimentos de otimização e avaliação dinâmica já que uma quantidade reduzida de parâmetros avaliados permite que seja possível uma interpretação mais eficiente e direcionada e um grau de liberdade a mais dificultaria significativamente a análise modal. Berry *et al.* (1994) concluíram que dois graus de liberdade (*i.e.*, modos) são capazes de reproduzir aproximadamente 95% dos movimentos de vibração das PV. Em outro trabalho (Zhang *et al.*, 2007) os autores descobriram que aproximadamente 99% da energia do movimento das PV pode estar associada aos dois primeiros modos das pregas e que por isto um modelo com dois graus de liberdade é capaz de representar este nível de energia do movimento. Estes resultados implicam que um modelo simples com dois graus de liberdade pode reproduzir suficientemente bem os padrões de vibração das pregas onde detalhes menores não são requeridos.

Outra importante característica do método que repercute diretamente na qualidade de seus resultados está relacionado à sua avaliação espectral na segunda etapa do procedimento de otimização. Como visto nos resultados, um bom encaixe de curvas de área glotal nem sempre repercute na produção de um espectro com mesmas características e é capaz de produzir componentes de frequência suficientes para excitar o trato e produzir um sinal de qualidade. Neste sentido, a fase de fechamento do pulso glotal desempenha papel fundamental na produção de altas frequências e por ser uma etapa de curta duração no pulso glotal, tem sua contribuição na função objetiva da primeira etapa (equação (3.17)) decrementada. Para corrigir esta distorção, a segunda etapa do procedimento de otimização tem como função objetiva (equação (3.18)) um método de parametrização da distância espectral entre o SGE e o SGS. Embora nem

todos os espectros simulados apresentassem uma forma adequada, os resultados mostraram que muitos preservaram as componentes harmônicas principais dentro da faixa espectral analisada. Em especial, o IF72 mostrou-se significativamente mais eficiente na medida em que produz uma fase de fechamento mais acentuada. Nos casos onde houve grande diferença entre os espectros dos sinais extraídos e simulados, acredita-se que os fatores que potencialmente tenham contribuído para isto estejam vinculados a muitas distorções no sinal extraído, inabilidade do modelo em reproduzir certos padrões ou o procedimento de otimização que não encontrou um mínimo global. Ainda, as próprias imagens produzidas pela videolaringoscopia nem sempre revelam uma fase de fechamento mais acentuada.

O uso de imagens clínicas pelo método viabiliza análises de diferentes padrões de vozes dando uma maior relevância ao método na medida em que possibilita o estudo *in vivo* de diversos grupos de indivíduos. Em outros trabalhos foram usadas imagens de pregas excisadas (Tao *et al.*, 2007), onde as condições de aquisição de imagens eram rigidamente controladas em laboratório, ou se limitaram a avaliar um único indivíduo (Mergell *et al.*, 2000; Qin *et al.*, 2009) ou dois (Dollinger *et al.*, 2002). No caso desta pesquisa, todos os vinte indivíduos analisados puderam ser processados e passaram pelo procedimento de extração.

No que tange as limitações do método desenvolvido, considerou-se como fixo o valor do comprimento da fenda glotal (*i.e.*, L_g) e que os indivíduos produziram vibrações estacionárias e uma fonação sustentada. Ainda, o método é susceptível a mínimos locais e a grande quantidade de descontinuidades dos modelos mecânicos dificulta a aplicação de métodos mais sofisticados de otimização baseados na análise do gradiente das funções.

Os resultados exibidos na Tabela 4.1 mostram que para um mesmo indivíduo, pode haver grande variação nos valores de parâmetros mecânicos entre os modelos SH95 e IF72. Mesmo embora um dos modelos empregados seja considerado uma

simplificação do outro, a diferença entre as constantes calculadas é bastante considerável e devido ao elevado grau de simplificação, tratam-se em termos práticos de modelos essencialmente distintos e por isto a comparação direta entre suas constantes pode ser inadequada. Dadas estas questões e considerando que os parâmetros encontrados pela otimização podem não ser únicos ou não corresponderem a melhor solução, a avaliação isolada destes pode ser falha. Ainda, como destacam Gómez-Vilda *et al.* (2007), estes parâmetros nem sempre podem ser usados como medidas diretas das propriedades das pregas e sim como constantes que, quando usadas por um determinado modelo, são capazes de reproduzir seu comportamento dinâmico e refletir algumas importantes características do oscilador que constituem as PV. Naturalmente, quando se faz esta afirmação, considera-se que a aderência do(s) modelo(s) ao sistema (*i.e.*, pregas vocais) é forte. Esta pesquisa faz esta mesma consideração.

Como indicam Childers e Lee (1991), a avaliação da qualidade da voz passa por uma análise conjunta e indissociável dos parâmetros que representam a tensão das pregas, massa e a pressão sub-glotal e juntos, todos estes parâmetros determinam o tipo de som produzido e reflete diretamente nos padrões de coeficiente de abertura³⁶, razão de tempo de abertura/fechamento e outras propriedades do ciclo glotal e seu espectro. Em (Jiang *et al.*, 2000) são discutidos os efeitos destes parâmetros no sinal de voz.

O uso de um modelamento mecânico das PV é também fundamental para obter um modelo determinístico de produção da voz capaz de tornar possível o estudo destas estruturas sob as novas perspectivas trazidas pela dinâmica não-linear de sistemas. Como anteriormente exposto, isto acontece porque alguns métodos desta área exigem um comportamento totalmente determinístico do sistema (Ott *et al.*, 1994; Abarbanel, 1995; Kaplan & Glass, 1995) e a análise direta de sinais experimentais pode ser inadequada, haja visto que estes quase sempre apresentam componentes estocásticas. Esta é uma justificativa adicional ao emprego do método aqui desenvolvido para

³⁶ Percentagem do ciclo onde as pregas estão abertas.

reproduzir de maneira totalmente determinística o comportamento dinâmico das pregas. Ainda, a extração de parâmetros usando modelos com poucos graus de liberdade (como os aqui empregados) pode servir para estimar parâmetros que podem ser usados em modelos com muitos graus de liberdade onde este processo de extração seria consideravelmente mais difícil.

4.2 – Avaliação da dinâmica das pregas vocais

4.2.1 – Análise dos modos não-lineares

A fonação é essencialmente um processo de transferência de energia e por isto é importante analisar o efeito que esta energia pode ter no sistema quando assume valores suficientes para excitar suas não-linearidades e quais impactos estas podem causar naquela que talvez seja uma das características mais básicas da voz: sua frequência fundamental. Como descrito no capítulo anterior, o modelo IF72 dos vinte indivíduos processados são avaliados por uma técnica numérica de cálculo de modos não-lineares normais (MNNs).

Para ilustrar os diversos MNNs calculados, foi usada uma representação gráfica baseada na relação entre frequência-energia (GFE) onde a energia total conservada (soma da energia cinética das massas e a potencial das molas) é computada a partir das condições iniciais que geraram o MNN. A frequência é diretamente calculada a partir do período T estimado pelo método. Como o sistema é não-linear, níveis diferentes de energia introduzidos no modelo através de suas condições iniciais produzem modos diferentes (ao contrário dos MLN que são constantes em sistemas lineares). Como são testados variados níveis de energia, pois as condições iniciais são incrementadas a cada iteração, é associado a cada patamar de energia do IF72 dois tipos de MNNs: em fase e fora de fase. A um conjunto de MNNs do mesmo tipo (*i.e.*, em fase ou fora de fase) em diferentes níveis de energia atribui-se o termo “famílias” ou “ramo” de MNNs que geram uma curva (ou ramo) que mostra visualmente como os MNNs de um mesmo tipo

Figura 4.9 – Modos não-lineares normais do indivíduo 18. Em (a) são exibidos a família de MNNs em fase (curva inferior) e fora de fase (curva superior). Estas curvas são mostradas individualmente em (b) e (c). Em cada uma das três curvas são ilustrados três MNNs nos gráficos (d) a (i). Nestes gráficos é possível ver o período de movimento de x_1 (curva sólida) e x_2 (curva pontilhada), as curvas modais (x_1 versus x_2) e a transformada de Fourier de x_1 .

evoluem através do incremento de energia do sistema. As Figuras 4.9 a 4.11 ilustram

estas curvas para o caso de três indivíduos que representam qualitativamente os vinte processados. O anexo A ilustra os resultados estimados para os demais indivíduos.

Na Figura 4.9(a), um MMN é representado por um ponto desenhado na frequência correspondente ao mínimo período que satisfaz a condição de movimento periódico e a um nível de energia correspondente à energia total do sistema que é introduzida pelas condições iniciais. O aumento destas condições introduz uma maior amplitude de deslocamento das massas e consequentemente uma maior energia mecânica. Como mencionado, o GFE é formado por duas curvas (ou ramos) que representam os MNs cujos movimentos de x_1 e x_2 estão em fase ou fora de fase. Como revelam os resultados, os modos não-lineares podem ser grosseiramente interpretados como a continuação dos correspondentes MLN em baixas amplitudes (*i.e.*, baixos níveis de energia). Note que nos ramos em fase (Figura 4.9(b)) e fora de fase (Figura 4.9(c)) em condições de energia menores que 1.10^{-6} J, os MNs possuem a mesma frequência dos MLNs³⁷ do modelo linearizado e mesmo com incrementos moderados de energia, esta frequência é fixa durante um determinado intervalo. Os gráficos das Figuras 4.9(d) e 4.8(g) revelam que a energia não é suficiente para excitar as não-linearidades do sistema e por isto este apresenta uma relação linear entre os deslocamentos x_1 e x_2 (curva modal) e não há criação de componentes harmônicos no sinal.

A análise gráfica dos resultados revela que a medida que incrementamos a energia do ramo em fase (ver pontos (e) e (f) da Figura 4.9(b)) o sistema passa a variar a frequência fundamental de vibração. Isto mostra que o sistema que representa o comportamento das PV é sensível ao nível de energia presente e existe uma relação de dependência entre frequência e energia exclusiva de sistemas não-lineares (Worden & Tomlinson, 2001). Em consequência disto, a resposta em frequência do modelo não é invariante, diferentemente de sistemas lineares cuja frequência natural e movimentos MLN são invariantes ao nível de energia do sistema. Os gráficos (e) e (f) da Figura 4.9

³⁷ Para calcular a frequência do MLN de um indivíduo deve ser usada a expressão (3.24).

também revelam que o incremento da energia produz curvas modais que não possuem uma relação linear entre as coordenadas do movimento³⁸ devido às distorções produzidas pelas não-linearidades nos deslocamentos. A criação de uma segunda componente harmônica em x_l é percebida e é outra marca característica do comportamento dinâmico não-linear. No ramo fora de fase, este efeito é menor conforme mostram as Figuras 4.9(h) e 4.9(i) e as distorções são consideravelmente menores ou até mesmo insignificantes. A maior parte dos indivíduos analisados nesta pesquisa apresentou um comportamento similar ao visto no exemplo da Figura 4.9. A Tabela 4.2 ilustra alguns

Tabela 4.2 – Modos não-lineares dos indivíduos que compõem a amostra.

<i>Ind.</i>	<i>Baixa energia²</i>		<i>Alta energia³</i>		<i>Ocorrência fenômeno</i>
	$F_{EF}(Hz)^1$	$F_{FF}(Hz)^1$	$F_{EF}(Hz)$	$F_{FF}(Hz)$	
1	114	206	145	229	<i>Não</i>
2	141	261	200	297	<i>Não</i>
3	168	418	207	438	<i>Não</i>
4	131	160	161	195	<i>Não</i>
5	85	258	122	273	<i>Não</i>
6	151	276	217	315	<i>Não</i>
7	132	219	185	251	<i>Não</i>
8	112	262	141	287	<i>Sim</i>
9	70	166	88	180	<i>Não</i>
10	138	268	196	298	<i>Não</i>
11	116	285	151	312	<i>Não</i>
12	168	271	232	313	<i>Não</i>
13	113	175	156	202	<i>Não</i>
14	202	382	229	458	<i>Sim</i>
15	133	242	192	280	<i>Não</i>
16	195	843	263	989	<i>Não</i>
17	141	278	204	311	<i>Não</i>
18	84	152	121	173	<i>Não</i>
19	128	326	*	353	<i>Sim</i>
20	200	822	268	973	<i>Não</i>

¹ Estimadas pela expressão (3.24). ² Valores calculados ao nível de energia de 10^{-6} J.

³Valores calculados ao nível de energia de 10^{-3} J. * Pela ocorrência de um possível fenômeno neste caso, não foi possível determinar este valor exatamente.

³⁸ Segundo Lacarbonara *et al.* (2003), esta relação não-linear entre as coordenadas do movimento inviabiliza a ortogonalidade entre os modos e assim a sua dissociação.

resultados referentes ao grupo processado.

Na Tabela 4.2 os resultados indicam que todos os indivíduos incrementaram a frequência de seus modos com o acréscimo de energia mecânica ao seu sistema. Em altas energias, não é possível estabelecer uma relação linear ponderada entre os dois modos que compõem os movimentos das pregas. O motivo disto é que o teorema da superposição não é mais válido nestas condições onde as não-linearidades tomam efeito significativo no sistema e o levam a vibrar com mais alta frequência e em alguns casos até a entrar em regime de produção de harmônicos.

Segundo Berry *et al.* (1996, 2001) mesmo embora o movimento das PV seja uma composição destes dois tipos de modos, o movimento em fase está diretamente associado ao movimento lateral das pregas que determina o ciclo de abertura e fechamento e a frequência fundamental da voz. Por esta questão, o modo com menor frequência geralmente está mais próximo da frequência fundamental da voz do indivíduo e nesta pesquisa o modo em fase foi o que mostrou menor frequência em todos os casos. Resultado similar em baixas frequências foi encontrado por Zhang *et al.* (2006) para um modelo linearizado. Os dados da Tabela 4.2 mostram que em situações de energia próximas a 1.10^{-3}J , todos os modos vibravam em frequências significativamente superiores as frequências com que estes mesmos sistemas trabalhavam com energias em níveis próximos a 1.10^{-6}J , ilustrando o impacto que as não-linearidades têm na frequência fundamental da voz.

O campo “*ocorrência de fenômeno*” exibido na Tabela 4.2 tem por propósito indicar a ocorrência de algum efeito ou resultado atípico percebido entre todas as amostras processadas e que desviam significativamente do padrão ilustrado na Figura 4.9. De todos vinte indivíduos processados, três deles apresentaram fenômenos adicionais sendo dois casos (indivíduos 8 e 19) de natureza similar. As Figuras 4.10 e 4.10 indicam os ramos de MNNs dos indivíduos 14 e 19, respectivamente.

Para o caso do indivíduo 14, aparentemente este apresenta um comportamento similar ao padrão ilustrado na Figura 4.9. Contudo, quando o ramo fora de fase é exibido com uma frequência correspondente à metade de sua original (curva cinza da Figura 4.10(a) e 4.10(b)), percebe-se que há uma interseção entre o ramo em fase e este ramo representado metade da frequência do ramo fora de fase. Em termos matemáticos, isto implica dizer que uma solução periódica com período T é também periódica com período $2T$ (ramo cinza). Logo, o que se vê é uma convergência de soluções que pode ser interpretada como uma espécie de ressonância interna³⁹ entre os ramos em fase e fora de fase criando uma interação entre estes e produzindo componentes de frequências dominantes nos movimentos MNNs (fenômeno similar foi percebido em (Kerschen *et al.*, 2009)). Deste modo, o que se tem para níveis de energia superiores a 1.10^{-4}J é uma transição suave entre o ramo em fase e o fora de fase. Os pontos (d) a (h) ilustram o desenvolvimento desta dinâmica.

A evolução da interação entre os modos começa pelo MNN (e) (Figura 4.10(e)). Antes disto (Figura 4.10(d)), o sistema apresenta comportamento predominantemente linear com uma única componente espectral. No MNN (e), a componente fundamental persiste como a de maior amplitude mesmo embora apareça uma segunda harmônica que tem sua contribuição aumentada até que esta assuma amplitude igual à frequência fundamental (ver Figura 4.10(f)). Deste ponto em diante, a segunda harmônica torna-se dominante (Figura 4.10(g)) na medida em que a fundamental decrementa até que em (h) torna-se praticamente a única componente de frequência. Como as demais componentes harmônicas desaparecem, a curva modal volta a tomar aspecto linear e muda sua inclinação indicado a presença dominante de um movimento fora de fase. Alguns autores (Shaw & Pierre, 1993; Peeters *et al.*, 2009) interpretam este processo como uma troca de energia entre os modos e as suas componentes harmônicas. De outra forma, este fenômeno remete a ideia de sincronização dos modos.

³⁹ Este termo assume outros nomes que têm significados iguais ou bastantes semelhantes como, por exemplo, emparelhamento dos modos ou sincronismo.

Figura 4.10 – Modos não-lineares normais do indivíduo 14. Em (a) são exibidos a família de MNNs em fase (curva inferior) e fora de fase (curva superior). A curva em cinza indica a mesma família de MNNs fora de fase mas com frequência duas vezes menor. Estes ramos de MNNs são mostrados individualmente em (b) e (c). Na curva de movimentos em fase, são ilustrados cinco MNNs que são descritos nos gráficos (d) a (h). Em (i) é ilustrado um MNN com movimento fora de fase. Nestes gráficos é possível ver o período de movimento de x_1 (curva sólida) e x_2 (curva pontilhada), as curvas modais (x_1 versus x_2) e a transformada de Fourier de x_1 .

O comportamento produzido nas Figuras 4.10(f) a 4.10(h) indicam uma grande e

irreal⁴⁰ amplitude de deslocamento para x_2 enquanto a amplitude de x_1 é pouco alterada e apresenta valores coerentes e aceitáveis para o deslocamento de uma PV. Esta ‘anomalia’ pode ser resultado do procedimento de cálculo de MNNs que busca encontrar condições iniciais (e um período T) que satisfaçam a equação (3.28) e para isto ajusta seus valores segundo a relação (3.27).

Como o método deve incrementar estas condições iniciais como forma de introduzir energia no sistema, algumas vezes este incremento é feito de forma desigual⁴¹ refletindo nas diferenças vistas entre as amplitudes de x_1 e x_2 . Apesar deste problema, percebe-se pela avaliação espectral de x_1 na Figura 4.10 que este mesmo efeito é visto nesta variável que possui valores coerentes de amplitude. Logo a interpretação deste fenômeno não é totalmente prejudicada por esta má definição de valores iniciais e o efeito dinâmico produzido aqui não é resultado integral do problema relativo às condições iniciais. De acordo com Kerschen *et al.*, (2009) é também possível excitar um modo de frequência maior e produzir uma resposta de larga amplitude através da ressonância entre estes modos. Tal dinâmica complexa foi observada em (Lee *et al.*, 2005) para um oscilador essencialmente não-linear com poucos graus de liberdade e foi comprovada em laboratório envolvendo estruturas metálicas com propriedades não-lineares.

Outro caso de fenômeno encontrado é ilustrado Figura 4.11 para o indivíduo 19 (resultado similar também foi percebido para o indivíduo 8). Como revela a Figura 4.11(b), novamente há uma interação entre o modo em fase com o modo fora de fase com frequência equivalente à metade da frequência original (*i.e.*, com período $2T$). Mesmo que o ponto (f) seja pertencente ao ramo em fase (*i.e.*, os valores de condições iniciais possuem mesmo sinal), o movimento produzido entre as duas massas dos

⁴⁰ Uma faixa típica de amplitude máxima de vibração das pregas vocais é de 0.5 a 2.0mm (Titze, 1988).

⁴¹ Considerando todas as amostras processadas, esta foi a única que aparentemente mostrou tal discrepância entre os valores iniciais.

osciladores é fora de fase. Isto acontece porque neste ponto há uma completa convergência com o modo fora de fase e a então segunda harmônica torna-se a frequência fundamental criando uma forma de onda que produz movimentos com aparência fora de fase (ver Figura 4.11(f)).

A novidade apresentada pela Figura 4.11 está na descoberta de uma “bifurcação” próxima ao ponto (e) que a partir deste gera dois ramos com características distintas. Nos casos onde a dinâmica do sistema segue para o ponto (f), há uma interação modal semelhante à descrita anteriormente (sincronismo entre os modos em fase e fora de fase). Contudo, há a possibilidade do sistema tomar uma dinâmica rumo aos pontos (g) e (h) onde a segunda harmônica tem presença importante, mas pelo fato de não interagir com nenhum outro modo, não consegue captar energia e por isto não assume amplitude superior à frequência fundamental. Contudo, a segunda harmônica gera grandes distorções nos movimentos como revelam os gráficos de deslocamentos x_1 e x_2 e a curva modal para estes dois MNNs.

No caso do ramo fora de fase (Figura 4.11(i)), embora exista uma segunda componente harmônica gerada pelas não-linearidades do IF72, não é visto grandes distorções no movimento das massas e na curva modal mesmo embora o nível de energia neste ponto seja relativamente alto. Isto revela que este modo é menos susceptível às não-linearidades. De maneira geral, os resultados dos indivíduos processados apontam que o modo fora de fase é geralmente menos susceptível às não-linearidades mesmo que estes também exibam uma grande variação entre frequência de vibração e energia.

Figura 4.11 – Modos não-lineares normais do indivíduo 19. Em (a) são exibidos a família de MNNs em fase (curva inferior) e fora de fase (curva superior). Estas curvas são mostradas individualmente em (b) e (c). A curva em cinza indica a mesma família de MNNs fora de fase mas com frequência duas vezes menor. Na curva de movimentos em fase, são ilustrados cinco MNNs que são descritos nos gráficos (d) a (h). Em (i) é ilustrado um MNN com movimento fora de fase. Nestes gráficos é possível ver o período de movimento de x_1 (curva sólida) e x_2 (curva pontilhada), as curvas modais (x_1 versus x_2) e a transformada de Fourier de x_1 .

Outra importante questão está ligada ao sincronismo dos MNNs em fase e fora de fase. Estes primeiros estão associados ao movimento lateral das pregas (responsável pela abertura e fechamento das mesmas) enquanto que os fora de fase reproduzem o movimento que determina uma forma glotal convergente ou divergente. Contudo, os resultados revelam que estes ciclos têm frequências divergentes para um mesmo nível de energia. Por exemplo, para o nível 10^{-3} J na Figura 4.9(b) e 4.9(c), o movimento em fase tem frequência aproximada de 121 Hz e o fora de fase 173 Hz. Segundo Berry *et al.* (2001), em uma “fonação modal” teoricamente estes dois modos devem se aproximar ou assumirem valores de frequência próximos ou iguais. Contudo, os resultados alcançados por estes autores em um modelo padrão IF72 (considerando uma resposta conservativa) usando técnicas lineares, mostram um largo espaçamento entre as frequências do modo em fase e fora de fase (120 e 201 Hz, respectivamente) dificultando o emparelhamento ou sincronismo entre estes dois tipos de modos. De maneira igual, os resultados ilustrados na Tabela 4.2 também revelaram um considerável espaçamento entre estas duas frequências para todos os indivíduos analisados aqui. Como questionaram os autores, “quais mecanismos existem para alinhar as duas frequências modais?”. Os próprios cogitaram a hipótese que as não-linearidades associadas à excitação, ao impacto entre as pregas e/ou dos tecidos poderiam produzir esta convergência⁴². Contudo, os resultados mostram nesta pesquisa que somente três indivíduos processados produziram um emparelhamento entre os modos na proporção de 1:2. Isto leva a crer que apenas as não-linearidades estruturais das pregas não são suficientes para produzir o sincronismo entre os modos na proporção de 1:1 e que as forças aerodinâmicas devem ter grande contribuição neste fundamento. A análise modal não-linear empregada aqui avalia de maneira individual cada modo e com esta técnica não é possível associar a um determinado movimento das pregas a contribuição individual de cada modo neste movimento. Mesmo assim, é possível esboçar o

⁴² Como foi utilizado um modelo mecânico de pregas vocais linear neste trabalho assim como modos lineares, não foi possível saber o efeito destas não-linearidades nos modos.

comportamento individual de cada modo segundo o nível de energia que o sistema possui e assim entender o comportamento das não-linearidades nos modos e no sistema.

4.2.2 – Avaliação da dinâmica usando diagramas de bifurcação e retratos de fase

Segundo Mergell e Herzel (1997), instabilidades na voz podem em alguns casos estarem relacionadas às bifurcações na dinâmica não-linear das PV. Geralmente elas são percebidas em vozes patológicas e experimentos práticos envolvendo pregas excisadas têm confirmado esta associação (Herzel, 1993; Berry *et al.*, 1996; Tokuda *et al.*, 2008). As bifurcações podem estar ligadas a pulsos alternantes de período duplo (*period-doubling*) ou triplo (*period-tripling*) ou até mesmo casos que remetam a comportamentos caóticos ou pseudo-aleatórios (Jiang *et al.*, 2006).

Como o objetivo de reproduzir estes possíveis comportamentos dinâmicos e ilustrar o emprego destas técnicas não-lineares aplicadas aos modelos determinísticos aqui produzidos, são estimados os diagramas de bifurcação do SH95 para três indivíduos. Geralmente as bifurcações são mais comuns em sujeitos com vozes patológicas e por esta razão foi simulada uma patologia de desequilíbrio (*e.g.*, assimetria) entre as pregas esquerda e direita através do parâmetro Q definido pela relação (3.33). Os resultados são sintetizados na Figura 4.12.

Segundo Döllinger *et al.* (2002), muitas patologias podem ser clinicamente explicadas por estas condições de assimetria entre as pregas que são aqui simuladas pelo fator Q . Tendo em vista esta abrangência, este parâmetro foi definido no eixo horizontal dos gráficos da Figura 4.12 e variou no intervalo $0,5 \leq Q_d \leq 1,0$ para indicar diferentes condições de assimetria. Quando $Q_d=1,0$, por exemplo, as pregas são simétricas e normais. Por outro lado, no caso onde $Q_d=0,5$, a massa e mola da porção

inferior da prega direita correspondem à metade do valor da prega esquerda. Para ilustrar as possíveis alterações produzidas no modelo por esta assimetria, foi considerada como variável de estado no eixo vertical dos diagramas de bifurcação o máximo valor de deslocamento de x_1 da prega direita (*i.e.*, x_{1d}).

Para os três diagramas de bifurcação ilustrados na Figura 4.12 foram encontrados valores de bifurcação. Estes revelam sob quais condições de assimetria o sistema se torna estruturalmente instável. Para o indivíduo 1, os valores de bifurcação foram concentrados por volta de $Q_d = 0,80$ e $Q_d = 0,52$ enquanto que para o indivíduo 2 eles foram encontrados por volta de $Q_d = 0,75$. No caso do indivíduo 3, o gráfico mostrou valores de bifurcação espalhados durante o gráfico. Estas características revelam mais uma vez que as PV apresentam uma dinâmica não-linear rica e complexa.

Para uma melhor interpretação da dinâmica do sistema a partir dos diagramas de bifurcação, utilizaram-se gráficos de retratos de fase. Esta técnica descreve visualmente a evolução do sistema dinâmico no espaço de fase. Neste caso, é exibida na

Figura 4. 12 – Diagramas de bifurcação de três indivíduos. O parâmetro Q_d representa o índice de assimetria entre as pregas definido pela expressão (3.33). Para um dado valor de Q_d , um simples ponto indica um estado estável enquanto que um grupo de n pontos registrados para um mesmo valor de Q_d podem indicar um período- n . Há também casos onde se vê um conjunto de pontos espalhados para um mesmo valor de índice de assimetria. Nestes casos eles podem indicar comportamento *quasi*-periódico ou caos.

Figura 4.13 para cada indivíduo a dinâmica do sistema antes (coluna à esquerda) e durante (coluna à direita) um valor de Q_d onde acontece uma bifurcação.

Os gráficos aqui encontrados mostram a projeção da trajetória do sistema no plano x_{1d} - x_{1e} . Para os indivíduos 1 e 2, antes dos valores de bifurcação, as trajetórias eram ciclo limites (Figuras 4.13(a) e 4.13(c)) e, após a bifurcação (Figuras 4.13(b) e 4.13(d)), tornaram-se ciclo limites de período duplo e triplo, respectivamente. Este tipo de comportamento está associado a pulsos alternantes no sinal de fluxo glotal. Nestas

Figura 4.13 – Retratos de fase dos indivíduos 1 (gráficos (a) e (b)), 2 (gráficos (c) e (d)) e 3 (gráficos (e) e (f)) usando o modelo SH95. Os eixos horizontal e vertical representam os deslocamentos da prega direita (x_{1d}) e esquerda (x_{1e}) em centímetros. Para cada indivíduo foram representadas as trajetórias no espaço de fase do sistema dinâmico durante um determinado valor Q_d (gráficos (a), (c) e (e)) e outro após este em situações onde acontecem bifurcações.

Figura 4.14 – O gráfico superior ilustra uma forma de onda *quasi*-periódica do sinal de fluxo glotal U_g simulado pelo modelo mecânico usando os parâmetros do indivíduo 3 com coeficiente de assimetria $Q_d = 0,61$. O gráfico inferior exibe o espectro de potência deste sinal.

situações, as pregas direita e esquerda vibram em diferentes frequências podendo produzir frequências fundamentais diferentes. Tal comportamento pode estar associado a patologias de diplofonia (Mergell *et al.*, 2000).

Para o indivíduo 3, em $Q_d = 0,62$ (ver Figura 4.13(e)), a trajetória do sistema tem geometria relativamente simples. Para este mesmo indivíduo, é possível observar uma alteração abrupta no comportamento dinâmico do sistema quando $Q_d = 0,61$. Neste ponto, a geometria do sistema é similar a de um torus (Figura 4.13(f)). Este resultado indica um complexo padrão de movimento vibracional que pode ser interpretado como uma órbita periódica de comportamento instável (Kantz & Schreiber, 2004) que pode conduzir a regimes harmônicos no sinal produzido na laringe. A Figura 4.14 ilustra o sinal produzido neste ponto de assimetria e seu correspondente espectro de potência. Uma fonação irregular é tipicamente caracterizada por um espalhamento espectral e muitas componentes harmônicas conforme se observa nos gráficos.

4.2.3 – Discussões gerais sobre as avaliações da dinâmica das pregas vocais

O emprego de métodos da dinâmica não-linear de sistemas é crescente e vem se mostrando importante em aplicações científicas envolvendo o estudo de voz. As análises mostradas anteriormente ilustram algumas destas potencialidades aplicadas à avaliação do comportamento não-linear das PV utilizando os modelos determinísticos otimizados pelo método desenvolvido nesta pesquisa.

Em especial, a análise modal não-linear produz uma visão estratégica das não-linearidades e seus efeitos durante o comportamento vibracional das PV. Mesmo sendo uma área emergente, tem mostrado importantes contribuições na análise de estruturas de vibração não-linear onde a teoria modal clássica é ineficaz. Outras técnicas (Shaw & Pierre, 1993) também usadas para identificar MNNs a partir de suas variedades⁴³ podem tratar de sistemas amortecidos. Contudo, sua complexidade dificulta a computação numérica direta dos MNNs. Graças à produção de um modelo determinístico de PV, é possível analisar esta interação entre os modos ilustrando mais uma vez a importância da produção de modelos determinísticos específicos para um determinado indivíduo (*i.e.*, extração de sua dinâmica determinística).

Alguns autores (Zhang *et al.*, 2006) questionam as diversas origens dos fenômenos não-lineares relacionando estas às origens acústicas, biomecânicas ou até mesmo um forte acoplamento entre o trato vocal, laringe e traquéia. Embora nenhuma destas alternativas possa ser descartada e devem apresentar significativas contribuições nos fenômenos essencialmente não-lineares, os resultados alcançados pela análise modal não-linear mostram que existe um comportamento não-linear estrutural nas propriedades intrínsecas das PV suficiente para produzir tais fenômenos, independentemente da natureza linear ou não-linear da excitação das PV ou de seu acoplamento. Assim, confirma-se que algumas das complexidades observadas na voz

⁴³ Tradução dada ao termo *manifolds*.

podem não ser causadas apenas por componentes estocásticos presentes na excitação do aparato vocal, mas sim por características não-lineares intrínsecas da estrutura dos tecidos das pregas e/ou o impacto entre elas. Outras não-linearidades conhecidas como, por exemplo, a relação não-linear entre a pressão do fluxo de ar nas pregas e a interação entre trato vocal e a laringe também devem possuir contribuição significativa, mas acredita-se que não são unicamente determinantes. Outros trabalhos (Luoa & Xie, 2001; Lee *et al.*, 2009) envolvendo a análise de sistemas vibro-impacto⁴⁴ têm confirmado que a colisão entre duas massas pode apresentar uma forte não-linearidade embutida e consequentemente uma dinâmica rica e complexa. Contudo, este é um assunto pouco explorado em pesquisas de vozes. Com a metodologia empregada aqui, não foi possível estabelecer as contribuições individuais das duas principais fontes de não-linearidades estruturais das pregas (*i.e.*, molas cúbicas representando as propriedades elásticas de seus tecidos e o impacto entre as massas).

O surgimento de uma bifurcação em alguns gráficos que relacionam frequência-energia dos MNNs (*i.e.*, GFE) é fruto essencialmente das não-linearidades do modelo e mesmo embora não seja uma característica muito comum em análises modais não-lineares, é um fenômeno previsto em sistemas com baixa dimensão e fortes não-linearidades (Shaw & Pierre, 1993; Lee *et al.*, 2005; Kerschen *et al.*, 2009) e confirma a complexidade e riqueza da dinâmica dos modelos de PV. Os resultados mostram que as não-linearidades assim como os seus fenômenos típicos têm efeito considerável em regiões de maior energia. Em termos práticos, quando as PV vibram sob a influência de uma alta pressão sub-glotal, elas absorvem uma maior energia gerando grandes deslocamentos. As conclusões de Herzel e Knudsen (1995) confirmam estes achados na medida em que revelam que comportamentos complexos atribuídos as não-linearidades das PV são mais frequentes quando existe uma pressão sub-glotal relativamente alta. Neste sentido, é possível relacionar a dinâmica observada nos GFE às observações

⁴⁴ Termo empregado para designar osciladores mecânicos que sofrem algum tipo de impacto entre seus elementos.

experimentais (Zhang & Jiang, 2005; Zhang *et al.*, 2007; Tokuda *et al.*, 2008). A relação entre frequência e energia é especialmente interessante para descrever como as não-linearidades agem nas PV em diferentes amplitudes de vibração e como elas podem alterar a frequência do sinal de fluxo glotal U_g . Neste sentido, trabalhos futuros envolvendo experimentos laboratoriais com pregas excisadas poderão confirmar os resultados teóricos alcançados aqui para os MNNs. De qualquer maneira, muitos dos experimentos teóricos envolvendo MNNs têm sido confirmados em laboratório (Lacarbonara *et al.*, 2003; Kerschen *et al.*, 2009; Peeters *et al.*, 2009) na área de análise mecânica de estruturas e vêm fortalecendo a validade e potencialidade deste método de análise de sistemas osciladores.

No que diz respeito aos diagramas de bifurcação e retratos de fase produzidos, estes podem ser usados para avaliar as condições de estabilidade das vibrações das PV e produzir atratores que podem ser empregados para uma classificação dos sinais produzidos como fez Giovanni *et al.* (1999). Os gráficos da Figura 4.13 mostraram que pequenas variações nos parâmetros elásticos das PV podem induzir a repentinos saltos para diferentes regimes dinâmicos. Foram também observados complicados padrões de bifurcação tais como períodos duplos e *torus*. Embora alguns trabalhos envolvendo laringes excisadas (Berry *et al.*, 1996; Jiang *et al.*, 2003) têm apontado a existência de comportamento caótico em circunstâncias de desequilíbrio entre as pregas, aqui, os valores de bifurcação e análises em gráficos de retrato de fase não indicaram comportamento caótico. De acordo com Steinecke e Herzel (1995), caos raramente acontece e é encontrado em baixos valores de Q e valores de pressão sub-glotal altos.

4.3 – Resumo

Este capítulo ilustrou os principais resultados alcançados no que diz respeito à extração da dinâmica das PV usando imagens de videolaringoscopia de alta velocidade e suas correspondentes avaliações dinâmicas empregando diagramas de bifurcação,

reconstruções em espaço de fase e, principalmente, a análise modal não-linear dos modelos dinâmicos reconstruídos.

O método de extração da dinâmica mostrou ser capaz de incorporar os principais padrões de vibrações extraídos das imagens em praticamente todos os vinte casos de indivíduos analisados. Quando comparados, o SH95 e IF72 mostram que este último é capaz de produzir um espectro mais rico em componentes de altas frequências que o SH95. Contudo, os sinais temporais de área glotal possuem geralmente um encaixe melhor para o SH95 que incorpora com menor erro os padrões extraídos das imagens.

As análises dinâmicas realizadas nos modelos SH95 e IF72 produziram uma série de fenômenos típicos de sistemas não-lineares e com isto reforçam o comportamento não-linear das PV e exibem como se comportam os modos de vibração de diferentes indivíduos. Principalmente, ilustram uma forte relação de dependência entre frequência fundamental de vibração e a energia mecânica contida nos tecidos das pregas.

Capítulo 5

Conclusões e considerações finais

J á há pelo menos meio século (van-den-Berg, 1958) tem-se investido esforços científicos para tentar desvendar as características dinâmicas das PV, sua interação aerodinâmica com o fluxo de ar e estruturas adjacentes e os complexos fenômenos oriundos da laringe responsáveis pela produção primária da voz. De forma similar, a comunidade científica também tem se dedicado a produzir modelos físicos e matemáticos capazes de simular, para diversos propósitos, o comportamento das pregas. Estas são iniciativas que perduram até os dias de hoje e por mais que se tenham desenvolvido avanços, há muito o que explorar. Prova disto são as recentes descobertas e aplicações trazidas da dinâmica não-linear de sistemas às pesquisas envolvendo vozes. Outra grande contribuição de fundamental importância veio através da tecnologia de videolaringoscopia de alta velocidade que possibilitou o registro do real ciclo de vibração das PV em detalhes nunca antes vistos.

O método desenvolvido aqui para extrair o comportamento vibracional das PV a partir de registros de videolaringoscopia de alta velocidade mostrou-se capaz de simular, com relativa precisão, os dados extraídos das imagens. Estes dados simulados são de natureza completamente determinística e por isto podem ser avaliados por métodos da dinâmica não-linear que impõem esta condição de análise. Ainda, o uso de um modelo otimizado permite um completo controle de toda simulação do comportamento dinâmico das pregas o que torna possível uma análise entre causa e consequência mais apurada. Sob a luz destas questões, é possível concluir que o método

desenvolvido é uma abordagem válida para uma descrição qualitativa (ou até mesmo quantitativa) e por isto é um meio de investigação compreensivo do comportamento das PV e suas propriedades mecânicas e acústicas. Por meio da técnica proposta, também foi possível simular uma ampla diversidade de cenários envolvendo vibrações normais (e algumas vezes patológicas), e graças a esta extração dinâmica determinística foi possível avaliar aspectos teóricos fundamentais das vibrações que estão vinculadas aos modos do sistema e como estes agem diante de algumas não-linearidades e têm influência no sinal produzido na laringe.

Os resultados indicaram que mesmo embora os modelos SH95 e IF72 provenham de um mesmo modelamento de duas massas, estes possuem características distintas sendo que o primeiro mostra maior facilidade em reproduzir os movimentos de vibração das pregas vocais enquanto o último (IF72) aparentemente tem mais facilidade em reproduzir um espectro mais rico em componentes.

As avaliações dinâmicas realizadas nos modelos otimizados apontaram para uma potencial existência de uma série de comportamentos complexos que podem ser produzidos pelas pregas que vão desde a produção de duas frequências fundamentais na voz, regime sub-harmônico ou até mesmo interações modais e a produção de sinais pseudo-aleatórios na laringe. Todos estes fenômenos são resultados da presença de não-linearidades, haja visto que sistemas lineares não são capazes de produzir tal comportamento (Kantz & Schreiber, 2004).

A análise modal não-linear revelou que os efeitos das não-linearidades são dependentes do nível de energia mecânica que o sistema possui e que por esta questão a resposta em frequência do sistema torna-se dependente do nível de excitação aplicada (*i.e.*, a dinâmica é dependente da entrada do sistema). Adicionalmente, na maior parte dos casos onde os níveis de energia eram relativamente altos, houve a produção de regimes harmônicos que introduziram distorções adicionais ao sinal. Além destas, foram produzidos em alguns casos complexas interações entre os modos que resultaram em

uma dinâmica relativamente sofisticada. Isto mostra que mesmo desconsiderando a interação não-linear do fluxo de ar com as pregas e o acoplamento destas com as outras estruturas adjacentes na laringe, as PV por si só são capazes de produzir determinados fenômenos no sinal devido às características visco-elásticas de seus tecidos e ao impacto produzido durante a vibração. Ainda, a não-linearidade se comporta de diferentes modos para cada indivíduo como era de se esperar haja visto que cada um possui uma dinâmica diferente. Ademais, as não-linearidades são importantes mecanismos para moldar a forma das curvas modais de tal forma a permitir ou viabilizar o emparelhamento modal.

Trabalhos futuros deverão explorar a produção de modelos determinísticos de PV com dois ou três graus de liberdade que apresentem uma maior suavidade em suas equações, um modelo de impacto mais realístico que leve em consideração coeficientes de restituição e uma maior flexibilidade na representação da fase de fechamento das pregas. Em algumas situações, a adição de uma componente estocástica à parte determinística pode ajudar a reproduzir componentes importantes no sinal e a modelá-lo com maior realismo. Ainda, o uso de sinais de videolaringoscopia de alta velocidade associados a sinais de filtragem inversa podem contribuir para a produção de referências mais sólidas para a otimização e por isto merecem uma investigação adicional. Por fim, uma comprovação adicional das interações modais simuladas aqui podem contribuir significativamente para a consolidação e entendimento dos mecanismos dinâmicos básicos das PV.

Apêndice A

Gráficos das amostras processadas

Neste anexo são ilustrados os principais gráficos gerados pelas análises de todos os indivíduos estudados nesta pesquisa. Em especial, são exibidos na primeira parte os gráficos dos sinais extraídos das imagens (SGE) e os simulados (SGS) pelos modelos SH95 e IF72. Para análises comparativas e avaliação de características destes sinais, são exibidos seus respectivos espectros de potência a derivada de área glotal. Na segunda parte deste anexo são mostrados os gráficos de frequência-energia (GFE) associados aos modos não-lineares processados para cada um dos indivíduos.

No caso dos gráficos da Figura A.1 a A.20 são mostrados os espectros de potência dos sinais extraídos e simulados usando uma janela de Hanning com largura média de 60 ms. As curvas na cor cinza indicam os sinais extraídos das imagens e os de cor preta os simulados pelo modelo SH95 (colunas à esquerda) e IF72 (colunas à direita).

Figura A.1 – Amostra do indivíduo 1. Sinais extraídos das imagens (sinais cinzas), simulados pelo modelo SH95 (sinais pretos; coluna à esquerda) e pelo IF72 (sinais pretos; coluna à direita).

Figura A.2 – Amostra do indivíduo 2. Sinais extraídos das imagens (sinais cinzas), simulados pelo modelo SH95 (sinais pretos; coluna à esquerda) e pelo IF72 (sinais pretos; coluna à direita).

Figura A.3 – Amostra do indivíduo 3. Sinais extraídos das imagens (sinais cinzas), simulados pelo modelo SH95 (sinais pretos; coluna à esquerda) e pelo IF72 (sinais pretos; coluna à direita).

Figura A.4 – Amostra do indivíduo 4. Sinais extraídos das imagens (sinais cinzas), simulados pelo modelo SH95 (sinais pretos; coluna à esquerda) e pelo IF72 (sinais pretos; coluna à direita).

Figura A.5 – Amostra do indivíduo 5. Sinais extraídos das imagens (sinais cinzas), simulados pelo modelo SH95 (sinais pretos; coluna à esquerda) e pelo IF72 (sinais pretos; coluna à direita).

Figura A.6 – Amostra do indivíduo 6. Sinais extraídos das imagens (sinais cinzas), simulados pelo modelo SH95 (sinais pretos; coluna à esquerda) e pelo IF72 (sinais pretos; coluna à direita).

Figura A.7 – Amostra do indivíduo 7. Sinais extraídos das imagens (sinais cinzas), simulados pelo modelo SH95 (sinais pretos; coluna à esquerda) e pelo IF72 (sinais pretos; coluna à direita).

Figura A.8 – Amostra do indivíduo 8. Sinais extraídos das imagens (sinais cinzas), simulados pelo modelo SH95 (sinais pretos; coluna à esquerda) e pelo IF72 (sinais pretos; coluna à direita).

Figura A.9 – Amostra do indivíduo 9. Sinais extraídos das imagens (sinais cinzas), simulados pelo modelo SH95 (sinais pretos; coluna à esquerda) e pelo IF72 (sinais pretos; coluna à direita).

Figura A.10 – Amostra do indivíduo 10. Sinais extraídos das imagens (sinais cinzas), simulados pelo modelo SH95 (sinais pretos; coluna à esquerda) e pelo IF72 (sinais pretos; coluna à direita).

Figura A.11 – Amostra do indivíduo 11. Sinais extraídos das imagens (sinais cinzas), simulados pelo modelo SH95 (sinais pretos; coluna à esquerda) e pelo IF72 (sinais pretos; coluna à direita).

Figura A.12 – Amostra do indivíduo 12. Sinais extraídos das imagens (sinais cinzas), simulados pelo modelo SH95 (sinais pretos; coluna à esquerda) e pelo IF72 (sinais pretos; coluna à direita).

Figura A.13 – Amostra do indivíduo 13. Sinais extraídos das imagens (sinais cinzas), simulados pelo modelo SH95 (sinais pretos; coluna à esquerda) e pelo IF72 (sinais pretos; coluna à direita).

Figura A.14 – Amostra do indivíduo 14. Sinais extraídos das imagens (sinais cinzas), simulados pelo modelo SH95 (sinais pretos; coluna à esquerda) e pelo IF72 (sinais pretos; coluna à direita).

Figura A.15 – Amostra do indivíduo 15. Sinais extraídos das imagens (sinais cinzas), simulados pelo modelo SH95 (sinais pretos; coluna à esquerda) e pelo IF72 (sinais pretos; coluna à direita).

Figura A.16 – Amostra do indivíduo 16. Sinais extraídos das imagens (sinais cinzas), simulados pelo modelo SH95 (sinais pretos; coluna à esquerda) e pelo IF72 (sinais pretos; coluna à direita).

Figura A.17 – Amostra do indivíduo 17. Sinais extraídos das imagens (sinais cinzas), simulados pelo modelo SH95 (sinais pretos; coluna à esquerda) e pelo IF72 (sinais pretos; coluna à direita).

Figura A.18 – Amostra do indivíduo 18. Sinais extraídos das imagens (sinais cinzas), simulados pelo modelo SH95 (sinais pretos; coluna à esquerda) e pelo IF72 (sinais pretos; coluna à direita).

Figura A.19 – Amostra do indivíduo 19. Sinais extraídos das imagens (sinais cinzas), simulados pelo modelo SH95 (sinais pretos; coluna à esquerda) e pelo IF72 (sinais pretos; coluna à direita).

Figura A.20 – Amostra do indivíduo 20. Sinais extraídos das imagens (sinais cinzas), simulados pelo modelo SH95 (sinais pretos; coluna à esquerda) e pelo IF72 (sinais pretos; coluna à direita).

Nas próximas figuras são ilustrados os modos não-lineares normais (MNNs) calculados através dos modelos mecânicos de PV estimados para os 20 indivíduos que participaram desta pesquisa. No diagrama que relaciona frequência de vibração com a energia mecânica do sistema (*i.e.*, GFE), a linha tracejada indica MNNs com movimento fora de fase (180°) e a contínua quando os movimentos estão em fase. Em cada curva, são ilustrados 3 MNNs rotulados como os pontos A, B e C para os movimentos em fase, e D, E e F para aqueles que estão fora de fase. A relação entre x_1 e x_2 é traçada para ilustrar o quanto as não-linearidades deformam esta relação de coordenadas entre as massas. As imagens também revelam o espectro de x_1 .

Figura A.21 –Gráfico frequência-energia dos modos não-lineares normais do indivíduo 1. A linha sólida indica os modos em fase e a pontilha os fora de fase. Para ilustrar o comportamento de alguns modos são exibidas as curvas modais (x_1 versus x_2) e a transformada de Fourier de x_1 para seis diferentes modos rotulados como A, B, C, D e F.

Figura A.22 –Gráfico frequência-energia dos modos não-lineares normais do indivíduo 2. A linha sólida indica os modos em fase e a pontilha os fora de fase. Para ilustrar o comportamento de alguns modos são exibidas as curvas modais (x_1 versus x_2) e a transformada de Fourier de x_1 para seis diferentes modos rotulados como A, B, C, D e F.

Figura A.23 –Gráfico frequência-energia dos modos não-lineares normais do indivíduo 3. A linha sólida indica os modos em fase e a pontilha os fora de fase. Para ilustrar o comportamento de alguns modos são exibidas as curvas modais (x_1 versus x_2) e a transformada de Fourier de x_1 para seis diferentes modos rotulados como A, B, C, D e F.

Figura A.24 –Gráfico frequência-energia dos modos não-lineares normais do indivíduo 4. A linha sólida indica os modos em fase e a pontilha os fora de fase. Para ilustrar o comportamento de alguns modos são exibidas as curvas modais (x_1 versus x_2) e a transformada de Fourier de x_1 para seis diferentes modos rotulados como A, B, C, D e F.

Figura A.25 –Gráfico frequência-energia dos modos não-lineares normais do indivíduo 5. A linha sólida indica os modos em fase e a pontilha os fora de fase. Para ilustrar o comportamento de alguns modos são exibidas as curvas modais (x_1 versus x_2) e a transformada de Fourier de x_1 para seis diferentes modos rotulados como A, B, C, D e F.

Figura A.26 –Gráfico frequência-energia dos modos não-lineares normais do indivíduo 6. A linha sólida indica os modos em fase e a pontilha os fora de fase. Para ilustrar o comportamento de alguns modos são exibidas as curvas modais (x_1 versus x_2) e a transformada de Fourier de x_1 para seis diferentes modos rotulados como A, B, C, D e F.

Figura A.27 –Gráfico frequência-energia dos modos não-lineares normais do indivíduo 7. A linha sólida indica os modos em fase e a pontilha os fora de fase. Para ilustrar o comportamento de alguns modos são exibidas as curvas modais (x_1 versus x_2) e a transformada de Fourier de x_1 para seis diferentes modos rotulados como A, B, C, D e F.

Figura A.28 –Gráfico frequência-energia dos modos não-lineares normais do indivíduo 8. A linha sólida indica os modos em fase e a pontilha os fora de fase. Para ilustrar o comportamento de alguns modos são exibidas as curvas modais (x_1 versus x_2) e a transformada de Fourier de x_1 para seis diferentes modos rotulados como A, B, C, D e F.

Figura A.29 –Gráfico frequência-energia dos modos não-lineares normais do indivíduo 9. A linha sólida indica os modos em fase e a pontilha os fora de fase. Para ilustrar o comportamento de alguns modos são exibidas as curvas modais (x_1 versus x_2) e a transformada de Fourier de x_1 para seis diferentes modos rotulados como A, B, C, D e F.

Figura A.30 –Gráfico frequência-energia dos modos não-lineares normais do indivíduo 10. A linha sólida indica os modos em fase e a pontilha os fora de fase. Para ilustrar o comportamento de alguns modos são exibidas as curvas modais (x_1 versus x_2) e a transformada de Fourier de x_1 para seis diferentes modos rotulados como A, B, C, D e F.

Figura A.31 –Gráfico frequência-energia dos modos não-lineares normais do indivíduo 11. A linha sólida indica os modos em fase e a pontilha os fora de fase. Para ilustrar o comportamento de alguns modos são exibidas as curvas modais (x_1 versus x_2) e a transformada de Fourier de x_1 para seis diferentes modos rotulados como A, B, C, D e F.

Figura A.32 –Gráfico frequência-energia dos modos não-lineares normais do indivíduo 12. A linha sólida indica os modos em fase e a pontilha os fora de fase. Para ilustrar o comportamento de alguns modos são exibidas as curvas modais (x_1 versus x_2) e a transformada de Fourier de x_1 para seis diferentes modos rotulados como A, B, C, D e F.

Figura A.33 –Gráfico frequência-energia dos modos não-lineares normais do indivíduo 13. A linha sólida indica os modos em fase e a pontilha os fora de fase. Para ilustrar o comportamento de alguns modos são exibidas as curvas modais (x_1 versus x_2) e a transformada de Fourier de x_1 para seis diferentes modos rotulados como A, B, C, D e F.

Figura A.34 –Gráfico frequência-energia dos modos não-lineares normais do indivíduo 14. A linha sólida indica os modos em fase e a pontilha os fora de fase. Para ilustrar o comportamento de alguns modos são exibidas as curvas modais (x_1 versus x_2) e a transformada de Fourier de x_1 para seis diferentes modos rotulados como A, B, C, D e F.

Figura A.35 –Gráfico frequência-energia dos modos não-lineares normais do indivíduo 15. A linha sólida indica os modos em fase e a pontilha os fora de fase. Para ilustrar o comportamento de alguns modos são exibidas as curvas modais (x_1 versus x_2) e a transformada de Fourier de x_1 para seis diferentes modos rotulados como A, B, C, D e F.

Figura A.36 –Gráfico frequência-energia dos modos não-lineares normais do indivíduo 16. A linha sólida indica os modos em fase e a pontilha os fora de fase. Para ilustrar o comportamento de alguns modos são exibidas as curvas modais (x_1 versus x_2) e a transformada de Fourier de x_1 para seis diferentes modos rotulados como A, B, C, D e F.

Figura A.37 –Gráfico frequência-energia dos modos não-lineares normais do indivíduo 17. A linha sólida indica os modos em fase e a pontilha os fora de fase. Para ilustrar o comportamento de alguns modos são exibidas as curvas modais (x_1 versus x_2) e a transformada de Fourier de x_1 para seis diferentes modos rotulados como A, B, C, D e F.

Figura A.38 –Gráfico frequência-energia dos modos não-lineares normais do indivíduo 18. A linha sólida indica os modos em fase e a pontilha os fora de fase. Para ilustrar o comportamento de alguns modos são exibidas as curvas modais (x_1 versus x_2) e a transformada de Fourier de x_1 para seis diferentes modos rotulados como A, B, C, D e F.

Figura A.39 –Gráfico frequência-energia dos modos não-lineares normais do indivíduo 19. A linha sólida indica os modos em fase e a pontilha os fora de fase. Para ilustrar o comportamento de alguns modos são exibidas as curvas modais (x_1 versus x_2) e a transformada de Fourier de x_1 para seis diferentes modos rotulados como A, B, C, D e F.

Figura A.40 –Gráfico frequência-energia dos modos não-lineares normais do indivíduo 20. A linha sólida indica os modos em fase e a pontilha os fora de fase. Para ilustrar o comportamento de alguns modos são exibidas as curvas modais (x_1 versus x_2) e a transformada de Fourier de x_1 para seis diferentes modos rotulados como A, B, C, D e F.

Referências bibliográficas

- ABARBANEL, H. D. I. Analysis of Observed Chaotic Data. Spriger-Verlag, Nwe York, 1995.
- AHMAD, K.; Y. YAN, *et al.* Analysis of Vocal-fold Vibrations from High-speed Laryngeal Images Using a Hilbert Transform-Based Methodology. *Journal of Voice*, v. 19, p. 161-175, 2005.
- ALIPOUR, F.; D. A. BERRY, *et al.* A finite-element model of vocal-fold vibration. *Journal of the Acoustical Society of America*, v. 108, p. 3003-3012, 2000.
- ALKU, P.; E. VILKMAN, *et al.* Estimation of amplitude features of the glottal flow by inverse filtering speech pressure signals. *Speech Communication*, v. 24, p. 123-132, 1998.
- BEHRMAN, A.; R. J. BAKEN. Correlation dimension of electroglottographic data from healthy and pathologic subjects. *Journal of the Acoustical Society of America*, v. 102, p. 2371-2379, 1997.
- BERG, J. V. D. Myoelastic-Aerodynamic Theory of Voice Production. *J. Speech Hear. Res.*, v. 1, p. 227-243, 1958.
- BERG, J. V. D.; J. ZANTEMA, *et al.* On the Air Resistance and Bernoulli Effect of the Juman Larynx. *Journal of the Acoustical Society of America*, v. 29, p. 626-631, 1957.
- BERRY, D. A. Mechanics of modal and nonmodal phonation. *Journal of Phonetics*, v. 29, p. 431-450, 2001.
- BERRY, D. A.; H. HERZEL, *et al.* Interpretation of biomechanical simulations of normal and chaotic vocal fold oscillations with empirical eigenfunctions. *Journal of the Acoustical Society of America*, v. 95, p. 3595-3604, 1994.
- BERRY, D. A.; H. HERZEL, *et al.* Bifurcations in Excised Larynx Experiments. *Journal of Voice*, v. 10, p. 129-138, 1996.
- BERRY, D. A.; I. R. TITZE. Normal modes in a continuum model of vocal fold tissues. *Journal of the Acoustical Society of America*, v. 100, p. 3345-3354, 1996.

- BOEK, W.; G. H. WIENEKE, *et al.* Clinical relevance of the fractal dimension of Fo perturbations computed by the box-counting method. *Journal of Voice*, v. 11, p. 437–442, 1997.
- BONNANS, J. F.; J. C. GILBERT, *et al.* Numerical optimization, theoretical and numerical aspects. Springer, 2006.
- CHEN, C.-T. Linear system theory and design. Oxford University Press, New York, 1999.
- CHILDERS, D.; D. HICKS, *et al.* A model for vocal fold vibratory motion, contact area, and the electroglottogram. *Journal of the Acoustical Society of America*, v. 80, p. 1309-1320, 1986.
- CHILDERS, D. G.; C. K. LEE. Vocal quality factors: analysis, synthesis, and perception. *Journal of the Acoustical Society of America*, v. 90, p. 2394-2410, 1991.
- CITARDI, M. J.; E. YANAGISAWA, *et al.* Videoendoscopic analysis of laryngeal function during laughter. *Ann. Otol. Rhinol. Laryngol.* , v. 105, p. 545-549, 1996.
- DECKER, G. Z.; S. L. THOMSON. Computational Simulations of Vocal Fold Vibration: Bernoulli Versus Navier–Stokes. *Journal of Voice*, v. 21, p. 273-284, 2007.
- DEUFLHARD, P.; H.-J. PESCH, *et al.* A modified continuation method for the numerical solution of nonlinear two-point boundary value problems by shooting techniques. *Numer. Math.*, v. 26, p. 327-343, 1976.
- DOLLINGER, M.; U. HOPPE, *et al.* Vibration parameter extraction from endoscopic image series of the vocal folds. *IEEE Transactions on Biomedical Engineering*, v. 49, p. 773-781, 2002.
- ECKMANN, J.-P.; S. O. KAMPHORST, *et al.* Liapunov exponents from time series. *Physical Review E*, v. 34, p. 4971-4979, 1986.
- FINKELHOR, B. K.; I. R. TITZE, *et al.* The effect of viscosity changes in the vocal folds on the range of oscillation. *Journal of Voice*, v. 1, p. 320-325, 1998.

- FLANAGAN, J. L.; L. L. LANDGRAF. Self-oscillating source for vocal-tract synthesizers. *IEEE Transactions on Audio and Electroacoustics*, v. 16, p. 57-63, 1968.
- FLANAGAN, L. J.; L. CHERRY. Excitation of vocal-tract synthesizers. *Journal of the Acoustical Society of America*, v. 45, p. 764-769, 1969.
- FRASER, A. M.; H. L. SWINNEY. Independent coordinates for strange attractors from mutual information. *Phys Rev Lett.*, v. 33, p. 1134-1140, 1986.
- GIOVANNI, A.; M. OUAKNINE, *et al.* Nonlinear Behavior of Vocal Fold Vibration: The Role of Coupling Between the Vocal Folds. *Journal of Voice*, v. 13, p. 465-476, 1999.
- GIOVANNI, A.; M. OUAKNINE, *et al.* Determination of largest Lyapunov exponents of vocal signal: application to unilateral laryngeal paralysis. *Journal of Voice*, v. 13, p. 341-354, 1998.
- GOLDBERG, D. E. Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley, 1989.
- GÓMEZ-VILDA, P.; R. FERNANDEZ-BAILLO, *et al.* Evaluation of Voice Pathology Based on the Estimation of Vocal Fold Biomechanical Parameters. *Journal of Voice*, v. 21, p. 450-476, 2007.
- GÓMEZ-VILDA, P.; R. FERNÁNDEZ-BAILLO, *et al.* Glottal source biometrical signature for voice pathology detection. *Speech Communication*, v. 51, p. 759-781, 2009.
- GONZALEZ, R. C.; R. E. WOODS. Digital Image Processing. 2002.
- GRASSBERGER, P.; I. PROCACCIA. Measuring the strangeness of strange attractors. *Physica D*, v. 9, p. 189-208, 1983.
- GUNTER, H. E. A mechanical model of vocal-fold collision with high spatial and temporal resolution. *Journal of the Acoustical Society of America*, v. 113, p. 994-1000, 2003.

- HAIWU, R.; X. WANG, *et al.* Subharmonic response of a single-degree-of-freedom nonlinear vibro-impact system to a narrow-band random excitation. *Physical Review E*, v. 80, p. 1-8, 2009.
- HERTRICH, I.; W. LUTZENBERGER, *et al.* Fractal dimension of sustained vowel productions in neurological dysphonias: an acoustic and electroglottographic analysis. *Journal of the Acoustical Society of America*, v. 102, p. 652-654, 1997.
- HERZEL, H. Bifurcations and chaos in voice signals. *Applied Mechanics Reviews*, v. 46, p. 399-413, 1993.
- HERZEL, H.; D. BERRY, *et al.* Nonlinear dynamics of the voice: signal analysis and biomechanical modeling. *Chaos*, v. 5, p. 30-34, 1995.
- HERZEL, H.; D. BERRY, *et al.* Analysis of vocal disorders with methods from nonlinear dynamics. *Journal of Speech and Hearing Research*, v. 37, p. 1008-1019, 1994.
- HERZEL, H.; C. KNUDSEN. Bifurcations in a vocal fold model. *Nonlinear Dynamics*, v. 7, p. 53-64, 1995.
- HIRANO, M. The vocal cord during phonation. *Igaku no Ayumi*, v. 80, p. 1-14, 1968.
- HIRANO, M. Morphological Structure of the Vocal Cord as a Vibrator and its Variations. *Folia phoniatic.*, v. 26, p. 89-94, 1974.
- HONG, K. H.; H. K. KIM, *et al.* Laryngeal gestures during stop production using high speed digital images. *Journal of Voice*, v. 16, p. 207-214, 2002.
- ISHIZAKA, K.; J. L. FLANAGAN. Synthesis of voiced sounds from a two-mass model of vocal cords. *Bell Syst. Tech. J.*, v. 51, p. 1233-1268, 1972.
- JIANG, J.; E. LIN, *et al.* Vocal folds physiology. *Otolaryngologic Clinics of North America*, v. 33, p. 699-718, 2000.
- JIANG, J. J.; Y. ZHANG. Chaotic vibration induced by turbulent noise in a two-mass model of vocal folds. *Journal of the Acoustical Society of America*, v. 112, p. 2127-2133, 2002.

- JIANG, J. J.; Y. ZHANG. Nonlinear dynamic analysis of speech from pathologic subjects. *Electronics Letters*, v. 38, p. 294–295, 2002.
- JIANG, J. J.; Y. ZHANG, *et al.* Nonlinear dynamics of phonations in excised larynx experiments. *Journal of the Acoustical Society of America*, v. 114, p. 1–8, 2003.
- JIANG, J. J.; Y. ZHANG, *et al.* Chaos in Voice, From Modeling to Measurement. *Journal of Voice*, v. 20, p. 2-17, 2006.
- JIANG, J. J.; Y. ZHANG, *et al.* Modeling of chaotic vibrations in symmetric vocal folds. *Journal of the Acoustical Society of America*, v. 110, p. 2120–2128, 2001.
- KAKITA, Y.; H. OKAMOTO. Visualizing the characteristics of vocal fluctuation from the viewpoint of chaos: an attempt toward qualitative quantification. Singular Publishing Group, San Diego, CA, 1995.
- KANTZ, H.; T. SCHREIBER. Nonlinear time series analysis. Cambridge Press, Cambridge 2004.
- KAPLAN, D.; L. GLASS. Understanding Nonlinear Dynamics. Springer-Verlag, New York, 1995.
- KENNEL, M. B.; R. BROWN, *et al.* Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical Review A*, v. 45, p. 3403-3411, 1992.
- KERSCHEN, G.; M. PEETERS, *et al.* Nonlinear normal modes, Part I: A useful framework for the structural dynamicist. *Mechanical Systems and Signal Processing*, v. 23, p. 170-194, 2009.
- KUMAR, A.; S. K. MULLICK. Nonlinear dynamical analysis of speech. *Journal of the Acoustical Society of America*, v. 100, p. 615–629, 1996.
- LACARBONARA, W.; G. REGA, *et al.* Ressonant non-linear normal modes, part I: analytical treatment for structural one-dimensional systems. *International Journal of Non-Linear Mechanics*, v. 38, p. 851–872, 2003.
- LEE, Y. S.; G. KERSCHEN, *et al.* Complicated dynamics of a linear oscillator with a light, essentially nonlinear attachment. *Physica D*, v. 204, p. 41–69, 2005.

- LEE, Y. S.; F. NUCERA, *et al.* Periodic orbits, damped transitions and targeted energy transfers in oscillators with vibro-impact attachments. *Physica D*, v. 238, p. 1868-1896, 2009.
- LINDEBERG, T. Detecting salient Blob-like image structures and their scales with a scale-space primal sketch: a method for focus-of-attention. *International Journal of Computer Vision*, v. 11, p. 283-318, 1993.
- LOHSCHELLER, J.; M. DOLLINGER, *et al.* Quantitative investigation of the vibratory pattern of the substitute voice generator. *IEEE Trans Biomed Eng.*, v. 51, p. 1394-1400, 2004.
- LOHSCHELLER, J.; U. EYSHOLDT. Phonovibrography: Mapping High-Speed Movies of Vocal Fold Vibrations Into 2-D Diagrams for Visualizing and Analyzing the Underlying Laryngeal Dynamics. *IEEE Transactions on Medical Imaging*, v. 27, p. 300-310, 2008.
- LORENZ, E. N. Deterministic nonperiodic flow. *Journal of Atmospheric Sciences*, v. 20, p. 130-141, 1963.
- LOUS, N. J. C.; G. C. J. HOFMANS, *et al.* A symmetrical two-mass vocal-fold model coupled to vocal tract and trachea, with application to prosthesis design. *Acta Acoustica*, v. 84, p. 1135-1150, 1998.
- LUCERO, J. C.; L. L. KOENIG. Simulations of temporal patterns of oral airflow in men and women using a two-mass model of the vocal folds under dynamic control *Journal of the Acoustical Society of America*, v. 117, p. 1362-1372, 2005.
- LUOA, G.; J. XIE. Bifurcations and chaos in a system with impacts. *Physica D*, v. 148, p. 183-200, 2001.
- MANEVITCH, L.; O. GENDELMAN. Oscillatory models of vibro-impact type for essentially non-linear systems. *Journal of Mechanical Engineering Science*, v. 222, p. 2007-2043, 2008.
- MAY, R. M. Simple mathematical models with very complicated dynamics. *Nature*, v. 261, p. 459-475, 1976.
- MENDE, W.; H. HERZEL, *et al.* Bifurcations and chaos in newborn infant cries. *Physics Letters A* v. 145, p. 418-424, 1990.

- MERGELL, P.; H. HERZEL. Modelling biphonation—the role of the vocal tract. *Speech Communication*, v. 22, p. 141-154, 1997.
- MERGELL, P.; H. HERZEL, *et al.* Irregular vocal-fold vibration - High-speed observation and modeling. *Journal of the Acoustical Society of America*, v. 108, p. 2996-3002, 2000.
- MUL, F. F. M. D.; N. A. GEORGE, *et al.* Depth-kymography of vocal fold vibrations: part II. Simulations and direct comparisons with 3D profile measurements. *Physics in Medicine and Biology*, v. 54, p. 3955–3977, 2009.
- OLIVEIRA, M. R.; J. C. PEREIRA. Aerodynamic study of three-dimensional larynx models using finite element methods. *Journal of Sound and Vibration*, v. 311, p. 39-55, 2008.
- OTT, E.; T. SAUER, *et al.* *Coping with Chaos*. Wiley, New York, 1994.
- PARKER, T. S.; L. O. CHUA (1989). Bifurcation diagrams. *Practical Numerical Algorithms for Chaotic Systems*, Springer-Verlag.
- PEETERS, M.; R. VIGUIÉ, *et al.* Nonlinear normal modes, Part II: Toward a practical computation using numerical continuation techniques. *Mechanical Systems and Signal Processing*, v. 23, p. 195-216, 2009.
- PELORSON, X. Theoretical and experimental study of quasisteady-flow separation within the glottis during phonation. Application to a modified two-mass model. *Journal of the Acoustical Society of America*, v. 96, p. 3416-3431, 1994.
- PINCUS, S. M. Approximate entropy as a measure of system complexity. *Proc. Natl. Acad. Sci.*, v. 88, p. 2297-2301, 1991.
- PINHEIRO, A. P.; J. C. PEREIRA, *et al.* (2010). Biomechanical modeling of laryngeal dynamics using inverse filtering of speech signals. *17th International Conference on Systems, Signals and Image Processing*. Rio de Janeiro, 1: 109-112.
- PINHEIRO, A. P.; D. E. STEWART, *et al.* Analysis of nonlinear dynamics of vocal folds using high-speed video observation and biomechanical modeling. *Digital Signal Processing*, (em impressão), 2011.

- QIN, X.; S. WANG, *et al.* Improving reliability and accuracy of vibration parameters of vocal folds based on high-speed video and electroglottography. *IEEE Transactions on Biomedical Engineering*, v. 56, p. 1744-1754, 2009.
- ROSENBERG, R. M. On nonlinear vibrations of systems with many degrees of freedom. *Advances in Applied Mechanics*, v. 9, p. 155-242, 1966.
- ROUBEAU, B.; N. HENRICH, *et al.* Laryngeal Vibratory Mechanisms: The Notion of Vocal Register Revisited. *Journal of Voice*, v. 23, p. 425-438, 2009.
- SAUER, T.; J. A. YORKE, *et al.* Embedology. *Journal of Statistical Physics*, v. 65, p. 579-616, 1991.
- SCHWARZ, R.; M. DOLLINGES, *et al.* Spatio-temporal quantification of vocal fold vibrations using high-speed videoendoscopy and a biomechanical model. *Journal of the Acoustical Society of America*, v. 123, p. 2717-2732, 2008.
- SHAW, S. W.; C. PIERRE. Normal modes for non-linear vibratory systems. *Journal of Sound and Vibration*, v. 164, p. 85-124, 1993.
- STAM, C. J. Nonlinear dynamical analysis of EEG and MEG: Review of an emerging field. *Clin. Neurophysiol.*, v. 116, p. 2266-2301, 2005.
- STEINECKE, I.; H. HERZEL. Bifurcations in an asymmetric vocal-fold model. *J. Acoust. Soc. Am.*, v. 97, p. 1874-1884, 1995.
- STORY, B. H. An overview of the physiology, physics and modeling of the sound source for vowels. *Acoustical Science and Technology*, v. 23, p. 195-206, 2002.
- STORY, B. H.; I. R. TITZE. Voice simulation with a body-cover model of the vocal folds. *J. Acoust. Soc. Am.*, v. 97, p. 1249-1260, 1995.
- TAO, C.; J. J. JIANG. A self-oscillating biophysical computer model of the elongated vocal fold. *Computers in Biology and Medicine*, v. 38, p. 1211-1217, 2008.
- TAO, C.; Y. ZHANG, *et al.* Extracting physiologically relevant parameters of vocal folds from high-speed video image series. *IEEE Transactions on Biomedical Engineering*, v. 54, p. 794-801, 2007.

- THOMSON, S. L.; L. MONGEAU, *et al.* Aerodynamic transfer of energy to the vocal folds. *Journal of the Acoustical Society of America*, v. 118, p. 1689-1700, 2005.
- TITZE, I. R. The human vocal cords: a mathematical model, part I. *Phonetica*, v. 28, p. 129-170, 1973.
- TITZE, I. R. The human vocal cords: a mathematical model, part II. *Phonetica*, v. 29, p. 1-21, 1974.
- TITZE, I. R. On the mechanics of vocal fold tissues. *Journal of the Acoustical Society of America*, v. 60, p. 1366-1380, 1976.
- TITZE, I. R. The physics of small-amplitude oscillation of the vocal folds. *Journal of the Acoustical Society of America*, v. 83, p. 1536-1552, 1988.
- TITZE, I. R. Regulating glotta airflow in phonation: application of the maximum power transfer theorem to a low dimensional phonation model. *Journal of the Acoustical Society of America*, v. 111, p. 367-376, 2002.
- TITZE, I. R.; R. BAKEN, *et al.* Evidence of chaos in vocal fold vibration. Singular Publishing Group, San Diego, 1993.
- TITZE, I. R.; B. H. STORY. Acoustic interactions of the voice source with the lower vocal tract. *Journal of the Acoustical Society of America*, v. 101, p. 2234-2243, 1997.
- TOKUDA, I.; H. HERZEL. Detecting synchronizations in an asymmetric vocal fold model from time series data. *Chaos*, v. 15, p. 013702, 2005.
- TOKUDA, I. T.; J. HORACEK, *et al.* Bifurcations and chaos in register transitions of excised larynx experiments. *Chaos*, v. 18, p. 013102, 2008.
- VAN-DEN-BERG, J. Myoelastic-aerodynamic theory of voice production. *Journal of Speech and Hearing Research*, v. 3, p. 227-244, 1958.
- VAZIRI, G.; F. ALMASGANJ, *et al.* Pathological assessment of patients' speech signals using nonlinear dynamical analysis. *Computers in Biology and Medicine*, v. 40, p. 54-63, 2010.

- WANG, Z.; X. NING, *et al.* Nonlinear dynamic characteristics analysis of synchronous 12-lead ECG signals. *IEEE Eng. Med. Biol. Mag.*, v. 19, p. 110-115, 2000.
- WOOD, S. A.; J. D. HOFORD, *et al.* A method for measurement of cross sectional area, segment length, and branching angle of airway tree structures in situ. *Comput. Med. Imag. Graph.*, v. 19, p. 145-152, 1995.
- WORDEN, K.; G. R. TOMLINSON. Nonlinearity in structural dynamics: detection, identification and modeling. IOP Publishing, Philadelphia, 2001.
- WURZBACHER, T.; I. VOIGT, *et al.* Calibration of laryngeal endoscopic high-speed image sequences by an automated detection of parallel laser line projections. *Medical Image Analysis*, v. 12, p. 300-317, 2008.
- XUE, Q.; R. MITTAL, *et al.* A computational study of the effect of vocal-fold asymmetry on phonation. *Journal of the Acoustical Society of America*, v. 128, p. 818-827, 2010.
- YAN, Y.; X. CHEN, *et al.* Automatic Tracing of Vocal-Fold Motion From High-Speed Digital Images. *IEEE Transaction on Biomedical Engineering*, v. 53, p. 1394-1400, 2006.
- YAN, Y.; E. DANROSE, *et al.* Functional Analysis of Voice Using Simultaneous High-speed Imaging and Acoustic Recordings. *Journal of Voice*, v. 21, p. 604-616, 2007.
- YANG, A.; J. LOHSCHELLER, *et al.* Biomechanical modeling of the three-dimensional aspects of human vocal fold dynamics. *Journal of the Acoustical Society of America*, v. 127, p. 1014-1031, 2010.
- YANG, A.; M. STINGL, *et al.* Computational of physiological human vocal fold parameters by mathematical optimization of a biomechanical model. *Journal of the Acoustical Society of America*, v. 130, p. 949-964, 2011.
- ZHANG, Y.; E. BIEGING, *et al.* Efficient and Effective Extraction of Vocal Fold Vibratory Patterns from High-Speed Digital Imaging. *Journal of Voice*, v. 24, p. 21-29, 2010.
- ZHANG, Y.; J. J. JIANG. Nonlinear dynamic analysis in signal typing of pathological human voices. *Electronics Letters*, v. 39, p. 1021-1023, 2003.

- ZHANG, Y.; J. J. JIANG. Chaotic vibrations of a vocal fold model with a unilateral polyp. *Journal of the Acoustical Society of America*, v. 115, p. 1266-1269, 2004.
- ZHANG, Y.; J. J. JIANG. Spatiotemporal chaos in excised larynx vibrations. *Physical Review E*, v. 72, p. 035201, 2005.
- ZHANG, Y.; J. J. JIANG. Nonlinear dynamic mechanism of vocal tremor from voice analysis and model simulations. *Journal of Sound and Vibration*, v. 316, p. 248-262, 2008.
- ZHANG, Y.; J. J. JIANG, *et al.* Quantifying the complexity of excised larynx vibrations from high-speed imaging using spatiotemporal and nonlinear dynamic analyses. *Chaos*, v. 17, p. 043114, 2007.
- ZHANG, Y.; J. SHAO, *et al.* High-speed image analysis reveals chaotic vibratory behaviors of pathological vocal folds. *Chaos, solitons & Fractals*, v. 44, p. 169-177, 2011.
- ZHANG, Y.; C. TAO, *et al.* Parameter estimation of an asymmetric vocal-fold system from glottal area time series using chaos synchronization. *Chaos*, v. 16, p. 023118, 2006.
- ZHANG, Z.; J. NEUBAUER, *et al.* Aerodynamically and acoustically driven modes of vibration in a physical model of the vocal folds. *Journal of the Acoustical Society of America*, v. 150, p. 2841-2849, 2006.
- ZHANG, Z.; J. NEUBAUER, *et al.* The influence of subglottal acoustics on laboratory models of phonation. *Journal of Acoustical Society of America*, v. 120, p. 1558-1569, 2006.