

# **Cirurgia Virtual da Laringe**

ARLINDO NETO MONTAGNOLI

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

Orientador: Prof. Dr. José Carlos Pereira

São Carlos  
2006

Aos meus pais, Pierina e Arlindo (in memoriam) e ao meu filho Pedro.

# Agradecimentos

Agradeço ao Prof. Dr. José Carlos Pereira por me orientar durante este período.

À Unesp por possibilitar e incentivar meu aperfeiçoamento.

Aos Dr. Marcos Grellet, Dr. Domingos H. Tsuji e Dr. Walther Jaeger por terem fornecido os exames de laringoscopia.

À fonoaudióloga Maria Eugenia Dajer por realizar a avaliação perceptiva das vozes.

Ao amigo Prof. Dr. Ricardo José Ferrari, pelas discussões pertinentes ao processamento de imagens.

À minha esposa Marley, à minha família e amigos por me apoiarem durante a realização dessa tese.

À todas as pessoas que colaboraram no desenvolvimento dessa tese.

# Resumo

Neste trabalho foram desenvolvidas técnicas computacionais que auxiliam os especialistas da fonologia a prever os efeitos produzidos na voz de um paciente quando alterações físicas e mecânicas são efetuadas na laringe, especificamente, nas pregas vocais após a realização da cirurgia.

O objetivo principal deste trabalho consiste na utilização de contornos ativos para desenvolver um modelo da laringe baseado no movimentos das pregas vocais e glote. A técnica é baseada em imagens obtidas do exame de endoscopia da laringe. A partir do sinal de voz do paciente é filtrado as características da glote obtidas da imagem da estrobo-laringoscopia, assumindo que as disfonias são causadas por irregularidades nas pregas vocais. As modificações feitas na glote são utilizadas para estimar um novo filtro glotal que é adicionada ao sinal previamente filtrado. Utilizando este método, pode-se obter uma nova voz sintetizada que mantém as características individuais dos pacientes após as alterações realizadas nas prega oriundas da cirurgia virtual da laringe.

Palavras-chave: cirurgia virtual; laringe; modelos deformáveis; síntese de voz; pregas vocais; glote.

# Abstract

The present study presents a computational method designed to assist phonologists to anticipate the effects produced in the voice, as a result of physical and mechanical alterations of the larynx model, due to a surgery to correct dysphonia. The main objective of the study was to use the active contours method to develop a larynx model based on vocal fold and glottis movements. The technique is based on images obtained from endoscopic exams of the larynx. Starting with the recorded voice signal of subject, we filter the characteristics of the glottal pulse obtained from the image exam, assuming that the dysphonia is caused by irregularities of the larynx. The modifications made in the glottis model are used to estimate a new glottal filter, which is then added to the filtered signal. Using this method, one can obtain a new voice that maintains the individual's personal characteristics after virtual surgery modifications.

Keywords: virtual surgery; larynx; deformable models; voice synthesis; vocal fold; glotte.

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Estruturas do aparelho fonador . . . . .	2
1.2	Fisiologia do aparelho fonador . . . . .	5
1.3	Objetivos . . . . .	9
1.4	Justificativas . . . . .	9
1.5	Organização desta tese . . . . .	10
<b>2</b>	<b>Revisão da literatura</b>	<b>11</b>
2.1	Modelos deformáveis . . . . .	11
2.1.1	Contorno ativo . . . . .	11
2.2	Modelos matemáticos para a produção da voz . . . . .	13
2.2.1	Modelo da excitação . . . . .	14
2.2.2	Modelo da glote . . . . .	14
2.2.3	Modelo do Trato Vocal . . . . .	15
2.2.4	Modelo da radiação . . . . .	15
2.3	Modelos mecânicos da laringe . . . . .	16
2.4	Processamento digital de sinais . . . . .	24
2.4.1	Processamento digital do sinal de voz . . . . .	24
2.5	Processamento digital de imagens médicas . . . . .	28
<b>3</b>	<b>Cirurgia virtual da laringe</b>	<b>31</b>
3.1	Processamento de Imagens . . . . .	31
3.1.1	Vídeoestroboscopia da laringe . . . . .	33
3.1.2	Processamento digital de imagens . . . . .	33
3.1.3	Redução de ruído na imagem . . . . .	34
3.1.4	Alteração no histograma da imagem . . . . .	34

3.1.5	Segmentação da imagem . . . . .	34
3.1.6	Quantificação . . . . .	35
3.1.7	Modelos deformáveis . . . . .	35
3.1.8	Observações práticas para construção do modelo . . . . .	40
3.1.9	Simulando o movimento glotal . . . . .	46
3.1.10	Pulso e filtro glotal a partir do modelo . . . . .	50
3.2	Processamento de áudio . . . . .	51
3.2.1	Gravação do sinal de voz . . . . .	51
3.2.2	Síntese da voz após modificações no modelo . . . . .	51
3.3	Programa para auxílio da vídeolaringoscopia e cirurgia virtual . . . . .	52
3.3.1	Banco de dados . . . . .	53
3.3.2	Gravação e edição de vídeo e extração do áudio . . . . .	53
3.3.3	Exame visual da laringoscopia . . . . .	53
3.3.4	Medidas da laringe . . . . .	53
3.3.5	Quimografia . . . . .	54
<b>4</b>	<b>Resultados</b>	<b>55</b>
4.1	Cirurgia Virtual . . . . .	55
4.1.1	Caso: Normal . . . . .	55
4.1.2	Caso: Normal com diminuição da rigidez da prega no lado direito . . . . .	55
4.1.3	Caso: Normal com eliminação da viscosidade das pregas . . . . .	58
4.1.4	Caso: Normal com adição de uma fenda na comissura superior . . . . .	58
4.1.5	Caso: Normal com adição de uma fenda no terço médio . . . . .	62
4.1.6	Caso: Normal com adição de massa na prega vocal esquerda . . . . .	62
4.1.7	Caso: Normal com adição de massa bilateralmente . . . . .	62
4.1.8	Caso: Cisto . . . . .	66
4.1.9	Outros Casos . . . . .	73
<b>5</b>	<b>Discussões</b>	<b>75</b>
5.1	União das amostras de voz . . . . .	75
5.2	Estroboscopia versus câmera de alta velocidade . . . . .	77
5.3	Mecanismos compensatórios . . . . .	77
5.4	Determinação da frequência fundamental . . . . .	77
5.5	Análise perceptiva . . . . .	78

5.6	Fechamento imperfeito da glote . . . . .	78
<b>6</b>	<b>Conclusões</b>	<b>79</b>
6.1	Trabalhos futuros . . . . .	80
<b>7</b>	<b>Referências Bibliográficas</b>	<b>82</b>
<b>A</b>	<b>Apêndice</b>	<b>91</b>
A.1	Estudo de casos . . . . .	91
A.1.1	Caso: Nódulo bilateral . . . . .	91
A.1.2	Caso: Leocoplasia . . . . .	94
A.1.3	Caso: Fenda no terço médio e sulco . . . . .	96
A.1.4	Caso: Paralisia . . . . .	105
A.1.5	Caso: Pólipo . . . . .	105
A.2	Programa Desenvolvido . . . . .	110
A.2.1	Banco de dados . . . . .	110
A.2.2	Captura e edição de vídeo . . . . .	110
A.2.3	Análise da endoscopia . . . . .	115
A.2.4	Medidas da laringe . . . . .	115
A.2.5	Quimografia . . . . .	115
A.2.6	Cirurgia virtual . . . . .	116
A.3	Programa Análise de Voz 5.0 . . . . .	116
A.3.1	Anamnese . . . . .	116
A.3.2	Análise acústica . . . . .	117
A.3.3	Espectrograma . . . . .	117
A.3.4	Outras funções do programa . . . . .	117

# Lista de Figuras

1.1	Aparelho fonador. . . . .	3
1.2	Estruturas da laringe. . . . .	4
1.3	Movimento das cartilagens da laringe, AP: ântero-posterior, ML: médio-lateral, AB: abdução, AD: adução. . . . .	6
1.4	Músculos da laringe, TA: tireoaritenóideo, CAP e CAL: cricoaritenóideo posterior e lateral, Ao e At: aritenóideo oblíquo e transversal, CT: crico-tireóideo. . . . .	6
1.5	Representação esquemática das camadas teciduais das pregas vocais. . . . .	6
1.6	Diagrama esquemático do sistema vocal. . . . .	8
1.7	Vista por cima e corte transversal das pregas vocais em um ciclo glótico. . . . .	8
2.1	Representação espectral do modelo linear da produção vocal. . . . .	17
2.2	Comparação entre um sinal de voz real e sintetizado para a vogal /a/. . . . .	18
2.3	Modelo uma massa. . . . .	20
2.4	Modelo duas massas. . . . .	20
2.5	Modelo duas massas. . . . .	21
2.6	Modelo multi-massas. . . . .	22
2.7	Modelo das tiras. . . . .	23
2.8	Modelo corpo-cobertura, ou três massas, das pregas vocais. . . . .	23
2.9	Modelo linear para produção da fala. . . . .	27
3.1	Diagrama de blocos: estágios desenvolvidos pelo método proposto. . . . .	32
3.2	Representação paramétrica de uma curva aberta e fechada. . . . .	36
3.3	Representação de uma curva discretizada. . . . .	37
3.4	Vetor tangente. . . . .	37
3.5	Curva se movendo somente sob ação da força elástica. . . . .	40

3.6	Curva se movendo somente sob ação de forças normais aos vértices. . . . .	41
3.7	Curva se movendo sob ação das derivadas de primeira e segunda ordem. . .	41
3.8	Expansão da curva com a remoção dos laços. . . . .	42
3.9	Sequência de quadros de um ciclo glotal. . . . .	43
3.10	Área de contato das pregas vocais (invertida). . . . .	43
3.11	Comparações entre area de contato das pregas vocais, eletroglotografia e pulso glotal. . . . .	44
3.12	Pulso glotal e sua derivada. . . . .	45
3.13	Expansão da glote devido a pressão do ar. . . . .	46
3.14	Expansão da glote devido a pressão do ar, com acréscimo da derivada de segunda ordem. . . . .	47
3.15	Simulação numérica do efeito 'zíper'. . . . .	47
3.16	Fase de fechamento ação da força elástica e de Bernoulli. . . . .	48
3.17	Relação entre a área glotal e o pulso de ar glotal. . . . .	50
3.18	Modelo linear para a fonte glotal. . . . .	50
3.19	Image estroboscópica, a linha selecionada para o corte e a respectiva videoestroboquimografia projetada abaixo. . . . .	54
4.1	Imagem de uma laringe normal obtida por um endoscópio rígido. . . . .	56
4.2	Exame quimográfico de um caso normal. . . . .	57
4.3	ciclo glotal do exame estroboscópico de um caso normal e o respectivo ciclo simulado. . . . .	58
4.4	simulação numérica do caso normal. . . . .	59
4.5	Ciclo glotal da simulação numérica com diminuição da rigidez no lado direito. . . . .	59
4.6	simulação numérica do caso normal com diminuição da rigidez no lado direito. . . . .	60
4.7	Ciclo glotal da simulação numérica sem viscosidade das pregas. . . . .	60
4.8	simulação numérica do caso normal sem viscosidade das pregas. . . . .	61
4.9	Ciclo glotal da simulação numérica com fenda na comissura superior. . .	62
4.10	simulação numérica do caso normal com fenda na comissura superior. . .	63
4.11	Ciclo glotal da simulação numérica com fenda no terço médio. . . . .	63
4.12	simulação numérica do caso normal com fenda no terço médio. . . . .	64

4.13	Ciclo glotal da simulação numérica com adição de massa na prega vocal esquerda. . . . .	64
4.14	simulação numérica do caso normal com adição de massa na prega vocal esquerda. . . . .	65
4.15	Ciclo glotal da simulação numérica com adição de massa bilateral nas pregas vocais. . . . .	65
4.16	simulação numérica do caso normal com adição de massa bilateral nas pregas vocais. . . . .	66
4.17	Posições dos cortes para os exames de quimografia. . . . .	67
4.18	Quimografia da região da fenda. . . . .	67
4.19	Quimografia da região do cisto. . . . .	68
4.20	Exame de endoscopia e simulação numérica do caso com cisto. . . . .	68
4.21	Simulação numérica do caso cisto com fenda na comissura posterior. . . . .	69
4.22	Simulação numérica da cirurgia virtual do caso cisto com fenda na comissura posterior. . . . .	70
4.23	Sinais de vozes antes e após a cirurgia virtual do caso com cisto. . . . .	71
4.24	Espectograma antes e depois da cirurgia virtual, caso cisto. . . . .	72
5.1	Espectograma com descontinuidade nas uniões das amostras do sinal de voz. . . . .	76
A.1	Exame de endoscopia e simulação numérica do caso com nódulo bilateral. . . . .	91
A.2	Simulação numérica do caso nódulo bilateral. . . . .	92
A.3	Simulação numérica da cirurgia virtual do caso nódulo bilateral. . . . .	93
A.4	Sinais de vozes antes e após a cirurgia virtual, caso nódulo. . . . .	94
A.5	Espectograma antes e depois da cirurgia virtual, caso nódulo. . . . .	95
A.6	Exame de endoscopia e simulação numérica do caso com leocoplasia. . . . .	96
A.7	Simulação numérica do caso leocoplasia. . . . .	97
A.8	Simulação numérica da cirurgia virtual do caso leocoplasia. . . . .	98
A.9	Sinais de vozes antes e após a cirurgia virtual, caso leocoplasia. . . . .	99
A.10	Espectograma antes e depois da cirurgia virtual, caso leocoplasia. . . . .	100
A.11	Exame de endoscopia e simulação numérica do caso com fenda e sulco. . . . .	101
A.12	Simulação numérica do caso fenda e sulco. . . . .	101
A.13	Simulação numérica da cirurgia virtual do caso fenda e sulco. . . . .	102

A.14	Sinais de vozes antes e após a cirurgia virtual, caso fenda e sulco. . . . .	103
A.15	Espectograma antes e depois da cirurgia virtual, caso fenda e sulco. . . . .	104
A.16	Exame de endoscopia e simulação numérica do caso com paralisia. . . . .	105
A.17	Simulação numérica do caso paralisia. . . . .	106
A.18	Simulação numérica da cirurgia virtual do caso paralisia. . . . .	107
A.19	Sinais de vozes antes e após a cirurgia virtual, caso paralisia. . . . .	108
A.20	Espectograma antes e depois da cirurgia virtual, caso paralisia. . . . .	109
A.21	Exame de endoscopia e simulação numérica do caso pólipó na prega vocal direita. . . . .	110
A.22	Simulação numérica do caso pólipó na prega vocal direita. . . . .	111
A.23	Simulação numérica da cirurgia virtual do caso pólipó na prega vocal direita.	112
A.24	Sinais de vozes antes e após a cirurgia virtual, caso pólipó. . . . .	113
A.25	Espectograma antes e depois da cirurgia virtual, caso pólipó. . . . .	114
A.26	Janela relativa ao gerenciamento do banco de dados. . . . .	115
A.27	Janela relativa à captura e edição das imagens de vídeo. . . . .	116
A.28	Janela relativa ao exame da endoscopia. . . . .	117
A.29	modelo do relatório. . . . .	118
A.30	Medidas das dimensões das pregas vocais. . . . .	119
A.31	Janela relativa à execução da quimografia. . . . .	120
A.32	Janela relativa à cirurgia virtual. . . . .	121
A.33	Janela relativa à anamnese. . . . .	122
A.34	Janela relativa à análise acústica do sinal de voz. . . . .	123
A.35	Janela relativa ao gerenciamento do banco de dados. . . . .	124
A.36	Janelas complementares do programa Análise de Voz 5.0 . . . . .	125

## Lista de Tabelas

4.1	Medidas das pregas vocais - caso normal . . . . .	56
4.2	Avaliação da dinâmica das pregas vocais - caso normal . . . . .	57
4.3	Avaliação perceptiva da voz - escala GRBAS. . . . .	73
4.4	Análise acústica dos casos estudados. . . . .	74

# Capítulo 1

## Introdução

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

Na maioria das vezes, quando ocorre alguma disfonia, ou seja, alguma alteração no comportamento vocal de uma pessoa, cuja causa é de origem orgânica, geralmente, essas alterações ocorrem na laringe, especificamente nas pregas vocais. Dentre as diversas patologias da laringe, as mais comuns são os pólipos, edemas, úlceras, granulomas e leoplasia. Muitas vezes a intervenção cirúrgica torna-se necessária, visando a melhoria da qualidade vocal, ou como uma medida profiláctica ao agravamento da patologia.

O mecanismo de funcionamento da laringe é bastante elaborado. Diversas teorias foram desenvolvidas para explicar seu funcionamento, como as teorias: mioelástica, aerodinâmica, das cordas vibrantes, do caos (HERZEL, 1993), entre outras. Uma das teorias mais aceitas é a teoria mioelástica-aerodinâmica. O correto movimento de vibração da mucosa depende de um perfeito equilíbrio de todas as estruturas que constituem a laringe. Esse equilíbrio é bastante crítico, e qualquer alteração em uma das partes desse conjunto é suficiente para comprometer ou, simplesmente, fazer com que a mucosa deixe de vibrar. O fluxo de ar produzido após a vibração da mucosa é chamado de pulso glotal, que depois de ser equalizado pelo trato vocal e nasal, e ao sair pela boca e nariz produz a voz humana.

A amplitude, frequência e forma são três parâmetros que descrevem um sinal temporal unidimensional, observando-se que estão implícitos na forma, a quantidade de harmônicos e ruídos contidos nesse sinal. Portanto, uma patologia na laringe que cause alguma disfonia, inevitavelmente altera um ou mais desses três parâmetros básicos que caracte-

rizam um sinal de voz. Uma laringe normal, para os sons vocálicos, produz vibrações praticamente constante em frequência e amplitude das pregas vocais, elas possuem uma pequena alteração chamadas de *jitter* e *shimmer*, respectivamente, que nos causam a sensação de voz humana e não de máquina.

A correlação entre a voz e o pulso glotal é muito estreita, pois sua forma correta é a responsável direta pela qualidade vocal, definindo as amplitudes e a quantidade de harmônicos encontrados no sinal de voz.

Quando há uma intervenção cirúrgica na laringe, a intenção inicial consiste em remover ou reparar alguma anomalia ou lesão, fazendo com que a laringe recupere sua forma saudável, exceto em casos em que o comprometimento orgânico é maior que o funcional, como exemplo, em casos em que se torna necessária a laringectomia.

As lesões na laringe modificam a forma do pulso glotal e, conseqüentemente, causam as disfonias devido a diversos fatores. O ar pode escapar pela glote devido à presença da lesão ou devido às alterações no tecido das pregas vocais, que modificam as condições mioelásticas das pregas, podendo inibir ou suprimir a vibração da mucosa.

Geralmente, as lesões de massa, como os nódulos, pólipos e edemas de Reinke, provocam uma perturbação no movimento normal das pregas vocais e, conseqüentemente, causam perturbações na frequência, amplitude e forma do sinal glotal que, por sua vez, também afeta o sinal de voz.

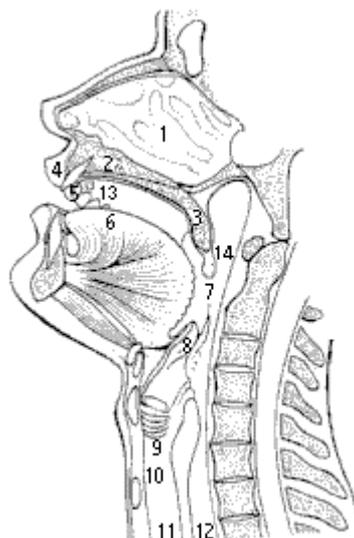
As alterações nas condições mioelástica das pregas causam modificações na amplitude dos harmônicos, geralmente diminuindo-os em relação a um sinal de voz normal. Ambas as alterações, de massa e mioelástica, podem ocasionar o escape de ar pelo não fechamento completo da glote, acarretando um acréscimo na razão sinal ruído da voz.

## 1.1 Estruturas do aparelho fonador

Os sons de nossa fala resultam da ação do sistema de órgãos que constituem o aparelho fonador e que criam as três condições necessárias para a produção da voz, como a geração da corrente de ar pelos pulmões, um obstáculo a esta corrente de ar feita pelas pregas vocais, e uma caixa de ressonância feita pela cavidade nasal e bucal. O aparelho fonador envolve estruturas anatômicas que se iniciam nos pulmões e terminam nos lábios e nariz, por onde o som é irradiado, (MARTINS, 2005; BEHLAU, 2001; BEHLAU; PONTES, 1995). Na Figura 1.1 são apresentadas algumas das estruturas mais importantes

que fazem parte do aparelho fonador.

Os pulmões, brônquios e traquéia, que constituem os órgãos respiratórios, funcionam



1) fossas nasais, 2) palato duro, 3) véu palatino, 4) lábios, 5) dentes, 6) língua, 7) faringe, 8) epiglote, 9) pregas vocais, 10) laringe, 11) traquéia, 12) esôfago, 13) alvéolos e 14) úvula

Figura 1.1: Aparelho fonador.

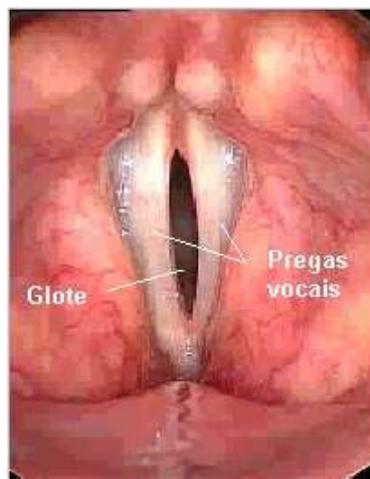
como dois foles, gerando o fluxo de ar, não apresentados na figura. A laringe, que possui nove cartilagens situadas na parte ântero-superior do pescoço, sob a base da língua, adiante da faringe e acima da traquéia, é detalhada na Figura 1.2. Na laringe fica localizada a glote, Figura 1.2(b), espaço produzido entre as duas pregas vocais. Na laringe encontra-se também a epiglote, válvula que obstrui a glote durante a deglutição, observando que proteger os pulmões dos alimentos e líquidos é a principal função da laringe. A faringe, que consiste em uma cavidade ligeiramente afunilada entre a boca e a parte superior do esôfago, conduz o ar para a boca e fossas nasais.

O palato mole, chamado também de véu palatino, quando levantado se cola à parte posterior da laringe e deixa livre apenas o conduto bucal, que associado aos movimentos da língua, lábios e maxilar são responsáveis pela articulação dos sons não-nasais. Quando o palato mole é abaixado, conecta-se as fossas nasais ao sistema fonatório, produzindo sons nasalados. Por último, encontra-se a boca e o nariz, que têm a função de irradiar os sons emitidos pelo aparelho fonador ao ambiente.

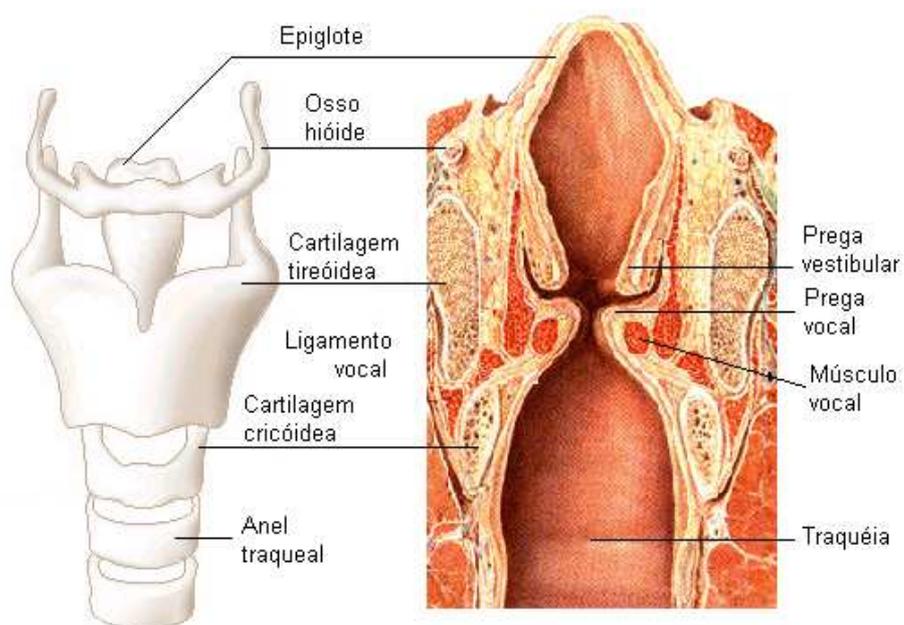
A musculatura da laringe é dividida em dois grupos: extrínsecos e intrínsecos. Os músculos extrínsecos são aqueles que suportam a laringe e sua função consiste em mover



(a) vista externa



(b) vista por cima



(c) cartilagens, ossos e corte longitudinal

Figura 1.2: Estruturas da laringe.

Fontes- <http://www.phon.ox.ac.uk/jcoleman/phonation.htm> e  
<http://www.gbmc.org/voice/anatomyphysiologyofthelarynx.cfm>

a laringe como um todo. Os músculos intrínsecos são responsáveis diretos pelo controle da função fonatória. Com exceção do músculo aritenóideo, todos os outros são encontrados aos pares. Os músculos da laringe são: tireoaritenóideo, cricoaritenóideo posterior, cricoaritenóideo lateral, aritenóideo, cricotireóideo, ariepiglótico e tireoepiglótico.

O músculo aritenóideo transversal é responsável por juntar as aritenóides em um movimento de deslizamento sob a cartilagem cricóidea, chamado de deslizamento médio-lateral. O movimento de rotação das aritenóides, complementa o movimento de adução, é realizado pelo músculo chamado de cricoaritenóideo lateral.

Os músculos tireoaritenóideos compõem o corpo das pregas vocais, possuindo dois feixes principais: o interno, chamado de músculo vocal, e o externo, músculo tireoaritenóideo. Esses músculos são responsáveis pelo deslizamento ântero-posterior das aritenóides, permitindo o afrouxamento das pregas vocais, conseqüentemente, espessando e encurtando, causando o movimento de adução.

O músculo cricoaritenóideo posterior, em forma de leque, é responsável pela abdução das pregas vocais, permitindo a respiração. O músculo cricotireóideo é responsável pelo movimento de rotação da cartilagem tireóidea, tendo um efeito secundário na ação de adução das pregas vocais.

Os movimentos das cartilagens podem ser observados na Figura 1.3 e os músculos responsáveis por esses movimentos estão apresentados na Figura 1.4.

A anatomia das pregas vocais proposta por Hirano e Kakita (1985), conhecida por corpo-cobertura, está ilustrado na Figura 1.5, sendo que, de modo geral, as pregas vocais são compostas de músculos e mucosa. A mucosa é composta pelo epitélio e lâmina própria e os músculos tireoaritenóideos fazem parte do corpo e sustentam a cobertura.

## 1.2 Fisiologia do aparelho fonador

Podemos dizer que os fonemas nascem na laringe e se completam na boca, e isso acontece graças ao concurso das arcadas dentárias, dos alvéolos do palato duro ou céu da boca, do palato mole e, sobretudo, devido à atividade da língua, dos lábios e das bochechas, os quais se movimentam para modificar a corrente de ar e produzir os diversos fonemas. As fossas nasais, cavidades situadas no maxilar superior, funcionam como caixa de ressonância dos fonemas nasais. O centro do cérebro que intervém na faculdade de falar encontra-se no córtex cerebral. Qualquer lesão, tanto nas extremidades das células do córtex ou das fibras que as unem, interfere nos processos da fala, (MARTINS, 2005;

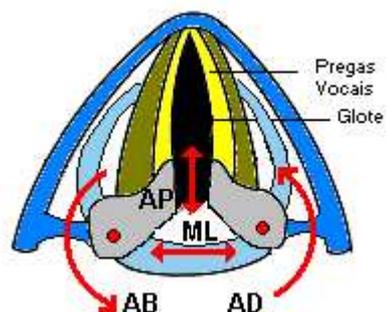


Figura 1.3: Movimento das cartilagens da laringe, AP: ântero-posterior, ML: médio-lateral, AB: abdução, AD: adução.

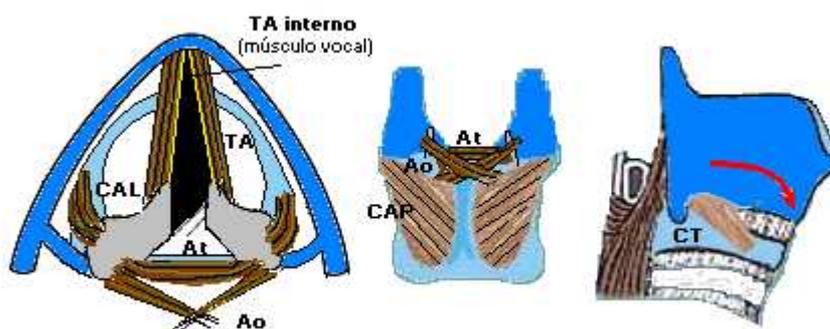


Figura 1.4: Músculos da laringe, TA: tíreoaritenóideo, CAP e CAL: cricoaritenóideo posterior e lateral, Ao e At: aritenóideo oblíquo e transversal, CT: cricotireóideo.



Figura 1.5: Representação esquemática das camadas teciduais das pregas vocais.

---

BEHLAU, 2001; ISSHIKI; TSUJI; SENNES, 1999).

O ar expelido dos pulmões, por via dos brônquios, penetra na traquéia e chega à laringe, onde, ao atravessar a glote, encontra o primeiro obstáculo à sua passagem. O fluxo de ar pode encontrá-las fechadas ou abertas, em virtude das bordas das pregas vocais estarem aproximadas ou afastadas. No primeiro caso, o ar força a passagem através das pregas vocais esticadas, fazendo-as vibrar e produzir os sons musicais chamados de vocálicos e também característicos das consoantes sonoras. No segundo caso, relaxando as pregas vocais, o ar escapa sem vibrações laríngeas, produzindo, assim, as consoantes denominadas de surdas.

Ao sair da laringe, a corrente expiratória entra na cavidade faríngea, que lhe oferece duas vias de acesso ao exterior: o canal bucal e o nasal. Suspenso no entrecruzar desses dois canais fica o véu palatino, órgão dotado de mobilidade capaz de obstruir ou não a passagem de ar na cavidade nasal e, conseqüentemente, de determinar a natureza oral ou nasal de um som. É, porém, na cavidade bucal que se produzem os movimentos fonadores variados, graças à maior ou menor separação dos maxilares, das bochechas e a mobilidade da língua e dos lábios.

Na produção dos sons da fala o trato vocal age como uma série de filtros acústicos que atenuam ou reforçam o tom original e fundamental. Esses filtros ou ressonadores são capazes de assumir uma infinidade de formas diferentes com orifício de tamanho variável, que produzem os sons característicos que distinguem um fonema do outro. Na Figura 1.6 é apresentado um resumo esquemático do sistema vocal.

Na produção dos sons vocálicos, o mecanismo vibratório é realizado basicamente pela passagem de ar pelas pregas vocais, que, inicialmente, encontram-se fechadas e com o aumento da pressão do ar infraglótico faz com que estas se abram e entrem em vibração, produzindo os ciclos glóticos, os quais são praticamente periódicos apresentando uma pequena flutuação em vozes reais, conforme pode ser visto no diagrama da Figura 1.7. Esses ciclos são usualmente divididos em quatro fases: fechada, de abertura, aberta e de fechamento. Durante a fase fechada a pressão do ar impele para cima e para os lados as pregas vocais, que começam a se abrir, dando início a fase de abertura e, posteriormente, permanecem por um curto intervalo de tempo na fase aberta. Devido à ação mioelástica e ao efeito Bernoulli, as pregas vocais começam a se fechar, iniciando a fase de fechamento

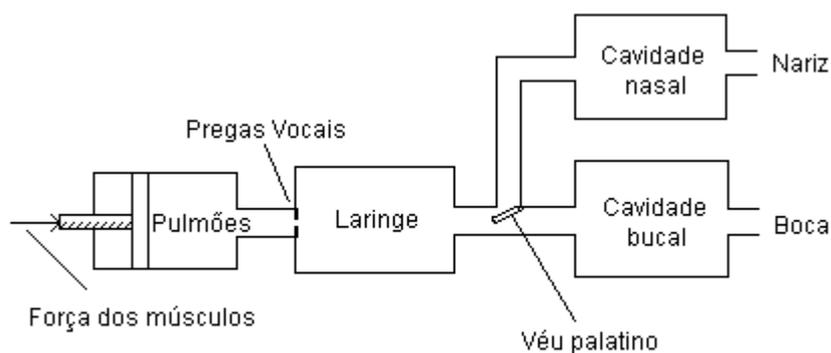


Figura 1.6: Diagrama esquemático do sistema vocal.

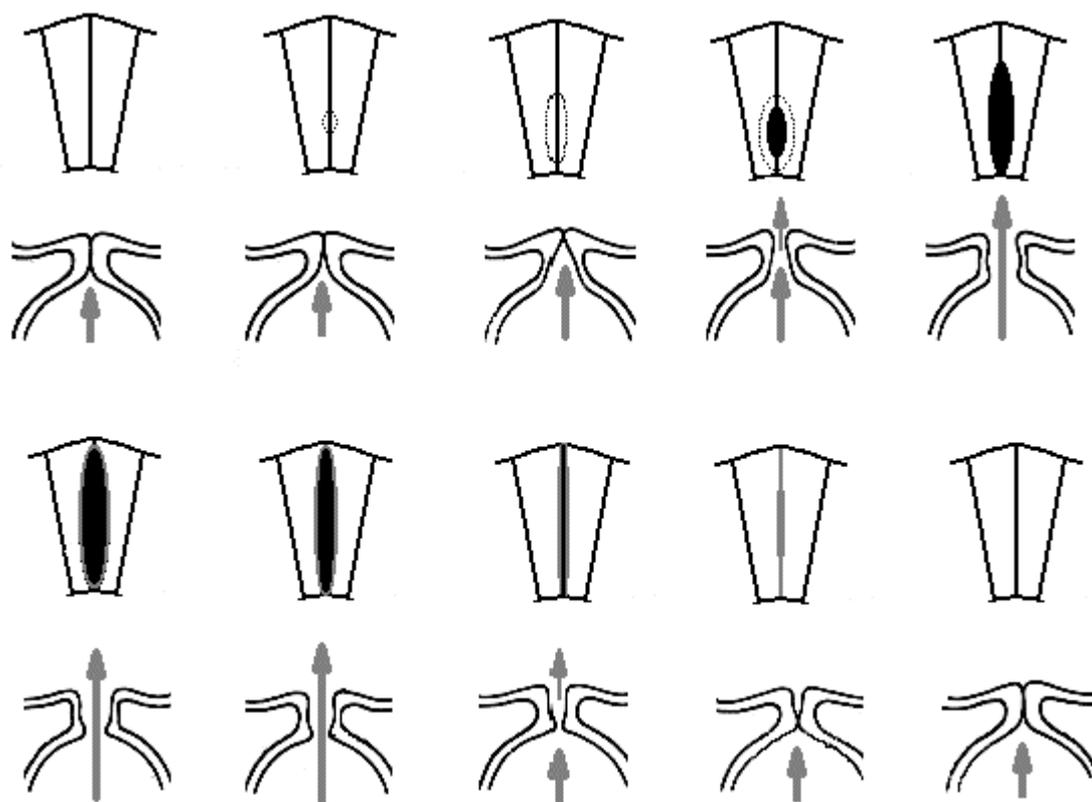


Figura 1.7: Vista por cima e corte transversal das pregas vocais em um ciclo glótico.

até atingir novamente a fase fechada, completando desta forma um ciclo glótico.

Segundo Sennes et al. (2000), para um completo entendimento do aparelho fonador é necessário o estudo exaustivo através de disseções, ressaltando, também, que devido à dificuldade na obtenção das peças anatômicas, em seu trabalho foi proposto a utilização de imagens tridimensionais para o estudo da laringe.

Outros detalhes da anatomia e fisiologia da produção vocal podem ser encontradas em diversas literaturas, (PEREIRA; MONTAGNOLI, 2005; BEHLAU, 2001; SENNES et al., 2000; ISSHIKI; TSUJI; SENNES, 1999; BEHLAU; PONTES, 1995; HIRANO; SATO, 1993; HIRANO, 1974).

### **1.3 Objetivos**

Os objetivos deste trabalho são:

1. Desenvolver um modelo de laringe utilizando técnicas de contornos ativos que simulem as condições da glote observadas pelo exame de videoestroboscopia.
2. Aplicar técnicas de processamento de sinais de voz, de tal forma a possibilitar a realização da cirurgia virtual da laringe.

### **1.4 Justificativas**

Muitos pacientes consideram uma operação na laringe desnecessária quando não há algesia e segundo os mesmos, convivem normalmente com suas disfonias. A preocupação com a qualidade vocal pós-operatória é um agravante que faz com que estes pacientes relutem e, muitas vezes, acabem desistindo de se submeterem a uma cirurgia, conseqüentemente, esta atitude pode levar a um agravamento do quadro clínico.

Além de desenvolver a cirurgia virtual da laringe, possibilitando ao paciente prever os resultados fonatórios antes da realização da cirurgia, este trabalho realizou uma revisão dos conceitos envolvidos nos processos mecânicos na produção da voz humana e os modelos matemáticos para sua síntese.

## 1.5 Organização desta tese

Esta tese foi dividida em 8 capítulos, juntamente com as referências bibliográficas e os anexos. Este trabalho foi escrito em Latex e segue as diretrizes para elaboração de teses da Escola de Engenharia de São Carlos-USP.

A seguir é apresentada uma breve descrição de cada capítulo:

- Capítulo 1: Introdução e apresenta os objetivos e justificativas desta pesquisa.
- Capítulo 2: Contém uma revisão da literatura apresentando o estado da arte em modelos deformáveis, modelos matemáticos para a produção da voz, modelos mecânicos da laringe e processamento digital de sinais.
- Capítulo 3: Descreve os materiais e métodos desenvolvidos para elaborar o modelo da glote, assim como à síntese da voz.
- Capítulo 4: Apresenta os resultados obtidos.
- Capítulo 5: Discute e analisa os resultados obtidos.
- Capítulo 6: Conclusões e propostas para trabalhos futuros.
- Capítulo 7: Contém as referências bibliográficas seguindo as normas ABNT NBR 6023.
- Apêndice : Estudos de casos e programas desenvolvidos.

# Capítulo 2

## Revisão da literatura

Neste capítulo serão descritos os principais trabalhos encontrados na literatura referentes a este projeto, os quais foram divididos em quatro grupos: modelos deformáveis, modelos matemáticos para a produção da voz, modelos mecânicos da laringe e processamento digital de sinais.

### 2.1 Modelos deformáveis

Os fundamentos matemáticos aplicáveis aos modelos deformáveis representam a confluência da geometria, da física e a teoria da aproximação. A geometria é utilizada para representar a forma do objeto; a física impõe restrições como a forma pode variar sobre o espaço e tempo e a teoria da aproximação prove o suporte formal para o mecanismo de ajustes do modelo por meio do conjunto de dados medidos.

Modelos deformáveis geométricos, usualmente, permitem que uma forma inicial convirja a uma nova representação geométrica, geralmente, envolvendo um problema de minimização de energias. A interpretação física de modelos deformáveis é vista como um corpo elástico que responde naturalmente às forças externas, impondo restrições à sua deformação.

#### 2.1.1 Contorno ativo

Um dos modelos deformáveis que atraiu a maioria das atenções é popularmente conhecido como contorno ativo: *snakes*, nome sugerido pelos autores (KASS; WITKIN; TERZOPOULOS, 1988), devido à forma com que os contornos moviam-se.

Uma *snake* tradicional foi inicialmente definida como uma curva  $v(s) = [x(s), y(s)]$ ,  $s \in$

$[0, 1]$ , que se move no domínio espacial de uma imagem para minimizar a função de energia ( $E$ ) dada pela equação eq.(2.1).

$$E_{snake} = \int_0^1 [E_{interna}(v(s)) + E_{externa}(v(s))] ds \quad (2.1)$$

A energia interna do contorno corresponde à energia inerente à própria curva, eq.(2.2).

$$E_{interna} = (\alpha(s)|v_s(s)|^2 + \beta(s)|v_{ss}(s)|^2)/2 \quad (2.2)$$

Na equação (2.2),  $\alpha$  e  $\beta$  são pesos que controlam a tensão e rigidez da *snake* e  $v_s(s)$  e  $v_{ss}(s)$  são as derivadas de primeira e segunda ordem da curva  $v(s)$ . O termo de primeira ordem,  $v_s(s)$ , está relacionado ao comportamento da curva similar a uma membrana elástica, sendo que a curva tende a expandir-se quando  $\alpha$  for positivo e a contrair-se quando o parâmetro  $\alpha$  for negativo. O termo de segunda ordem,  $v_{ss}(s)$ , sugere um comportamento semelhante a uma placa fina, em que a rigidez evita que o contorno se dobre, portanto, se  $\beta$  for nulo a curva poderia desenvolver uma quina, em que sua função seria descontínua. Por outro lado, se o parâmetro  $\beta$  for positivo, garante uma curva suave se desenvolvendo com uma curvatura positiva. A energia externa é responsável por atrair a *snake* para alguma característica de interesse da imagem, como os contornos de determinadas estruturas, em que dada uma determinada imagem  $Y(x, y)$ , a energia externa poderia ser formulada conforme a equação eq.(2.3).

$$E_{externa} = -|\nabla(G_\sigma(x, y) * Y(x, y))|^2 \quad (2.3)$$

Nessa equação,  $G_\sigma(x, y)$  é um filtro gaussiano bidirecional com desvio padrão  $\sigma$  convoluído com a imagem  $Y(x, y)$ , com a finalidade de reduzir o nível de ruído e provocar um efeito de suavização das bordas, aumentando o alcance de captura da *snake*. O operador  $\nabla$  realiza a operação gradiente da função imagem.

Para que a *snake* convirja aos valores corretos é necessário que o contorno inicial seja colocado próximo a borda de interesse da imagem. Durante um processo iterativo a *snake* é atraída para o contorno desejado devido a ação de forças que controlam a forma e localização do contorno.

O método original proposto por Kass, Witkin e Terzopoulos (1988), não é capaz de resolver todas as dificuldades encontradas na determinação das bordas de uma imagem.

---

Algumas delas são encontradas no modelo original, como a incapacidade do contorno mover-se em regiões côncavas, problemas com a descontinuidade da imagem e posição inicial da curva, mas já foram solucionadas por diversos autores (COHEN; COHEN, 1993; XU; JR.; PRINCE, 2000), os quais geralmente adicionam novos termos à componente de energia externa da *snake* forçando a se desenvolver em direção às bordas das estruturas de interesse.

## 2.2 Modelos matemáticos para a produção da voz

Um modelo matemático para a produção de sons vocálicos foi proposto por Fant (1960), no qual uma excitação inicial composta por um gerador de impulsos é filtrada pelo modelo da glote, posteriormente pelo trato vocal e, finalmente, pela radiação labial. O trato vocal atua como uma série de filtros acústicos que atenuam ou reforçam o tom original. Para um som vocálico com a sustentação de uma determinada vogal pode simplesmente ser modelado como uma onda de pressão acústica, resultado de uma fonte periódica, que corresponde à glote, e por um tubo acústico, que corresponde ao trato vocal e lábios. Como geralmente as patologias se encontram na laringe, a técnica de filtragem inversa aplicada a sinais de pressão sonora pode remover efeitos do tubo acústico e resultar em um sinal aproximado da fonte periódica, contendo informações suficientes para análise de efeitos acústicos de uma patologia na laringe (DAVIS, 1976). Medidas baseadas no sinal de voz filtrado são potencialmente mais informativas que as medidas baseadas em sinais sem filtragem. Há dois métodos para a filtragem inversa usada para se obter informações acústicas sobre a fonte glotal, (FRÖHLICH; MICHAELIS; STRUBE, 2001). A primeira consiste no filtro inverso glotal, nesse processo o inverso da radiação dos lábios e a contribuição do trato vocal são utilizados para estimar a velocidade volumétrica glotal em função do tempo, (WONG; MARKEL; GRAY, 1979) e a outra tem por objeto a filtragem inversa residual, em que é estimado o sinal de excitação apresentado no modelo matemático. Basicamente, o estudo do filtro inverso glotal usa técnicas analógicas e, usualmente, elimina somente o primeiro e o segundo harmônico. Posteriormente estuda a envoltória por meio de técnicas computacionais, (CHILDERS; HU, 1994; CHILDERS; AHN, 1995). As frequências e bandas formantes são estimadas e usadas para o ajuste do filtro inverso glotal até a forma de onda esperada surgir sob um aspecto triangular seguido por um intervalo em zero. Contudo, toda aparência do resultado da onda

não satisfaz conceitos intuitivos das pregas vocais fechadas. Outra técnica de predição linear do sincronismo do *pitch* para o filtro inverso glotal produz aceitáveis formas de onda, sem estimar os formantes, mas o ponto de fechamento glotal para cada período de *pitch* deve ser localizado pela inspeção visual.

### 2.2.1 Modelo da excitação

A excitação  $e(n)$  é modelada como uma série de impulsos unitários espaçados pelo período correspondente à frequência fundamental para sons vocálicos e pode ser expressa conforme a equação eq.(2.4):

$$e(n) = E_0 \sum_{j=0}^{\infty} \delta(n - jP), n = 0, 1, \dots \quad (2.4)$$

onde,  $E_0$  é o fator de escala,  $\delta(n)$  é a função delta. A transformada-z  $[E(z)]$  de  $e(n)$  é fornecida pela equação eq.(2.5).

$$E(z) = \frac{E_0}{1 - z^{-P}} \quad (2.5)$$

onde  $P$  é o período correspondente a frequência fundamental  $F_0$ , que tem seu valor igual a  $1/PT$  e a frequência de amostragem  $F_s = 1/T$ .

### 2.2.2 Modelo da glote

O filtro glotal pode ser representado por vários modelos tudo-pólo, (FLANAGAN, 1972), que são baseados em observações fisiológicas. Um dos modelos mais simples é o filtro em que a resposta ao impulso unitário é dado pela equação eq.(2.6).

$$g(n) = G_0(n + 1)e^{-cnT}, n = 0, 1, \dots \quad (2.6)$$

A transformada-z  $G(z)$  é dada pela equação eq.(2.7),

$$G(z) = \frac{G_0}{(1 - e^{-cT}z^{-1})^2} \quad (2.7)$$

onde,  $G_0$  é o fator escala e  $e^{-cT}$  é um pólo no círculo unitário do plano-z. Usando um

trem de pulso  $e(n)$  como entrada do filtro glotal  $G(z)$ , a velocidade volumétrica  $u(n)$  é modelada como a saída do filtro e é representada pela equação eq.(2.8).

$$u(n) = E_0 G_0 \sum_{j=0}^{\infty} (n + 1 - jP) e^{-c(n-jP)}, n = 1, 2, \dots \quad (2.8)$$

### 2.2.3 Modelo do Trato Vocal

O modelo do trato vocal é usado para modelar características acústicas de ressonância do espaço contido entre a glote e os lábios. Este modelo é baseado em aproximações acústicas do trato vocal. Essa aproximação do tubo acústico é composta de um número específico de seções, em que cada seção tem uma área da seção transversal variável, (SONDHI, 1979). A estimação da área do trato vocal baseia-se na predição linear da voz em que as limitações que envolvem o método estão relacionadas aos limites das bandas e frequências do sinal, características da fonte, condições limites, energia perdida com o trato vocal e o método de avaliação empregado (WAKITA, 1979). A transformada- $z$   $V(z)$  do filtro do trato vocal pode ser modelada por um número finito  $K$  de pólos complexos da banda-limite, na qual a frequência central estende-se de zero a  $1/2T$  Hz. A transformada- $z$  é dada pelas equações eq.(2.9) e eq.(2.10).

$$V(z) = \prod_{i=1}^K \frac{1}{(1 - z^{-1}z_i)(1 - z^{-1}z_i^*)} \quad (2.9)$$

ou

$$V(z) = \prod_{i=1}^K \frac{1}{1 - e^{-\pi b_i T} \cos(2\pi f_i T) z^{-1} + e^{-2\pi b_i T} z^{-2}} \quad (2.10)$$

Uma ressonância espectral da frequência central  $f_i$  e os dois lados da banda  $b_i$  são definidos por cada par de pólo complexo  $(z_i, z_i^*)$ , onde  $z_i = \exp[-\pi b_i T + j2\pi f_i T]$  e  $z_i^*$  é o conjugado complexo de  $z_i$  e  $T$  é o período de amostragem. Essas ressonâncias também são chamadas de formantes, sendo que no modelo elas têm uma frequência central geralmente menor que 5000 Hz e a banda menor que 150 Hz.

### 2.2.4 Modelo da radiação

O filtro da radiação dos lábios representa a transformação do sinal da velocidade volumétrica nos lábios para o sinal da pressão sonora. A função de transferência é bastante

simplificada em relação à Flanagan (1972), sendo que a transformada-z  $L(z)$  é aproximada por um filtro transversal dado pela equação eq.(2.11).

$$L(z) = L_0(1 - \mu z^{-1}) \quad (2.11)$$

onde  $L_0$  é o fator de escala e  $\mu$  é um pólo com modulo entre 0,95 e 0,98.

O modelo linear de produção da fala para uma pressão sonora acústica  $S(z)$  é produto da transformada-z do filtro da radiação  $L(z)$ , do filtro do trato vocal  $V(z)$ , do filtro glotal  $G(z)$  e a fonte de excitação  $E(z)$ , e é dado pela equação eq.(2.12).

$$S(z) = L(z)V(z)G(z)E(z) \quad (2.12)$$

ou seja,

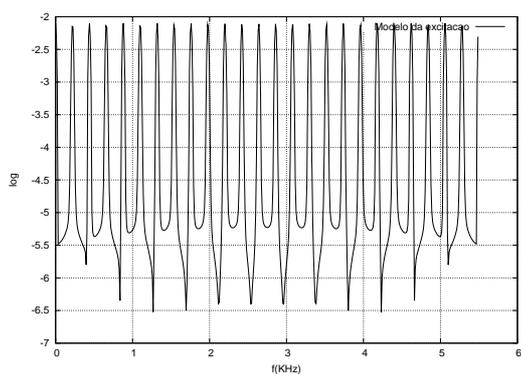
$$S(z) = \frac{L_0 G_0 E_0 (1 - \mu z^{-1}) (1 - e^{cT} z^{-1})^{-2} (1 - z^{-P})^{-1}}{\prod_{i=1}^K 1 - e^{-\pi b_i T} \cos(2\pi f_i T) z^{-1} + e^{-2\pi b_i T} z^{-2}} \quad (2.13)$$

Uma simulação numérica dos modelos descritos nos ítems anteriores foram desenvolvidas para realizar uma comparação com um sinal real de voz, no caso, a vogal /a/ sustentada de uma voz feminina. A representação espectral individual para os vários segmentos do modelo está apresentada na Figura 2.1.

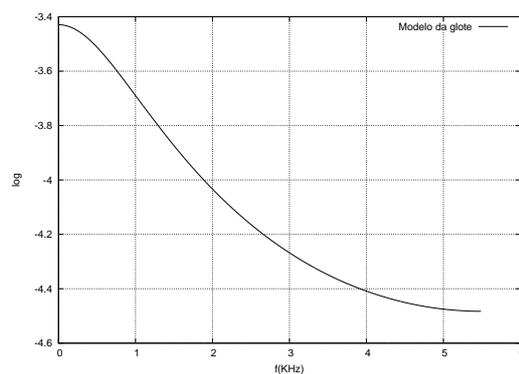
A Figura 2.2 apresenta uma comparação entre um sinal de voz real que possui uma perturbação de *jitter* igual a 0.39% e *shimmer* de 4.49% com um sinal sintetizado, utilizando os modelos descritos e com as mesmas características de frequência fundamental e formantes, porém, com *jitter* e *shimmer* iguais a zero.

## 2.3 Modelos mecânicos da laringe

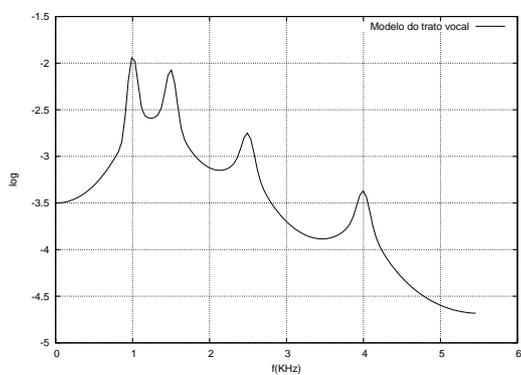
Para a simulação numérica da produção da voz humana é necessário um método numérico que envolva um modelo fisiológico elaborado, que inclui o controle neural dos pulmões, traquéia e trato vocal, o que torna este modelo demasiadamente complexo e de difícil solução. As equações dos movimentos do fluxo de ar e de tecidos, geralmente, são descritas por equações não-lineares. As colisões das pregas vocais geram uma onda elástica que se propaga na mucosa das pregas vocais e envolve também uma mecânica não-linear complexa. Na literatura são encontrados numerosos métodos aplicados a solução



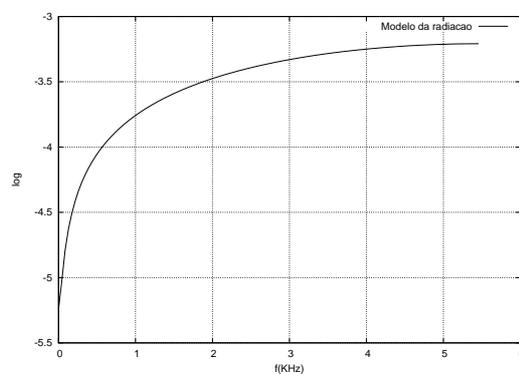
(a) Modelo da excitação:  $F_0 = 220.5\text{Hz}$



(b) Modelo da glote:  $c = 200/\pi$

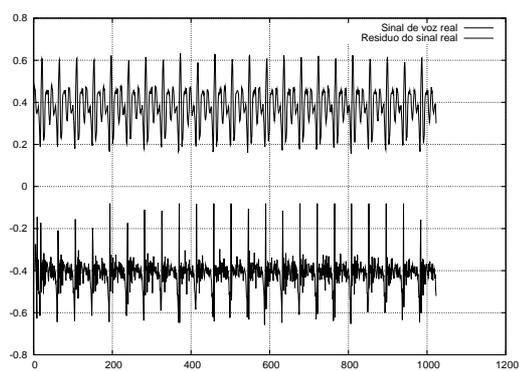


(c) Modelo do trato vocal: Formantes=1000, 1500, 2500 e 4000 Hz

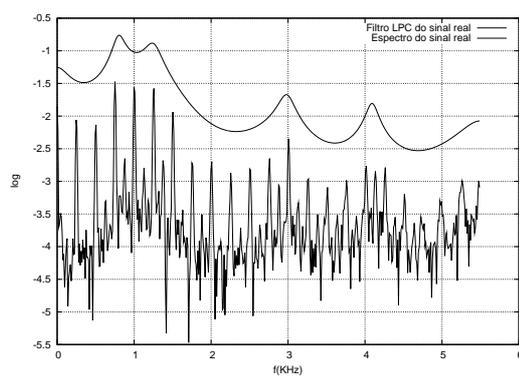


(d) Modelo da excitação labial:  $\mu = 0.98$

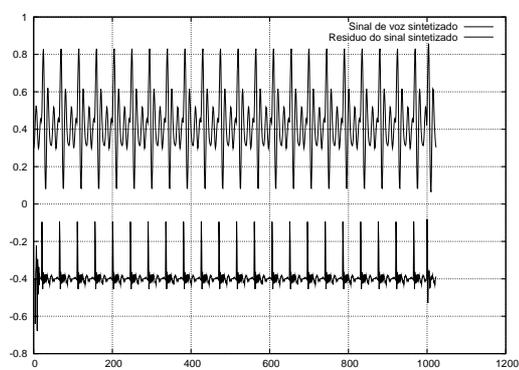
Figura 2.1: Representação espectral do modelo linear da produção vocal.



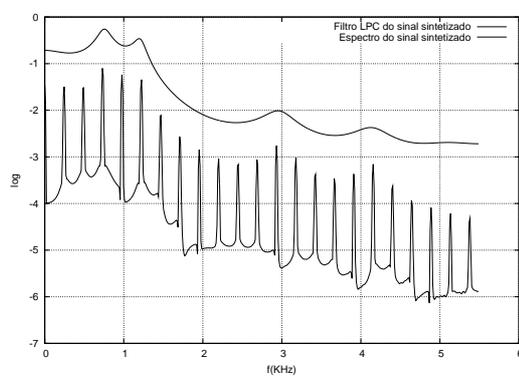
(a) Sinal e resíduo de uma voz real



(b) Espectro e filtro LPC da voz real



(c) Sinal e resíduo de uma voz sintetizada



(d) Espectro e filtro LPC da voz sintetizada

Figura 2.2: Comparação entre um sinal de voz real e sintetizado para a vogal /a/.

deste problema, contudo, inevitavelmente são necessárias diversas simplificações para que o modelo se torne tratável.

O modelo de uma massa de Flanagan e Landgraf (1968) foi o primeiro modelo massa-mola para simular a auto-oscilação das pregas vocais. Posteriormente o modelo duas massas de Ishizaka e Flanagan (1972), devido à sua melhor equivalência aos movimentos das pregas vocais, foi bastante aceito e, geralmente, é o precursor de métodos mais elaborados. Esse modelo considera a glote como duas massas interligadas por uma mola vibrando sobre duas molas e dois amortecedores presos a uma estrutura.

Os modelos massa-mola nos fornecem informações referentes ao comportamento físico das pregas vocais e ao fluxo de ar na glote em um sistema simulado. A dinâmica das pregas vocais tem sido extensivamente estudada durante várias décadas e tendo sido desenvolvido um grande número de modelos das pregas vocais. Os modelos mais citados como precedentes dos métodos utilizados atualmente incluem o modelo de uma massa de Flanagan e Landgraf (1968), Flanagan e Cherry (1969), o modelo duas massas de Ishizaka e Flanagan (1972), o modelo múltiplas massas de Titze (1973), o modelo contínuo de Titze e Strong (1975), o modelo de quatro parâmetros de Titze (1984) e o modelo corpo-cobertura de Hirano e Kakita (1985), desenvolvidos por Story e Titze (1995).

Resumidamente, pode-se classificar os modelos em quatro categorias distintas, os modelos mecânicos, que incluem os modelos de uma, duas e múltiplas massas, o modelo contínuo, o modelo das tiras, *ribbon model*, com quatro parâmetros e o modelo corpo-cobertura.

No modelo uma massa, as vibrações das pregas vocais são modeladas por um único oscilador massa-mola, modulando o fluxo de ar vindo dos pulmões, conforme a Figura 2.3. Este modelo relativamente simples tem a vantagem do baixo custo computacional, permitindo simular a área e velocidade volumétrica glotal, contudo não permite a obtenção da diferença de fase no movimento das pregas vocais.

O modelo duas massas permite dividir a prega vocal em duas partes possibilitando a análise da diferença de fase observada no movimento das pregas vocais. Cada parte consiste em um oscilador massa-mola-amortecedor, como na Figura 2.4. Sendo que a mola representa a propriedade elástica das pregas e o amortecedor representa forças dissipativas como viscosidade e fricção. Há, também, uma interação entre as duas massas, representada pelo acoplamento  $S_c$ , que representa a força de restauração de equilíbrio que ocorre entre as duas massas quando estas estão deslocadas. Este modelo permite simulações mais realísticas das propriedades glotais com pouco esforço computacional.

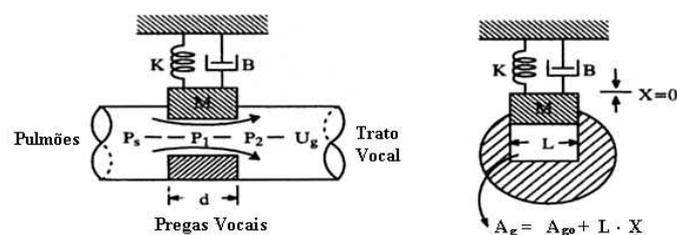


Figura 2.3: Modelo uma massa.

fonte - Desenho original, (FLANAGAN; LANDGRAF, 1968).

O modelo de massas múltiplas, apresentado por Titze (1973) e discutido em Titze

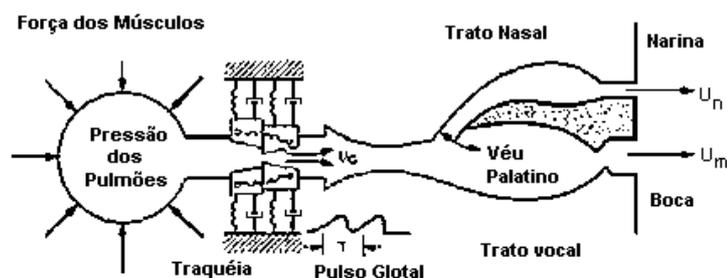


Figura 2.4: Modelo duas massas.

fonte - Desenho original, (ISHIZAKA; FLANAGAN, 1972).

(1974, 1976), propôs 16 massas compostas de duas linhas com oito massas cada uma, Figura 2.6, sendo que a linha superior representa a mucosa e a linha inferior representa os ligamentos e músculos vocais. Esse modelo pode simular o movimento da onda de superfície da mucosa e seus parâmetros têm uma correlação direta com as medições fisiológicas.

O modelo contínuo proposto por Titze e Strong (1975), Titze e Talkin (1979) é altamente informativo na relação entre as estruturas das pregas vocais e os modos de vibração das mesmas, contudo, a forma das pregas vocais nesse modelo é restrita à forma retangular. Trabalhos como de Alipour-Haghighi e Titze (1988), Alipour-Haghighi, Berry e Titze (2000), Rosa et al. (2003) utilizam o método de elementos finitos para solucionar as matrizes de equações diferenciais que são obtidas da análise do sistema de equações de movimento.

No modelo de quatro parâmetros de Titze (1984, 1989), também chamado modelo

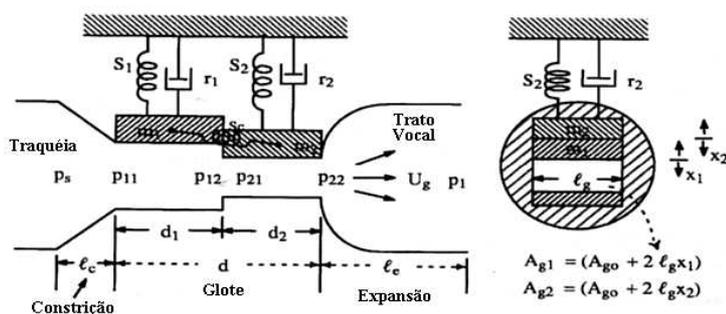


Figura 2.5: Modelo duas massas.

fonte - Desenho original, (ISHIZAKA; FLANAGAN, 1972).

das tiras, *Ribbon model*, apresentado na Figura 2.7, consiste um modelo em três dimensões relacionando o fluxo glotal, com a área de abertura da glote e a área de contato das pregas vocais, sendo que, o movimento da glote é controlado por quatro parâmetros: quociente de abdução, quociente da forma, quociente de inchamento e o quociente de fase.

O conceito corpo-cobertura, (HIRANO; KAKITA, 1985), é geralmente usado para descrever a estrutura em camadas das pregas vocais. Esse modelo sugere que as pregas vocais podem ser divididas em duas camadas de tecidos com propriedades mecânicas diferentes. A camada do corpo consiste em fibras musculares presas firmemente nas fibras de colágeno do ligamento vocal. A camada de cobertura consiste em tecido flexível que envolve a camada do corpo. O movimento da camada de cobertura pode ser observado como uma onda de superfície originando-se na parte inferior para a parte superior das pregas vocais, propagando-se nas direções laterais e verticais. Um modelo para representar a estrutura corpo-cobertura das pregas vocais foi proposto por Story e Titze (1995), que é uma extensão do modelo duas massas para um modelo três massas conforme apresentado na Figura 2.8.

A cobertura, representada por duas massas, é acoplada lateralmente ao corpo por um par de molas e amortecedores, que, por sua vez, estão acoplados a uma parede rígida, que representa a cartilagem tireóidea, por uma mola e um amortecedor. Todos os elementos desse sistema possuem características não-lineares.

Completando o estado da arte, a seguir serão citados alguns dos trabalhos mais recentes relativos à modelagem da laringe. Alipour-Haghighi e Titze (1991) apresentam um modelo elástico dos tecidos das pregas vocais. Alipour-Haghighi, Berry e Titze

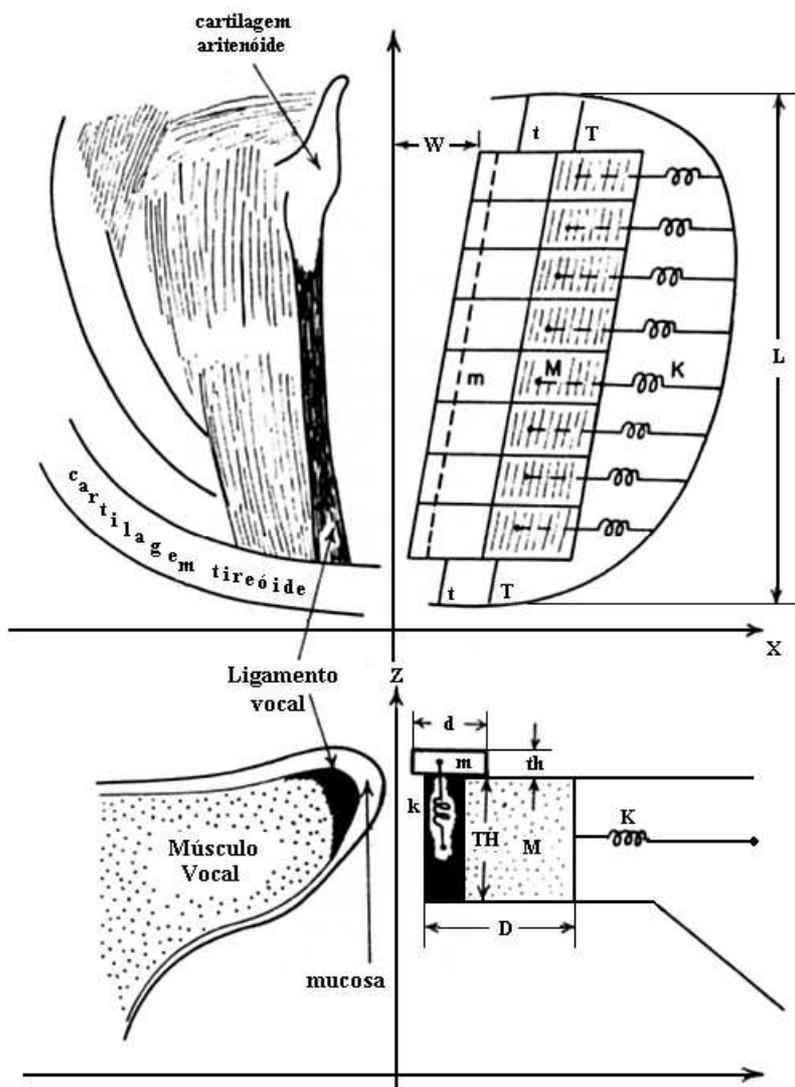


Figura 2.6: Modelo multi-massas.  
 fonte - Desenho original, (TITZE, 1976).

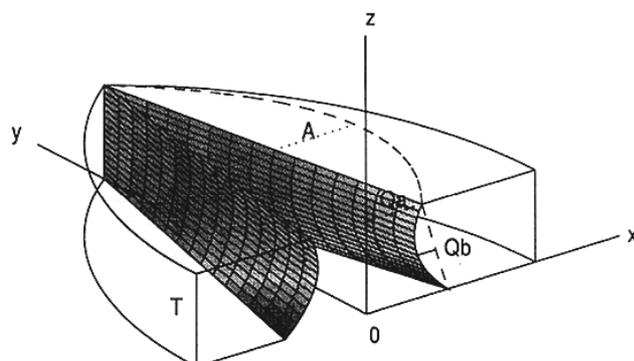


Figura 2.7: Modelo das tiras.  
 fonte - Desenho original, (TITZE, 1989).

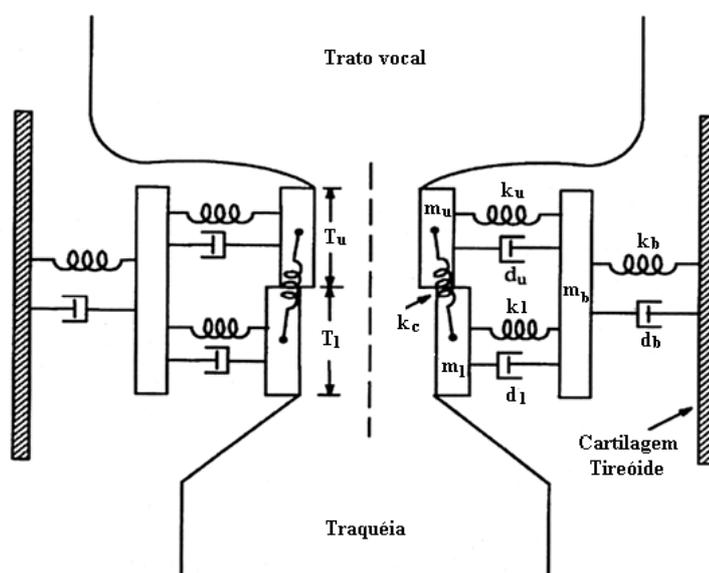


Figura 2.8: Modelo corpo-cobertura, ou três massas, das pregas vocais.  
 fonte - Desenho original, (STORY; TITZE, 1995).

(2000) e Rosa et al. (2003), simulam um modelo da laringe em três dimensões usando o método de elementos finitos. Jiang e Zhang (2001) apresentam um modelo das pregas vocais utilizando o estudo de vibrações caóticas e, posteriormente, acrescentam um póliplo unilateral em seu modelo. Chan (2004) apresentou um método empírico para medir a visco elasticidade das pregas vocais. Titze (2002) apresenta um modelo corpo-cobertura com baixa dimensionalidade utilizando um sistema massa-mola e uma placa rotativa. Döllinger et al. (2002) apresentaram um método para extração de parâmetros a partir de imagens endoscópicas das pregas vocais. Alipour-Haghighi (2002) desenvolve um modelo para calcular o perfil da velocidade e pressão em um modelo estático da laringe. Titze (2002) propôs um modelo para regular o fluxo glotal por meio do teorema da máxima transferência de potência. Gunter (2003) apresenta um modelo mecânico com alta resolução espacial e temporal das colisões das pregas vocais. LaMar, Qi e Xin (2003), modela o movimento das pregas vocais com um modelo semicontínuo hidrodinâmico. Alipour-Haghighi e Scherer (2004) estuda a separação do fluxo de ar em um modelo computacional das pregas vocais. Titze e Hunter (2004) estudam as frequências de vibração normal do ligamento vocal. Titze et al. (2004b), Titze, Broadhead e Gray (2005) projetam um bioreator para estudar o comportamento vibratório dos tecidos das pregas vocais. Lucero e Koenig (2005) apresentam um modelo duas massas com controle dinâmico e, finalmente, um exaustivo trabalho realizado por Holzrichter et al. (2005), apresentou várias medidas da estrutura glotal, produzindo as imagens e sinais da laringe de uma forma precisa, pois utilizaram equipamentos modernos como medidas de movimento a *laserdoppler*, câmeras de alta velocidade e um sensor eletromagnético de micro-potência glotal, possibilitando-nos esclarecer dúvidas ou corroborar com maior exatidão as diversas teorias até então apresentadas.

## **2.4 Processamento digital de sinais**

### **2.4.1 Processamento digital do sinal de voz**

O processamento digital de sinais consiste na representação de um sinal como uma seqüência ordenada de números e o subsequente processamento computacional desta seqüência. Todos os sinais que podem ser digitalizados são divididos em diversos grupos específicos, que possuem tecnologias, algoritmos e especializações matemáticas próprias, entre elas, podemos citar como exemplo um pequeno grupo destas áreas: o processamento

de imagens, sonar, radar, comunicação, música, televisão, vídeo, informática e a área específica referente a este tópico denominada de processamento digital do sinal de voz.

A digitalização de um sinal consiste na transformação de um sinal contínuo no tempo ou espaço em um sinal discreto no tempo ou espaço, ou seja, para sinais unidimensionais, como o de áudio, são realizadas amostragens em intervalos regulares de tempo, chamado de taxa de amostragem. Posteriormente, através do processo de quantificação as amostras são convertidas em uma seqüência de números digitais. Em casos bidimensionais, como as imagens, o espaço é segmentado em intervalos regulares chamados de *pixel* e a intensidade de cada elemento é convertida em uma seqüência de números binários.

Geralmente, o processamento de sinais de voz consiste em obter a representação do sinal baseado em um dado modelo e a aplicação de alguma transformada para colocar o sinal em uma forma mais conveniente para seu tratamento. A transformação envolvida no processamento de voz normalmente possui uma entrada  $x(n)$  e uma saída  $y(n) = T[x(n)]$ , sendo que, muitas vezes são realizadas algumas simplificações, por exemplo, considerar que o modelo da produção da voz é caracterizado como um sistema linear e invariante ao deslocamento e pode ser completamente determinado pela sua resposta a uma entrada de impulso unitário. A saída pode ser calculada utilizando o sinal de entrada,  $x(n)$  convoluído com a resposta ao impulso unitário,  $h(n)$ , usando a expressão da soma de convolução:  $y(n) = x(k)h(n - k) = x(n) * h(n)$ .

Uma das transformadas mais utilizada em processamento de voz é a transformada-z. A representação de uma seqüência é definida pelo par de equações eq.(2.14) e eq.(2.15).

$$X(z) = \sum_{n=-\infty}^{\infty} x(n)z^{-n} \quad (2.14)$$

$$x(n) = \frac{1}{2\pi j} \oint_C X(z)z^{n-1} dz \quad (2.15)$$

A transformada-z de  $x(n)$  é definida pela equação eq.(2.14) e pode ser vista como uma série de potências infinitas na variável  $z^{-1}$ , em que a seqüência de valores,  $x(n)$ , têm a função dos coeficientes na série de potências. Em geral, cada série de potências converge para um único valor finito para certos valores de  $z$ . O conjunto de valores para que uma série convirja define a região no plano complexo-z conhecida como região de convergência.

Outra transformada muito utilizada em processamento de sinais de voz é a Transformada de Fourier na qual a representação de um sinal discretizado no tempo é dado pelas

seguintes equações, eq.(2.16) e eq.(2.17).

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x(n)e^{-j\omega n} \quad (2.16)$$

$$x(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\omega}) e^{j\omega n} d\omega \quad (2.17)$$

Estas equações podem ser vistas como um caso particular da transformada-z, especificamente, a representação de Fourier é obtida pela restrição da transformada-z para o círculo unitário do plano-z.

Um filtro digital consiste em um sistema linear discreto no tempo e invariante ao deslocamento que tem a propriedade da saída ser uma função da entrada que satisfaça a equação linear da diferença, eq.(2.18).

$$y(n) - \sum_{k=1}^N a_k y(n-k) = \sum_{r=0}^M b_r x(n-r) \quad (2.18)$$

Aplicando-se a transformada-z em ambos os lados da equação pode-se demonstrar que a função  $H(z)$  é a razão entre a saída pela entrada, conforme a equação eq.(2.19).

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{r=0}^M b_r z^{-r}}{1 - \sum_{k=1}^N a_k z^{-k}} \quad (2.19)$$

A função sistema,  $H(z)$ , chamada de função de transferência é em geral uma função racional de  $z^{-1}$ , a qual é caracterizada pela localização dos pólos e zeros no plano-z.

Nos estudos acústicos da voz, geralmente, são utilizados os sons vocálicos devido à produção de vibrações das pregas vocais que por sua vez torna mais fácil a obtenção de característica da laringe. Segundo Behlau e Pontes (1995) a vogal /a/ é uma vogal oral, central e aberta, que por sua natureza articulatória evidencia as mínimas articulações no equilíbrio mioelástico da laringe. Um modelo linear para a produção da fala não-nasalada foi proposto por Fant (1960) e, subsequentemente, este modelo foi analisado com maiores detalhes por Flanagan (1972). Conforme a Figura 2.9, o modelo é composto por três filtros acústicos. A justificativa para utilizar esse modelo está baseada na teoria de tubos acústicos, nas medidas da velocidade volumétrica e pressão sonora e nos resultados de circuitos elétricos para síntese de voz, outra característica importante desse modelo é a

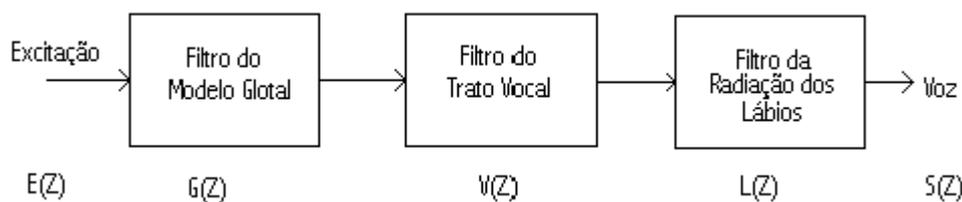


Figura 2.9: Modelo linear para produção da fala.

separabilidade dos segmentos de filtros.

Para um sinal de voz analógico  $s_a(t)$  amostrado a cada  $T$  segundos temos um sinal discreto no tempo  $s(n)$  que pode ser representado no domínio da frequência como o produto de convolução da excitação com os três filtros de cada modelo conforme descrito na equação eq.(2.21).

$$s(n) = s_a(nT), -\infty < n < \infty \quad (2.20)$$

$$s(n) = e(n) * g(n) * v(n) * l(n) \quad (2.21)$$

$$S(z) = E(z)G(z)V(z)L(z) \quad (2.22)$$

Na equação eq.(2.22)  $E(z)$  é o modelo para a excitação,  $G(z)$  é o modelo da glote,  $V(z)$  é o modelo do trato vocal,  $L(z)$  é o modelo da radiação labial.

Outra vantagem em se trabalhar no domínio da frequência consiste em transformar a convolução da eq.(2.21), em um simples produto conforme a eq.(2.22), facilitando a manipulação dos termos e, teoricamente, possibilitando a remoção de qualquer parte do modelo pela filtragem inversa. O modelo linear do sistema fonador permitiu, também, o desenvolvimento de técnicas de reconhecimento e síntese de voz, (MONTAGNOLI; FONSECA, 2005; MONTAGNOLI; PEREIRA, 2004; CATALDO et al., 2004; CHILDERS, 1990; CHILDERS; LEE, 1991; CHILDERS; WU, 1991)

Outra área de estudo que se beneficiou dos trabalhos de modelagem do sistema de produção de voz foi a do auxílio computadorizado a diagnósticos das condições do sistema fonador, (GÓMES et al., 2004; MITEV; HADJITODOROV, 2003; NIEDZIELSKA; GLIJER; NIEDZIELSKI, 2001; PARSA; JAMIESON, 2000; MARTINEZ; RUFINER, 2000; MANFREDI et al., 2000; GERRATT; KREIMAN, 2000; MONTAGNOLI, 1998; BOYANOV; HADJITODOROV, 1997; CHILDERS, 1990; ANANTHAPADMANABHA; YEGNANARAYANA,

1979; DAVIS, 1976; KOIKE; MARKEL, 1975).

Atualmente a teoria do caos tem sido aplicada com intenção de estudar os movimentos caóticos que as duas pregas vocais, como dois sistemas auto-oscilantes, podem gerar e, conseqüentemente, atribuir esse comportamento aos modelos da produção vocal. A utilização dessas informações podem ser utilizadas para a análise das condições laríngeas, (MACIEL; PEREIRA; DAJER, 2005; MONTAGNOLI, 2005; ZHANG; PAYANDEH; DILL, 2004; JIANG; ZHANG, 2001; STEINECKE; HERZEL, 1995; BERRY et al., 1994; HERZEL, 1993).

## **2.5 Processamento digital de imagens médicas**

A utilização de equipamentos como o raio-x, ultra-som, tomografia, ressonância magnética e instrumentos de aquisição de imagens e vídeo, como endoscopia, e especificamente na área laringológica o exame de estroboscopia, propiciaram o desenvolvimento de uma nova área na medicina, que consiste no diagnóstico auxiliado por imagens. Com o aprimoramento computacional, a digitalização das imagens médicas permite a utilização de técnicas de processamento digital de imagens, as quais vem proporcionando um grande avanço à medicina.

O processamento de imagens médicas tem sido utilizado para dois propósitos básicos: a melhoria visual da aparência da imagem para a visualização humana, também chamado de realce da imagem, e preparar a imagem para realizar medidas e extração de características nas estruturas presentes.

Imagens médicas são muitas vezes deterioradas por ruídos devido a várias fontes de interferências e outros fenômenos que afetam o processo de aquisição de imagens. Algoritmos de realce são utilizados para reduzir o ruído da imagem e melhorar o contraste das estruturas de interesse. Em imagens em que a distinção entre tecidos normais e anormais são sutis, a interpretação precisa pode se tornar difícil se níveis significativos de ruídos são encontrados na imagem. Em muitos casos, o realce de imagens implica em melhoria na qualidade da imagem facilitando o diagnóstico. Técnicas de realce de imagens são geralmente usadas para prover uma clareza na imagem para um observador humano, porém, elas podem ser utilizadas como uma das etapas intermediárias em processos automáticos para análises de imagens.

A borda é o limite entre um objeto e um determinado fundo, podendo indicar também a sobreposição de objetos. A detecção da borda faz parte de um processo mais

---

abrangente chamado de segmentação. Na visão computacional, o processo que envolve a identificação e classificação de objetos em uma determinada imagem, inclui o processo de detecção das bordas o qual é apontado como uma das ferramentas mais importantes. Isto significa que quando a borda pode ser identificada com precisão, algoritmos de quantificação são aplicados em estruturas segmentadas para extrair informações como a forma, tamanho, área, ângulo, textura e movimento, as quais são fundamentais para um correto diagnóstico.

A segmentação de imagens médicas tem como objetivo delinear alguma estrutura de interesse, discriminando-a dos tecidos ao seu redor. Essa separação, que é geralmente fácil para o sistema de visão humana, pode se tornar uma enorme dificuldade para sistema computacionais. Em muitos casos o método de segmentação é o fator mais importante responsável pelo resultado de toda a análise.

Algoritmos de segmentação operam nas variações de intensidade ou textura de uma determinada imagem usando técnicas que incluem, limiarização, contornos ativos, máscaras deformáveis, técnicas estatísticas de reconhecimento de padrões e redes neurais

Posterior a quantificação, em muitos sistemas automáticos, é necessário a etapa de classificação. O objetivo da classificação consiste em agrupar atributos provenientes da imagem constituindo classes que possuam alguma propriedade em comum. As redes neurais e classificadores estatísticos são os métodos mais comuns encontrados para realizar a classificação de uma imagem.(JAIN; DUIN; MAO, 2000; MONTAGNOLI; PEREIRA; FERRARI, 2003).

A visualização de imagens digitalizadas é uma área relativamente nova que contribui significativamente à medicina. Enquanto sistemas automáticos são bons para realizar uma análise quantitativa e precisa, para que um exame de imagens médicas seja completo, ele deve ser acompanhado pela inspeção visual de um observador humano. Em alguns casos, algoritmos de modelagem multidimensional permitem a visualização das imagens em três dimensões, possibilitando translações, rotações e cortes na imagem.

Para a compressão e armazenamento de imagens médicas são necessárias técnicas que permitam a diminuição da quantidade de memória necessária para conter as informação em uma imagem ou vídeo e que sejam confiáveis quanto a segurança da informação. Algoritmos de compressão de imagens sem perdas estão cada vez mais eficientes em relação ao seu propósito, o que também contribui para uma rápida comunicação entre os dispositivos de armazenagem com os clientes que as utilizam, pois os volumes de informação trafegados nos meios de comunicação possuem um tamanho menor.

Essa seção ressaltou apenas alguns tópicos pertinentes ao desenvolvimento deste trabalho, sendo que, a área de processamento digital de sinais está bastante sedimentada e possui uma literatura básica bastante ampla. Mais detalhes podem ser encontrados nos seguintes livros: (BANKMAN, 2000; SMITH, 1999; PARKER, 1996; MASTERS, 1994; RUSS, 1994; STEARNS; DAVID, 1988; GONZALES; WOODS, 1992; RABINER; SCHAFER, 1978; MARKEL; GRAY, 1976; OPPENHEIM; SCHAFER, 1975).

## Capítulo 3

# Cirurgia virtual da laringe

O método proposto para a realizar a cirurgia virtual da laringe pode ser dividido em duas etapas: a primeira consiste no processamento digital de imagem da obtida através da vídeolaringoscopia e a segunda, no processamento do sinal de voz. O diagrama de blocos da Figura 3.1 resume as fases necessárias em cada etapa, as quais serão descritas com mais detalhes nos ítems a seguir. Ao contrário de todas as rotinas utilizadas na realização deste trabalho, que foram desenvolvidas utilizando a linguagem C++ padrão, as rotinas para aquisição e manipulação de vídeo foram desenvolvidas utilizando a linguagem C++, porém, valendo-se da interface para programação de aplicação (*API Windows*®). As interfaces gráficas foram desenvolvidas utilizando o programa C++ *Builder* 4.0 da Borland®.

### 3.1 Processamento de Imagens

Uma das dificuldades encontradas no processamento das imagens de vídeo consistiu na obtenção dos quadros de imagens a partir dos arquivos de vídeo no formato *avi* compactado. Mesmo utilizando uma resolução 320 x 240, a quantidade de memória exigida para a captura de vídeo é bastante grande. Para reduzir esse tamanho foi utilizado o codificador e decodificador da Intel, *codec Indeo vídeo 5.04* ®.

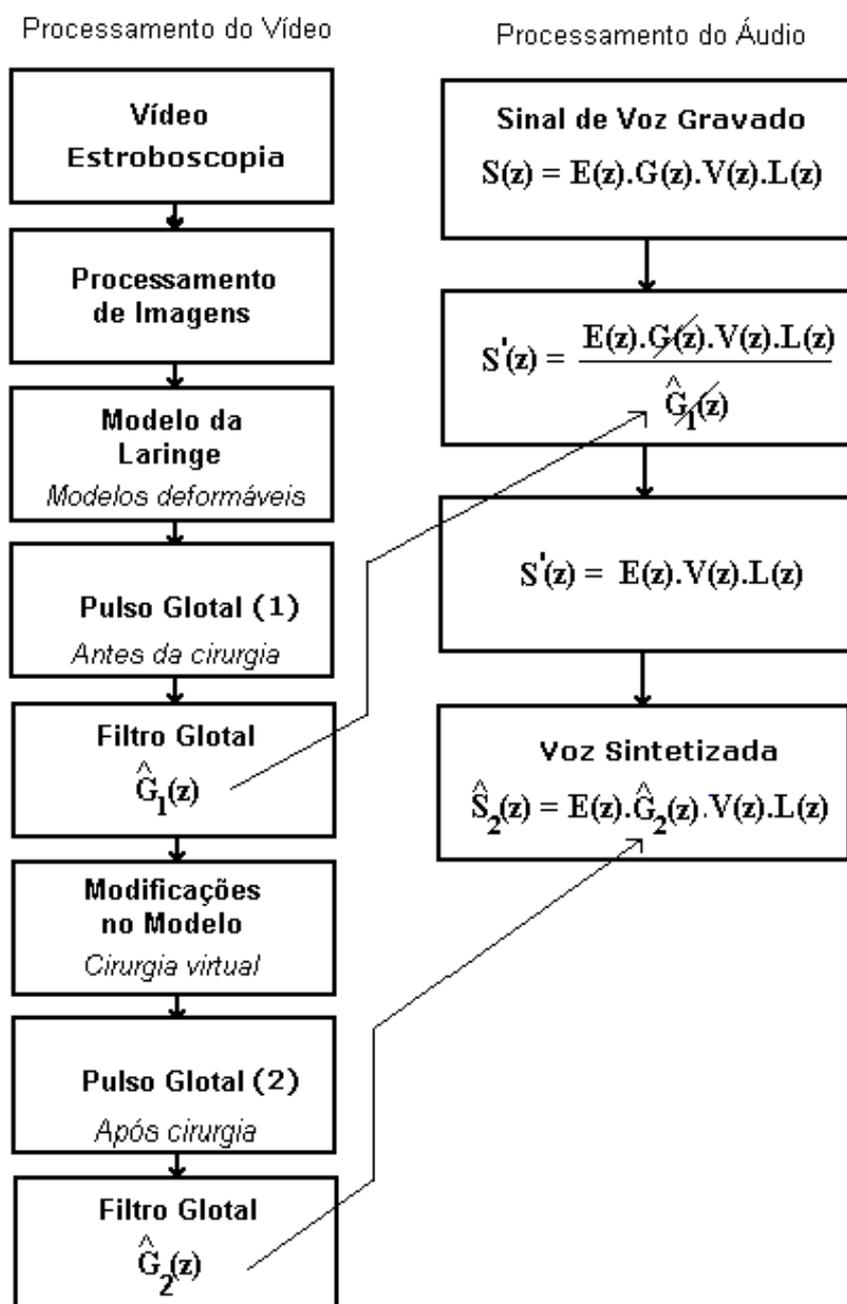


Figura 3.1: Diagrama de blocos: estágios desenvolvidos pelo método proposto.

### 3.1.1 Vídeostroboscopia da laringe

A laringoscopia utilizando fibras óticas é uma importante ferramenta utilizada para visualizar as pregas vocais paradas, e a estroboscopia complementa o exame permitindo a observação das pregas vocais em movimento.

As imagens obtidas com o exame de endoscopia ainda são bastante pobres devido à resolução baixa das câmeras de vídeo instaladas nos equipamentos de videostroboscopia. Este exame apresenta inúmeras dificuldades inerentes ao sistema, podendo destacar entre elas as distorções geométricas, o borramento por movimentação da fibra, a ausência de auto-foco, o qual é ajustado manualmente apenas uma vez antes do exame e geralmente a iluminação não é uniforme. O ruído introduzido pelo sistema também é bastante significativo, pois várias são as etapas intermediárias entre a fibra até a obtenção da imagem digital.

Os exames da laringe foram obtidos no núcleo de otorrinolaringologia (NOG), Angra dos Reis, RJ. realizadas pelo cirurgião Walther Luiz Garcia Jaeger, utilizando um aparelho de estroboscopia *Strobview 2000*. Outro conjunto de exames foi realizado pelo cirurgião Dr. Domingos Hiroshi Tsuji da Faculdade de Medicina, USP, São Paulo, utilizando o aparelho da *Kay Elemetrics: RLS 9100*

Para a captura de vídeo foi utilizada uma placa captura de *s-vídeo* com resolução 320 x 240 conectada ao PC e ligada à saída de vídeo do estroboscópio, e posteriormente, posteriormente gravadas em *compact disc (CD)*. Outro método utilizado para obtenção das imagens foi através da digitalização de exames gravados em fitas de vídeo.

### 3.1.2 Processamento digital de imagens

As etapas para o processamento de imagens já estão bastante evoluídas e solidificadas e podem ser encontradas nas literaturas Bankman (2000), Russ (1994), Parker (1996), Gonzales e Woods (1992). As etapas normalmente seguidas com o intuito de processar imagens de um modo geral podem ser resumidas em pré-processamento, segmentação, quantificação e classificação. Os algoritmos para o processamento de imagens utilizados neste trabalho foram desenvolvidos utilizando os pacotes *OpenCv®*, da Intel e rotinas desenvolvidas e transcritas dos livros supracitados, utilizando a linguagem C++ padrão.

### 3.1.3 Redução de ruído na imagem

Observando as imagens da laringe adquiridas pelo sistema, notou-se a presença de ruídos. Para a redução do ruído na imagem foi realizado um pré-processamento inicial através da aplicação de uma máscara de convolução no domínio espacial, em que a média de uma pequena região da imagem é determinada por meio da implementação generalizada mostrada na equação eq.(3.1), em que o valor do *pixel* na região é multiplicado pelo conjunto de pesos internos.

$$P_{x,y} = \frac{\sum_{i=-m}^m \sum_{j=-m}^m W_{i,j} P_{x+i,y+j}}{\sum_{i=-m}^m \sum_{j=-m}^m W_{i,j}} \quad (3.1)$$

Modificando-se a máscara pode-se obter vários filtros na imagem, como borramento, suavização ou realce das bordas, podendo-se empregar as máscaras mais comuns como a gaussiana, mediana, laplaciana entre outras, (RUSS, 1994).

### 3.1.4 Alteração no histograma da imagem

Técnicas utilizadas para alterar o histograma da imagem modificam-na de uma forma global, pois a informação localizada de cada *pixel* não se encontra no histograma. Geralmente alterações na distribuição do histograma são empregadas para realçar ou segmentar a imagem, possibilitando uma melhor visualização em vídeo ou impresso.

As principais operações que podem ser efetuadas no histograma da imagem são alterações globais no brilho, limiarização (*thresholding*), quantização, segmentação e equalização do histograma. Para realçar a visualização das imagens de vídeo da laringe pode-se aplicar a equalização de seus histogramas.

### 3.1.5 Segmentação da imagem

Numerosos algoritmos já foram desenvolvidos destinados à segmentação de imagens. O objetivo principal da segmentação consiste em separar ou destacar as estruturas de interesse em uma determinada imagem por possuírem alguma característica em comum. Particularmente em imagens médicas o interesse maior está na delimitação do contorno de diversas estruturas anatômicas, tipos de tecidos ou regiões patológicas.

Entre as diversas técnicas de segmentação de estruturas em imagens médicas, destacam-se os contornos ativos, que utilizam curvas definidas no domínio da imagem, as quais

podem mover-se por ação de forças internas, inerentes à própria curva e à forças externas derivadas das características da imagem.

Dois tipos principais de modelos de contornos ativos são encontrados na literatura, os modelos paramétricos e os modelos geométricos, sendo que um dos trabalhos mais significativos que originou as técnicas mais recentes de segmentação foi a proposta do modelo paramétrico denominado de *snakes*, desenvolvida por Kass, Witkin e Terzopoulos (1988). Os modelos geométricos aplicados à segmentação de imagens têm como trabalhos precursores aqueles propostos por Sethian (1997). Esses métodos foram intitulados de *Fast Marching Methods* e *Level Set Methods*.

### 3.1.6 Quantificação

O processo pelo qual uma estrutura de uma imagem é convertido em medidas ou atributos é chamado de quantificação.

Após a segmentação da glote, a medida de sua área é um dos parâmetros mais importantes, pois a forma do pulso glotal estimado será proporcional a este valor.

O comprimento longitudinal da prega foi utilizado como referência métrica e de escala. Na captura das imagens do exame de laringoscopia, o tamanho real da laringe é perdido devido à falta de quaisquer referências reais na imagem.

As medidas da abertura central máxima da glote e a espessura das pregas vocais também fazem parte dos valores que são medidos diretamente na imagem e, posteriormente, especificados no modelo proposto.

### 3.1.7 Modelos deformáveis

Uma das principais características dos modelos deformáveis ou contornos ativos está na possibilidade de especificar um amplo conjunto de propriedades por meio de uma função energia, que possui analogia a sistemas físicos. A energia do contorno geralmente é dividida em duas partes: a energia interna ( $E_{int}$ ) e externa ( $E_{ext}$ ). A energia interna depende de parâmetros intrínsecos do contorno, como comprimento e curvatura. A energia externa depende de outros fatores, como a estrutura da imagem em que o contorno se move ou alguma outra restrição externa ao seu movimento.

Pode-se estender uma analogia física ao movimento do contorno considerando a ação

de forças que atuam no sistema. A energia interna à curva confere propriedades físicas que possibilitam o contorno se comportar como uma membrana elástica ou uma placa rígida. Resumidamente, pode-se observar essas propriedades através da energia interna  $E_{int}$ , eq.(3.3), pertencentes à equação de energia, eq.(3.2), proposta por Kass, Witkin e Terzopoulos (1988), conforme descrito no capítulo anterior.

$$E_{snake} = \int_0^1 E_{snake}(v(s))ds = \int_0^1 E_{int}(v(s)) + E_{ext}(v(s))ds, \quad (3.2)$$

onde,

$$E_{int} = (\alpha(s)|v_s(s)|^2 + \beta(s)|v_{ss}(s)|^2)/2 \quad (3.3)$$

Nas equações eq.(3.2) e eq.(3.3)  $\alpha$  e  $\beta$  são pesos que controlam a tensão e rigidez da *snake* e  $v_s(s)$  e  $v_{ss}(s)$  são as derivadas de primeira e segunda ordem da curva  $v(s)$ . A energia externa ( $E_{ext}$ ) geralmente é derivada de uma imagem que traz algumas características de interesse, como no caso do gradiente da imagem para a obtenção de contornos.

Uma curva pode ser representada parametricamente por:  $v(s) = (x(s), y(s))$ , onde  $0 \leq s \leq 1$ , sendo que para uma curva aberta temos  $v(0) \neq v(1)$  e para uma curva fechada,  $v(0) = v(1)$ , conforme mostrado na Figura 3.2.

Um método normalmente encontrado na literatura utilizado para discretizar o modelo

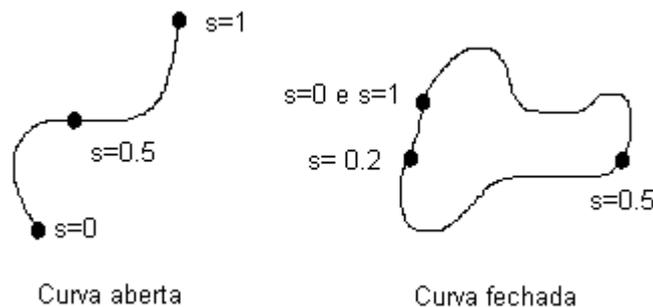


Figura 3.2: Representação paramétrica de uma curva aberta e fechada.

de contornos ativos pode ser visto na Figura 3.3, na qual o contorno é constituído por um conjunto de vértices que são conectados entre si por segmentos de retas chamados de borda. A posição do vértice  $v_i$  é representada por um vetor  $p_i$  e a borda entre  $v_i$  e  $v_{i+1}$  pelo vetor  $d_i$ .

A deformação é causada pela combinação das forças atuantes no vértice que resultam

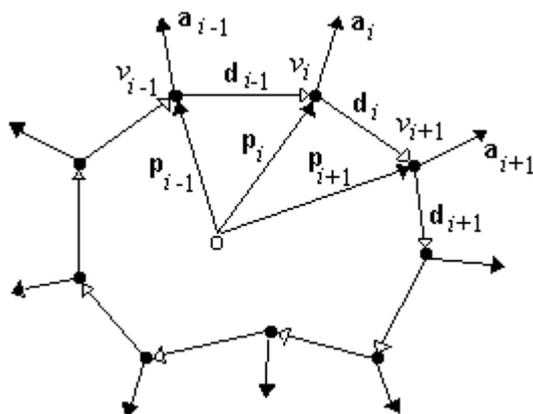


Figura 3.3: Representação de uma curva discretizada.

em uma aceleração representada pelo vetor  $a_i$  e uma velocidade  $vel_i$  não representada na figura.

O comprimento  $d_i$  de um segmento representa a resolução local do modelo e se este valor for muito grande o modelo não será capaz de se conformar em detalhes pequenos. Esse tamanho pode ser modificado durante a deformação, causando uma variação local na resolução do modelo. Para manter esta variação limitada, geralmente é necessário a inserção ou remoção de alguns vértices.

O vetor tangente  $t_i$  a um vértice  $v_i$  pode ser definido como a direção da junção dos segmentos de retas entre os dois vértices vizinhos, (Figura 3.4 e eq.(3.4)).

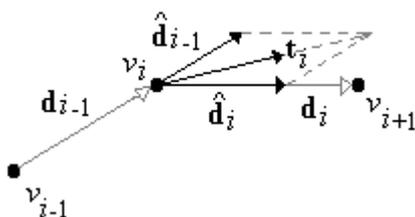


Figura 3.4: Vetor tangente.

$$\vec{t}_i = \frac{\hat{d}_i + \hat{d}_{i-1}}{\|\hat{d}_i + \hat{d}_{i-1}\|} \quad (3.4)$$

O vetor normal  $n_i$ , é definido como um vetor ortogonal a tangente  $t_i$  :  $n_i = t_i^\perp$ , sendo que,  $(x, y)^\perp = (y, -x)$ .

Supondo que a curva se mova ao longo de sua normal com velocidade  $F$  que depende somente da curvatura local  $\kappa$  da curva  $s$ , sendo que, o vetor posição  $p_i$  parametriza a curva  $s$  no tempo  $t$ , onde  $0 \leq s \leq 1$ , temos que  $F(\kappa) = n \cdot p_t$ . A equação do movimento pode ser escrita em termos das componentes individuais  $p = (x, y)$  conforme as equações eq.(3.5) e eq.(3.6)

$$x_t = F \left[ \frac{y_{ss}x_s - x_{ss}y_s}{(x_s^2 + y_s^2)^{3/2}} \right] \left( \frac{y_s}{(x_s^2 + y_s^2)^{1/2}} \right), \quad (3.5)$$

$$y_t = -F \left[ \frac{y_{ss}x_s - x_{ss}y_s}{(x_s^2 + y_s^2)^{3/2}} \right] \left( \frac{x_s}{(x_s^2 + y_s^2)^{1/2}} \right), \quad (3.6)$$

onde a parametrização da curvatura da função de velocidade  $F(\kappa)$  é dada pela equação  $\kappa = \frac{y_{ss}x_s - x_{ss}y_s}{(x_s^2 + y_s^2)^{3/2}}$  e a normal é dada por  $n = (y_s, -x_s)/(x_s^2 + y_s^2)^{1/2}$ . Essa é a representação *lagrangiana*, porque o limite de  $(x(s, t), y(s, t))$  descreve a frente de movimento, (SETHIAN, 2002).

Um método convencional, chamado de método da corda ou método nodal, é utilizado para modelar frentes de propagação de movimento originadas da discretização da forma lagrangiana das equações de movimento, conforme apresentado nas equações eq.(3.5) e eq.(3.6). Nesse método, a parametrização é discretizada dentro de um conjunto de partículas cuja posição, em qualquer tempo, pode ser obtida para reconstruir a frente. Em duas dimensões a frente pode ser reconstruída como segmentos de linha formando uma curva e em três dimensões, geralmente, são utilizadas malhas triangulares.

Esse método pode ser ilustrado por meio de um esquema direto a partir de uma aproximação por diferenciação simples das equações de movimentos lagrangianas. O intervalo de parametrização  $[0, S]$  é dividido dentro de  $M$  intervalos de tamanho  $\Delta s$ , produzindo  $M + 1$  pontos de malha, portanto,  $s_i = i\Delta s, i = 0, \dots, M$ . Similarmente, o tempo pode ser dividido dentro de intervalos iguais de tamanho  $\Delta t$ . A imagem de cada ponto da malha  $i\Delta s$  para cada passo do tempo  $n\Delta t$  é marcado o ponto  $(x_i^n, y_i^n)$  na frente de movimento. O objetivo consiste em elaborar um algoritmo numérico que produzirá novos

valores  $(x_i^{n+1}, y_i^{n+1})$ , originados dos valores das posições anteriores conforme discutido em Sethian (2002).

A aproximação por diferenças centrais baseadas na série de Taylor produz a derivada de cada ponto pela diferença dos pontos vizinhos da malha, conforme apresentado nas equações eq.(3.7) e eq.(3.8).

$$\frac{dx_i^n}{ds} \approx \frac{x_{i+1}^n - x_{i-1}^n}{2\Delta s}, \quad \frac{dy_i^n}{ds} \approx \frac{y_{i+1}^n - y_{i-1}^n}{2\Delta s}, \quad (3.7)$$

$$\frac{d^2x_i^n}{ds^2} \approx \frac{x_{i+1}^n - 2x_i^n + x_{i-1}^n}{\Delta s^2}, \quad \frac{d^2y_i^n}{ds^2} \approx \frac{y_{i+1}^n - 2y_i^n + y_{i-1}^n}{\Delta s^2}. \quad (3.8)$$

Similarmente, a derivada no tempo pode ser obtida pela aproximação por diferenças conforme a eq.(3.9)

$$\frac{dx_i^n}{dt} \approx \frac{x_i^{n+1} - x_i^n}{\Delta t}, \quad \frac{dy_i^n}{dt} \approx \frac{y_i^{n+1} - y_i^n}{\Delta t} \quad (3.9)$$

Substituindo essas aproximações na equação do movimento dada pela equação eq.(3.5) e eq.(3.6), tem-se a seguinte equação da frente de movimento:

$$(x_i^{n+1}, y_i^{n+1}) = (x_i^n, y_i^n) + \Delta t F(\kappa_i^n) \frac{((y_{i+1}^n - y_{i-1}^n), -(x_{i+1}^n - x_{i-1}^n))}{((x_{i+1}^n - x_{i-1}^n)^2 + (y_{i+1}^n - y_{i-1}^n)^2)^{1/2}}, \quad (3.10)$$

onde a curvatura é dada por,

$$\kappa_i^n = 4 \frac{(y_{i+1}^n - 2y_i^n + y_{i-1}^n)(x_{i+1}^n - x_{i-1}^n) - (x_{i+1}^n - 2x_i^n + x_{i-1}^n)(y_{i+1}^n - y_{i-1}^n)}{((x_{i+1}^n - x_{i-1}^n)^2 + (y_{i+1}^n - y_{i-1}^n)^2)^{3/2}}. \quad (3.11)$$

A seguir serão apresentados alguns exemplos de contornos movendo-se sob a ação de diversas forças. Conforme o exemplo da Figura 3.5, um contorno fechado que se move somente sob a ação da força elástica tende a chegar a uma condição de menor energia, ou seja, para este caso em um círculo. Para um contorno aberto a curva tende a se estabilizar em uma reta.

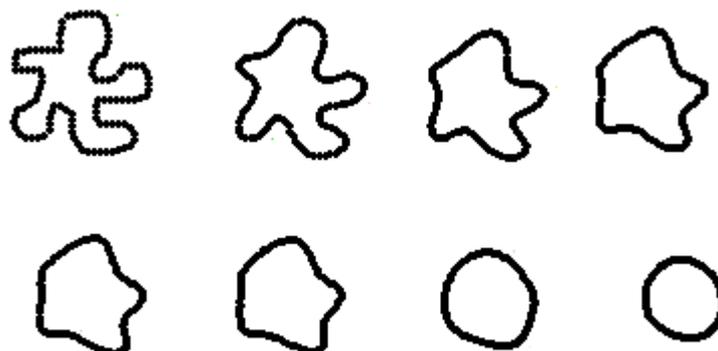


Figura 3.5: Curva se movendo somente sob ação da força elástica.

Um contorno que se move somente sob a ação de forças normais aos seus vértices conforme a equação eq.(3.10), tende a formar laços, chamados de *rabos de andorinha*, que devido as características da equação tendem a evoluir rapidamente a valores impróprios, conforme mostrado na Figura 3.6.

Segundo Kass, Witkin e Terzopoulos (1988), os parâmetros da eq.(3.3) quando  $\alpha(s) = \alpha$  e  $\beta(s) = \beta$  são constantes, implica que a força interna limita a deformação e dobramento da curva, tornando-a suave e homogênea, conforme pode ser visto na Figura 3.7. Para a remoção dos laços foi desenvolvido um algoritmo que detecta os vértices que estão dentro da curva e os remove. O resultado desse método pode ser observado no exemplo da Figura 3.8.

### 3.1.8 Observações práticas para construção do modelo

Por meio da análise das seqüências dos quadros obtidos a partir do exame de videoes-troboscopia, observa-se um efeito na fase de abertura da glote, em que as pregas vocais, inicialmente fechadas e unidas pela mucosa, iniciem o processo de abertura como se fossem um "zíper", partindo de um ponto inicial intermediário até as comissuras superior e inferior. Esse efeito pode ser visto em uma seqüência de quadros de um ciclo glótico da Figura 3.9.

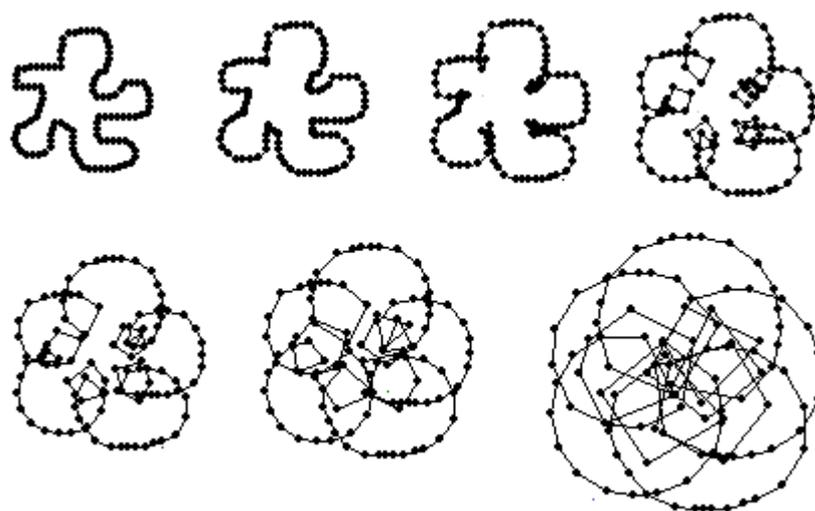


Figura 3.6: Curva se movendo somente sob ação de forças normais aos vértices.

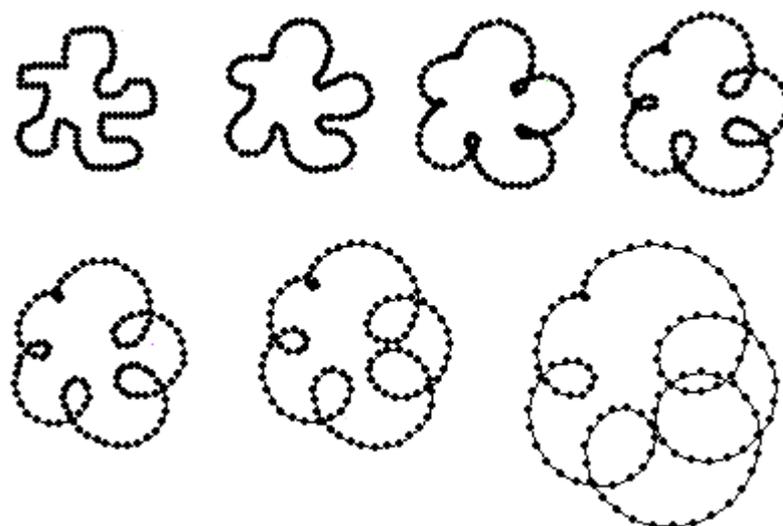


Figura 3.7: Curva se movendo sob ação das derivadas de primeira e segunda ordem.

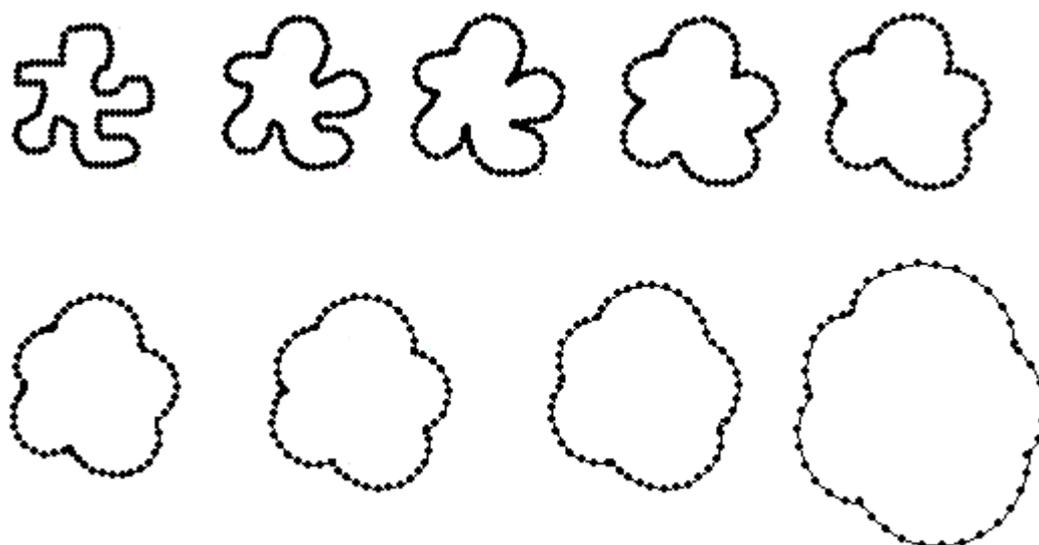


Figura 3.8: Expansão da curva com a remoção dos laços.

Essa análise é condizente com as observadas no gráfico da área de contato das pregas vocais da Figura 3.10, em que as pregas vocais que estão inicialmente fechadas, apresentam neste instante a área de contato máxima e indicada na parte inferior do gráfico, sendo que, o mesmo foi apresentado invertido apenas para ficar condizente com os artigos de Rothenberg (1979).

Conforme apresentado no modelo da Figura 3.10, após o instante  $t_1$ , em que as pregas estão completamente fechadas, ocorre o início da abertura em um determinado ponto, não necessariamente no centro. Geralmente, essa posição depende do formato das pregas vocais e das condições da viscosidade da mucosa e ocorre, normalmente, no terço médio das pregas vocais. Conforme apresentado no exemplo, o início da abertura ocorre na parte inferior e as pregas começam a abrir em direção às comissuras como se fosse um zíper. Esse fenômeno é responsável pela mudança da taxa de abertura das pregas vocais observadas no intervalo  $t_2$  para o  $t_3$ . No momento em que as frentes de abertura das pregas encontram um dos lados da comissura, sua abertura é interrompida, contudo, o outro lado ainda se encontra em movimento, porém a taxa de crescimento da área de contato das pregas vocais cai praticamente pela metade.

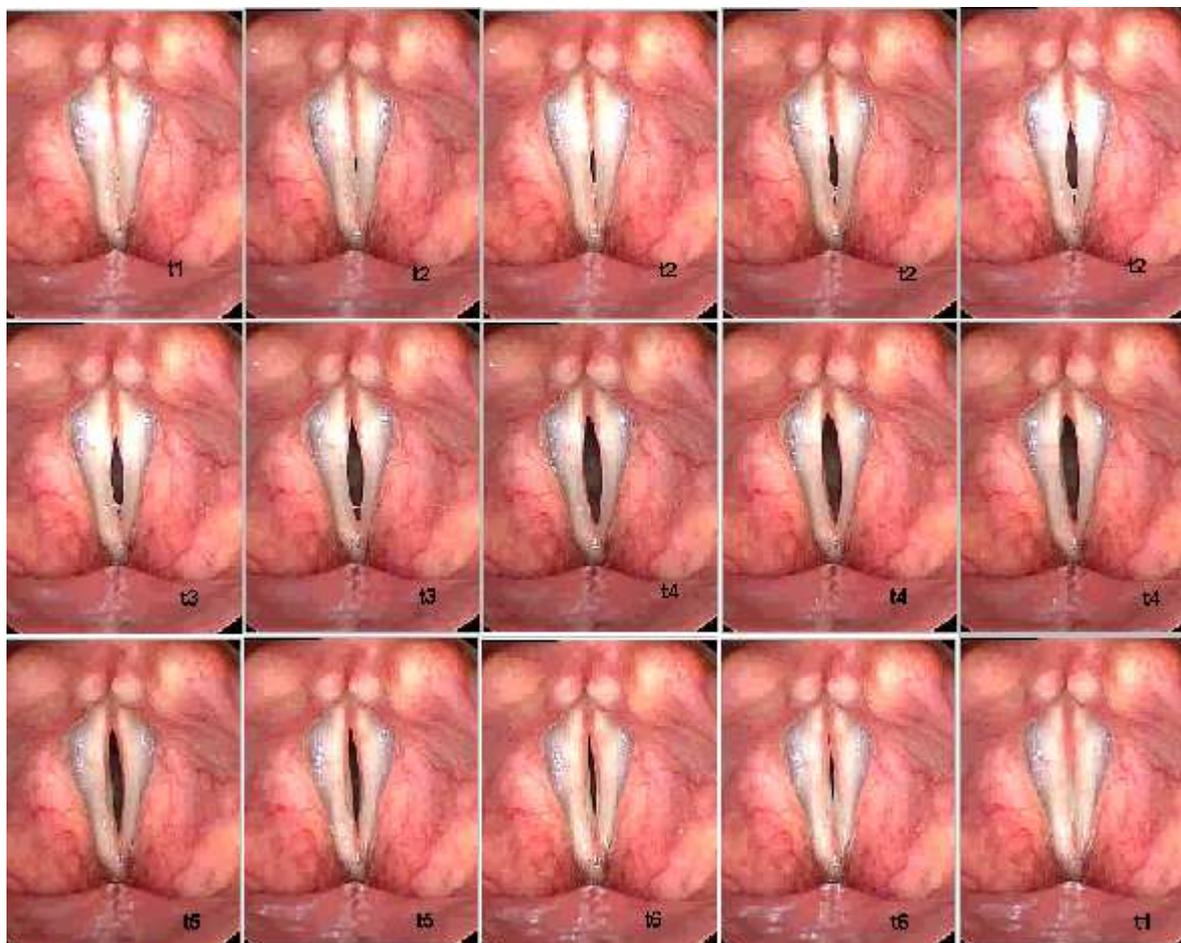


Figura 3.9: Sequência de quadros de um ciclo glotal.

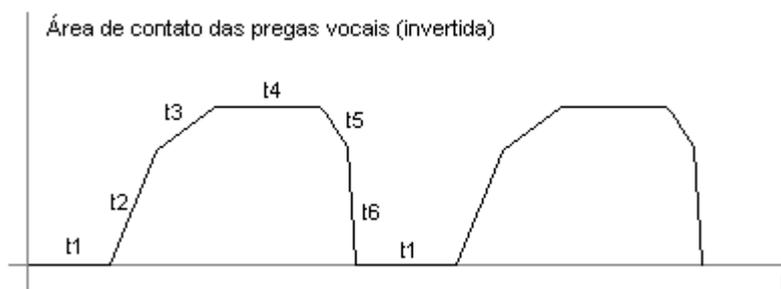


Figura 3.10: Área de contato das pregas vocais (invertida).

Mesmo quando ambos os lados de abertura encontram as comissuras inferior e superior o processo de abertura das pregas ainda permanece por um determinado instante, porém a área e contato das pregas vocais permanecerão constante em seu valor mínimo, indicado pelo trecho t4.

Após a fase aberta, o ciclo glotal entra em fase de fechamento que ocorre em duas etapas: as pregas começam a retornar à sua posição de repouso e próximo a elas o efeito Bernoulli acentua-se, fazendo com que estas se unam em uma taxa mais acelerada.

Outro exame que corrobora esses resultados consiste na eletroglotografia, que é realizado por meio de dois eletrodos presos no pescoço perto da glote por uma tira elástica. É importante observar que o sinal proveniente deste exame não deve ser confundido com o pulso glotal. O sinal da eletroglotografia é diretamente proporcional à área de contato das pregas vocais e devido à natureza do exame as bordas são suavizadas pela ação da filtragem inerente ao método. Por outro lado, o pulso glotal está relacionado ao ar que sai através da glote. Esses sinais estão correlacionados entre si, conforme mostrado na Figura 3.11, porém, são informações distintas.

Poucos são os trabalhos apresentados na literatura em que é observado o estudo

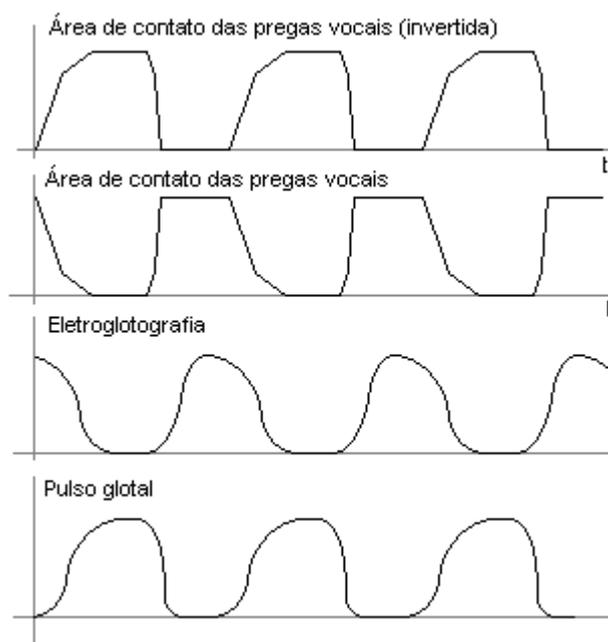


Figura 3.11: Comparações entre area de contato das pregas vocais, eletroglotografia e pulso glotal.

da área de contato das pregas vocais, assim como, uma análise mais detalhada do pulso glotal e que contenha as informações de sua derivada (FANT; LILJENCRAINTS; LIN, 1985; FANT, 1995; VELDHUIS, 1988). A análise direta do sinal de voz, assim como o sinal do pulso glotal não é capaz de fornecer subsídios suficiente para garantir se o sinal sintetizado manteve características essenciais de um sinal real. A análise da derivada do sinal glotal, a qual possui características bem definidas quanto a sua forma, facilita a inspeção visual para determinar a qualidade do pulso glotal.

Trabalhos recentes quantificam algumas particularidades encontradas na derivada do pulso glotal, (GUERRA, 2005). A derivada do pulso glotal pode ser vista na Figura 3.12, existindo cinco etapas bem definidas. Na primeira o pulso glotal possui uma variação lenta ascendente até o valor  $t_i$ , trecho em que a derivada chega ao valor máximo, posteriormente o pulso chega ao cume, instante  $t_p$ , em que a derivada tem valor igual a zero, neste instante o pulso glotal começa a descer e torna-se nítida as duas fase  $t_e$  e  $t_c$ , sendo que a primeira faz com que a derivada chegue ao seu valor mínimo e a segunda fase, a derivada retorna ao valor igual a zero em uma ascensão exponencial, permanecendo com este valor até o instante seguinte  $t_0$ , período correspondente ao tempo em que as pregas encontram-se fechadas.

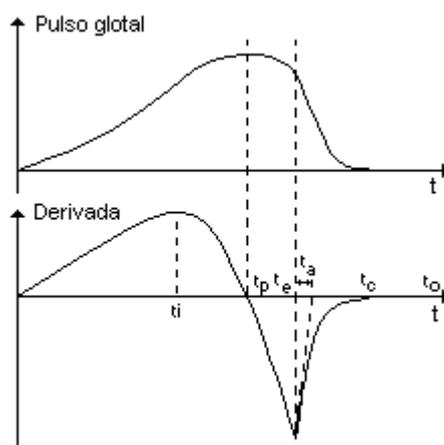


Figura 3.12: Pulso glotal e sua derivada.

### 3.1.9 Simulando o movimento glotal

A simulação inicial dos movimentos da glote foi realizada em duas dimensões, levando-se em conta as observações realizadas anteriormente e, principalmente, feitas no exame de endoscopia.

O comprimento vertical de abertura da glote no exame é medido em números de *pixels*. Essa medida foi considerada como um valor de referência métrica na imagem, pois não é possível estabelecer uma referência métrica precisa em um exame da laringoscopia. Para produzir o contorno da glote foram utilizadas duas curvas abertas com 100 vértices cada uma. Todas as outras dimensões são proporcionais à referência estabelecida.

A prega vocal foi simulada em sua posição fechada por uma curva aberta, representada por uma reta que possui as extremidades presas, simulando as comissuras posterior e anterior. As pregas vocais que estão sob ação da pressão do ar, originado dos pulmões, faz com que as pregas se abram. No exemplo da Figura 3.13 gera-se uma frente de propagação constante em direção a sua normal, conforme as equações eq.(3.10) e considerando a função da curvatura  $F(k) = 1$ . Para que o movimento da glote simule o compor-

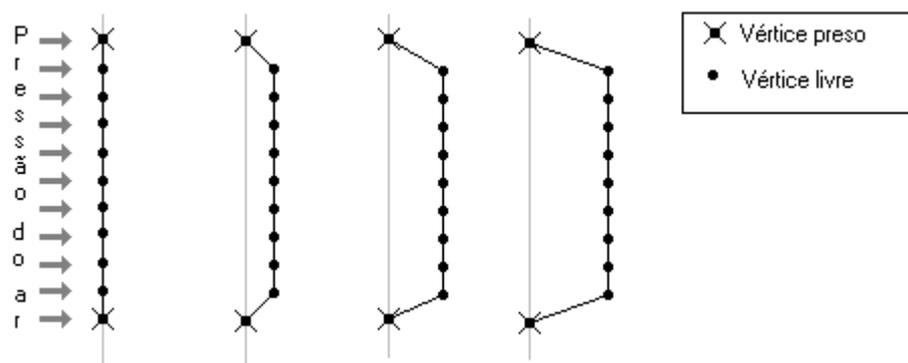


Figura 3.13: Expansão da glote devido a pressão do ar.

tamento de uma membrana elástica foi acrescentado à eq.(3.10) uma derivada de segunda ordem, conforme proposto por Kass, Witkin e Terzopoulos (1988). Na Figura 3.14 temos os efeitos ocorridos devido a estas alterações.

No exame de estroboscopia, conforme descrito anteriormente, observa-se que as pregas se abrem, desunindo as paredes da mucosa como se fosse um zíper. Esse efeito foi obtido por meio das propriedades das forças que regem o movimento. A força que atua sobre a vizinhança de um determinado vértice atrai os mesmos para uma posição de

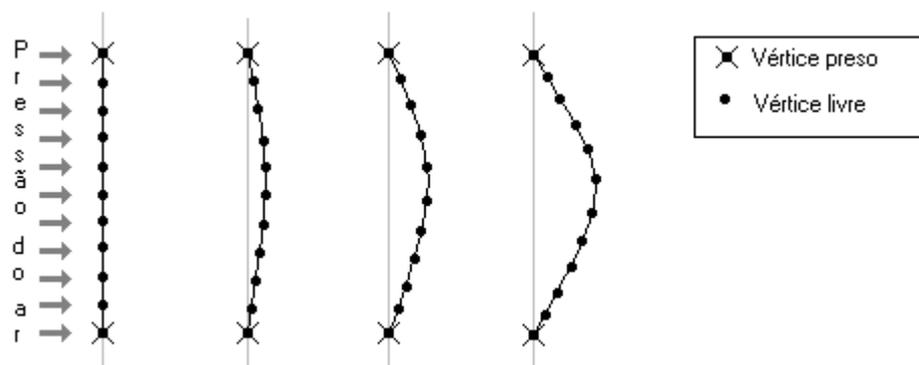


Figura 3.14: Expansão da glote devido a pressão do ar, com acréscimo da derivada de segunda ordem.

menor energia ou curvatura. Valendo-se desse princípio, foi introduzido um limiar que não permite que os vértices vizinhos se movimentem antes de ultrapassar um determinado valor. No início da fase de abertura escolhe-se uma determinada posição para começar o processo, similar à observada no exame de estroboscopia. Se estabelece que o vértice relativo a esta posição tenha o limiar igual a zero, portando, ele se move livremente, exercendo uma força sobre os vértices vizinhos proporcional ao seu deslocamento. A Figura 3.15 apresenta este efeito.

Após a fase aberta, inicia-se a fase de fechamento, na qual o efeito zíper não ocorre

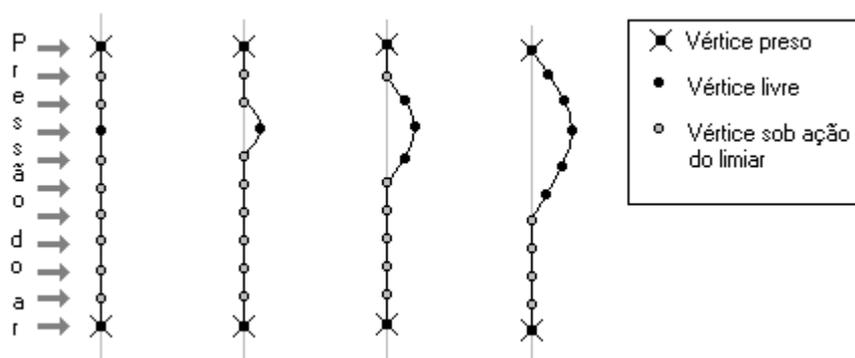


Figura 3.15: Simulação numérica do efeito 'zíper'.

mais. O fechamento das pregas ocorre devido a elasticidade das pregas e de forças aerodinâmicas advindas do efeito Bernoulli. Esse efeito faz com que as pregas se fechem devido

ao fluxo de ar que as atravessam e também devido à sua geometria, conforme apresentado na Figura 3.16. O termo de segunda ordem, o qual confere à curva propriedades de membrana, também é responsável pelo fechamento, pois é considerado que o movimento realizado pela fase de abertura armazene uma certa energia e o termo de segunda ordem é responsável pela restauração da curva em sua posição original.

Quando a glote se fecha, ocorre um choque elástico entre as mucosas das pregas

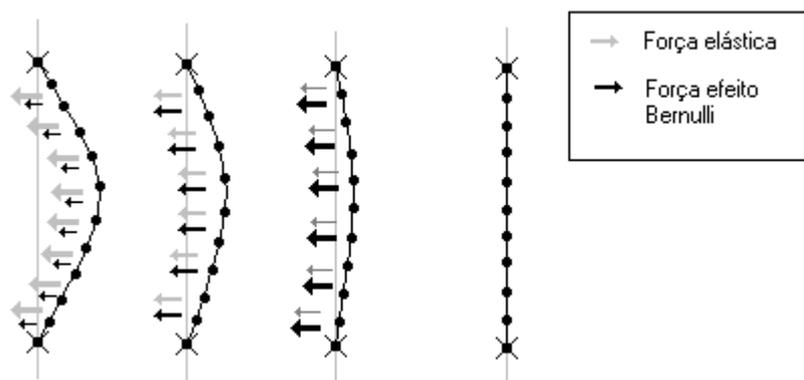


Figura 3.16: Fase de fechamento ação da força elástica e de Bernoulli.

vocais, parte dessa energia é dissipada e a outra parte é convertida novamente em movimento, por meio da adição de uma força de colisão no modelo.

Levando em conta todas as considerações anteriores, pode-se estabelecer o modelo que tem como equação de movimento da glote a equação eq.(3.12):

$$(x_i^{n+1}, y_i^{n+1}) = (x_i^n, y_i^n) + (A_o + C_{o_i} - F_r(\kappa_i^n) - F_b(S))\vec{n}_i - K_m v_{ss} \quad (3.12)$$

onde,

$A_o$  corresponde ao fator de escala relativo à pressão do ar na glote.

$C_{o_i}$  corresponde ao fator de escala relativo ao choque elástico das mucosa, dependente da posição  $i$ .

$F_r(\kappa_i^n)$  Força de restauração em função da curvatura ( $\kappa_i^n$ ).

$F_b(S)$  Força do efeito Bernoulli em função da área da glote.

$K_m$  Constante de membrana.

sendo que,

$$\vec{n} = \frac{((y_{i+1}^n - y_{i-1}^n), -(x_{i+1}^n - x_{i-1}^n))}{((x_{i+1}^n - x_{i-1}^n)^2 + (y_{i+1}^n - y_{i-1}^n)^2)^{1/2}}, \quad (3.13)$$

e,

$$v_{ss} = ((x_{i+1}^n - 2x_i^n + x_{i-1}^n)/2), (y_{i+1}^n - 2y_i^n + y_{i-1}^n)/2)) \quad (3.14)$$

A função de restauração é proporcional à curvatura, conforme a equação eq.(3.15).

$$F_r(\kappa_i^n) = Ro \frac{(y_{i+1}^n - 2y_i^n + y_{i-1}^n)(x_{i+1}^n - x_{i-1}^n) - (x_{i+1}^n - 2x_i^n + x_{i-1}^n)(y_{i+1}^n - y_{i-1}^n)}{((x_{i+1}^n - x_{i-1}^n)^2 + (y_{i+1}^n - y_{i-1}^n)^2)^{3/2}} \quad (3.15)$$

E a função relativa ao efeito Bernoulli é inversamente proporcional ao quadrado da área da glote, conforme a equação eq.(3.16), se a área for diferente de zero, caso contrário, seu valor é considerado nulo.

$$F_b(S) = \frac{Bo}{S_g^2} \quad (3.16)$$

O valores  $Ro$  e  $Bo$  são fatores de escala das forças de restauração e Bernoulli, respectivamente.

A área da glote  $S_g$  é calculada pelo método trapezoidal. Esse método utiliza um algoritmo que encontra a área de vários trapezóides, a soma das áreas de todos eles definiram a área do polígono, conforme apresentado na equação eq.(3.17).

$$S = \sum_{i=1}^n \frac{(x_{i+1} - x_i)(y_{i+1} + y_i)}{2} \quad (3.17)$$

observando que,  $(x_{n+1}, y_{n+1}) = (x_1, y_1)$  e quando  $(x_{i+1} < x_i)$  a contribuição para área é negativa, permitindo que os trapezóides formados pela porção mais baixa do polígono sejam subtraídos.

### 3.1.10 Pulso e filtro glotal a partir do modelo

Segundo Flanagan (1972), Rothenberg (1981), Titze (1988), o fluxo de ar glotal é proporcional à área de abertura da glote e devido à inércia da massa de ar o pulso glotal possui uma forma não-simétrica, conforme mostrado na Figura 3.17.

Na Figura 3.18 é mostrado um modelo linear simplificado para a fonte glotal uti-

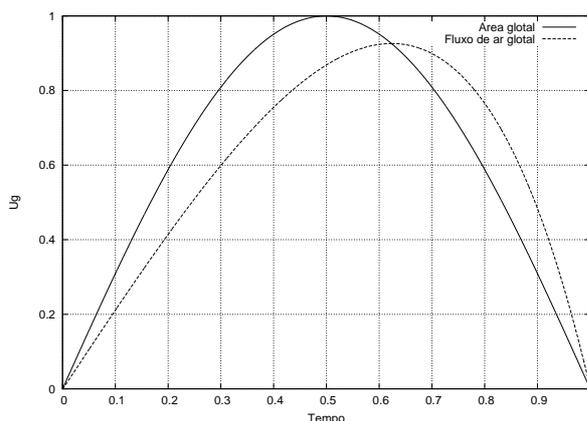


Figura 3.17: Relação entre a área glotal e o pulso de ar glotal.

lizando um circuito elétrico, em que a área da glote é proporcional à condutância  $Y_g$ , a inércia do ar devido ao acoplamento com o trato vocal é proporcional à  $L_t$  e o fluxo de ar é representado por  $U_g$ .

O equacionamento do modelo leva a uma equação diferencial eq.(3.18).

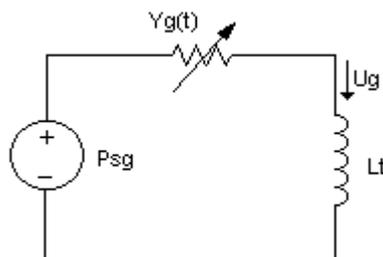


Figura 3.18: Modelo linear para a fonte glotal.

$$P_{sg} = \frac{U_g}{Y_g(t)} + L_t \dot{U}_g \quad (3.18)$$

A solução algébrica para algumas funções conhecidas da área da glote para o modelo descrito acima pode ser encontrada no trabalho de Rothenberg (1981). Devido a área da glote ser uma função do tempo obtida através da medida da área da glote em exames de estroboscopia, a solução algébrica desenvolvidas por Rothenberg (1981) não pode ser utilizada. Para isto a equação diferencial foi solucionada utilizando diferenças finitas e seu resultado é apresentado na equação eq.(3.19).

$$U_g^{t+1} = \frac{\Delta t Y_g(t) + Y_g(t) L_t U_g^t}{\Delta t + Y_g(t) L_t} \quad (3.19)$$

O filtro glotal é obtido pela transformada de Fourier do pulso glotal, e segundo o resultado de diversos modelos o espectro do filtro é semelhante a uma curva com decaimento exponencial.

## 3.2 Processamento de áudio

### 3.2.1 Gravação do sinal de voz

A digitalização, edição e análise acústica dos sinais de voz foram realizadas utilizando o programa Análise Acústica da Voz 5.0, (MONTAGNOLI; PEREIRA, 2005), desenvolvido a partir da dissertação de mestrado, (MONTAGNOLI, 1998).

As amostragens dos sinais de voz foram feitas utilizando a gravação digital pela placa de som do PC, através de uma taxa de conversão de 44 KHz e 16 bits de resolução. O sinal de voz foi particionado em janelas de tamanho 2048 pontos, devido as exigências da transformada rápida de Fourier por quadros de tamanho igual a potência de dois.

### 3.2.2 Síntese da voz após modificações no modelo

Conforme descrito anteriormente, o modelo do sistema fonador, quando considerado linear e invariante no tempo, possui propriedades de separabilidade. O sinal de voz gravado traz informações desde a excitação até a radiação, passando pelo modelo da glote e trato vocal. No domínio da frequência temos o produto de todos os termos conforme apresentado na equação eq.(3.20).

$$S(z) = E(z).G(z).V(z).L(z) \quad (3.20)$$

onde  $E(z)$  é o modelo para a excitação,  $G(z)$  é o modelo da glote,  $V(z)$  é o modelo do trato vocal,  $L(z)$  é o modelo da radiação labial.

A remoção das características da glote do sinal de voz é feita através da aplicação do filtro inverso glotal estimado:

$$S'(z) = E(z).G(z).V(z).L(z) \frac{1}{\hat{G}(z)} \quad (3.21)$$

$$S'(z) = E(z).V(z).L(z) \quad (3.22)$$

Realizando as alterações no modelo, um novo pulso glotal pode ser estimado a partir dos movimentos da glote no modelo e, posteriormente, é adicionado ao sinal  $S'$ , conforme a equação eq.(3.23). Dessa forma é estimado um novo sinal de voz que mantém as características da excitação, trato vocal e radiação.

$$\hat{S}(z) = E(z).\hat{G}(z).V(z).L(z) \quad (3.23)$$

Os ajustes nos parâmetros do modelo para uma melhor obtenção das características normais de uma voz foram realizados, principalmente, observando as características da derivada do pulso glotal e da forma do espectro de seu filtro. Uma discussão mais detalhada sobre estes procedimentos será realizada nos capítulos subsequentes.

### 3.3 Programa para auxílio da vídeolaringoscopia e cirurgia virtual

Para organizar as etapas envolvidas na elaboração da cirurgia virtual foi desenvolvido o programa para auxílio da vídeolaringoscopia e cirurgia virtual, que possui os seguintes módulos: Banco de dados, Gravação e edição de vídeo e extração do áudio, Exame visual da laringoscopia, Medidas da laringe, Quimografia e Cirurgia Virtual. Cada um destes

---

módulos será apresentado nos próximos itens, e as janelas de interface com o usuário serão apresentadas no Apêndice A.

### **3.3.1 Banco de dados**

Devido ao grande volume de informações interrelacionadas, como dados pessoais, imagens, vídeos e sinais de vozes dos pacientes tornou-se imprescindível a elaboração de um banco de dados para gerenciar estas tarefas.

O banco de dados utilizado foi o Paradox e as tabelas foram relacionadas utilizando o campo protocolo de cada paciente como chave de relacionamento. O cadastro de pacientes é o mesmo utilizado no Programa Análise de voz 5.0, (MONTAGNOLI; PEREIRA, 2005), Apêndice A. Nesse programa foram acrescentadas as tabelas para armazenar novas informações tais como as medidas nos exames de laringoscopia e quimografia, e auxiliar o gerenciamento das imagens e vídeos.

### **3.3.2 Gravação e edição de vídeo e extração do áudio**

As gravações de vídeos foram feitas através da captura direta do sinal da câmera ou pela digitalização de fitas de videocassete. Devido a grande quantidade de memória necessária para o armazenamento, foi fundamental o procedimento de edição e compactação das imagens.

Nesse módulo também é possível fazer a extração do sinal de áudio do exame de vídeo, para uma posterior análise acústica da voz.

### **3.3.3 Exame visual da laringoscopia**

Através do exame visual da videolaringoscopia é possível identificar diversos aspectos quanto à forma e movimento das pregas vocais tais como: fechamento glótico, fase de fechamento, amplitude e mobilidade das pregas, simetria de fase, onda mucosa, liberdade das bordas.

### **3.3.4 Medidas da laringe**

O módulo do programa que permite as medidas da laringe permite realizar medidas na laringe e utiliza o comprimento vertical como referência métrica. As medidas extraídas

deste módulo consistem na espessura das pregas vocais, largura máxima na comissura superior, largura mínima na comissura inferior, posição inicial da abertura das pregas, largura e comprimento de eventual fenda e posição e área das alterações de massas.

### 3.3.5 Quimografia

O exame de quimografia, conforme proposto por Lee et al. (2001), é realizado através da justaposição de uma seqüência de linhas obtidas dos quadros de vídeo de um exame de estroboscopia da laringe, Figura 3.19.

A partir deste exame são efetuadas medidas complementares como o tempo das fases de abertura, da fase de fechamento, da fase de aberta e da fase fechada, e as amplitudes máximas e mínimas da abertura das pregas vocais. Para isso deve-se selecionar a posição de corte dos quadros de vídeo que possua a maior amplitude de abertura das pregas vocais.

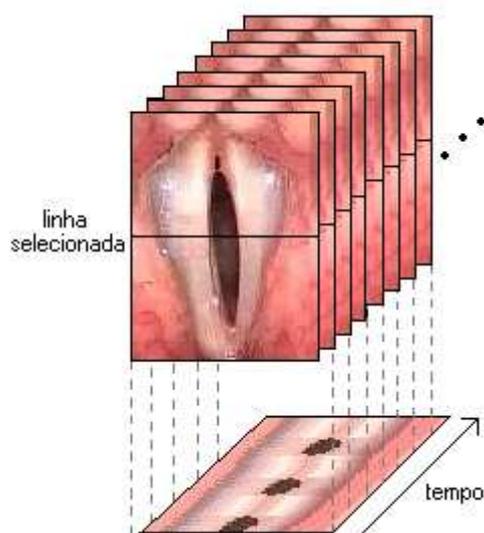


Figura 3.19: Image estroboscópica, a linha selecionada para o corte e a respectiva vídeo-estrobocinografia projetada abaixo.

# Capítulo 4

## Resultados

### 4.1 Cirurgia Virtual

A cirurgia virtual consiste em realizar alterações paramétricas e ajustes para que o modelo se comporte o mais próximo possível de um caso real.

Foram realizados alguns estudos de casos, iniciando pelo caso normal, onde alterações no modelo foram efetuadas com o objetivo de provocar as anomalias possíveis previstas pelo modelo e analisar os efeitos mecânicos e acústicos correlacionados. Posteriormente algumas patologias foram analisada, como cisto, nódulo, pólipos, fenda e paralisia.

#### 4.1.1 Caso: Normal

Na Tabela 4.1 são apresentadas as medidas efetuadas a partir da imagem das pregas vocais, conforme mostrado na Figura 4.1. Na Tabela 4.2 são apresentados as medidas da dinâmica das pregas vocais realizadas a partir do exame de quimografia, 4.2.

Na Figura 4.3 é apresentado um ciclo glotal de um exame de endoscopia estroboscópica de uma laringe normal e seu respectivo ciclo simulado pelo modelo proposto. Cada quadro apresentado é amostrado em uma razão de 1/10 do vídeo original.

O pulso glotal e sua derivada, o espectro de Fourier e a área de contato das pregas vocais para uma simulação de uma laringe normal são apresentados na Figura 4.4.

#### 4.1.2 Caso: Normal com diminuição da rigidez da prega no lado direito

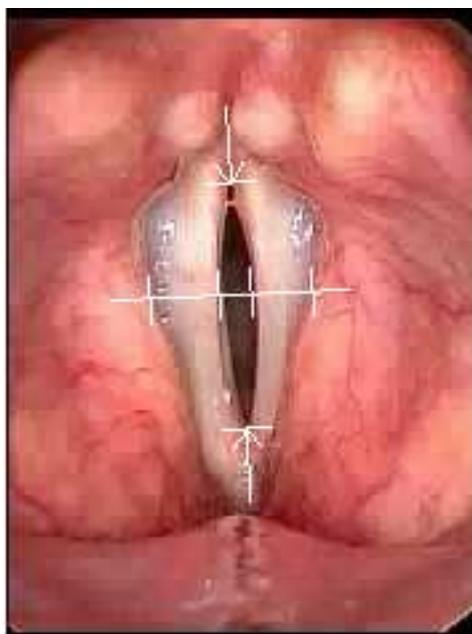


Figura 4.1: Imagem de uma laringe normal obtida por um endoscópio rígido.

Tabela 4.1: Medidas das pregas vocais - caso normal

Parâmetro	Valor
Comprimento Vertical	98,0
Espessura da prega vocal direita	28,6
Espessura da prega vocal esquerda	28,5
Largura máxima na comissura superior	72,4
Largura mínima na comissura inferior	30,6
Largura da Fenda	0,00
Comprimento da Fenda	0,00
Posição da Fenda	0,00
Posição inicial da abertura das pregas	63,0
Posição da alteração de massa 1	0,00
Área da massa 1	0,00
Posição da alteração de massa 2	0,00
Área da massa 2	0,00

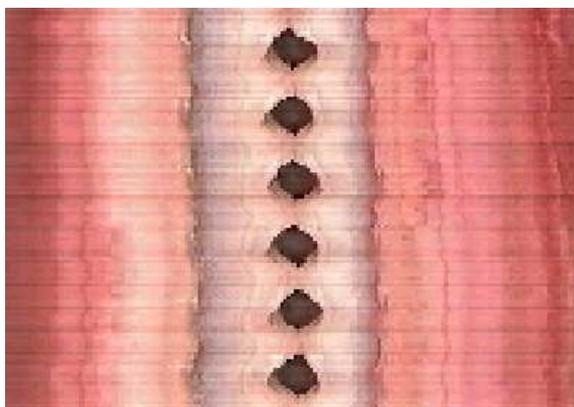


Figura 4.2: Exame quimográfico de um caso normal.

Tabela 4.2: Avaliação da dinâmica das pregas vocais - caso normal

Parâmetro	Valor
Fase aberta	0,385
Fase fechada	0,367
Fase de abertura	0,178
Fase de fechamento	0,207
Quociente de abertura	0,51
Amplitude máxima direita	7,35
Amplitude mínima direita	0,00
Amplitude máxima esquerda	7,35
Amplitude mínima esquerda	0,00

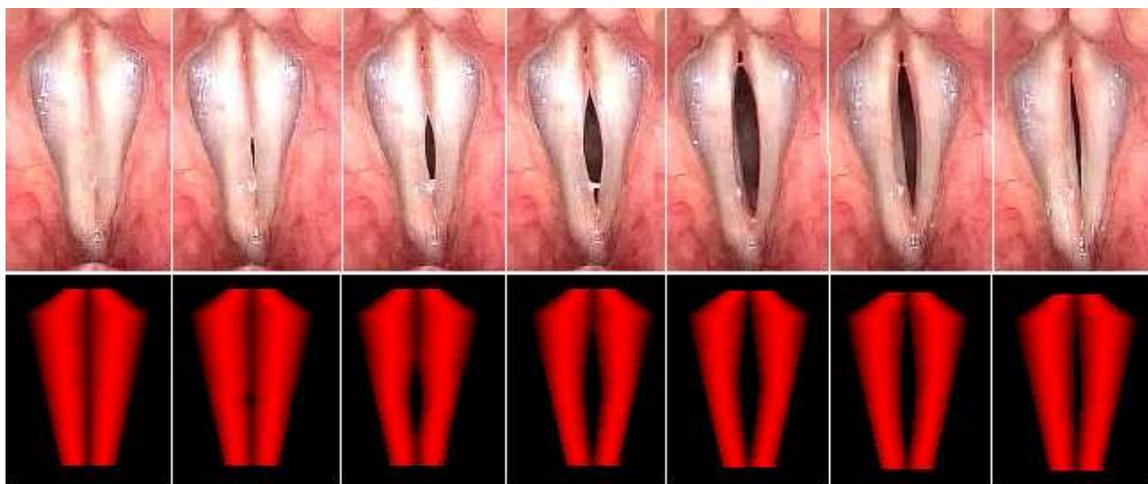


Figura 4.3: ciclo glotal do exame estroboscópico de um caso normal e o respectivo ciclo simulado.

Na Figura 4.5 é apresentado um ciclo glotal simulado pelo modelo proposto com a diminuição da rigidez da prega somente no lado direito, sendo que, o modelo pode modificar ambas as rigidez da pregas direita e esquerda, optou-se por alterar apenas a da direita com o intuito de evidenciar o efeito.

O pulso glotal e sua derivada, o espectro de Fourier e a área de contato das pregas vocais, para uma simulação de uma laringe normal, com a diminuição da rigidez do lado direito são apresentados na Figura 4.6.

#### 4.1.3 Caso: Normal com eliminação da viscosidade das pregas

Na Figura 4.7 é apresentado um ciclo glotal simulado pelo modelo proposto sem a viscosidade entre as pregas vocais. Nota-se que o efeito zipper é eliminado e a área de contato muda de totalmente em contato para nenhum contato quase que instantaneamente.

O pulso glotal e sua derivada, o espectro de Fourier e a área de contato das pregas vocais para uma simulação de uma laringe normal com eliminação da viscosidade das pregas vocais são apresentados na Figura 4.8.

#### 4.1.4 Caso: Normal com adição de uma fenda na comissura superior

Outra alteração possível no fechamento glótico consiste na adição de uma fenda nas pregas vocais. Na Figura 4.9 é apresentado um ciclo glotal simulado pelo modelo pro-

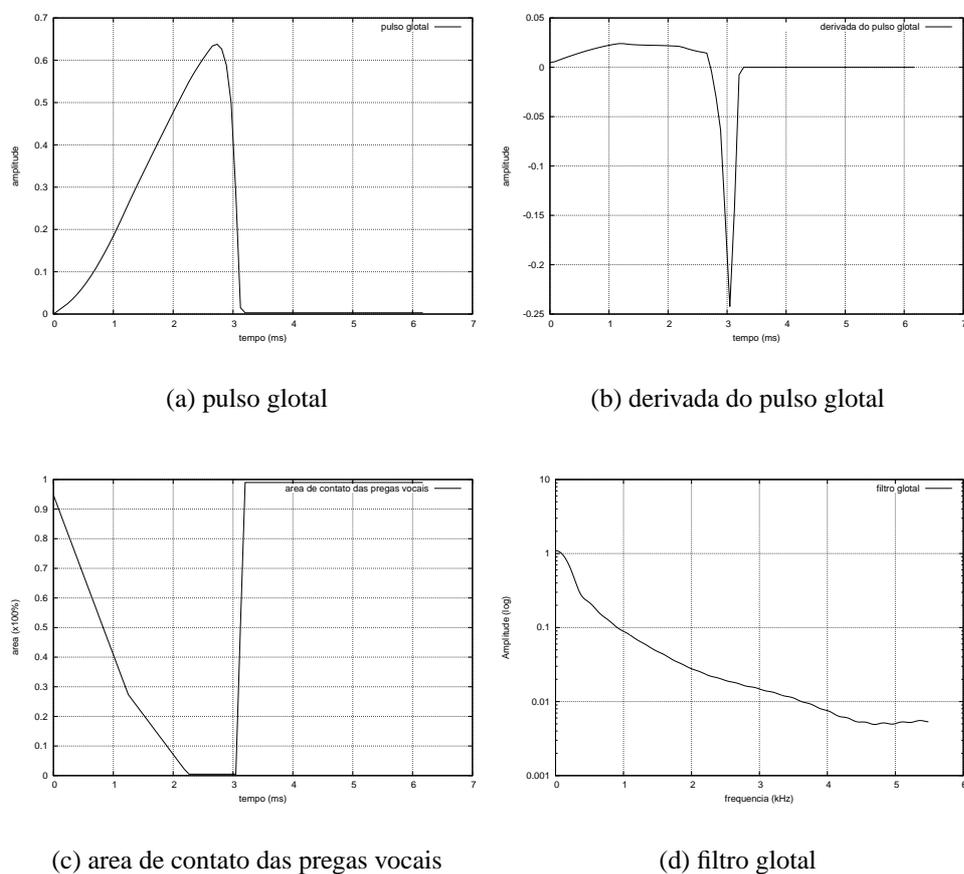


Figura 4.4: simulação numérica do caso normal.

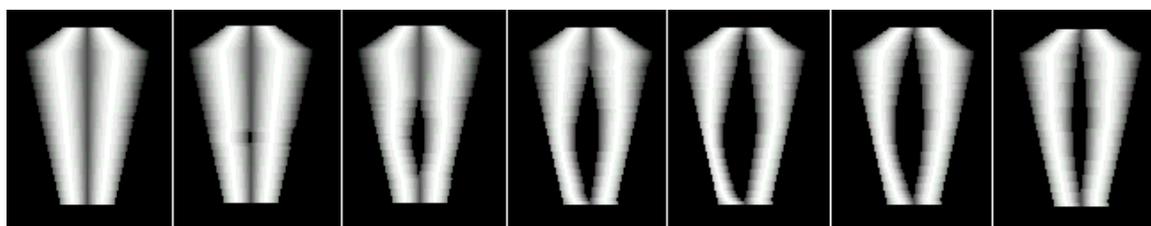


Figura 4.5: Ciclo glotal da simulação numérica com diminuição da rigidez no lado direito.

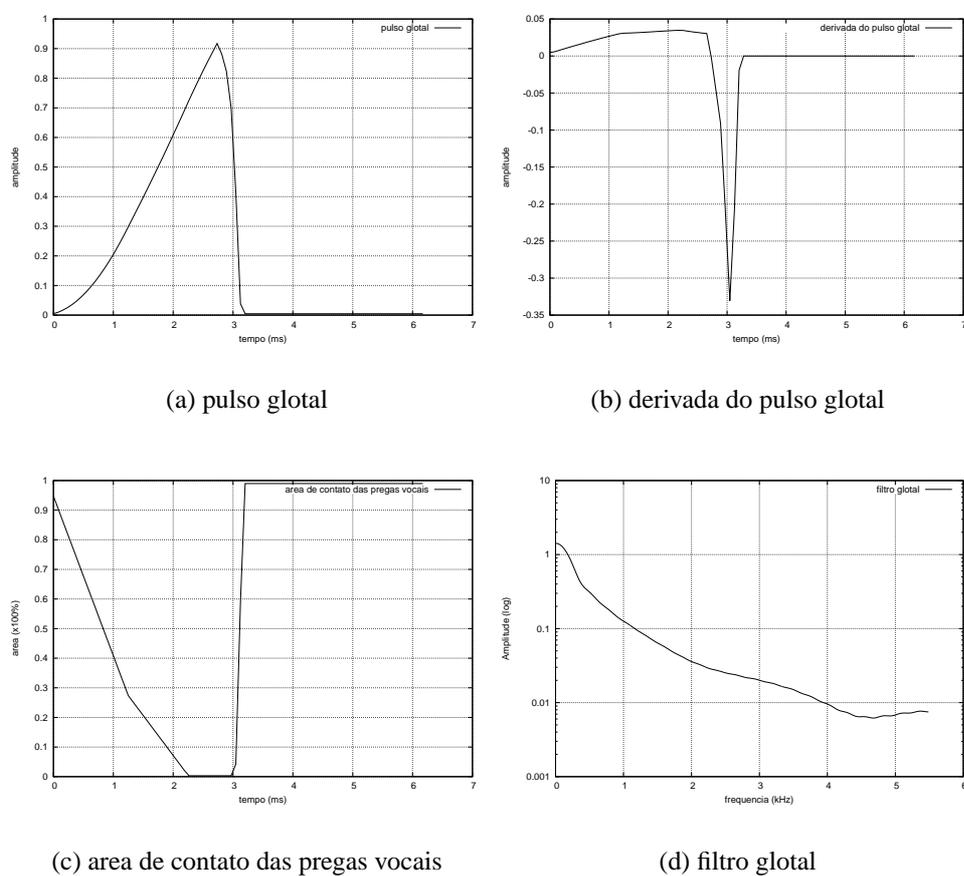


Figura 4.6: simulação numérica do caso normal com diminuição da rigidez no lado direito.

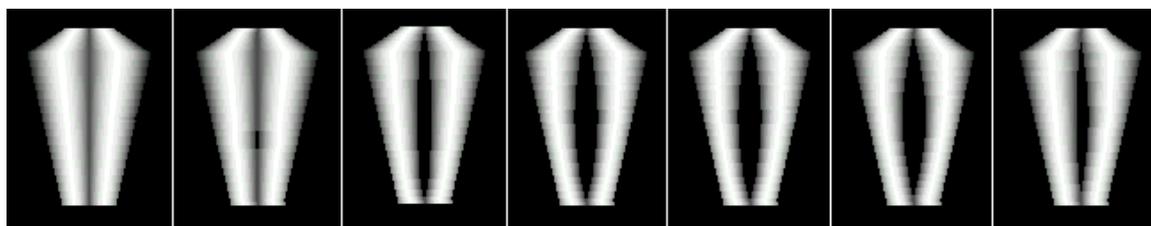
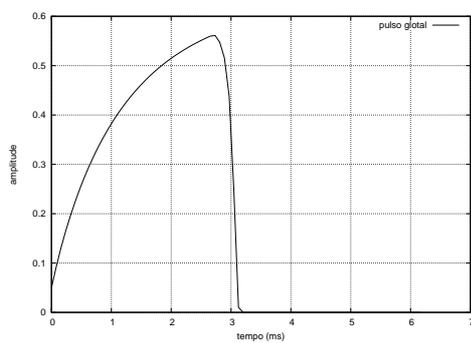
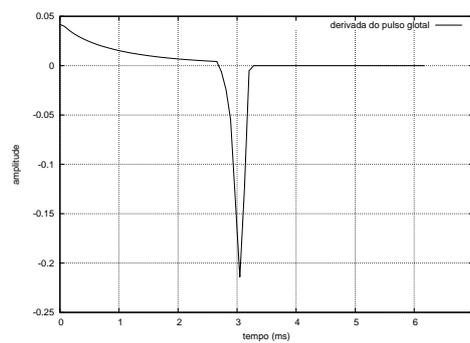


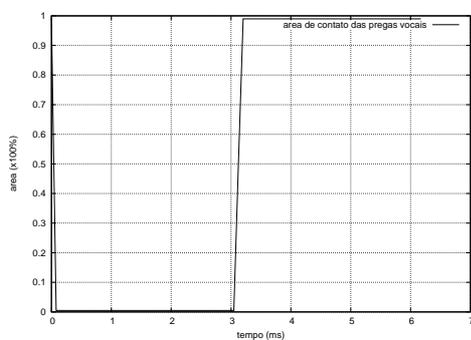
Figura 4.7: Ciclo glotal da simulação numérica sem viscosidade das pregas.



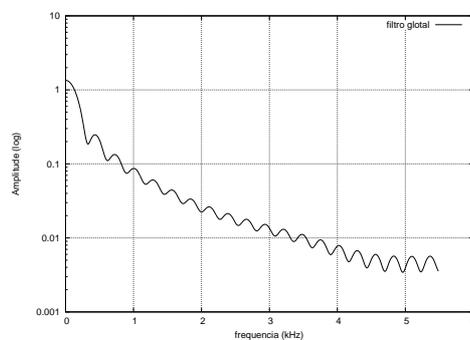
(a) pulso glotal



(b) derivada do pulso glotal



(c) area de contato das pregas vocais



(d) filtro glotal

Figura 4.8: simulação numérica do caso normal sem viscosidade das pregas.

posto com adição de uma fenda na comissura superior das pregas vocais.

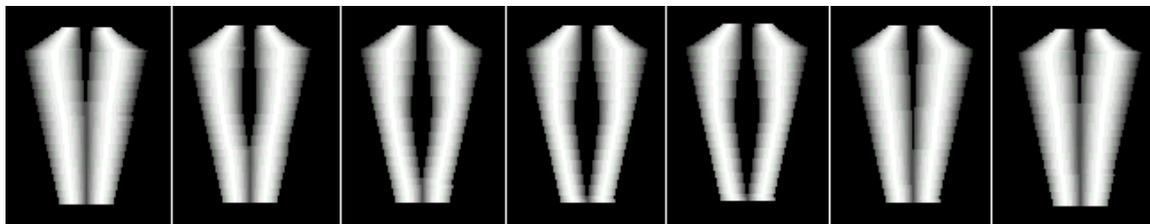


Figura 4.9: Ciclo glotal da simulação numérica com fenda na comissura superior.

O pulso glotal e sua derivada, o espectro de Fourier e a área de contato das pregas vocais para uma simulação de uma laringe normal com adição de uma fenda na comissura superior das pregas vocais são apresentados na Figura 4.10.

#### **4.1.5 Caso: Normal com adição de uma fenda no terço médio**

Na Figura 4.11 é apresentado um ciclo glotal simulado pelo modelo proposto com adição de uma fenda no terço médio das pregas vocais.

O pulso glotal e sua derivada, o espectro de Fourier e a área de contato das pregas vocais para uma simulação de uma laringe normal com adição de uma fenda no terço médio das pregas vocais são apresentados na Figura 4.12.

#### **4.1.6 Caso: Normal com adição de massa na prega vocal esquerda**

Muitas patologias de laringe consistem de alguma forma em uma alteração de massa nas pregas vocais, podendo-se encontrar bilateralmente como nos casos de nódulos ou unilateralmente como em casos de pólipos e cistos. Na Figura 4.13 é apresentado um ciclo glotal simulado pelo modelo proposto com adição de massa na prega vocal esquerda.

O pulso glotal e sua derivada, o espectro de Fourier e a área de contato das pregas vocais para uma simulação de uma laringe normal com adição de massa na prega vocal esquerda são apresentados na Figura 4.14.

#### **4.1.7 Caso: Normal com adição de massa bilateralmente**

Na Figura 4.13 é apresentado um ciclo glotal simulado pelo modelo proposto com adição de massa bilateral nas pregas vocais.

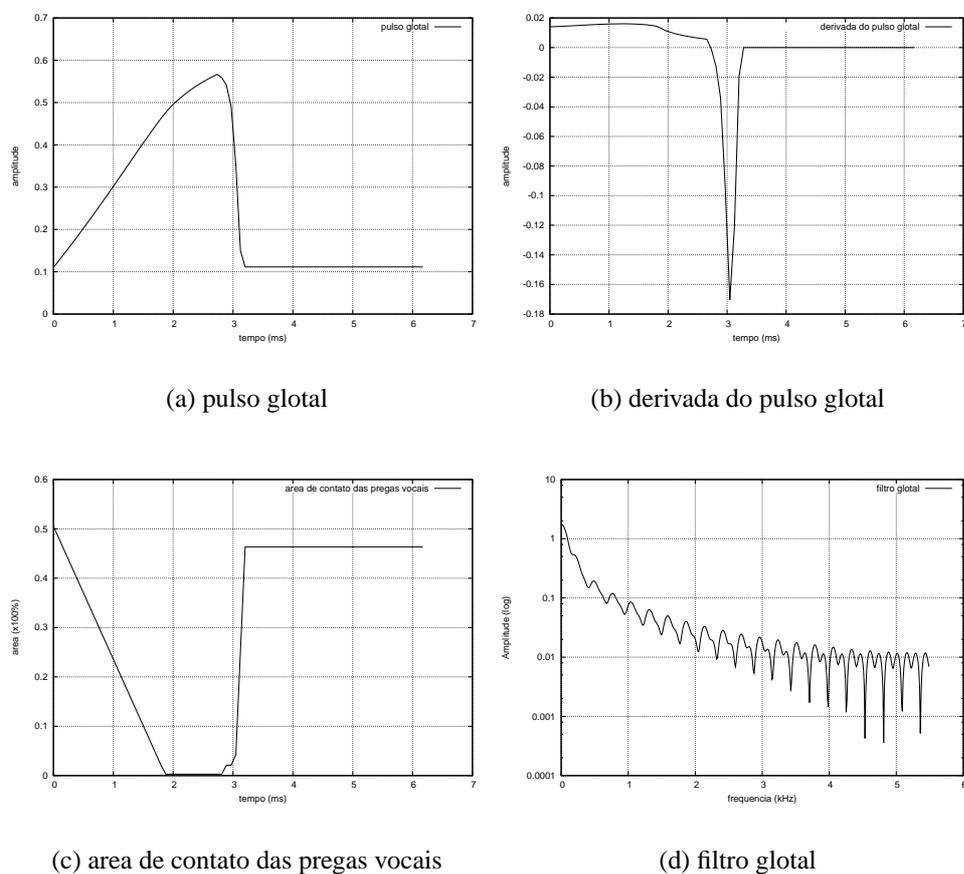


Figura 4.10: simulação numérica do caso normal com fenda na comissura superior.

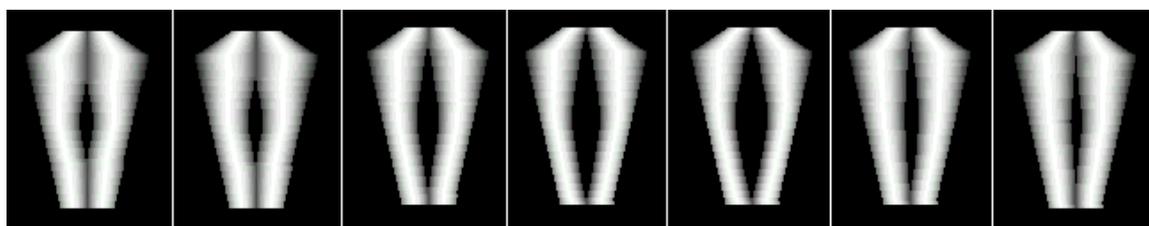


Figura 4.11: Ciclo glotal da simulação numérica com fenda no terço médio.

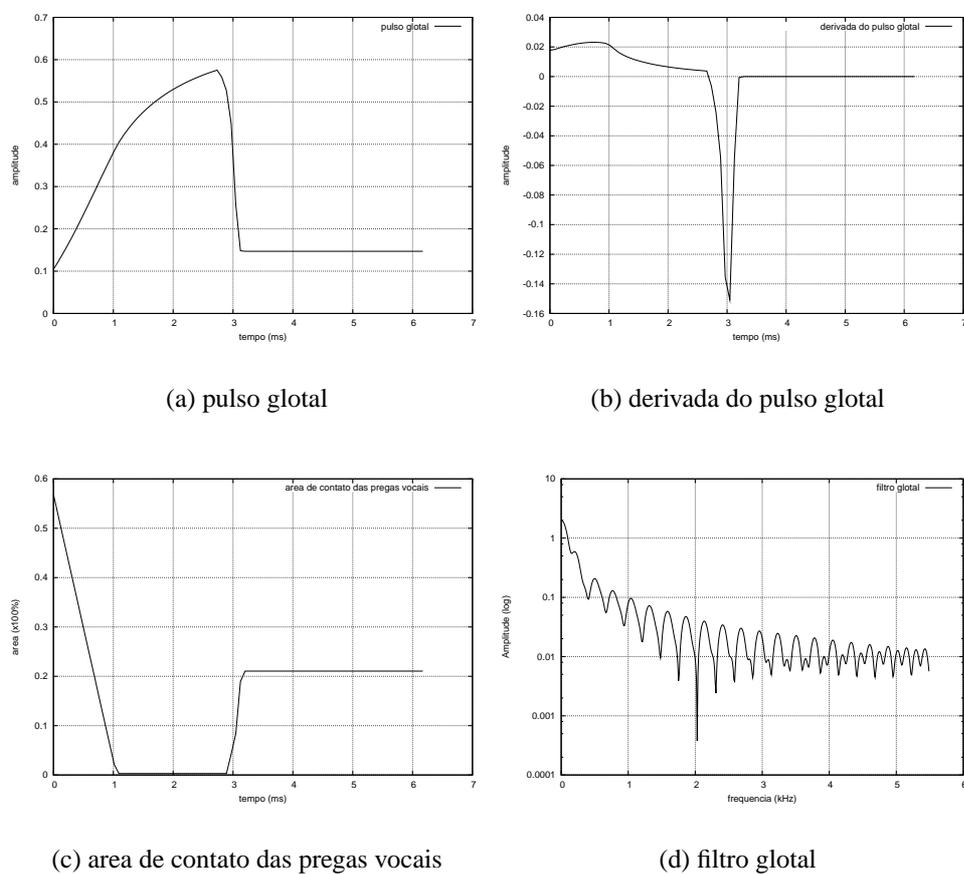


Figura 4.12: simulação numérica do caso normal com fenda no terço médio.

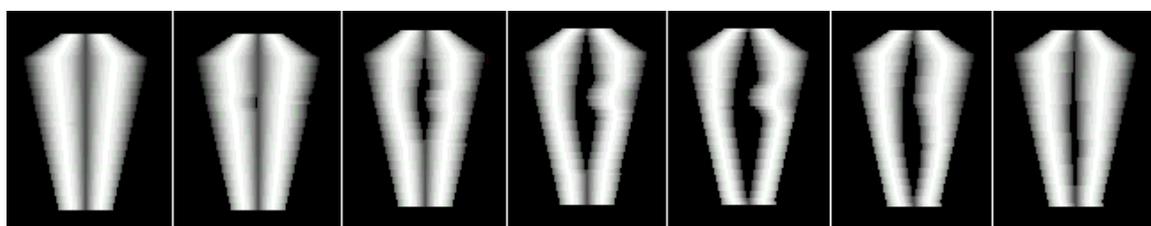


Figura 4.13: Ciclo glotal da simulação numérica com adição de massa na prega vocal esquerda.

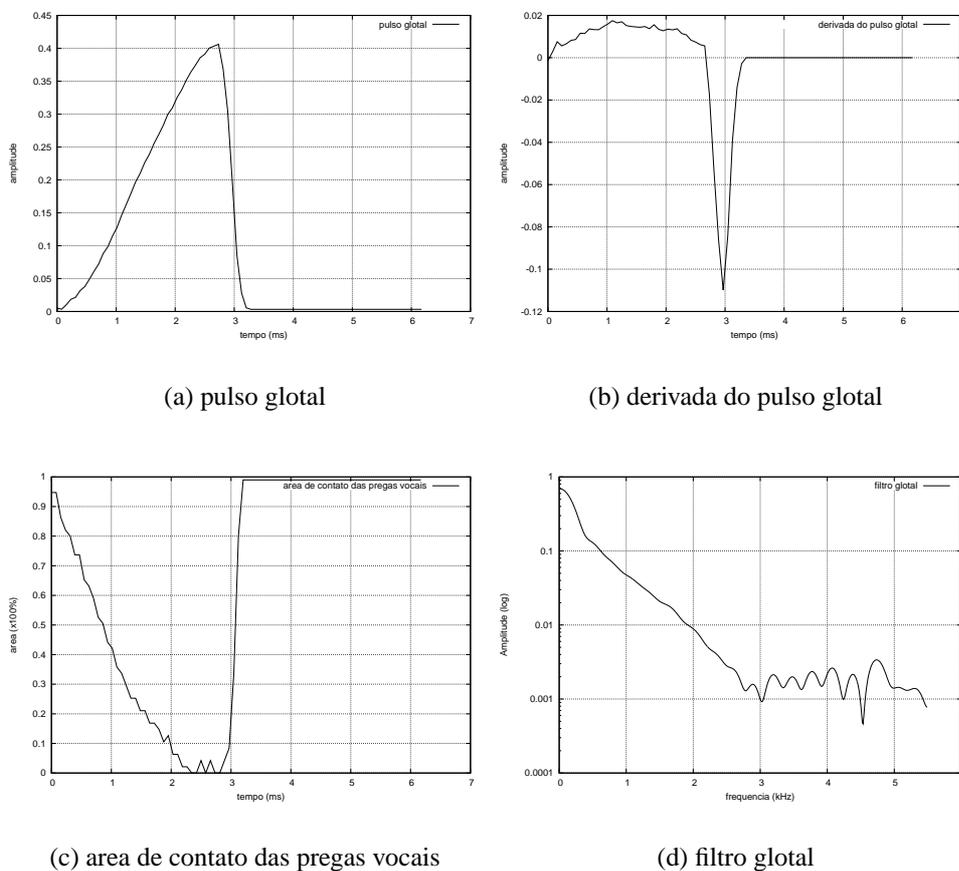


Figura 4.14: simulação numérica do caso normal com adição de massa na prega vocal esquerda.

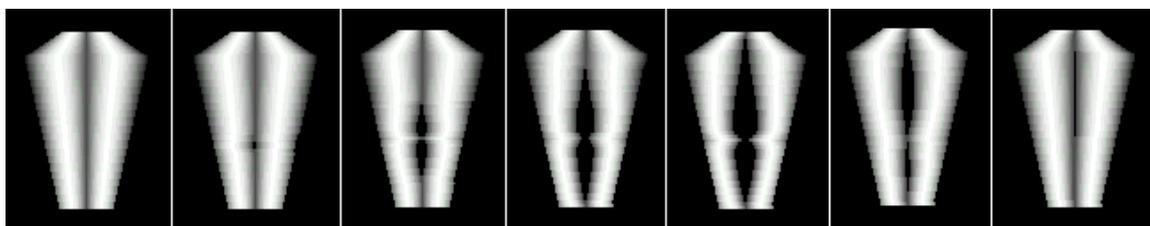


Figura 4.15: Ciclo glotal da simulação numérica com adição de massa bilateral nas pregas vocais.

O pulso glotal e sua derivada, o espectro de Fourier e a área de contato das pregas vocais para uma simulação de uma laringe normal com adição de massa bilateral nas pregas vocais são apresentados na Figura 4.16.

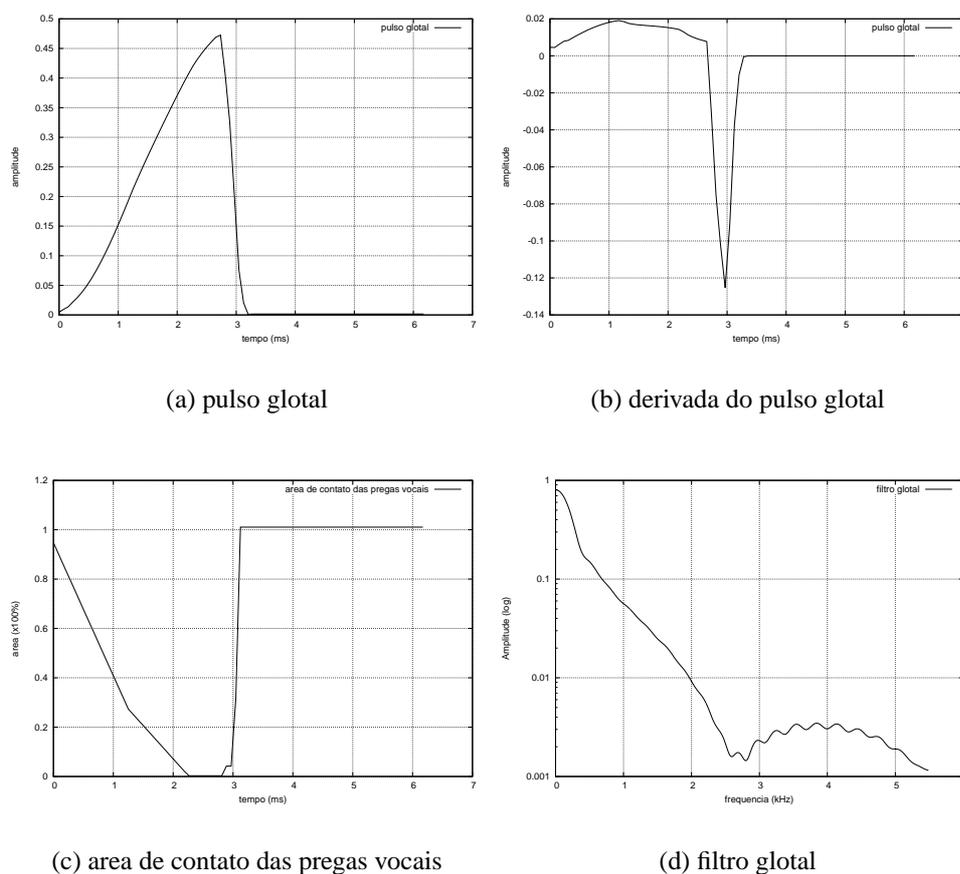


Figura 4.16: simulação numérica do caso normal com adição de massa bilateral nas pregas vocais.

### 4.1.8 Caso: Cisto

O caso real apresentado é um cisto unilateral da prega vocal direita com uma fenda triangular na comissura posterior. Na Figura 4.17 são apresentadas as duas posições onde são realizados os cortes para a elaboração das respectivas quimografias, Figuras 4.18 e 4.19. A primeira posição está localizada na região da fenda e a segunda na região da ocorrência do cisto. O exame de endoscopia e a simulação deste caso são apresentados na Figura 4.20.

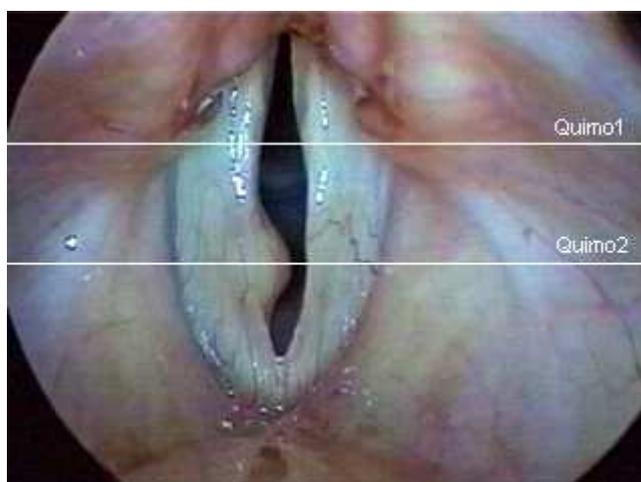


Figura 4.17: Posições dos cortes para os exames de quimografia.



Figura 4.18: Quimografia da região da fenda.



Figura 4.19: Quimografia da região do cisto.

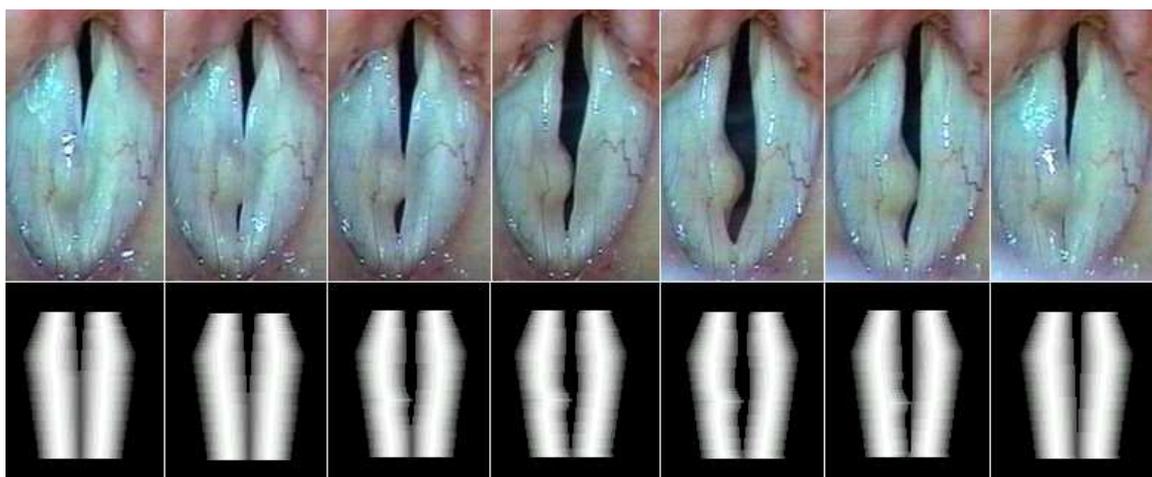


Figura 4.20: Exame de endoscopia e simulação numérica do caso com cisto.

O pulso glotal, sua derivada e transformada de Fourier juntamente com a área de contato das pregas vocais são mostrados na Figura 4.21.

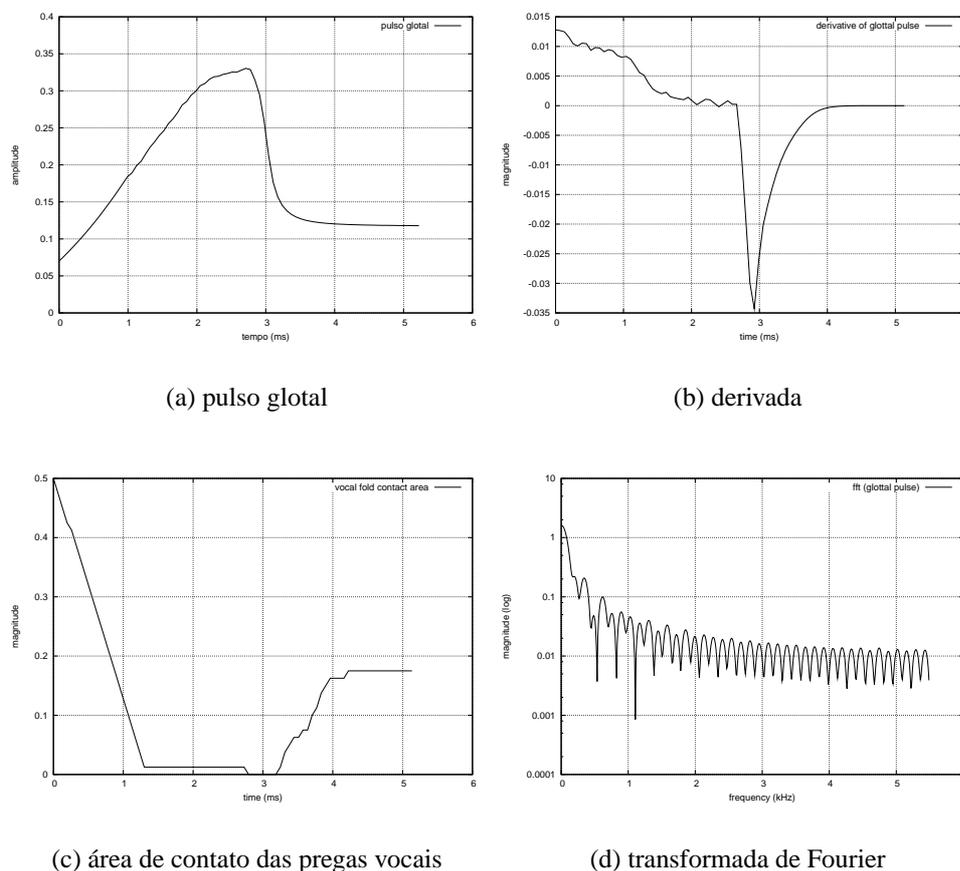


Figura 4.21: Simulação numérica do caso cisto com fenda na comissura posterior.

A cirurgia virtual é realizada com a eliminação do cisto e um melhor fechamento das aritenóides eliminando a fenda da comissura posterior. Os resultados dessas alterações são apresentados na Figura 4.22

Na Figura 4.23 são apresentados uma amostra da voz e seu respectivo espectro antes e depois da cirurgia e na Figura 4.24 são mostrados os dois espectrogramas da vogal /a/ sustentada, antes e depois da cirurgia. Nestes espectrogramas é possível observar a redução do ruído inter-harmônicos e do ruído em altas frequências e pode-se notar também um realce dos harmônicos em altas frequências.

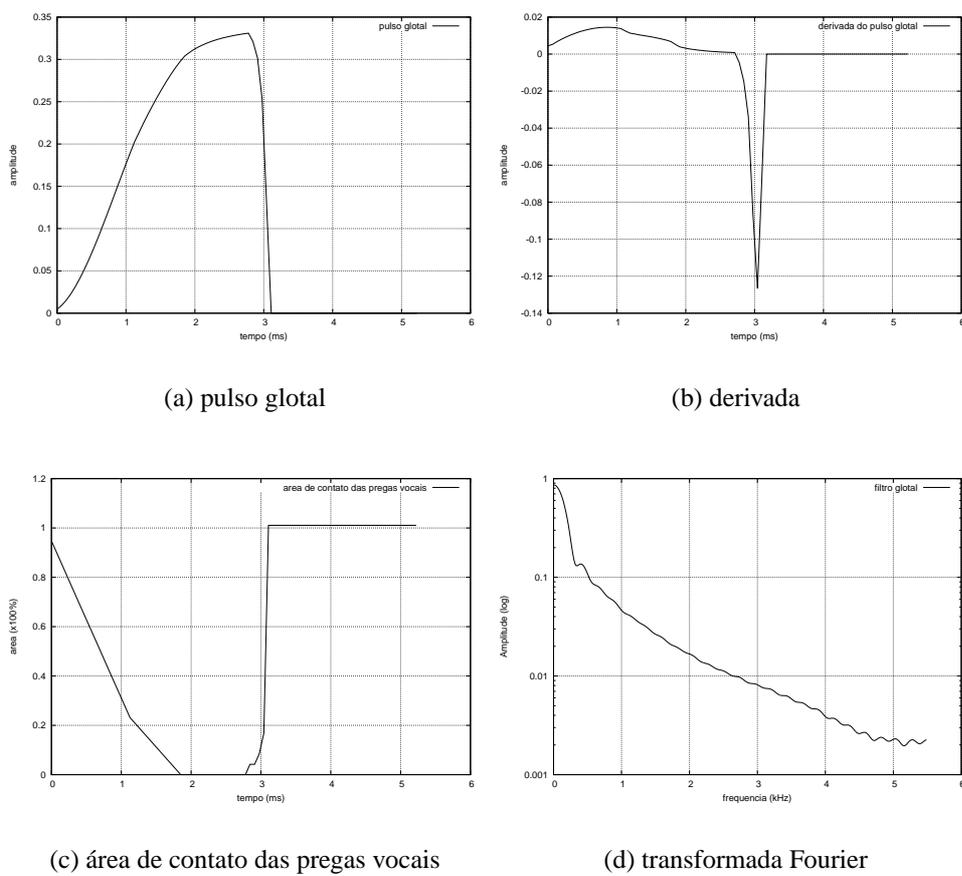
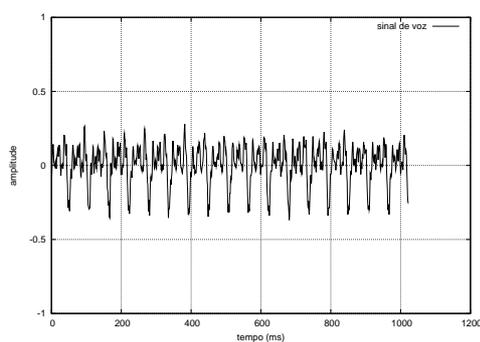
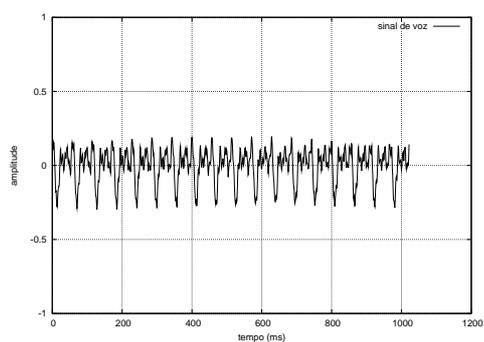


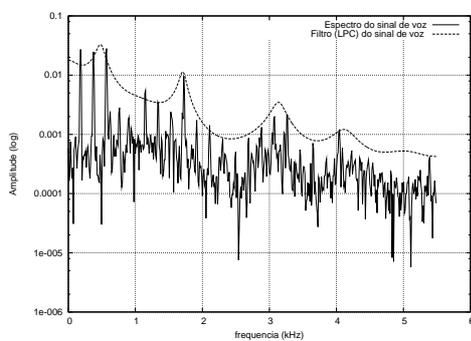
Figura 4.22: Simulação numérica da cirurgia virtual do caso cisto com fenda na comissura posterior.



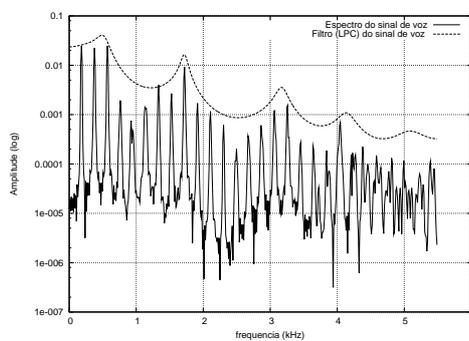
(a) Sinal de voz antes da cirurgia virtual



(b) Sinal de voz após a cirurgia virtual



(c) Espectro do sinal de voz antes da cirurgia virtual



(d) Espectro do sinal de voz após a cirurgia virtual

Figura 4.23: Sinais de vozes antes e após a cirurgia virtual do caso com cisto.

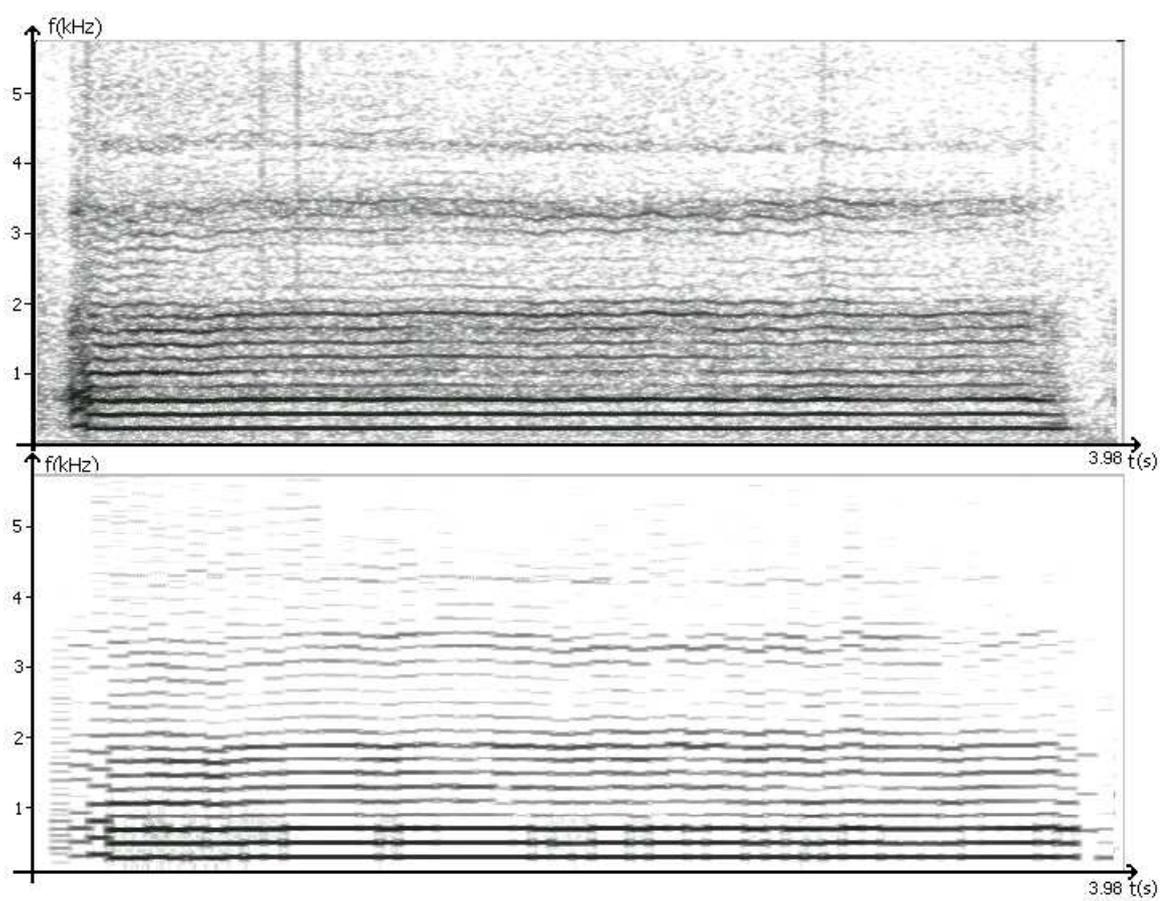


Figura 4.24: Espectrograma antes e depois da cirurgia virtual, caso cisto.

### 4.1.9 Outros Casos

Os demais casos analisados, por possuírem resultados análogos aos apresentados anteriormente, estão documentados no Apêndice A.

Na Tabela 4.3 é apresentado uma avaliação fonoaudiológica dos casos estudados pré e pós-operatório utilizando o método proposto. Foi escolhida a escala GRBAS por possibilitar uma quantificação da análise perceptiva, onde os fatores são denominados como (G) grau, (R) rouco-áspera, (B) soproso, (A) astênico e (S) tensão. Este método utiliza o sistema de pontuação: 0 = normal; 1 = leve; 2 = moderado; 3 = extremo, (ISSHIKI; TSUJI; SENNES, 1999).

A análise acústica foi realizada utilizando o programa Análise de Voz 5.0 e os resultados estão apresentados na Tabela 4.4

Tabela 4.3: Avaliação perceptiva da voz - escala GRBAS.

Paciente	sexo	G	R	B	A	S	classificação
Nódulo pré	fem.	2	2	2	0	0	$G_2R_2B_2A_0S_0$
Nódulo pós	fem.	1	1	0	0	0	$G_1R_1B_0A_0S_0$
Leocoplasia pré	mas.	2	2	0	0	0	$G_2R_2B_0A_0S_0$
Leocoplasia pós	mas.	2	2	0	0	0	$G_2R_2B_0A_0S_0$
Sulco e fenda pré	fem.	1	1	1	0	1	$G_1R_1B_1A_0S_1$
Sulco e fenda pós	fem.	0	0	0	0	1	$G_0R_0B_0A_0S_1$
Cisto pré	fem.	2	2	2	1	0	$G_2R_2B_2A_1S_0$
Cisto pós	fem.	1	0	0	1	0	$G_1R_0B_0A_1S_0$
Paralisia pré	fem.	3	1	3	1	0	$G_3R_1B_3A_1S_0$
Paralisia pós	fem.	1	0	0	1	0	$G_1R_0B_0A_1S_0$
Pólipo pré	mas.	2	2	1	0	1	$G_2R_2B_1A_0S_1$
Pólipo pós	mas.	1	0	0	0	1	$G_1R_0B_0A_0S_1$

Tabela 4.4: Análise acústica dos casos estudados.

Paciente	Sexo	Jitter	Shimmer	SFR	SFF	PA	Ex	Fo
Nódulo pré	fem.	7,02	15,67	-5,62	-14,64	0,29	5,65	153
Nódulo pós	fem.	5,23	11,82	-11,38	-17,42	0,45	5,08	153
Leocoplasia pré	masc.	0,74	9,31	-5,94	-9,03	0,35	2,18	98
Leocoplasia pós	masc.	0,41	9,33	-10,53	-11,28	0,50	2,06	98
Sulco e fenda pré	fem.	7,06	7,99	-8,10	-15,25	0,44	7,11	253
Sulco e fenda pós	fem.	4,85	4,71	-13,52	-18,27	0,52	48,19	252
Cisto pré	fem.	8,81	17,80	-6,98	-12,03	0,44	2,95	190
Cisto pós	fem.	6,15	13,26	-16,23	-16,05	0,64	1,37	190
Paralisia pré	fem.	1,69	10,15	-7,80	-7,02	0,47	1,60	211
Paralisia pós	fem.	2,75	7,37	-17,28	-11,23	0,63	1,46	211
Pólipo pré	masc.	3,34	8,09	-7,82	-17,74	0,40	14,46	165
Pólipo pós	masc.	2,01	8,05	-13,40	-18,76	0,62	4,29	165

# Capítulo 5

## Discussões

### 5.1 União das amostras de voz

A voz de uma vogal sustentada é um sinal quase periódico. Esta pequena variação durante a fonação é que nos emite uma sensação de características humanas e não de máquina. Portanto durante a síntese do sinal de voz, as características dos filtros da glote, do trato vocal e da radiação labial, assim como, a perturbação em frequência devem ser analisados a cada instante durante a fonação. Se esse tempo de amostragem for muito pequeno resulta numa transformada de Fourier pobre sem distinção de harmônicos. Se for grande demais as características momentâneas passam a ser globais recorrendo novamente no problema da sensação de voz sintetizada por máquina. Nesse trabalho chegou-se a um valor razoável do tamanho das amostras com 2.048 pontos, equivalentes a 92,8ms em uma taxa de amostragem de 22.050Hz. Porém, outro efeito indesejável acaba surgindo, que provoca um batimento entre a frequência da voz e a do tamanho das janelas, sendo que, as amostras devem ser justapostos para refazer todo o período de fonação. Desta forma, as uniões entre as janela acabam não sendo mais perfeitas, conforme mostrado na Figura 5.1, provocando um efeito de batimento audível indesejado que acaba deteriorando a qualidade do áudio.

Para solucionar esse problema foi realizada uma sobreposição do final de cada janela com o início da próxima e posteriormente através da correlação entre elas é determinado o melhor deslocamento para realizar as emendas. Este procedimento minimizou sensivelmente o efeito tornando-o quase imperceptível.

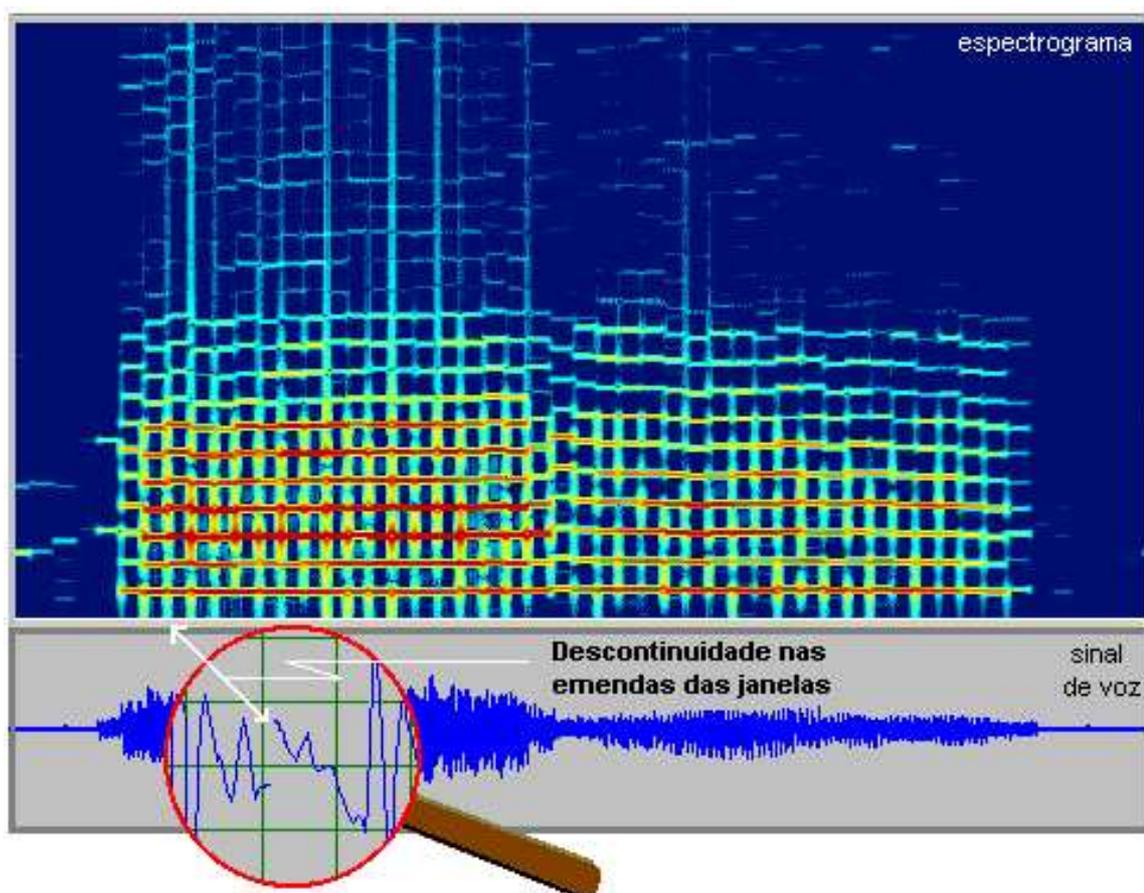


Figura 5.1: Espectrograma com descontinuidade nas uniões das amostras do sinal de voz.

## 5.2 Estroboscopia versus câmera de alta velocidade

No desenvolvimento dessa tese foram utilizados somente exames de estroboscopia os quais ainda são apenas obtidos em grandes centros devido aos custos. Já os exames com câmeras de alta velocidade são ainda mais raros, muito distantes da realidade clínica.

A estroboscopia é um processo adequado para sistemas periódicos ou que possua uma certa regularidade em seus ciclos vibratórios como nos casos das vozes normais. Para algumas vozes disfônicas os ciclos não são periódicos, portanto em exames de estroboscopia a imagem da borda das pregas vocais aparecem borradas e muitas vezes o sincronismo é perdido causando uma sequência aleatória dos quadros de vídeo.

Devido a este processo não é possível ter uma correlação direta entre o vídeo e o áudio, pois os quadros de vídeo são sub-amostrados, ou seja, durante um ciclo glótico visto pelo exame de estroboscopia vários ciclos reais foram produzidos pelas pregas vocais, dificultando a análise instantânea do momento que ocorrem as disfonias.

## 5.3 Mecanismos compensatórios

Devido ao sistema fonador não ser constituído por um único órgão, existe a possibilidade de que haja mecanismos de compensação do trato vocal à lesões apresentadas na laringe. Portanto após uma cirurgia é possível que o paciente reconfigure o trato vocal para uma melhor performance do sistema fonador ou em alguns casos a compensação inadequada é mantida causando uma disфонia mesmo após a eliminação da causa orgânica. Na cirurgia virtual não existe o mecanismo de compensação devido ao fato que a voz sintetizada mantém as características do trato vocal inalteradas.

## 5.4 Determinação da frequência fundamental

Para a simulação da fonte glotal é necessário a produção da excitação que são impulsos com período proporcional à frequência fundamental da voz. Para produzir as nuances da voz humana optou-se por quadros de tamanho de 2048, pois apresentaram empiricamente uma melhor relação entre a qualidade da voz e a resolução do espectro. Contudo, em algumas vozes disfônicas, a determinação da frequência fundamental utilizando o resíduo falha em alguns trechos do sinal, às vezes impossibilitando a análise auditiva de todo o

---

trecho continuamente. Nestes casos os trechos com melhor desempenho foram editados e gravados em novos arquivos para a análise fonoaudiológica.

## 5.5 Análise perceptiva

Os resultados da análise perceptiva realizados pela fonoaudióloga utilizando a escala GR-BAS corroboram a análise acústica das vozes antes e depois da cirurgia virtual. No caso da leocoplasia houve uma descontinuidade no cálculo da frequência fundamental, limitando o tamanho dos trechos de vozes o que impossibilitou a análise auditiva. Nos outros casos analisados, todos apresentaram sensível redução nos parâmetros grau (G), rouquidão (R) e sopro (B) e não houveram alterações nos parâmetros astenia (A) e força (S), isso se deve ao fato das alterações serem realizadas com o intuito de remover apenas as anomalias nas pregas vocais e glote principalmente reduzindo ou eliminando as fendas glóticas as quais são uma das principais causas da sopro.

## 5.6 Fechamento imperfeito da glote

Segundo as observações de Isshiki, Tsuji e Sennes (1999) a fonação normal ocorre somente sob certas condições de alguns parâmetros e entre eles temos a área glótica inicial ( $A_{g0}$ ), pressão subglótica, rigidez das pregas vocais, coaptação entre a parte superior e inferior das bordas das pregas vocais ou fisiologicamente, a mobilidade da prega vocal e massa da prega vocal. Estas observações confirmam os resultados obtidos principalmente em relação a área da glote inicial e a área da glote remanescente durante a fonação. Praticamente todas as patologias que levaram a formação de uma fenda na glote ocasionam um contínuo escape de ar modulado pela vibração das pregas vocais. Esse acréscimo no nível contínuo ( $DC$ ) do pulso glotal proporciona um significativo aumento no ruído em altas frequências, as quais podem ser observadas nas análises espectrais. O caminho inverso, no caso da cirurgia virtual, a eliminação das fendas levaram a vozes sintetizadas com consideráveis redução da sopro.

# Capítulo 6

## Conclusões

Neste trabalho foi comprovada a possibilidade da utilização de contornos ativos para a elaboração de um modelo da glote os quais produzem resultados muito similares aos observados por exame de estroboscopia.

Os pulsos glotais são estimados através da área da glote e seu respectivo filtro aplicado inversamente antes da cirurgia e diretamente após a cirurgia virtual resultam em um sinal de voz sintetizado com aumento significativo na qualidade vocal.

Titze (1989), Rothenberg (1979) apresentaram em seus trabalhos o parâmetro chamado de áreas de contato das pregas vocais, *VFCA*, as quais são somente simuladas corretamente se o efeito denominado 'ziper' for considerado na elaboração do modelo, portanto, os resultados obtidos como a área de contato das pregas vocais, velocidade volumétrica glotal e sua respectiva derivada são condizentes aos resultados observados em seus trabalhos.

O modelo proposto para a síntese da voz mantém as características do modelo do trato vocal e da radiação, levando a obtenção de uma voz sintetizada com as mesmas nuances da voz original, as quais identificam o fonador. Este fato é perceptível auditivamente em todos os casos que não houveram descontinuidade na determinação da frequência fundamental.

Foi demonstrado que em situações em que a laringe se encontra sem anomalias o acréscimo de massas e fendas nas pregas vocais produzem resultados acústicos seme-

---

lhantes aos reais.

As vozes sintetizadas após a cirurgia virtual foram analisadas quantitativamente pelo programa Análise de Voz 5.0, (MONTAGNOLI; PEREIRA, 2005) e qualitativamente por uma fonoaudióloga especializada que utilizou a escala GRBAS, (ISSHIKI; TSUJI; SENNES, 1999). Os resultados obtidos foram similares, evidenciando uma melhora significativa nos parâmetros acústicos, assim como nos valores da escala GRBAS.

## 6.1 Trabalhos futuros

Em trabalhos futuros pode-se utilizar algoritmos mais eficientes na determinação do período fundamental das vozes disfônicas os quais eliminariam as discontinuidades durante a síntese de algumas vozes.

A utilização de exames com referências métricas reais como em casos de laringes excisadas levaria a fatores de escala que melhor correlacionariam os parâmetros obtidos com valores reais, ou seja, as dimensões utilizadas no modelo tem como referência o comprimento vertical da glote, por não haver referências métricas reais na imagem, o tamanho das estruturas na laringe depende da distância da fibra e do ganho óptico (*zoom*) da câmera.

A qualidade do exame de quimografia é dependente do exame de estroboscopia, os quais muitas vezes são realizados sem esta preocupação, causando excessivo movimento da fibra durante o exame. Algoritmos para remover estes movimentos e alinhar verticalmente a imagem da glote, juntamente com algoritmos de contornos ativos aplicados à segmentação das pregas vocais, sugerem uma automatização de todo o processo de modelagem.

A capacidade do corpo humano em realizar controles através de realimentações é extremamente eficiente em alguns casos, como no caso da fonação, muitas vezes o processo é involuntário, acrescentando uma quantidade de variáveis difíceis de serem isoladas. Estudos com laringes excisadas podem dirimir dúvidas como por exemplo se quando há uma área inicial na glote e o seu posterior fechamento são causas da realimentação do sistema ou são efeitos dinâmicos das pregas vocais.

A aplicação de modelos deformáveis em três dimensões possibilitará a realização de uma laringe completa permitindo alterações em toda a caixa laríngea inclusive no arcabouço laríngeo, cirurgias as quais são chamadas de tireoplastias. Complementando este processo com técnicas de realidade virtual será possível executar as cirurgias em tempo real através da emulação de instrumental.

## Referências Bibliográficas

ALIPOUR-HAGHIGHI, F. Pressure and velocity profiles in a static mechanical hemilarynx model. *Journal Acoustical Society of America*, v. 112, n. 6, p. 2996–3003, Dec. 2002.

ALIPOUR-HAGHIGHI, F.; BERRY, D. A.; TITZE, I. R. A finite-element model of vocal-fold vibration. *Journal Acoustical Society of America*, v. 108, n. 6, p. 3003–3012, Dec. 2000.

ALIPOUR-HAGHIGHI, F.; SCHERER, R. C. Flow separation in a computational oscillating vocal fold model. *Journal Acoustical Society of America*, v. 116, n. 3, p. 1710–1719, Sep. 2004.

ALIPOUR-HAGHIGHI, F.; TITZE, I. R. A finite element simulation of vocal folds vibrations. In: *Fourteenth Annual Northeast IEEE Bioengineering Conference* Durham, NH: , 1988. p. 186–189.

ALIPOUR-HAGHIGHI, F.; TITZE, I. R. Elastic model of vocal fold tissues. *Journal Acoustical Society of America*, v. 90, n. 3, p. 1326–1331, Sep. 1991.

ANANTHAPADMANABHA, T. V.; YEGNANARAYANA, B. Epoch extraction from linear prediction residual for identification of closed glottis interval. *IEEE Transaction on Acoustics, Speech and Signal Processing*, v. 27, n. 4, p. 309–318, Aug. 1979.

BANKMAN, I. N. Handbook of medical imaging, processing and analysis. *Academic Press*, p. 901, 2000.

BEHLAU, M. *Voz: O livro do especialista*. Rio de Janeiro, Brazil: Revinter, 2001.

BEHLAU, M.; PONTES, P. *Avaliação e Tratamento das disfonias*. São Paulo, Brazil: Lovise, 1995.

---

BERRY, D. A.; HERZEL, H.; TITZE, I. R.; KRISCHER, K. Interpretation of biomechanical simulations of normal and chaotic vocal fold oscillations with empirical eigenfunctions. *Journal Acoustical Society of America*, v. 95, n. 6, p. 3595–3604, Jun. 1994.

BOYANOV, B.; HADJITODOROV, S. Acoustical analysis of pathological voices. *IEEE Engineering in Medicine and Biology*, v. 97, p. 74–81, Jul. 1997.

CATALDO, E.; LETA, F. R.; LUCERO, J.; NICOLATO, L. Voiced sounds synthesis using mechanical models. In: *International Workshop on Systems, Signals and Image Processing* Poznan, Poland: , 2004. p. 4.

CHAN, R. W. Measurements of vocal fold tissue viscoelasticity: Approaching the male phonatory frequency range. *Journal Acoustical Society of America*, v. 115, n. 6, p. 3461–3170, Jun. 2004.

CHILDERS, D. G. Speech processing and synthesis for assessing vocal disorders. *IEEE Engineering in Medicine and Biology Magazine*, p. 69–71, mar. 1990.

CHILDERS, D. G.; AHN, C. Modeling the glottal volume-velocity waveform for three voice types. *Journal Acoustical Society of America*, v. 97, n. 1, p. 505–519, Jan. 1995.

CHILDERS, D. G.; HU, H. T. Speech synthesis by glottal excited linear prediction. *Journal Acoustical Society of America*, v. 96, n. 4, p. 2026–2036, Oct. 1994.

CHILDERS, D. G.; LEE, C. K. Vocal quality factors: Analysis synthesis, and perception. *Journal Acoustical Society of America*, v. 90, n. 5, p. 2394–2410, Nov. 1991.

CHILDERS, D. G.; WU, K. Gender recognition from speech. part ii: Fine analysis. *Journal Acoustical Society of America*, v. 90, n. 4, p. 1841–1856, Oct. 1991.

COHEN, L. D.; COHEN, I. Finite-element method for active contour models and balloons for 2d and 3d images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1993.

DAVIS, S. B. *Acoustical characteristics of normal and pathological voices*. S. Barbara: Academic Press, 1976.

DÖLLINGER, M.; HOPPE, U.; HETTLICH, F.; LOHSCHELLER, J.; SCHUBERTH, S.; EYSHOLDT, U. Vibration parameter extraction from endoscopic image series of the vocal folds. *IEEE Transaction on Biomedical Engineering*, v. 49, n. 8, p. 773–781, Aug. 2002.

---

FANT, G. *Acoustic theory of speech production*. Gravenhage, Sweden: Moreton & Co, 1960.

FANT, G. The lf-model revisited. transformation and frequency domain analysis. *STL-Quarterly Progress and Status Report*, p. 119–156, 1995.

FANT, G.; LILJENCRANTS, J.; LIN, Q. A four-parameter model of glottal flow. *STL-Quarterly Progress and Status Report*, p. 1–13, 1985.

FLANAGAN, J. L. *Speech analysis, synthesis, and perception*. Berlin: Springer-Verlag, 1972.

FLANAGAN, J. L.; CHERRY, L. Excitation of vocal-tract synthesizers. *Journal Acoustical Society of America*, v. 45, n. 3, p. 764–769, Nov. 1969.

FLANAGAN, J. L.; LANDGRAF, L. Self-oscilating source for vocal-tract synthesizers. *IEEE Transactions on Audio and Electroacoustics*, v. 16, n. 1, p. 57–64, Mar. 1968.

FRÖHLICH, M.; MICHAELIS, D.; STRUBE, H. W. Sim - simultaneous inverse filtering and matching of a glottal flow model for acoustic speech signals. *Journal Acoustical Society of America*, v. 110, n. 1, p. 479–488, Jul. 2001.

GERRATT, B. R.; KREIMAN, J. Theoretical and methodological development in the study of pathological voice quality. *Journal of Phonetics*, v. 28, p. 335–342, 2000.

GÓMES, P.; GODINO, J. I.; RODRIGUES, F.; DIAZ, F.; NIETO, V.; ÁLVAREZ, A.; RODELLAR, V. Evidence of vocal cord pathology from the mucosal wave cepstral contents. In: *IEEE - ICASSP*, 2004.

GONZALES, R. C.; WOODS, R. E. *Digital image processing*. Reading, MA: Addison-Wesley, 1992.

GUERRA, A. C. *Estimação do sinal glotal para padrões acústicos de doenças da laringe*. Tese (Doutorado) — EESC-USP, São Carlos, 2005.

GUNTER, H. E. A mechanical model of vocal-fold collision with high spatial and temporal resolution. *Journal Acoustical Society of America*, v. 113, n. 2, p. 994–1000, Feb. 2003.

HERZEL, H. Bifurcations and chaos in voice signals. *Journal of Voice*, v. 7, n. 1, 1993.

---

HIRANO, M. Morphological structure of the vocal cord as a vibrator and its variations. *Folia Phoniatrica*, v. 26, p. 89–94, 1974.

HIRANO, M.; KAKITA, Y. Cover-body theory of vocal cord vibration. In: *Speech Science* San Diego: , 1985. p. 1–46.

HIRANO, M.; SATO, K. *Histological color atlas of the human larynx*. San Diego, CA: Singular Publishing Group, 1993.

HOLZRICHTER, J. F.; NG, L. C.; BURKE, J.; CHAMPAGNE, N. J.; KALLMAN, J. S.; SHARPE, R. M.; KOBLE, J. B.; HILLMAN, R. E.; ROSOWKI, J. J. Measurements of glottal structure dynamics. *Journal Acoustical Society of America*, v. 117, n. 3, p. 1373–1385, Mar. 2005.

ISHIZAKA, K.; FLANAGAN, J. Synthesis of voiced sounds from two-mass model of the vocal cords. *Bell System Technical Journal*, v. 51, p. 1233–1268, 1972.

ISSHIKI, N.; TSUJI, D. H.; SENNES, L. U. *Tireoplastia*. São Paulo: Bios Comunicação e Editora, 1999.

JAIN, A. K.; DUIN, R. P. W.; MAO, J. Statistical pattern recognition: A review. *IEEE Trans. Pattern Analysis and Machine Intelligence*, v. 22, n. 1, p. 4–37, Jan. 2000.

JIANG, J. J.; ZHANG, Y. Modeling of chaotic vibrations in symmetric vocal folds. *Journal Acoustical Society of America*, v. 11, n. 4, p. 2120–2128, Oct. 2001.

KASS, M.; WITKIN, A.; TERZOPOULOS, D. Snakes: Active contour models. *International Journal of Computer Vision*, v. 1, n. 4, p. 321–331, 1988.

KOIKE, Y.; MARKEL, J. D. Application of inverse filtering for detecting laryngeal pathology. In: *Annals of Otology, Rhinology, & Laryngology*, 1975. p. 117–124.

LAMAR, M. D.; QI, Y.; XIN, J. Modeling vocal fold motion with a hydrodynamic semi-continuum model. *Journal Acoustical Society of America*, v. 114, n. 1, p. 445–464, Jul. 2003.

LEE, J. S.; KIM, E.; SUNG, M. W.; KIM, K. H.; SUNG, M. W.; KIM, K. H.; SUNG, M. Y.; PARK, K. A method for assessing the regional pattern of vocal folds by analysing the video recording of stroboscopy. *Medical & Biological Engineering & Computing*, v. 39, 2001.

---

LUCERO, J. C.; KOENIG, L. L. Simulations of temporal patterns of oral airflow in men and women using a two-mass model of the vocal folds under dynamic control. *Journal Acoustical Society of America*, v. 117, n. 3, p. 1362–13372, Mar. 2005.

MACIEL, C. D.; PEREIRA, J. C.; DAJER, M. E. Análise exploratória de sinais de voz para discriminação de patologias. In: *4º Congresso Temático de Dinâmica, Controle e Aplicações* Bauru, Brasil: , 2005.

MANFREDI, C.; D' ANIELLO, M.; BRUSCAGLIONI, P.; ISMAELLI, A. Comparative analysis of fundamental frequency estimation methods with application to pathological voices. *Medical Engineering & Physics*, v. 20, p. 135–147, 2000.

MARKEL, J. D.; GRAY, A. H. *Linear prediction of speech*. New York, USA: Springer, 1976.

MARTINEZ, C. E.; RUFINER, H. L. Acoustic analysis of speech for detection of laryngeal pathologies. In: *22nd annual EMBS Conference, IEEE*. Chicago, IL: IEEE, 2000. p. 23–28.

MARTINS, R. H. G. *A Voz e seus distúrbios*. Botucatu, SP: Cultura Acadêmica, 2005.

MASTERS, T. *Signal and image processing with neural networks: a C ++ sourcebook*. New York, NY: John Wiley & Sons, 1994.

MITEV, P.; HADJITODOROV, S. Fundamental frequency estimation of vice of patients with laryngeal disorders. *Information Sciences*, v. 156, p. 3–19, 2003.

MONTAGNOLI, A. N. *Análise Residual do Sinal de Voz*. Dissertação (Mestrado) — EESC-USP, S. Carlos, 1998.

MONTAGNOLI, A. N. An approach to build phase spaces by assigning the time delay. In: *Workshop on Nonlinear Phenomena: Modeling and Their Applications*. Rio Claro: ICONNE Series, 2005.

MONTAGNOLI, A. N.; FONSECA, E. S. Larynx virtual surgery. In: *IEEE International Symposium on Multimedia*. California-USA: In press, 2005.

MONTAGNOLI, A. N.; PEREIRA, J. C. Controle por voz utilizando um classificador de máxima verossimilhança aproximado em microcontroladores de 8-bits. In: *3o Congresso*

---

*Temático de Aplicações de Dinâmica e Controle da Sociedade Brasileira de Matemática Aplicada e Computacional* Ilha Solteira: , 2004.

MONTAGNOLI, A. N.; PEREIRA, J. C. *Análise de Voz 5.0*. Software: sem registro, 2005.

MONTAGNOLI, A. N.; PEREIRA, J. C.; FERRARI, R. J. Uma nova aproximação para o classificador bayesiano. In: *2o Congresso Temático de Aplicações de Dinâmica e Controle da Sociedade Brasileira de Matemática Aplicada e Computacional*, 2003. p. 2877–2886.

NIEDZIELSKA, G.; GLIJER, A.; NIEDZIELSKI, A. Acoustic analysis of voice in children with noduli vocales. *Journal of Pediatric Otorhinolaryngology*, v. 60, p. 119–122, 2001.

OPPENHEIM, A. V.; SCHAFER, R. W. *Digital Signal Processing*. Englewood Cliffs, N.J.: Prentice-Hall, 1975.

PARKER, J. R. *Algorithms for image processing and computer vision*. New York, NY: John Wiley & Sons, 1996.

PARSA, V.; JAMIESON, D. G. Identification of pathological voices using glottal noise measures. *Journal of Speech, Language, and Hearing Research*, v. 43, p. 469–484, Apr. 2000.

PEREIRA, J. C.; MONTAGNOLI, A. N. Análise acústica da voz. In: *4o Congresso Temático de Aplicações de Dinâmica e Controle da Sociedade Brasileira de Matemática Aplicada e Computacional* Bauru: , 2005.

RABINER, L. R.; SCHAFER, R. W. *Digital processing of speech signals*. Englewood Cliffs, N. J.: Prentice-Hall, 1978.

ROSA, M. O.; PEREIRA, J. C.; GRELLET, M.; ALWAN, A. A contribution to simulating a three-dimensional larynx model using the finite element method. *Journal Acoustical Society of America*, v. 114, n. 5, p. 2893–2905, Nov. 2003.

ROTHENBERG, M. Some relations between glottal air flow and vocal fold contact area. In: *Assessment of Voal Pathology*. ASHA-Reports 11: National Institutes of Health, 1979. p. 88–96.

---

ROTHENBERG, M. Acoustic interaction between the glottal source and vocal tract. In: *Vocal Fold Physiology*. University of Tokyo: Sterens, K. N. Hirano, M., 1981. p. 305–323.

RUSS, J. C. *The image processing handbook*. Boca Raton, FL: CRC Press, 1994.

SENNES, L. U.; TSUJI, D. H.; BODAHARA, S.; BENTO, R. F.; RIBAS, C. G. O uso de imagens tridimensionais no ensino da anatomia da laringe. *Arquivos da Fundação Otorrinolaringologia*, v. 4, n. 3, 2000.

SETHIAN, J. A. Tracking interfaces with level sets. *American Scientist*, p. 254–263, May 1997.

SETHIAN, J. A. *Level set Methods and Fast Marching Methods*. New York, NY: Cambridge, 2002.

SMITH, S. W. *The Scientist and Engineer's Guide to Digital Signal Processing*. California, USA: California Technica Publishing, 1999.

SONDHI, M. M. Estimation of vocal-tract areas: The need for acoustical measurements. *IEEE Transaction on Acoustics, Speech and Signal Processing*, v. 27, n. 3, p. 268–273, Jun. 1979.

STEARNS, S. D.; DAVID, R. A. *Signal processing algorithms*. Prentice-Hall, 1988.

STEINECKE, I.; HERZEL, H. Bifurcations in an asymmetric vocal-fold model. *Journal Acoustical Society of America*, v. 97, n. 3, p. 1874–1884, Mar. 1995.

STORY, B. H.; TITZE, I. R. Voice simulation with a body-cover model of the vocal folds. *Journal Acoustical Society of America*, v. 97, n. 2, p. 1249–1260, 1995.

TITZE, I. R. The human vocal cords: a mathematical model, part 1. *Phonetica*, v. 28, p. 129–170, 1973.

TITZE, I. R. The human vocal cords: a mathematical model, part 2. *Phonetica*, v. 29, p. 1–21, 1974.

TITZE, I. R. On the mechanics of vocal-fold vibration. *Journal Acoustical Society of America*, v. 60, n. 6, p. 1366–1380, Dec. 1976.

---

TITZE, I. R. Parametrization of the glottal area, glottal flow, and vocal fold contact area. *Journal Acoustical Society of America*, v. 75, n. 2, p. 570–580, Feb. 1984.

TITZE, I. R. The physics of small-amplitude oscillation of the vocal folds. *Journal Acoustical Society of America*, v. 83, n. 4, p. 1536–1552, Apr. 1988.

TITZE, I. R. A four-parameter model of the glottis and vocal fold contact area. *Speech Communication*, v. 8, p. 191–201, 1989.

TITZE, I. R. Regulating glottal airflow in phonation: Application of the maximum power transfer theorem to a low dimensional phonation model. *Journal Acoustical Society of America*, v. 111, n. 1, p. 367–376, Jan. 2002.

TITZE, I. R.; BROADHEAD, K.; GRAY, T. S. Strain distribution in an elastic substrate vibrated in a bioreactor for vocal fold tissue engineering. *Journal of Biomechanics*, 2005.

TITZE, I. R.; HITCHCOCK, R. W.; BROADHEAD, K.; WEBB, K.; LI, W.; GRAY, S. D.; TRESKO, P. A. Design and validation of a bioreactor for engineering vocal fold tissues under combined tensile and vibrational stresses. *Journal of Biomechanics*, v. 37, p. 1521–1529, Jan. 2004b.

TITZE, I. R.; HUNTER, E. J. Normal vibration frequencies of vocal ligament. *Journal Acoustic Society of America*, v. 115, n. 5, p. 2264–2269, May 2004.

TITZE, I. R.; STRONG, W. J. Normal modes in vocal cords tissues. *Journal Acoustical Society of America*, v. 57, n. 3, p. 736–744, Mar. 1975.

TITZE, I. R.; TALKIN, D. T. A theoretical study of effects of various laryngeal configurations on the acoustics of phonation. *Journal Acoustical Society of America*, v. 66, n. 1, p. 60–74, Jul. 1979.

VELDHUIS, R. A computationally efficient alternative for the Liljencrants-fant model and perceptual evaluation. *Journal Acoustical Society of America*, v. 103, n. 1, p. 566–571, Jan. 1988.

WAKITA, H. Estimation of vocal-tract shapes from acoustical analysis of the speech wave: The state of the art. *IEEE Transaction on Acoustics, Speech and Signal Processing*, v. 27, n. 3, p. 281–285, 1979.

WONG, D. Y.; MARKEL, J.; GRAY, A. H. Least square glottal inverse filtering from the acoustic speech waveform. *IEEE Transaction on Acoustics, Speech and Signal Processing*, v. 27, n. 4, p. 350–355, Aug. 1979.

XU, C.; JR., A. Y.; PRINCE, J. L. On the relationship between parametric and geometric contours. In: *34th Asilomar Conference on Signals, Systems, and Computers*. sei la: xuxu, 2000. p. 438–439.

ZHANG, H.; PAYANDEH, S.; DILL, J. On cutting and dissection of virtual deformable objects. *International Conference on Robotics & Automation*, p. 3908–3913, Apr. 2004.

# Capítulo A

## Apêndice

### A.1 Estudo de casos

#### A.1.1 Caso: Nódulo bilateral

O caso real apresentado é um nódulo bilateral das pregas vocais. O exame de endoscopia e a simulação numérica deste caso são apresentados na Figura A.1.

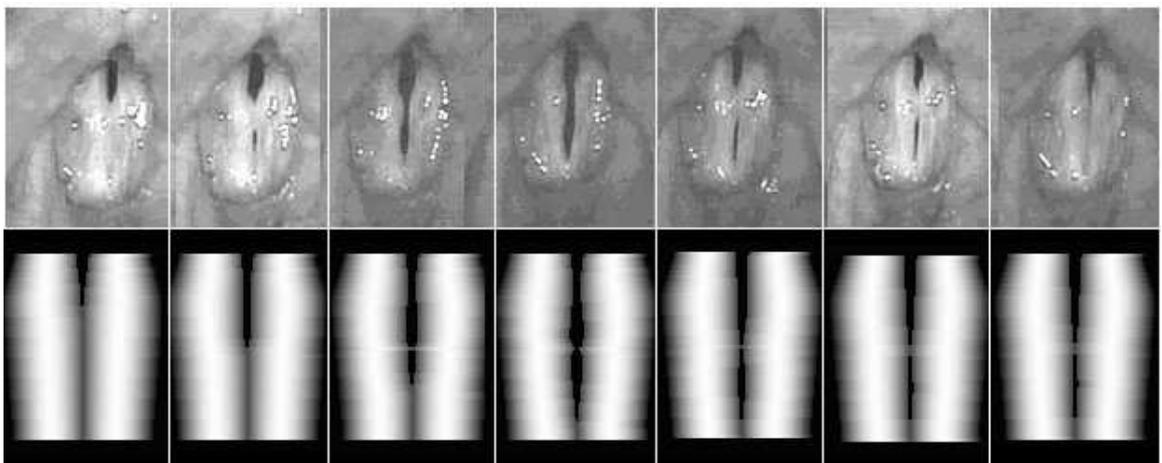
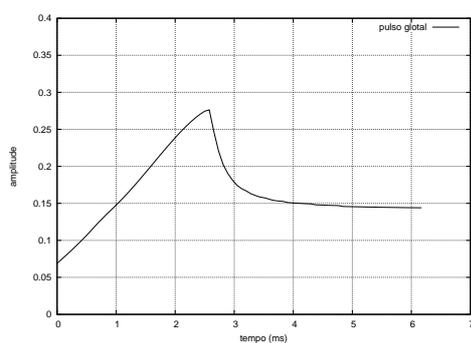


Figura A.1: Exame de endoscopia e simulação numérica do caso com nódulo bilateral.

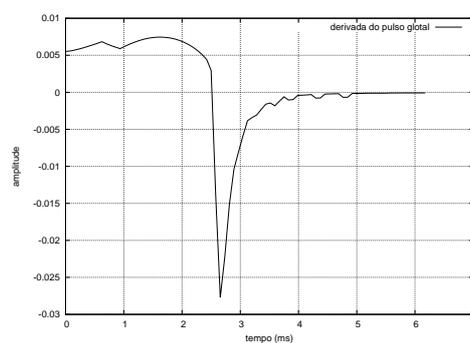
O pulso glotal, sua derivada e transformada de Fourier juntamente com a área de contato das pregas vocais são mostrados na Figura A.2.

A cirurgia virtual é realizada com a eliminação dos nódulos, apresentados na Figura A.3

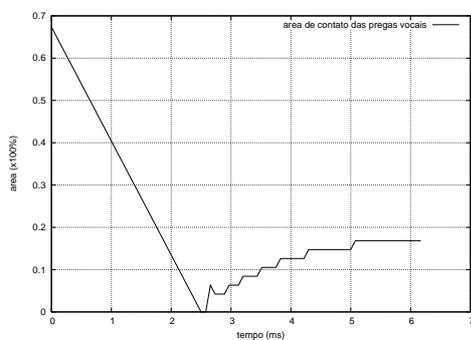
Na Figura A.4 são apresentados uma amostra da voz e seu respectivo espectro antes



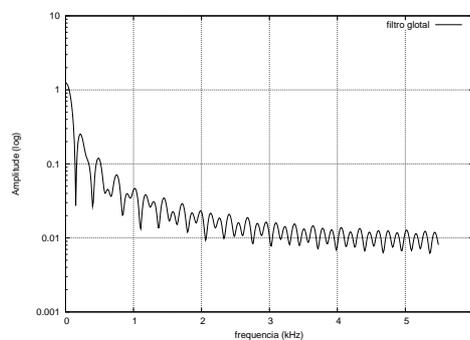
(a) Pulso glotal



(b) Derivada do pulso glotal

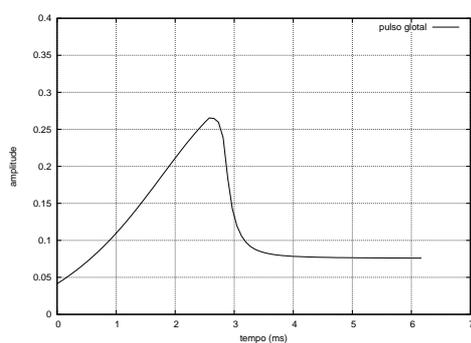


(c) área de contato das pregas vocais

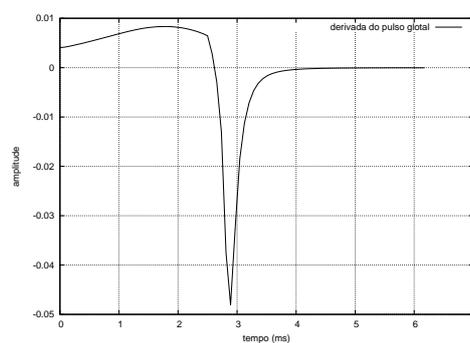


(d) Espectro do filtro glotal

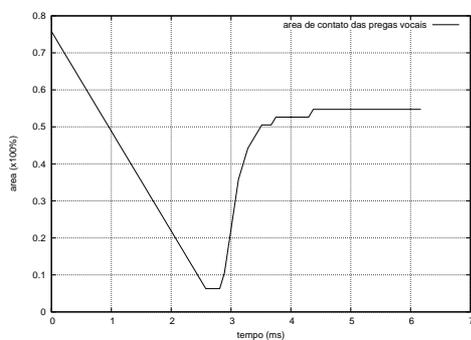
Figura A.2: Simulação numérica do caso nódulo bilateral.



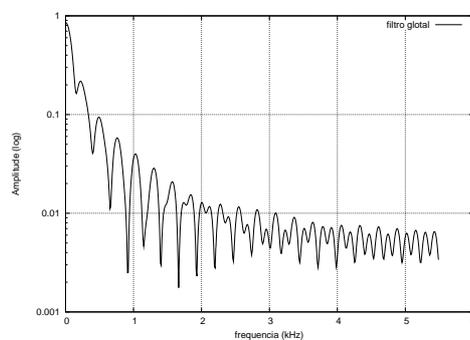
(a) Pulso glotal



(b) Derivada do pulso glotal



(c) área de contato das pregas vocais



(d) Espectro do filtro glotal

Figura A.3: Simulação numérica da cirurgia virtual do caso nódulo bilateral.

e depois da cirurgia e na Figura A.5 são mostrados os dois espectrogramas da vogal /a/ sustentada, antes e depois da cirurgia. Nestes espectrogramas é possível observar a redução do ruído inter-harmônicos e do ruído em altas frequências, pode-se notar também um realce dos harmônicos em altas frequências.

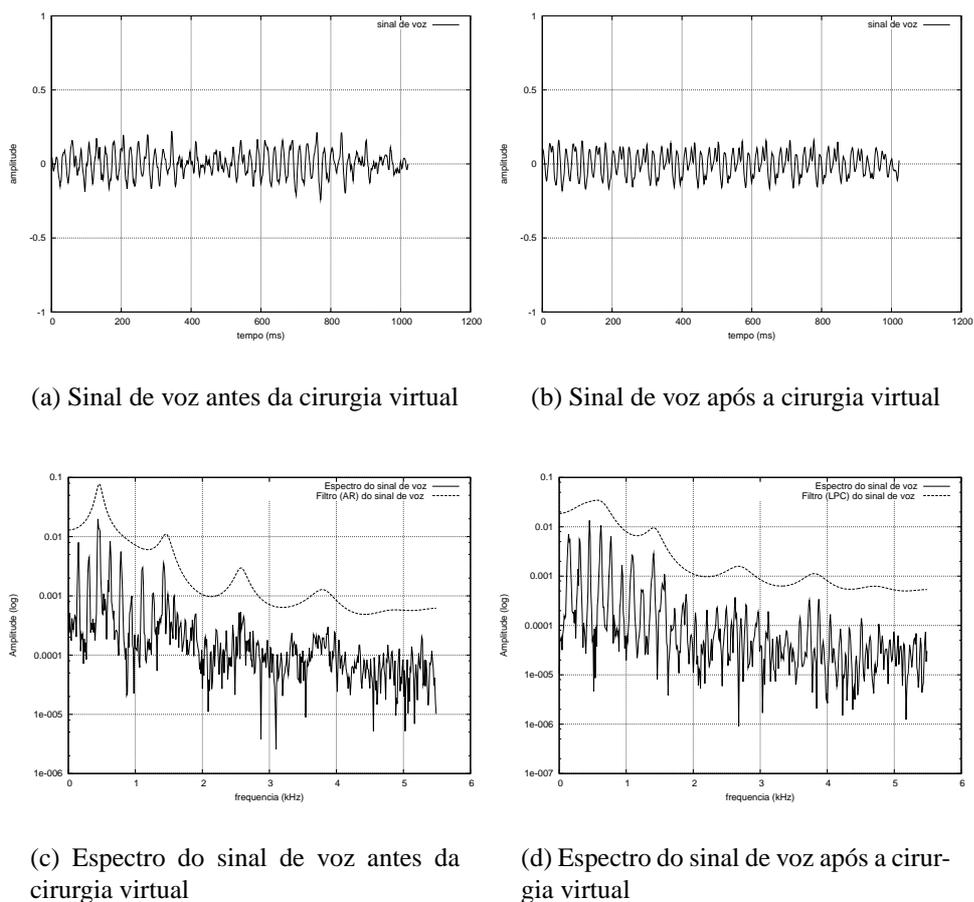


Figura A.4: Sinais de vozes antes e após a cirurgia virtual, caso nódulo.

## A.1.2 Caso: Leocoplasmia

O caso real apresentado é uma leocoplasmia das pregas vocais. O exame de endoscopia e a simulação numérica deste caso são apresentados na Figura A.6.

O pulso glotal, sua derivada e transformada de Fourier juntamente com a área de contato das pregas vocais são mostrados na Figura A.7.

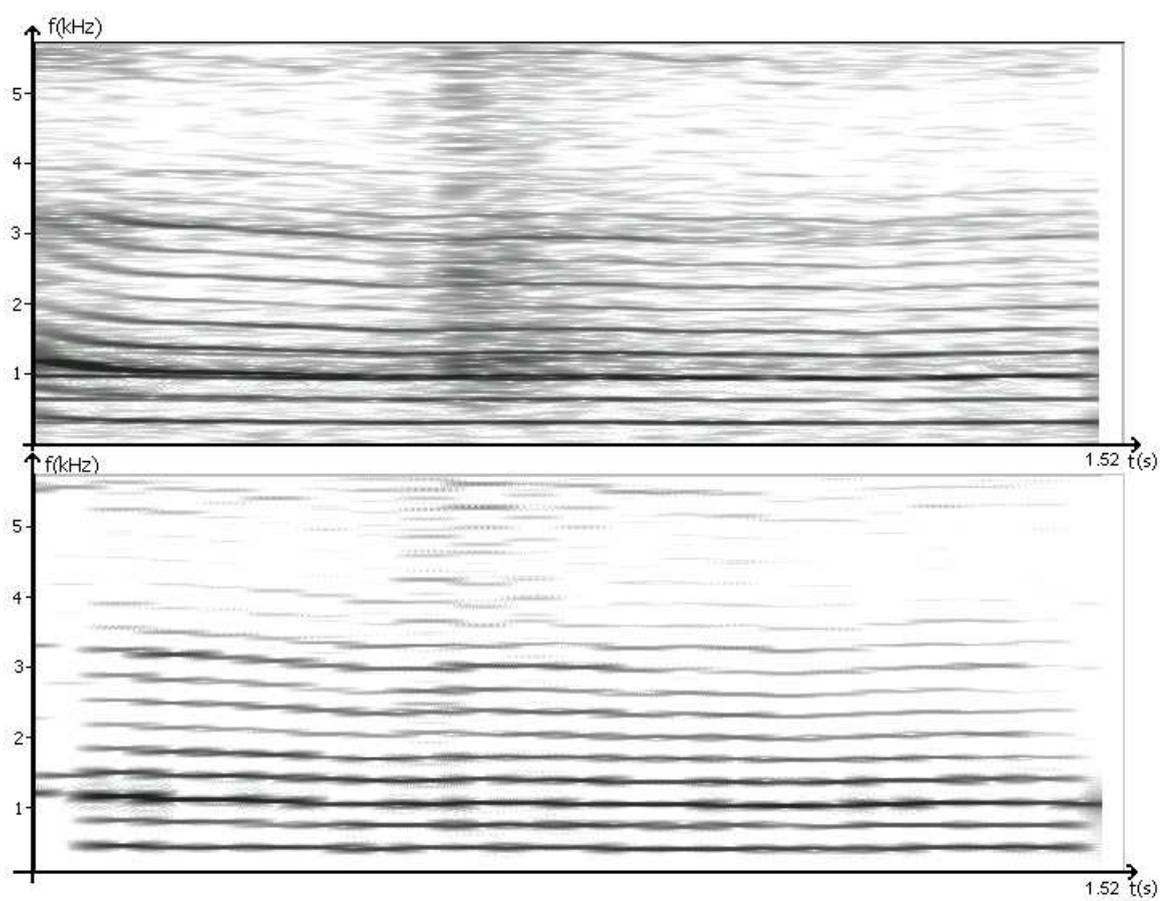


Figura A.5: Espectrograma antes e depois da cirurgia virtual, caso nódulo.

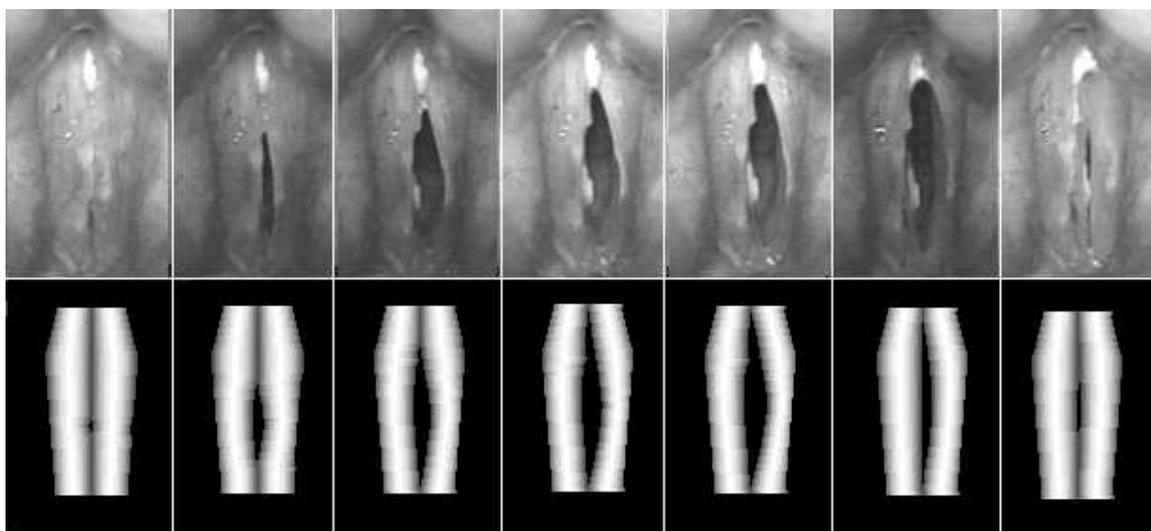


Figura A.6: Exame de endoscopia e simulação numérica do caso com leucoplasia.

A cirurgia virtual é realizada com a eliminação da leucoplasia e uma melhora na mobilidade das pregas vocais. Os resultados dessas alterações são apresentados na Figura A.8

Na Figura A.9 são apresentados uma amostra da voz e seu respectivo espectro antes e depois da cirurgia e na Figura A.10 são mostrados os dois espectrogramas da vogal /a/ sustentada, antes e depois da cirurgia.

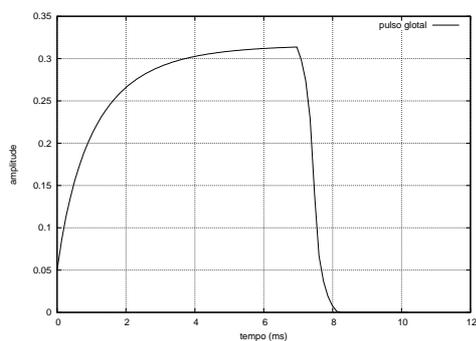
### A.1.3 Caso: Fenda no terço médio e sulco

O caso real apresentado é uma fenda no terço médio as pregas vocais. O exame de endoscopia e a simulação numérica deste caso são apresentados na Figura A.11.

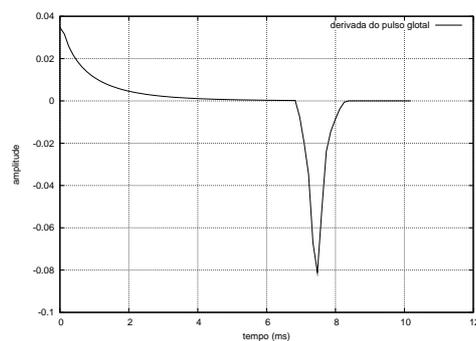
O pulso glotal, sua derivada e transformada de Fourier juntamente com a área de contato das pregas vocais são mostrados na Figura A.12.

A cirurgia virtual é realizada com a eliminação da fenda. Os resultados dessas alterações são apresentados na Figura A.13

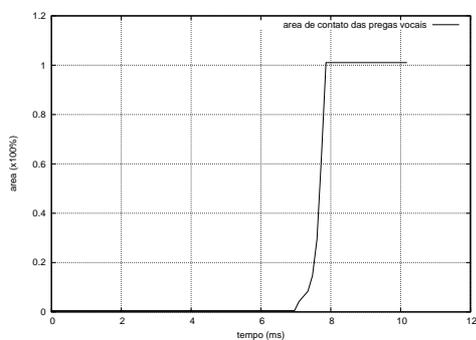
Na Figura A.14 são apresentados uma amostra da voz e seu respectivo espectro antes e depois da cirurgia e na Figura A.15 são mostrados os dois espectrogramas da vogal /a/ sustentada, antes e depois da cirurgia.



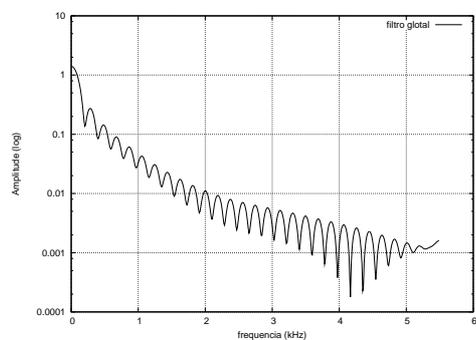
(a) Pulso glotal



(b) Derivada do pulso glotal

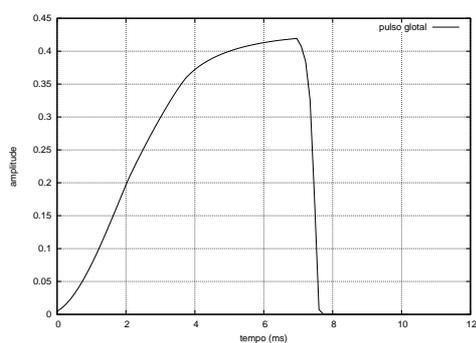


(c) área de contato das pregas vocais

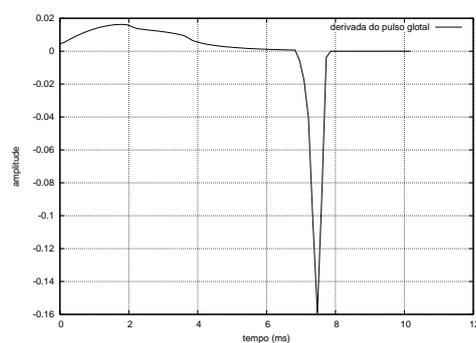


(d) Espectro do filtro glotal

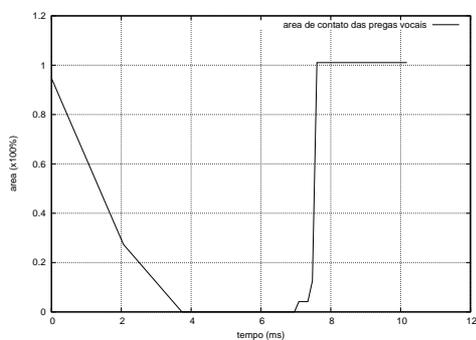
Figura A.7: Simulação numérica do caso leocoplasi.



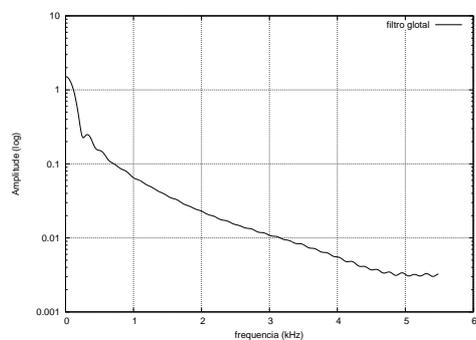
(a) Pulso glotal



(b) Derivada do pulso glotal

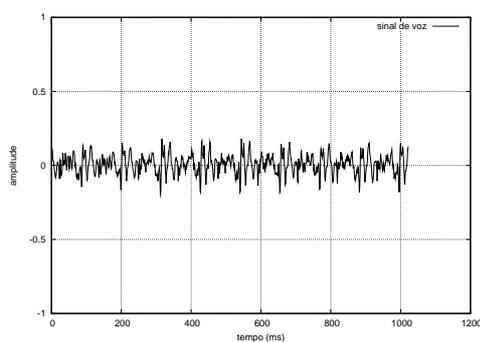


(c) área de contato das pregas vocais

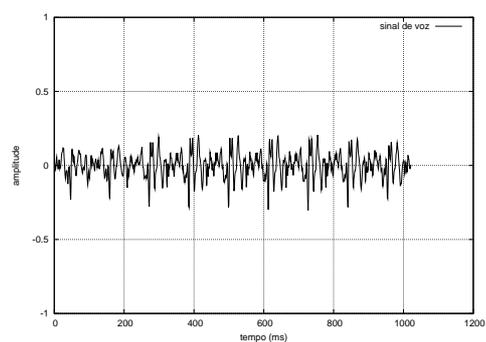


(d) Espectro do filtro glotal

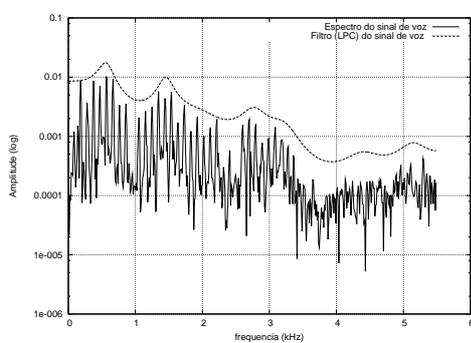
Figura A.8: Simulação numérica da cirurgia virtual do caso leocoplasia.



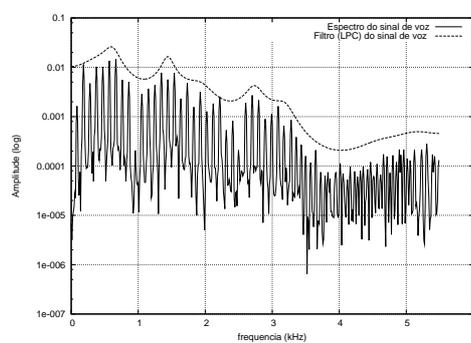
(a) Sinal de voz antes da cirurgia virtual



(b) Sinal de voz após a cirurgia virtual



(c) Espectro do sinal de voz antes da cirurgia virtual



(d) Espectro do sinal de voz após a cirurgia virtual

Figura A.9: Sinais de vozes antes e após a cirurgia virtual, caso leocoplusia.

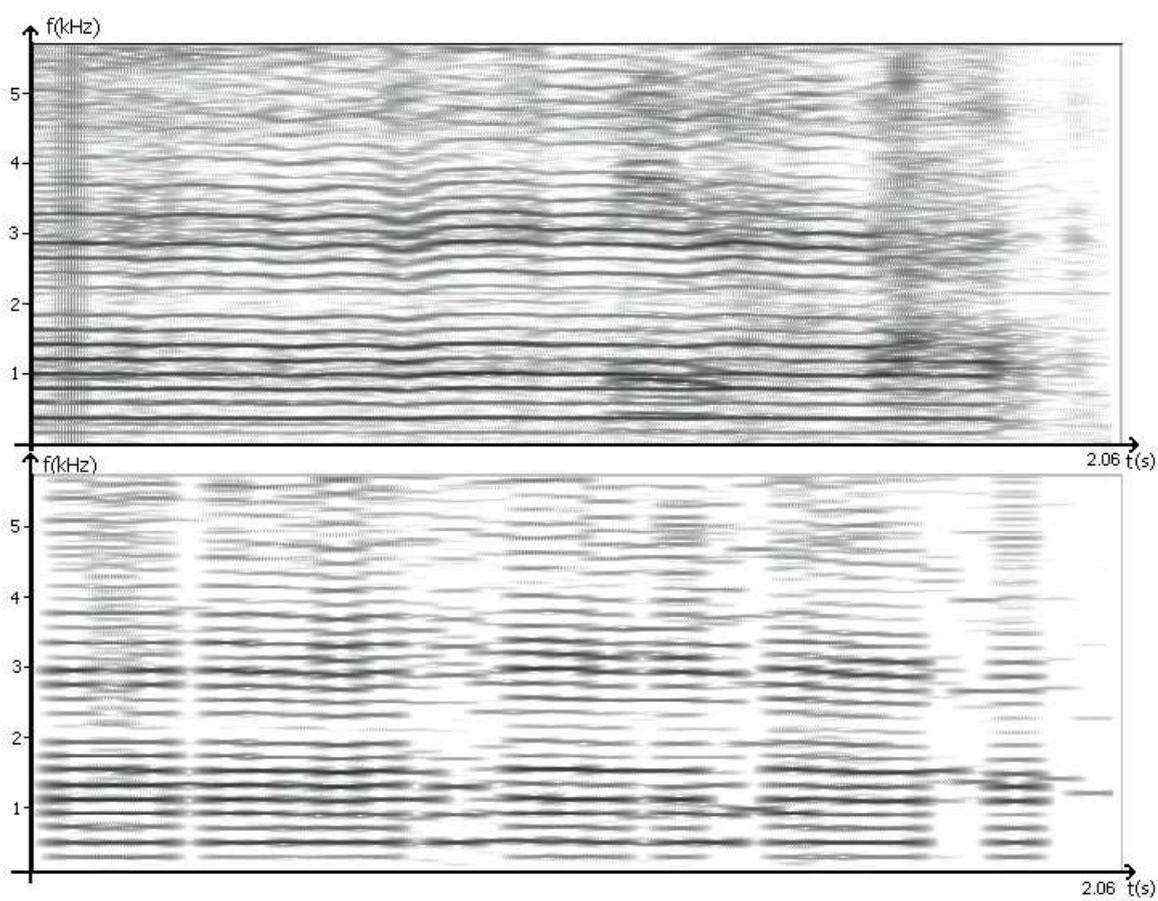


Figura A.10: Espectograma antes e depois da cirurgia virtual, caso leocoplasia.

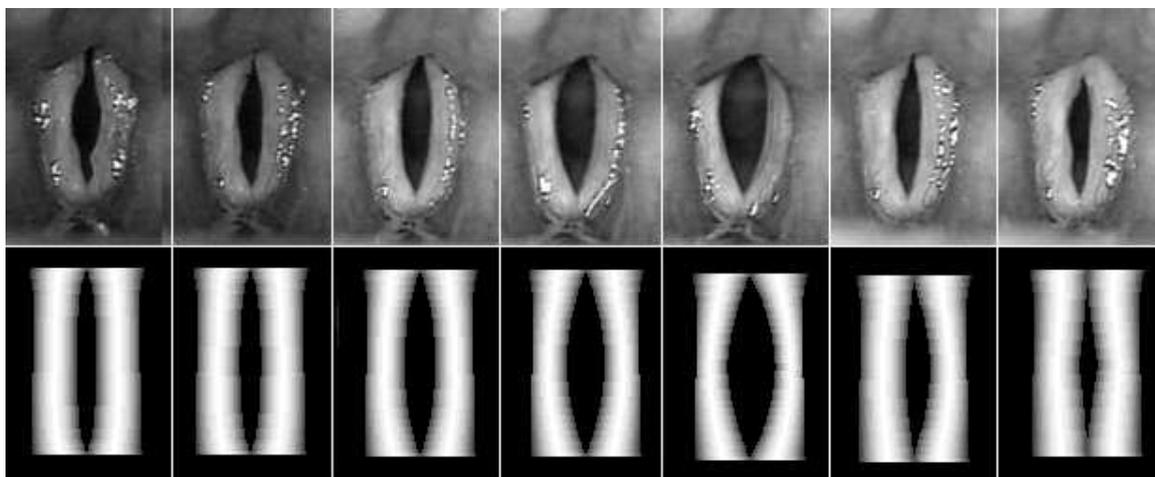


Figura A.11: Exame de endoscopia e simulação numérica do caso com fenda e sulco.

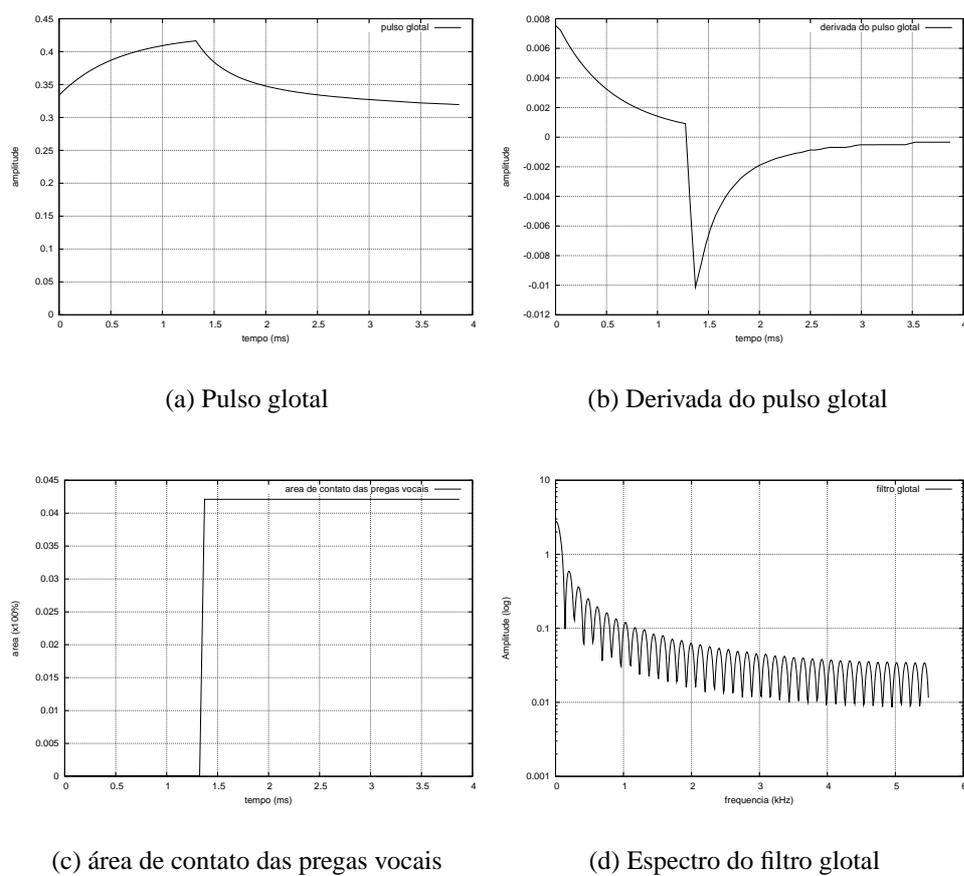
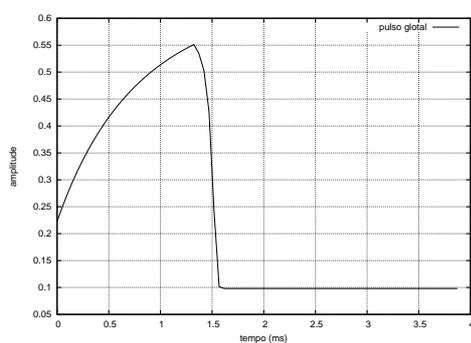
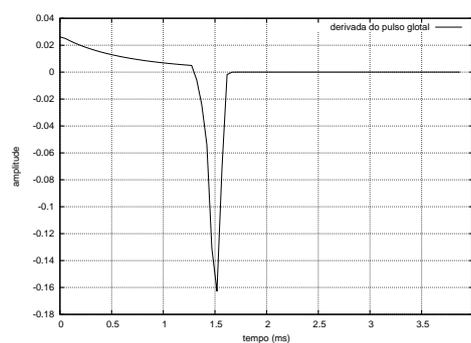


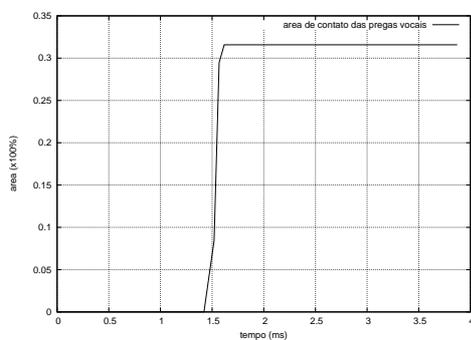
Figura A.12: Simulação numérica do caso fenda e sulco.



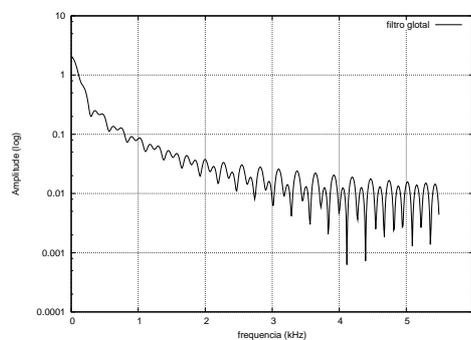
(a) Pulso glotal



(b) Derivada do pulso glotal

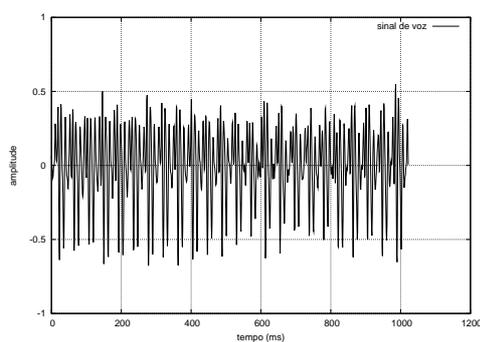


(c) área de contato das pregas vocais

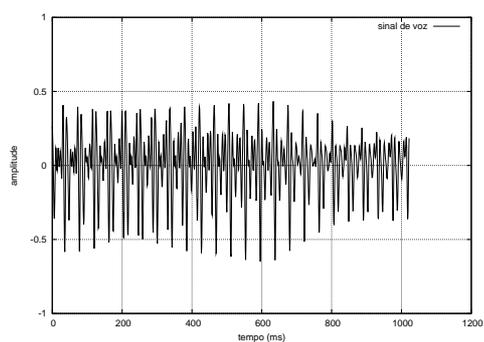


(d) Espectro do filtro glotal

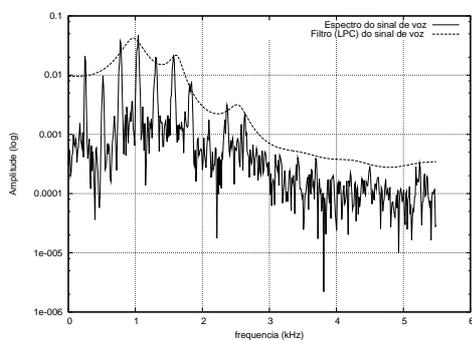
Figura A.13: Simulação numérica da cirurgia virtual do caso fenda e sulco.



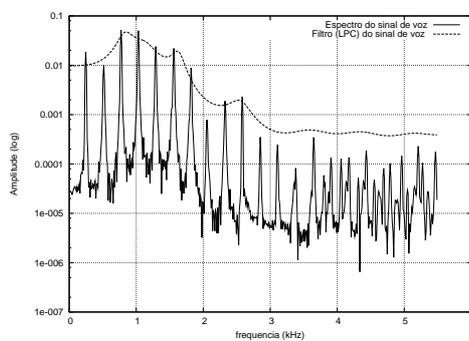
(a) Sinal de voz antes da cirurgia virtual



(b) Sinal de voz após a cirurgia virtual



(c) Espectro do sinal de voz antes da cirurgia virtual



(d) Espectro do sinal de voz após a cirurgia virtual

Figura A.14: Sinais de vozes antes e após a cirurgia virtual, caso fenda e sulco.

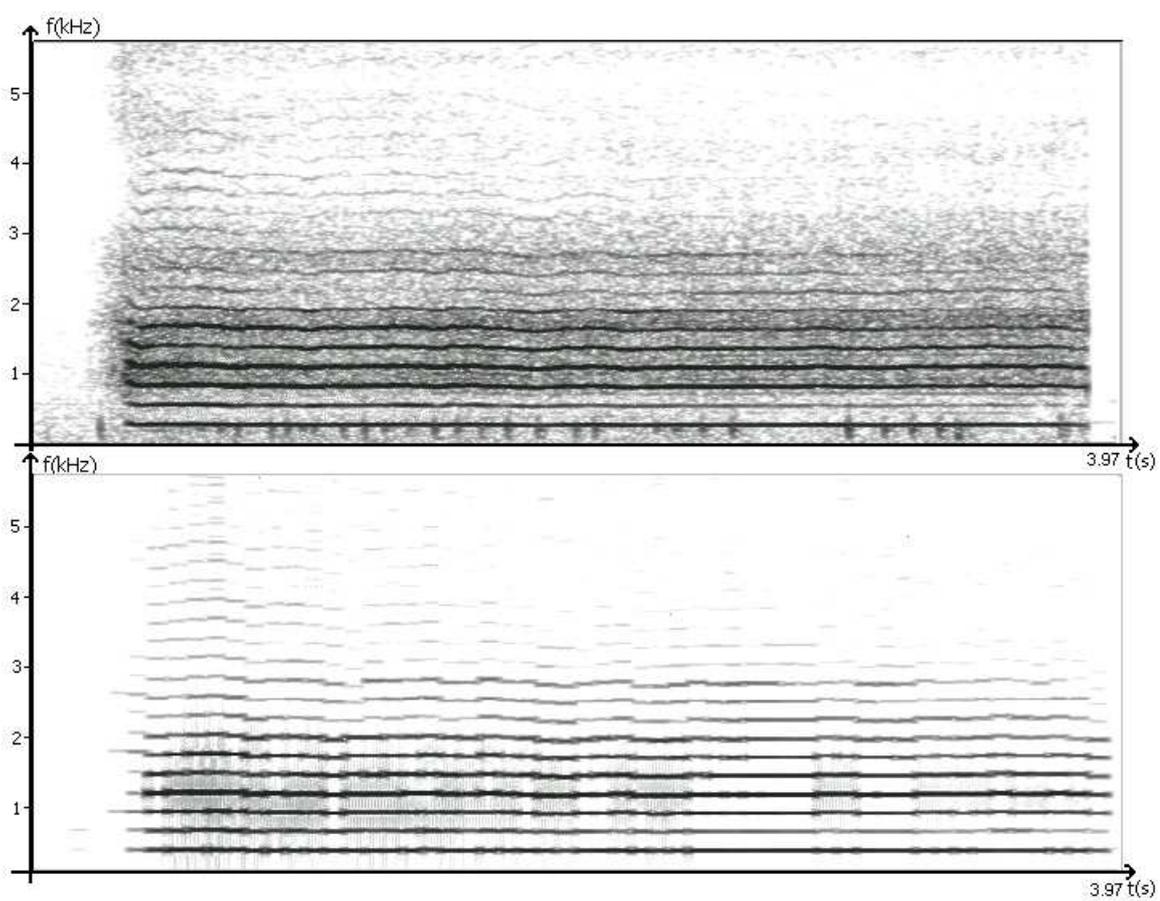


Figura A.15: Espectrograma antes e depois da cirurgia virtual, caso fenda e sulco.

### A.1.4 Caso: Paralisia

O caso real apresentado é uma paralisia do lado esquerdo da laringe. O exame de endoscopia e a simulação numérica deste caso são apresentados na Figura A.16.

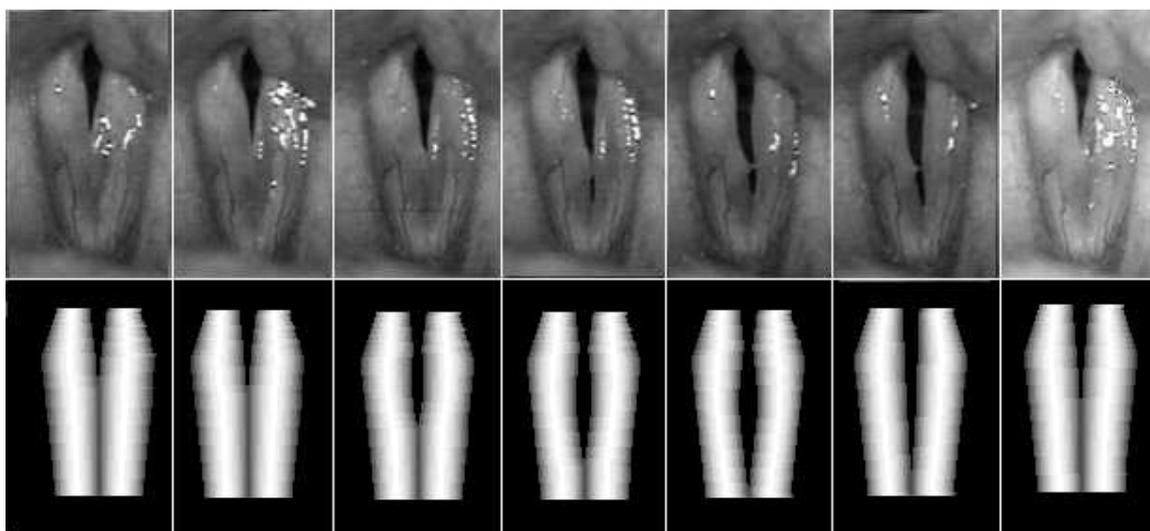


Figura A.16: Exame de endoscopia e simulação numérica do caso com paralisia.

O pulso glotal, sua derivada e transformada de Fourier juntamente com a área de contato das pregas vocais são mostrados na Figura A.17.

A cirurgia virtual é realizada com um melhor fechamento das aritenóides eliminando a fenda da comissura posterior. Os resultados dessas alterações são apresentados na Figura A.18

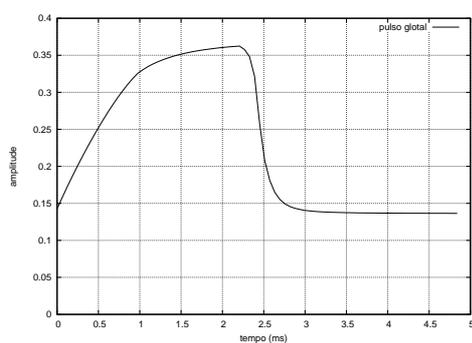
Na Figura A.19 são apresentados uma amostra da voz e seu respectivo espectro antes e depois da cirurgia e na Figura A.20 são mostrados os dois espectrogramas da vogal /a/ sustentada, antes e depois da cirurgia.

### A.1.5 Caso: Pólipo

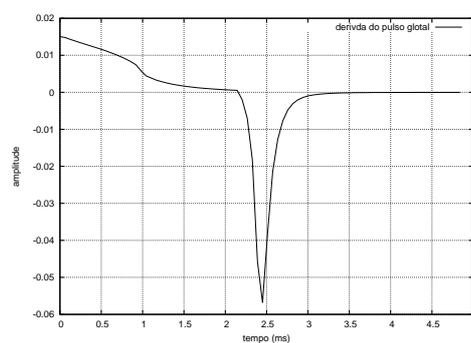
O caso real apresentado é um pólipo do lado direito da laringe. O exame de endoscopia e a simulação numérica deste caso são apresentados na Figura A.21.

O pulso glotal, sua derivada e transformada de Fourier juntamente com a área de contato das pregas vocais são mostrados na Figura A.22.

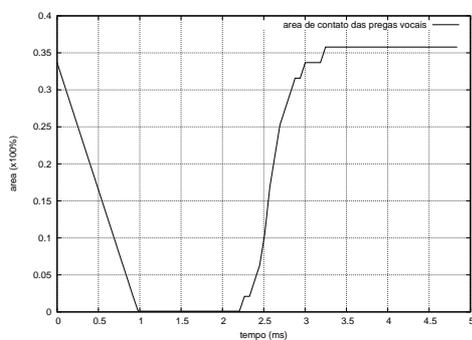
A cirurgia virtual é realizada com a eliminação do pólipo na prega vocal direita. Os resultados dessas alterações são apresentados na Figura A.23



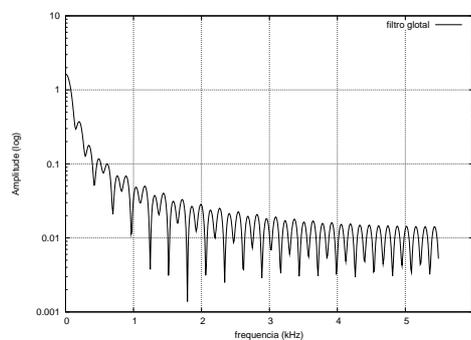
(a) Pulso glotal



(b) Derivada do pulso glotal

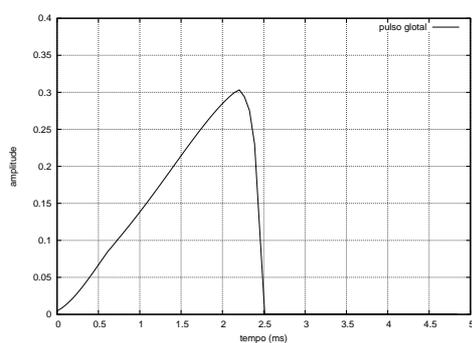


(c) área de contato das pregas vocais

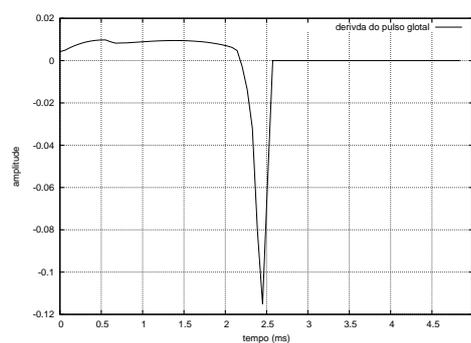


(d) Espectro do filtro glotal

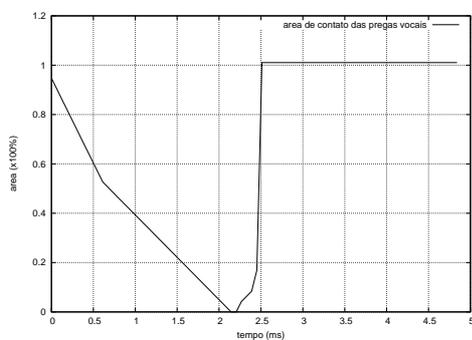
Figura A.17: Simulação numérica do caso paralisia.



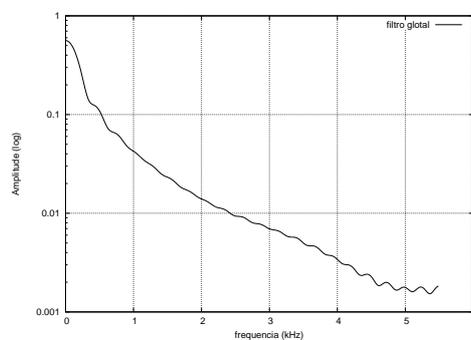
(a) Pulso glotal



(b) Derivada do pulso glotal

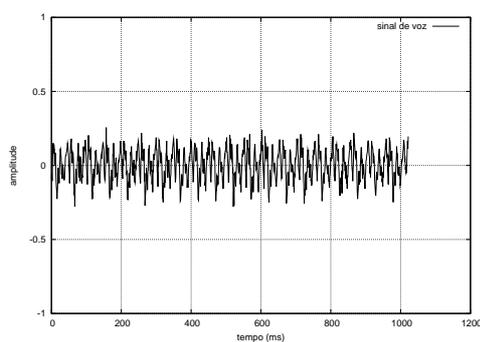


(c) área de contato das pregas vocais

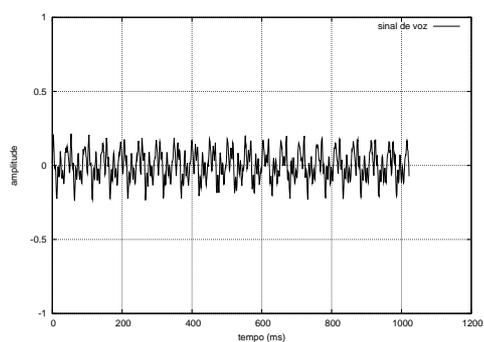


(d) Espectro do filtro glotal

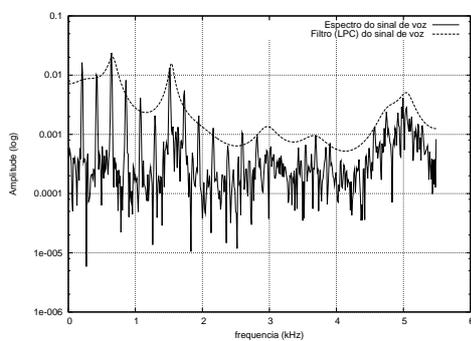
Figura A.18: Simulação numérica da cirurgia virtual do caso paralisia.



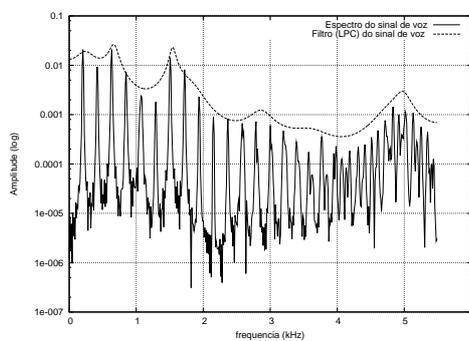
(a) Sinal de voz antes da cirurgia virtual



(b) Sinal de voz após a cirurgia virtual



(c) Espectro do sinal de voz antes da cirurgia virtual



(d) Espectro do sinal de voz após a cirurgia virtual

Figura A.19: Sinais de vozes antes e após a cirurgia virtual, caso paralisia.

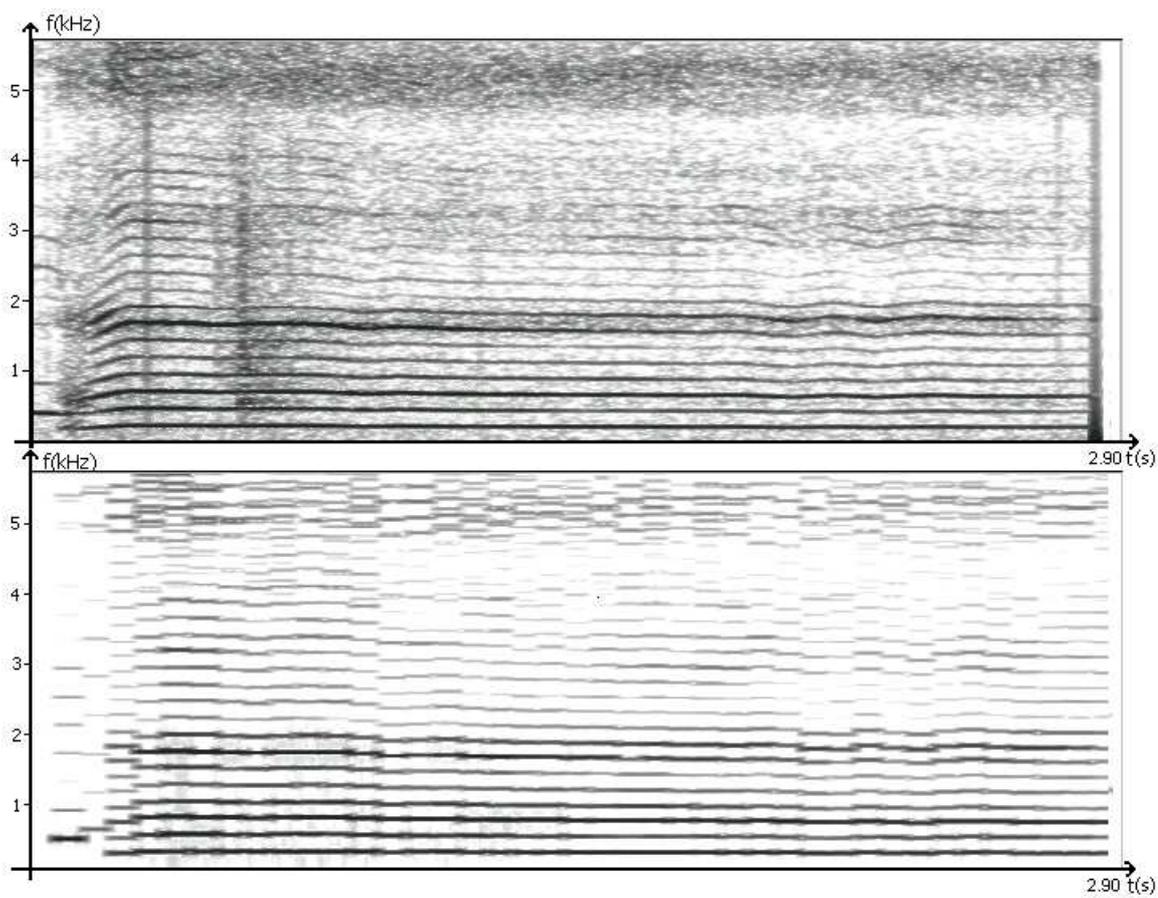


Figura A.20: Espectograma antes e depois da cirurgia virtual, caso paralisia.

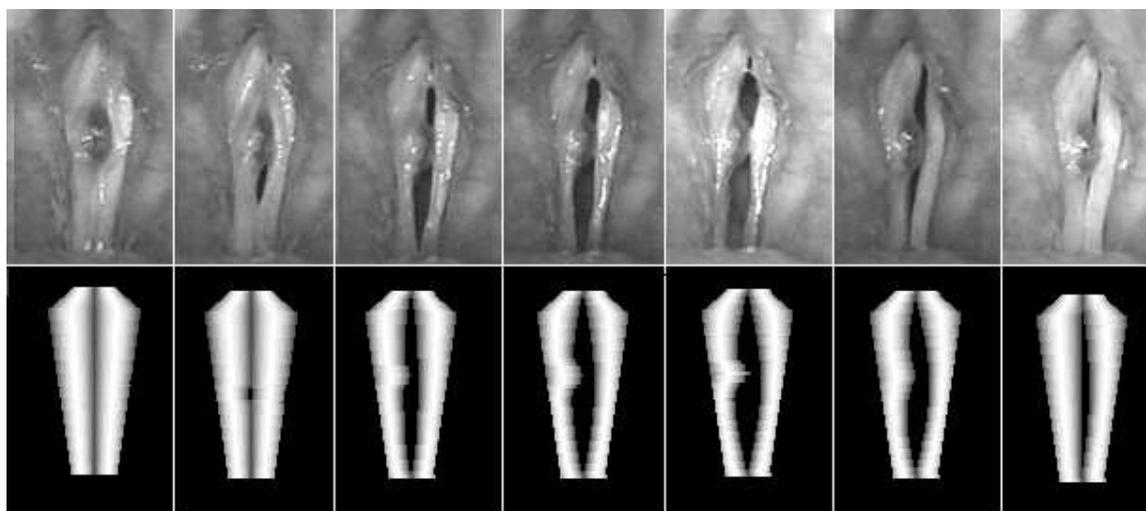


Figura A.21: Exame de endoscopia e simulação numérica do caso pólipso na prega vocal direita.

Na Figura A.24 são apresentados uma amostra da voz e seu respectivo espectro antes e depois da cirurgia e na Figura A.25 são mostrados os dois espectrogramas da vogal /a/ sustentada, antes e depois da cirurgia.

## A.2 Programa Desenvolvido

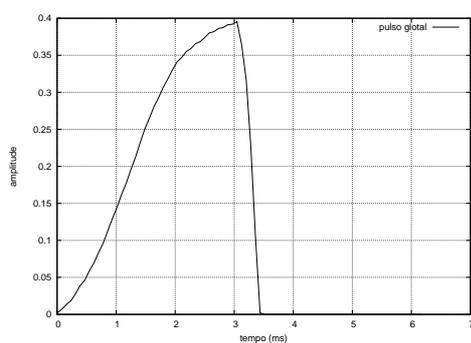
Nas Figuras a seguir serão apresentadas as janelas de interface com o usuário do programa desenvolvido para a elaboração desta tese.

### A.2.1 Banco de dados

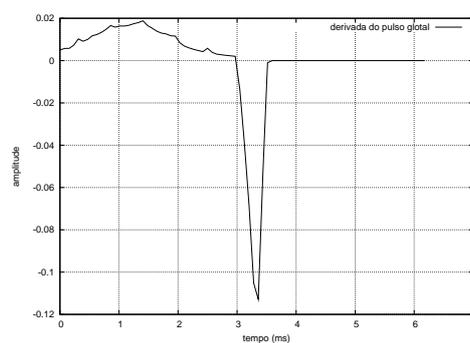
A janela do programa referente ao banco de dados foi importada do programa Análise de Voz com o acréscimo das tabelas para o gerenciamento das imagens, Figura A.26

### A.2.2 Captura e edição de vídeo

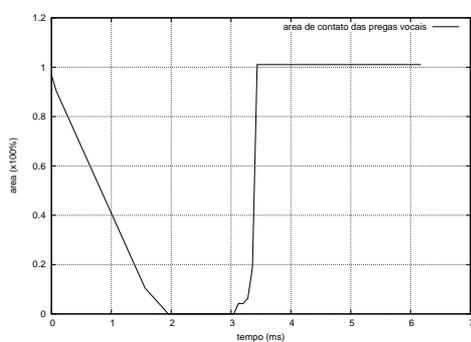
Para a captura e edição de vídeo foi desenvolvida a Janela apresentada na Figura A.27 onde é possível fazer a extração do áudio do vídeo e gravação de *CDs 'compact disks'*.



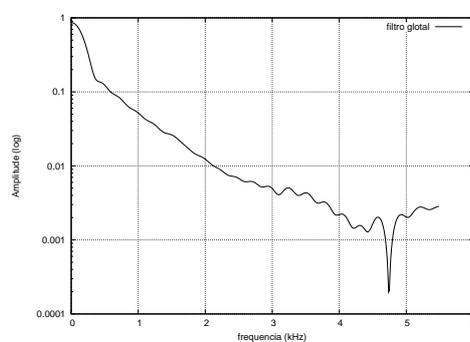
(a) Pulso glotal



(b) Derivada do pulso glotal



(c) área de contato das pregas vocais



(d) Espectro do filtro glotal

Figura A.22: Simulação numérica do caso pólipó na prega vocal direita.

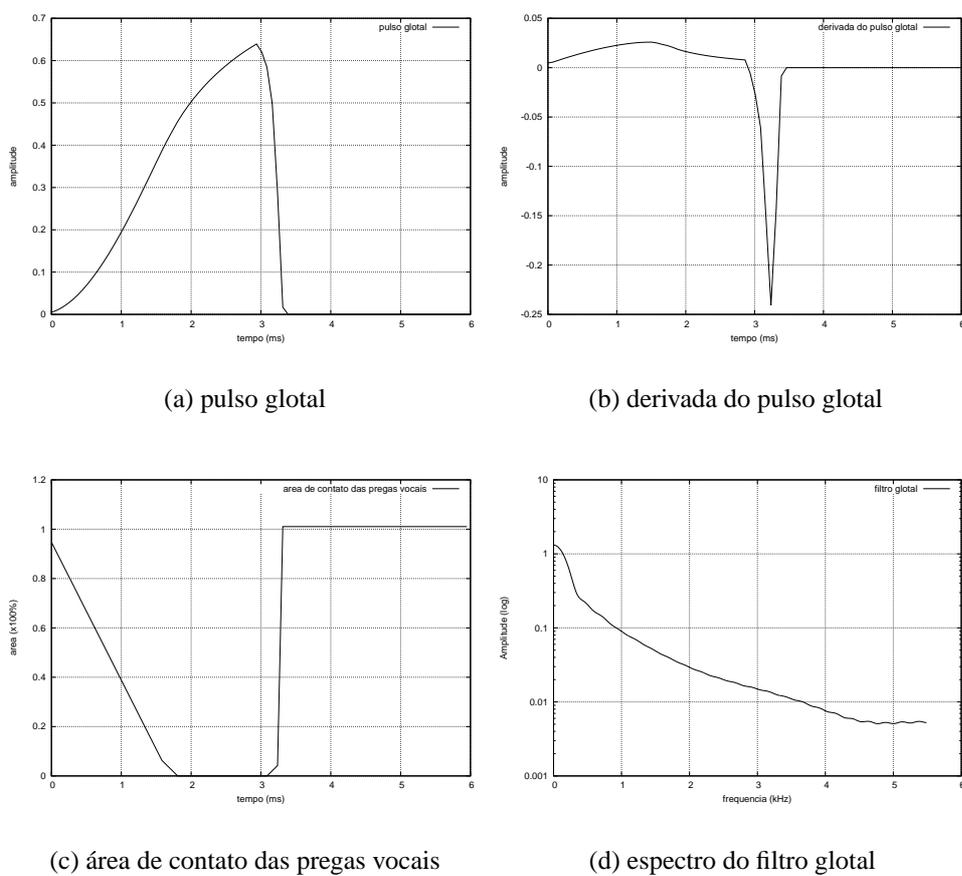
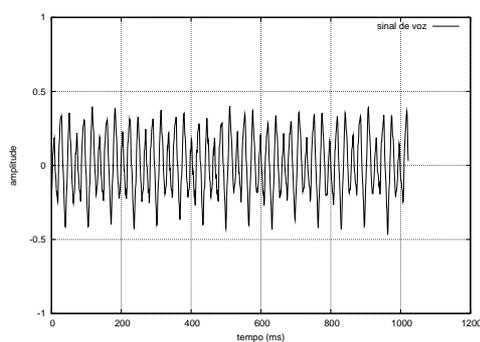
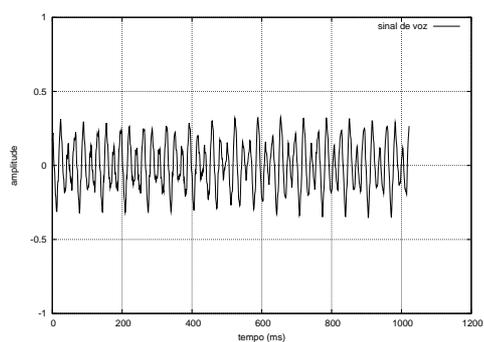


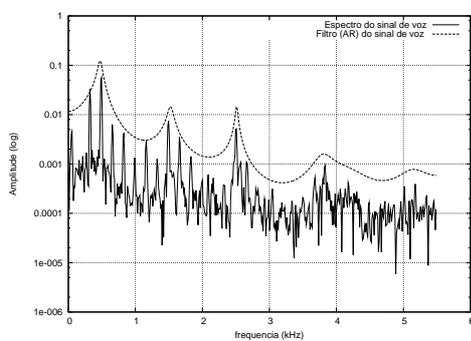
Figura A.23: Simulação numérica da cirurgia virtual do caso pólipos na prega vocal direita.



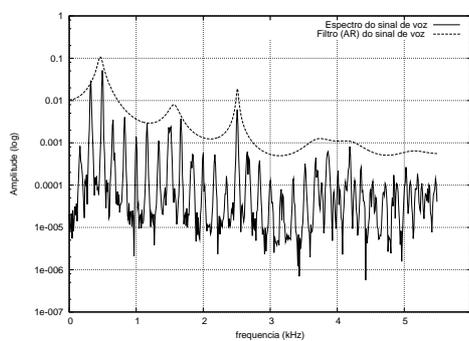
(a) Sinal de voz antes da cirurgia virtual



(b) Sinal de voz após a cirurgia virtual



(c) Espectro do sinal de voz antes da cirurgia virtual



(d) Espectro do sinal de voz após a cirurgia virtual

Figura A.24: Sinais de vozes antes e após a cirurgia virtual, caso pólipos.

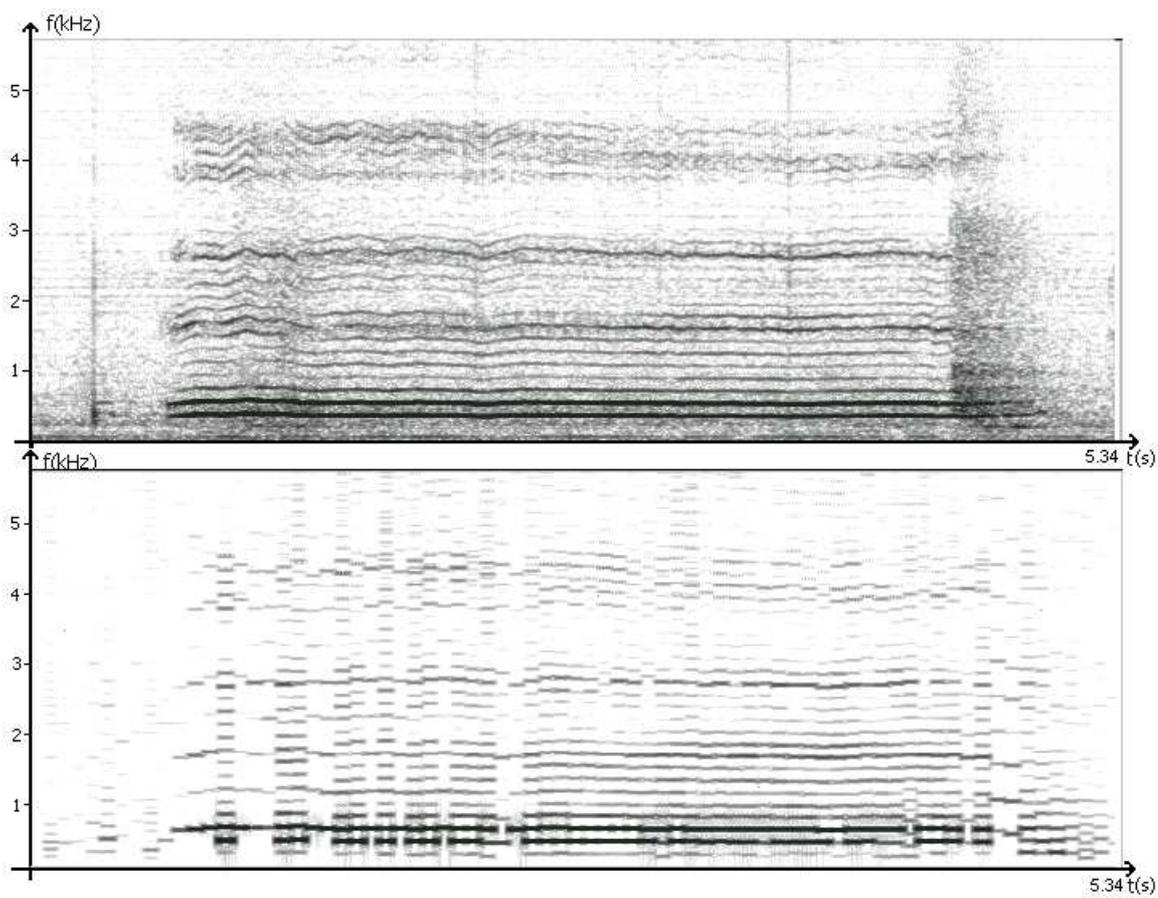


Figura A.25: Espectograma antes e depois da cirurgia virtual, caso pólipos.

Localizar:

Ordenar por:

Número
00000002
00000008
00000016
00000022
00000013
00000015
00000014
00000011
00000021
00000019
<b>00000009</b>
00000020
00000012

Comentários

Cisto na prega v

**Editar**

Número Prontuário Convênio

00000009

Nome Fulano de Tal Sexo M Nascimento 4/10/1980

Estado Civil Solteiro Profissão Estudante Idade 25

Endereço Rua Zero

Bairro  Cidade São Carlos Estado SP CEP 13560

Telefone  Tel. Contato

e-mail

Identidade  CPF  Encaminhado por

Informações adicionais sobre o paciente:

Cisto na prega vocal direita

OK Cancelar

Figura A.26: Janela relativa ao gerenciamento do banco de dados.

### A.2.3 Análise da endoscopia

Na janela mostrada na Figura A.28 é possível fazer um relatório rápido das observações da estroboscopia conforme apresentado na Figura A.29.

### A.2.4 Medidas da laringe

A dimensões da laringe assim como as de suas alterações são realizadas na Janela apresentada na Figura A.30.

### A.2.5 Quimografia

Para realizar o estudo e quantificação da dinâmica das pregas vocais o exame de quimografia é o métodos utilizados na janela mostrado na Figura A.31.

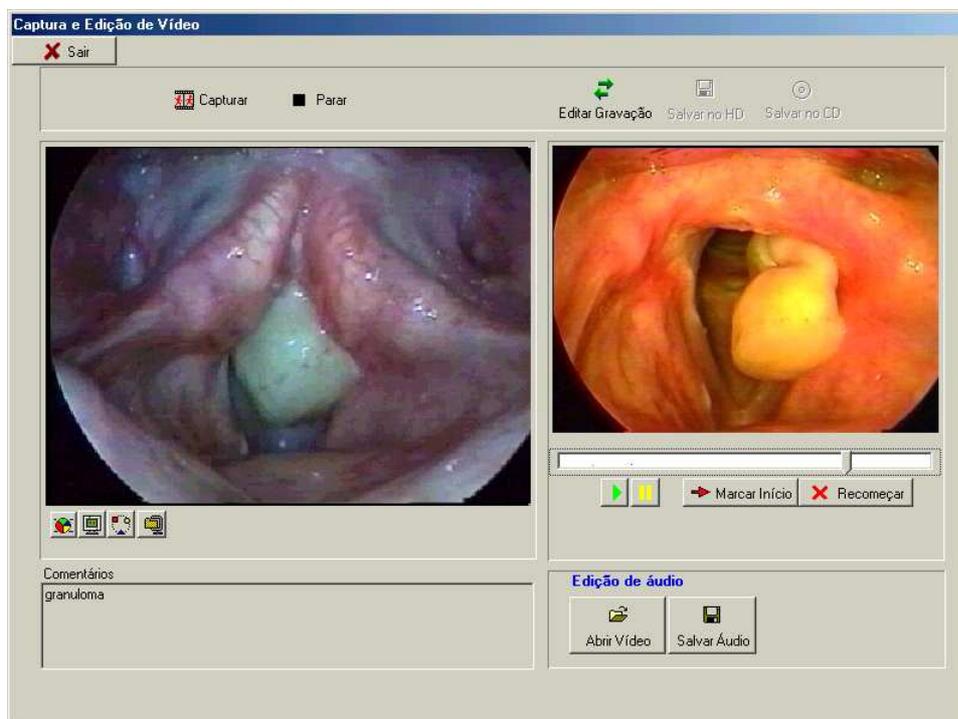


Figura A.27: Janela relativa à captura e edição das imagens de vídeo.

## A.2.6 Cirurgia virtual

A execução do modelo proposto juntamente com as análises dos sinais obtidos na glote são realizados na janela da cirurgia virtual apresentada na Figura A.32.

## A.3 Programa Análise de Voz 5.0

Nas Figuras a seguir serão apresentadas as janelas de interface com o usuário do programa desenvolvido para a análise acústica da voz.

### A.3.1 Anamnese

Na janela anamnese é possível armazenar as informações perceptivas e da entrevista ao paciente, Figura A.33.

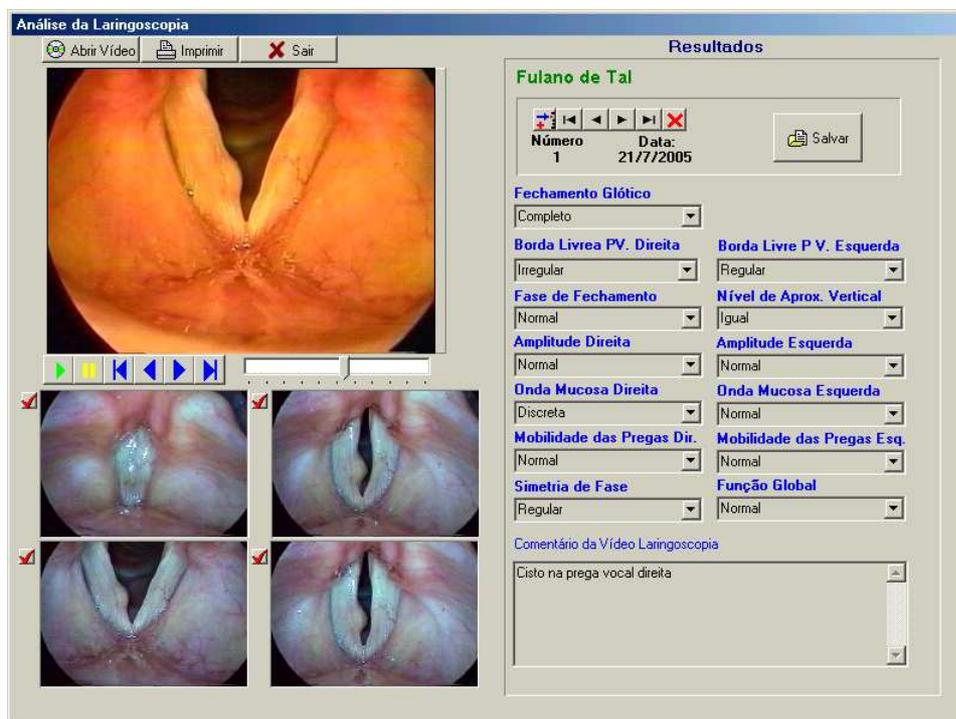


Figura A.28: Janela relativa ao exame da endoscopia.

### A.3.2 Análise acústica

A análise acústica é feita através da obtenção de nove parâmetros acústicos: *Jitter*, *Shimmer*, Nivelamento espectral do filtro e resíduo, coeficiente de excesso, amplitude de *pitch*, frequência fundamental, ataque e nasalização, conforme pode ser visto na Figura A.34.

### A.3.3 Espectrograma

O espectrograma é um gráfico tridimensional (frequência x amplitude x tempo) apresentado em duas dimensões onde a amplitude é quantificada por escala de cores, Figura A.35.

### A.3.4 Outras funções do programa

As janelas a seguir serão apresentadas na Figura A.36 devido ao fato de serem similares às apresentadas anteriormente.



EESC - Escola de Engenharia de São Carlos - USP  
Av. Trabalhador São-carlense, 400  
Fone: 3371 - 9330  
arlindonm@yahoo.com.br

Nome: Fulano de Tal  
Nasc: 4/10/1980

Data: 12/7/2005

#### Videolaringoscopia Computadorizada

**Fechamento Glótico:**  
Completo

**Borda Livre da prega Vocal**  
Esquerda: Regular  
Direita: Irregular

**Fase de Fechamento**  
Normal

**Nível de Aprox. Vertical**  
Igual

**Amplitude**  
Esquerda: Normal  
Direita: Normal

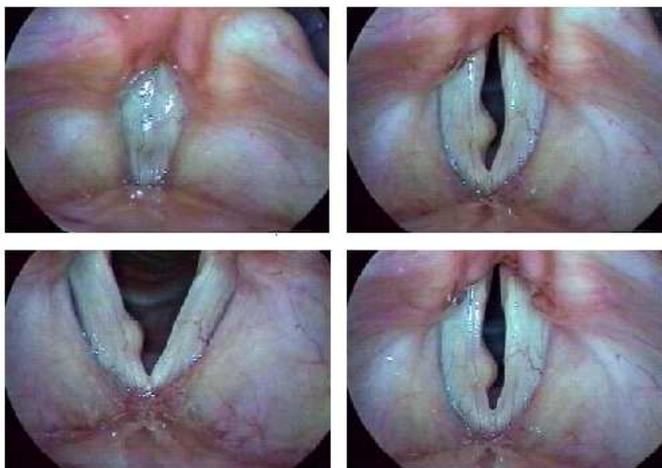
**Onda Mucosa**

Esquerda: Normal  
Direita: Discreta

**Mobilidade das Pregas**  
Esquerda: Normal  
Direita: Normal

**Simetria de Fase**  
Regular

**Função Global**  
Normal

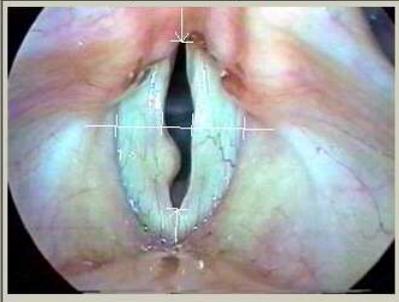


Comentários:

Cisto na prega vocal direita.

Figura A.29: modelo do relatório.

**Medidas**



**Fulano de Tal**      **Número:** 2      **Data:** 12/7/2005

Procurar o trecho de máxima abertura das pregas durante a fonação

Abertura Vertical (Ref.)	Ângulo
✓ 140 (100%)	88,36
Essessura PVD	Essessura PVE
✓ 27,1428571428571	30
Largura. Max. Pregas (C.S.)	Largura min. pregas (C.I.)
✓ 73,571428   15,7142857	✓ 55,7142857142857

Procurar o trecho com o máximo fechamento p/ medir os dois próximos valores:  
O retorno ao trecho com a máxima abertura das pregas é automático

Abertura comissura superior	Comp. Abertura
✓ 8,57142857142857	✓ 14,2857142857143

Procurar o trecho com o início da abertura das pregas

Posição inicial da abertura	
✓ 21,4285714285714	
Posição Alteração 1	Área1
✓ 63,37	✓ 9,69
Posição Alteração 2	Área2
✓	✓
Posição Alteração 3	Área3
✓	✓
Posição Alteração 4	Área4
✓	✓

**Terço Médio da Prega Vocal Direita**

Figura A.30: Medidas das dimensões das pregas vocais.

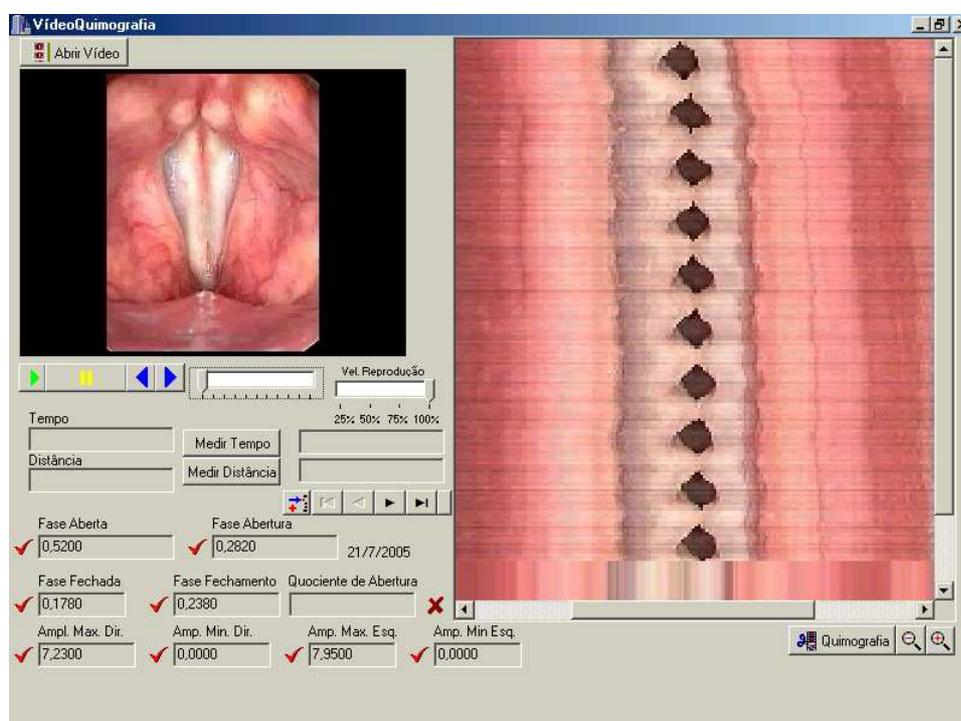


Figura A.31: Janela relativa à execução da quimografia.

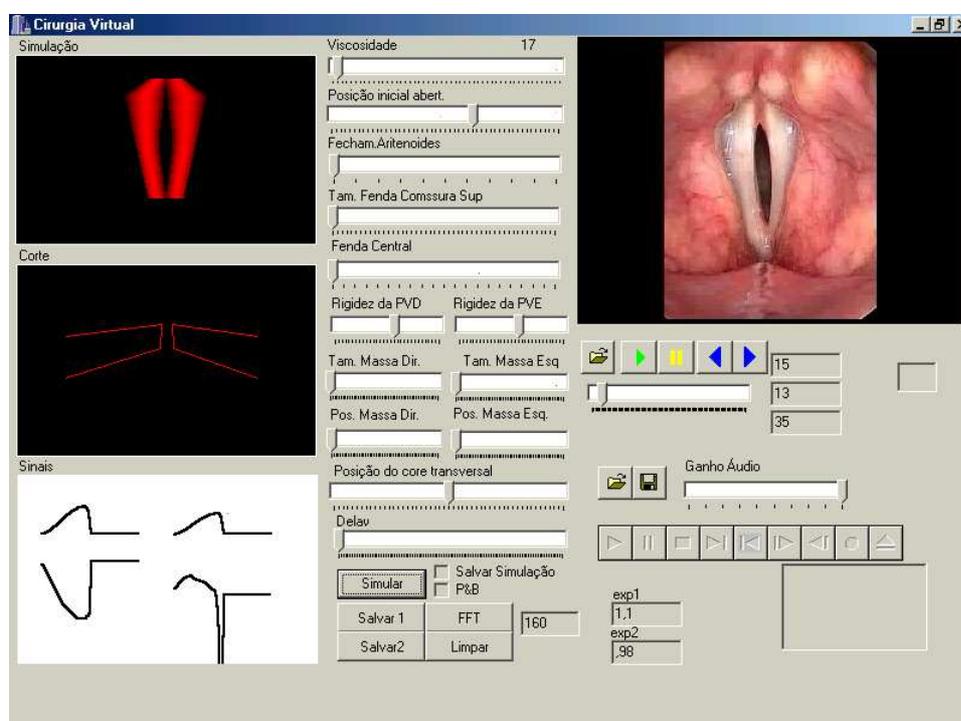


Figura A.32: Janela relativa à cirurgia virtual.

**Anamnese: Fulano de Tal**

**Anotações Gerais do Paciente** Imprimir Sair

Alergias: Nega  
 Alt. Digestivas:   
 Alt. Faringeas:   
 Medicção de uso contínuo:   
 Uso da Voz: Diário:   
 Profissão:   
 Canto:   
 Tabagismo: Nega  
 Alt. Hormonais:   
 Alt. Nasais:   
 Cirurgias e Intercorrências:   
 Tipo de Alteração: Cisto epidermóide  
 Conduta Terapêutica: Cirurgia e Fonoterapia  
 Etilismo: Nega  
 Alt. Pulmonares:   
 Alt. Auditivas:   
 Alt. Neurológicas:   
 Classificação: Distonia Orgânica  
 Outros:   
 Comentários sobre o paciente:

**Anotações Diárias do Paciente**

Inserir Salvar Excluir ← → **Qualidade Vocal**

Data	Tipo de Voz	Ressonância	Pitch	Intensidade	Articulação	Velocidade	Coord. pneumof.	Ataque	Respiração
17/10/2005	Soprosa	Equilibrada	Médio	Média	Precisa	Média	Presente	Brusco	Mista

**Recursos Vocais**

Ênfase	Pausas	Curva Melódica	Ritmo
Presente	Adequadas	Descendente	Repetitivo

**Recursos não verbais**

Exp. Faciais	Gestos	Post. Corporal
Adequadas	Adequados	Adequada

**Avaliação Miofuncional - (DFA)**

Função neurovegetativa: Sucção:   
 Mastigação:   
 Deglutição:   
 Respiração:   
 Tônus Cervical (palpação laringea):

**Análise acústica:**  
 (17/10/2005 )  
 Jitter(<2%): 5,92% Shimmer(<8%): 14,87% SFR(<5): -7 SFF(<5): -12,4 PA(>0,3): 0,54  
 Ex(>2): 1,34 Fo: 189,7 Ataque(<50): 84,54 % Nasalidade(<3): 2,28 S/Z:

Comentários do dia 17/10/2005 (queixas, exames solicitados e resultados ?)

Figura A.33: Janela relativa à anamnese.



Figura A.34: Janela relativa à análise acústica do sinal de voz.

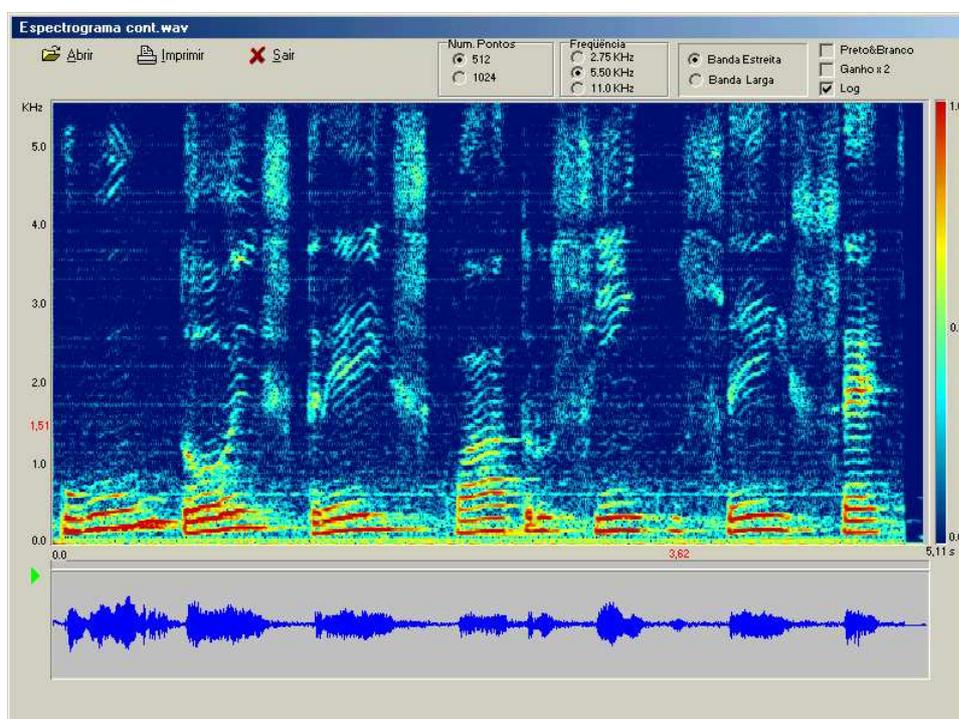
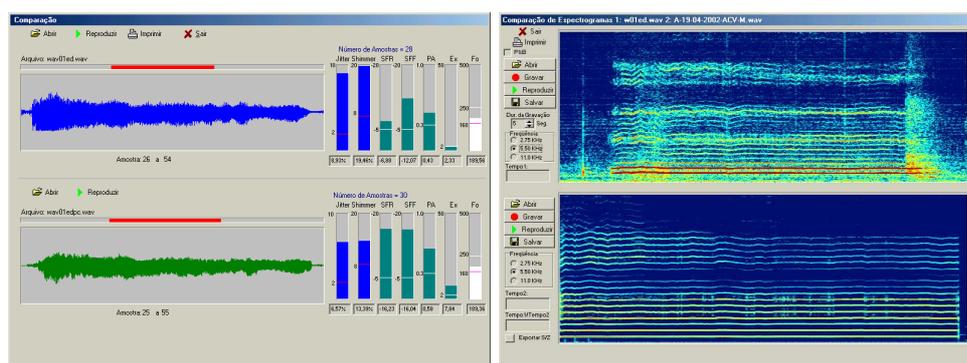
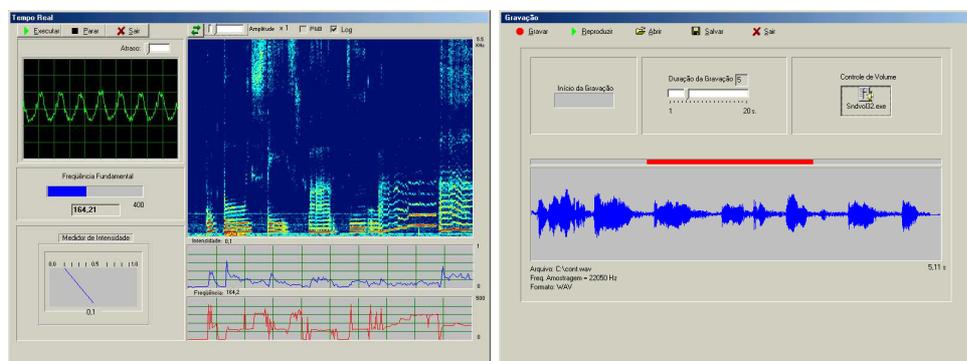


Figura A.35: Janela relativa ao gerenciamento do banco de dados.



(a) Comparação da análise acústica

(b) Comparação de espectrogramas



(c) Tempo real

(d) Gravação e edição de áudio

Figura A.36: Janelas complementares do programa Análise de Voz 5.0