Universidade de São Paulo
Departamento de Engenharia Elétrica

Fernando Araujo de Andrade Sobrinho

Medida da dispersão da periodicidade de um
sinal de voz normal e voz patológica através
da seção de Poincaré

São Carlos

2011

Fernando Araujo de Andrade Sobrinho

Medida da dispersão da periodicidade de um sinal de voz normal e voz patológica através da seção de Poincaré

Dissertação apresentada ao programa de pós graduação da Escola de Engenharia de São Carlos como parte dos requisitos para obtenção do título de Mestre em Ciências, Programa de Engenharia Elétrica.

Área de concentração: Processamento de Sinais e Instrumentação.

Orientador: Prof. Dr. José Carlos Pereira

São Carlos

2011

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

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

A553m	<p>Andrade Sobrinho, Fernando Araujo de</p> <p>Medida da dispersão da periodicidade de um sinal de voz normal e voz patológica através da seção de Poincaré / Fernando Araujo de Andrade Sobrinho; orientador José Carlos Pereira. São Carlos, 2011.</p> <p>Dissertação (Mestrado-Programa de Pós-Graduação em Engenharia Elétrica e Área de Concentração em Processamento de Sinais e Instrumentação) -- Escola de Engenharia de São Carlos da Universidade de São Paulo, 2011.</p> <p>1. Processamento digital de voz. 2. Seção Poincaré. 3. Dinâmica não linear. 4. Espaço de fase. 5. Análise de voz. I. Título.</p>
-------	--

FOLHA DE JULGAMENTO

Candidato(a): Engenheiro **FERNANDO ARAUJO DE ANDRADE SOBRINHO**.

Dissertação defendida e julgada em 06/05/2011 perante a Comissão Julgadora:

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

APROVADO

Profª. Drª. **LIDIA CRISTINA DA SILVA TELES**
(Faculdade de Odontologia de Bauru/USP)

Aprovado

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

Aprovado

Prof. Titular **DENIS VINICIUS COURY**
Coordenador do Programa de Pós-Graduação em
Engenharia Elétrica

Prof. Associado **PAULO CÉSAR LIMA SEGANTINE**
Presidente da Comissão da Pós-Graduação da EESC

Dedicatória

À minha esposa, Rose Helena Valério, com amor, admiração e gratidão por sua compreensão, carinho, presença e incansável apoio ao longo do período da elaboração deste trabalho.

Agradecimentos

Agradeço à Deus.

Aos meus pais Sérgio e Silvana por todo esforço que fizeram pra que eu chegasse até aqui.

Ao Prof. Dr. José Carlos Pereira por ter me concedido a oportunidade de realizar este trabalho, pela orientação e amizade, contribuindo para meu crescimento científico e intelectual.

À Doutora Maria Eugênia Dajer, pela amizade, companheirismo e as publicações geradas em conjunto que contribuíram para o enriquecimento desse trabalho.

As amigos da República dos Largados Farelo, Bicudo, Spin, Velhinho, Maru e agregados.

Ao amigo Cláudio Cyrino Duarte de Souza.

Resumo

Andrade Sobrinho, F.A.. Medida da dispersão da periodicidade de um sinal de voz através da seção de Poincaré. 2011. pp. 96
Dissertação Mestrado. Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2010.

A voz humana apresenta flutuações na frequência, amplitude e formato de onda. Esse comportamento característico pode ser estudado usando técnicas de análise não linear, além das técnicas convencionais. O objetivo desse trabalho é analisar sinais de vozes normais e patológicas (com nódulos e edemas) usando seção de Poincaré de vários trechos do espaço de fase reconstruído e calcular a dispersão em relação ao ponto médio da seção e em relação à distribuição dos pontos sobre os eixos coordenados. Essa dispersão será calculada utilizando o conceito estatístico de desvio padrão. Foram analisados 48 sinais de voz humanas, divididos em 3 grupos (16 normais, 16 com nódulo e 16 com edema de *Reinke*). Foram selecionados trechos de 500ms do sinal temporal nas regiões de maior estacionariedade, descartando os trechos iniciais e finais do sinal para evitar possíveis transitórios. A partir do espaço de fase bidimensional, a seção de Poincaré foi traçada em 10 trechos distintos da trajetória. Em seguida, foi gerado o espaço de fase em 3 dimensões contendo os pontos da seção. Foi feita uma rotação tridimensional dos pontos utilizando a reta tangente à trajetória de modo que a seção ficasse paralela ao plano $x = 0$. Da seção resultante foram extraídas as componentes principais e em seguida calculado o desvio padrão da dispersão e o desvio padrão dos pontos projetados no plano em relação aos eixos coordenados ($y; z$). A validação da ferramenta desenvolvida para esse estudo foi realizada utilizando um sinal senoidal inserindo gradativamente *Jitter* e *Shimmer*, onde se verificou uma variação proporcional da média da dispersão. Os resultados obtidos para esse conjunto de vozes mostraram que o desvio padrão da dispersão e o desvio padrão em relação aos eixos coordenados dos pontos de vozes normais é menor do que os encontrados para vozes com edema e com nódulo. Concluiu-se que a ferramenta proposta conseguiu diferenciar vozes normais das vozes patológicas. Portanto, a ferramenta é promissora para análise e avaliação desse grupo de vozes.

Palavras-chave: Seção de Poincaré, dinâmica não linear, espaço de fase, análise de voz.

Abstract

The human voice, normal or pathological, has fluctuations in the frequency, amplitude and waveform. This characteristic behavior can be studied using techniques of nonlinear analysis, in addition to conventional techniques. The aim of this study is to analyze signals of normal and pathological voices (with nodules and edema) using the Poincaré section of several parts of the reconstructed phase space and calculate the dispersion in relation to the midpoint of the section and in relation to the distribution of points on coordinate axes. This dispersion is calculated using the statistical concept of standard deviation. We analyzed 48 human voice signals divided into 3 groups (16 normal, 16 with nodules and 16 with Reinke's edema). It was selected 500ms signal frames presenting good stationarity, discarding the initial and final portions of the signal to avoid possible transient. From the two-dimensional phase space, the Poincaré section was drawn on 10 different stretches of the path. It was then generated the three-dimensional phase space containing the points of the section. We conducted a three-dimensional rotation of the points using the tangent to the trajectory so that the section stayed parallel to the plane. From the resulting section, principal components were extracted and then calculated the standard deviation of the dispersion and the standard deviation of the coordinate axes of the projected points of the section in the plan. The validation tool developed for this study was performed using a sinusoidal signal gradually inserting jitter and shimmer, where there was a proportional variation of the dispersion media. The results for this set of voices showed that the standard deviation of the dispersion and the standard deviation related to the coordinate axes of the points of normal voices is smaller than those found for voices with edema and nodule. It was concluded that the proposal was promising tool for analyzing and evaluating this group voices.

Key words: Poincaré section, Nonlinear dynamic, Phase space, voice analysis.

Sumário

Lista de Figuras	ix
Lista de Tabelas	xiii
1 Introdução	1
1.1 Objetivo	3
2 Formação da Voz	4
2.1 Classificação dos sons da fala	8
2.2 Tipos de Vozes	9
2.3 Voz normal e voz patológica	10
2.4 Nódulos vocais	11
2.5 Edema de Reinke	12
3 Análise acústica de sinais de voz	14
3.1 Análise Perceptivo auditiva	15
3.2 Análise espectrográfica de um sinal de voz	15
3.3 Medição da Frequencia Fundamental	16
3.4 Cálculo do <i>Shimmer</i>	17
3.5 Cálculo do <i>Jitter</i>	18
3.6 Relação Sinal Ruído	20
4 Técnicas de Análise não-linear	21
4.1 Reconstrução do espaço de fase	21

4.2	Tempo de atraso	22
4.3	Informação mútua	24
4.4	Seção de Poincaré	24
4.4.1	Determinação da seção de Poincaré	25
4.5	Análise das Componentes Principais - PCA	27
4.6	Caos na produção de voz	27
5	Materiais e Métodos	30
5.1	Materiais	30
5.2	Validação da ferramenta utilizando senóides	30
5.3	Procedimento	33
5.3.1	Cálculo de ϵ	39
5.4	Cálculo da Dispersão	41
6	Resultados e discussões	43
6.1	Seção de Poincaré de Vozes Normais	43
6.1.1	Voz normal n° 6	43
6.2	Seção de Poincaré de Vozes com nódulo	52
6.2.1	Voz com nódulo n°2	52
6.3	Voz com edema de Reinke	60
6.3.1	Voz com edema de Reinke n° 14	60
7	Teste ANOVA	71
8	Conclusão	75
	Referências	76

Lista de Figuras

1	Esquema do mecanismo vocal humano [Flanagan,1972 - Pg10].	5
2	Vista posterior e anterior dos principais componentes da laringe.	6
3	Esquema das pregas vocais em posição respiratória (aberta) e em posição fonatória (fechadas). Figura modificada de <i>Mayo Foundation for medical education e Research</i>	6
4	Forma de onda do pulso glotal	7
5	Modelo para o trato vocal modificada	8
6	Sistema Sistema fonte filtro para a formação da voz	9
7	Formas de onda típicas de acordo com a classificação de Ingo R. Titze	10
8	Nódulos vocais	12
9	Edema de <i>Reinke</i>	13
10	Sinal de voz normal e seu espectro	16
11	Sinal de voz com nódulo e seu espectro	16
12	Perturbação em amplitude (Shimmer)	18
13	Perturbação em Período (Jitter)	19
14	Vogal o	26
15	Espaço de fase reconstruído obtido através de 10000 amostras da vogal sustentada /o/. Os cortes em vermelho é o plano de Poincaré (Kubin, 1997).	26
16	Seção de Poincaré obtida a partir da reconstrução do espaço da fase da vogal /o/(Kubin, 1997).	26
17	Rotação dos eixos originais através das componentes principais	28
18	Média da dispersão dos pontos da seção em relação aos eixos (x; y)	31
19	Média da dispersão dos pontos da seção em relação ao ponto médio	32

20	Sinal de voz completo	33
21	Sinal de voz normalizado	34
22	Trecho de 500ms do sinal de voz normalizado	34
23	Trecho de um sinal de voz interpolado	35
24	Informação Mútua	36
25	Espaço de fase	36
26	Espaço de fase ampliado. As barras tracejadas delimitam a região do corte e o trecho de orientação da reta normal à região do corte	37
27	Espaço de fase ampliado. Observa-se a ocorrência de mais de um ponto no trecho em estudo	38
28	Determinação dos pontos pertencentes ao plano de Poincaré através do produto escalar	39
29	Trecho do espaço de fase de um sinal de 100ms. O resultado do produto escalar selecionou apenas os pontos pertencentes ao plano de Poincaré - indicado em vermelho	40
30	Seção de Poincaré do trecho (2,0-2,5ms) do sinal de voz	41
31	Dispersão dos pontos em relação ao ponto médio da seção	41
32	Dispersão dos pontos em relação ao eixo y	41
33	Dispersão dos pontos em relação ao eixo z	41
34	Voz normal 6	43
35	Trecho do sinal de voz normal 6	44
36	Espaço de fase do sinal de voz normal 6 com 10 cortes	44
37	Seção de Poincaré do Corte 1 e seus histogramas	45
38	Seção de Poincaré do Corte 2 e seus histogramas	45
39	Seção de Poincaré do Corte 3 e seus histogramas	46
40	Seção de Poincaré do Corte 4 e seus histogramas	46
41	Seção de Poincaré do Corte 5 e seus histogramas	47
42	Seção de Poincaré do Corte 6 e seus histogramas	47

43	Seção de Poincaré do Corte 7 e seus histogramas	48
44	Seção de Poincaré do Corte 8 e seus histogramas	48
45	Seção de Poincaré do Corte 9 e seus histogramas	49
46	Seção de Poincaré do Corte 10 e seus histogramas	49
47	Voz com nódulo nº 2	53
48	Espaço de fase da voz com nódulo nº 2 com 10 cortes	53
49	Seção de Poincaré do Corte 1 e seus histogramas	54
50	Seção de Poincaré do Corte 2 e seus histogramas	54
51	Seção de Poincaré do Corte 3 e seus histogramas	55
52	Seção de Poincaré do Corte 4 e seus histogramas	55
53	Seção de Poincaré do Corte 5 e seus histogramas	56
54	Seção de Poincaré do Corte 6 e seus histogramas	56
55	Seção de Poincaré do Corte 7 e seus histogramas	57
56	Seção de Poincaré do Corte 8 e seus histogramas	57
57	Seção de Poincaré do Corte 9 e seus histogramas	58
58	Seção de Poincaré do Corte 10 e seus histogramas	58
59	Voz com edema de Reinke	61
60	Espaço de fase do sinal de voz com edema de Reinke com 10 cortes	62
61	Seção de Poincaré do Corte 1 e seus histogramas	62
62	Seção de Poincaré do Corte 2 e seus histogramas	63
63	Seção de Poincaré do Corte 3 e seus histogramas	63
64	Seção de Poincaré do Corte 4 e seus histogramas	64
65	Seção de Poincaré do Corte 5 e seus histogramas	64
66	Seção de Poincaré do Corte 6 e seus histogramas	65
67	Seção de Poincaré do Corte 7 e seus histogramas	65
68	Seção de Poincaré do Corte 8 e seus histogramas	66

69	Seção de Poincaré do Corte 9 e seus histogramas	66
70	Seção de Poincaré do Corte 10 e seus histogramas	67
71	Teste anova para o DPE	72
72	Teste anova para o DPD	74

Lista de Tabelas

1	Média da dispersão dos pontos da seção de Poincaré em relação aos eixos coordenados.	31
2	Média da dispersão dos pontos da seção de Poincaré em relação ao ponto médio da seção	32
3	Desvio padrão da voz normal n° 6	50
4	Desvio padrão da dispersão das vozes normais em relação aos eixos xy . . .	51
5	Desvio padrão das vozes normais em relação ao ponto médio da seção . . .	52
6	Desvio padrão da voz com nódulo n° 2	59
7	Desvio padrão das vozes com nódulo em relação aos eixos cartesianos . . .	60
8	Desvio padrão das vozes com nódulo em relação ao ponto médio da seção .	61
9	Desvio padrão da voz com Edema de Reinke 14	68
10	Desvio padrão das vozes com edema de Reinke em relação aos eixos cartesianos	69
11	Desvio padrão das vozes com edema de Reinke em relação ao ponto médio da seção	69
12	Tabela comparativa entre as medidas DPE e DPD dos 3 tipos de vozes estudadas	70
13	DPE das vozes normais, com nódulo e com edema de Reinke	71
14	DPD das 16 vozes normais, com nódulo e com edema de Reinke	73
15	Valores de P para o teste Anova em grupos de dois dois	74

1 Introdução

Sistemas não lineares sempre estiveram presentes no estudo de fenômenos naturais, mas nas últimas décadas o interesse nas pesquisas desse tipo de sistemas aumentaram devido à redução dos custos de computadores pessoais e ao desenvolvimento de técnicas computacionais para o estudo. Antes, o estudo de sistemas não lineares estava restrito apenas à pesquisadores com acesso a computadores de grande porte; hoje em dia, qualquer pessoa que possua um computador de boa capacidade pode simular um sistema não linear

Ao contrário de um sistema linear, onde o conhecimento dos autovalores e autovetores são suficientes para a resolução do sistema, a grande maioria dos sistemas não lineares necessitam de algoritmos para simular o sistema. Outro motivo importante para o aumento do interesse em sistemas não lineares é devido à descoberta do caos. Caos é o nome dado à sistemas complexos, aparentemente randômico, e geralmente com comportamento imprevisível observados em sistemas dinâmicos não lineares.

Um conceito clássico da engenharia é que sistemas determinísticos são previsíveis ao longo do tempo. Com a descoberta do caos, um sistema determinístico, apesar de ser representado por uma equação determinística, é extremamente sensível à pequenas variações nas condições iniciais apresentando um comportamento imprevisível na saída desse sistema. Diversos estudos demonstraram que um sinal de voz normal ou patológica, apresenta flutuações na frequência, amplitude e formato de onda . Esse comportamento dá uma característica não linear ao sinal e por isso ele pode ser estudado usando técnicas de análise não linear, além das técnicas convencionais, inclusive sob o enfoque do caos.

A análise de sinais vocálicos para detecção de patologias é feita por profissionais treinados e que utilizam da percepção acústica. Através das técnicas de análise da dinâmica não linear é possível extrair parâmetros do sinal de voz, tanto visuais quanto matemáticos, facilitando a análise de vozes normais e patológicas, sendo possível inclusive classificar as patologias. Como técnicas de análise não linear pode-se citar seção de Poincaré, entropia de Kolmogorov, dimensão fractal, dimensão de correlação, expoente de Liapunov, e

outros.

Os exames usados na detecção de patologias vocais utilizando processamento digital de sinais são técnicas não invasivas que evitam o desconforto causado pelas técnicas tradicionais. As análises acústicas das características no tempo e na frequência do sinal de voz são técnicas auxiliares à laringoscopia, para pré-diagnósticos e também acompanhamento dos tratamentos.

Dentre alguns estudos recentes pode-se citar o trabalho de Scalassara (2009) que utilizou medidas de previsibilidade em sinais de voz para diferenciar sinais saudáveis e patológicos. Também recente foi o trabalho de Dajer, M.E., Andrade Sobrinho, F.A, Pereira, J.C (2010) que verificou padrões visuais em sinais de vozes normal e patológicas. Herzel, Berry, Titze, & Saleh (1994) estudaram vozes patológicas sob o ponto de vista da dinâmica não linear, além de utilizarem as técnicas tradicionais, e demonstraram que de fato os métodos são complementares. No estudo de Herzel, (1993); Mende, Herzel & Wermke (1990) e Titze, Baken, & Herzel (1993) foi demonstrado que existe uma estreita relação entre sistemas dinâmicos não lineares e a vibração das pregas vocais.

Mende, Herzel, & Wermke (1990) estudaram o choro de crianças recém-nascidas utilizando métodos de dinâmica não linear e verificaram que existe uma rica variação de bifurcações e comportamentos irregulares. Através da análise da seção de Poincaré e dos atratores concluíram que esses eventos são na verdade um comportamento caótico determinístico de baixa dimensão. Costa, Correia, Falcão, Almeida, & Assis (2007) utilizaram medidas de entropia em sinais de vozes disfônicas devido às patologias nas pregas vocais. Alonso et al (2006) também utilizou recursos da dinâmica não linear usando de dimensão de correlação e expoente de Lyapunov para detecção de patologias na laringe. Outros estudos também demonstraram que a aspereza observada em algumas vozes é causada pelo comportamento não linear do movimento das pregas vocais.

A análise de voz através da dinâmica não linear, de fato, tem-se se mostrado um método eficiente e não invasivo para detecção de patologias. Mais informações podem ser encontradas em outros trabalhos Banci, Monini, Falaschi, & Sario (1986), Boyanov & Chollet (1992), Boyanov, Doskov, Ivanov, & Hadjitodorov (1991), Herzel, Berry, Titze, & Saleh (1994), Titze (1993), Bergè, Pomeau, & Vidal (1986), Holden (1986), Glass & Mackey (1988), etc.

1.1 Objetivo

O objetivo desse trabalho é analisar e quantificar sinais de vozes normais e patológicas sob o enfoque da dinâmica não linear. O estudo foi feito sobre a dispersão das trajetórias do espaço de fase de vozes normais, vozes com nódulo e vozes com edema de Reinke através da técnica de análise não linear chamada seção de Poincaré. O cálculo da dispersão foi feito utilizando o conceito estatístico de desvio padrão em relação ao ponto médio da seção e em relação à distribuição dos pontos sobre os eixos coordenados y e z .

2 Formação da Voz

A fala humana é um fenômeno que existe desde o nascimento e se apresenta de diversas formas como choro, grito, riso e sons da fala. É um dos meios de comunicação do indivíduo com o exterior e particularmente com seus semelhantes.

A produção da voz é feita através de movimentos sincronizados e articulados que envolvem o sistema respiratório, a laringe - que contém pregas vocais - e o trato vocal, que representa o sistema de ressonância para produção dos sons. Os pulmões, brônquios e traquéia são a fonte de energia necessária para a produção da fala através do fluxo de ar produzido. Conforme descrito por Flanagan (1972), a formação da voz é um sistema controlado por realimentação acústica, através dos ouvidos, e realimentação cinestésica dos músculos da fala. Todo esse processo é controlado e coordenado pelo sistema nervoso central para a produção da voz. A Figura 1 representa o corte médio-sagital através do trato vocal de um adulto contendo os principais elementos envolvidos na produção da fala.

Figura 1: Esquema do mecanismo vocal humano [Flanagan,1972 - Pg10].

Os elementos do mecanismo vocal humano indicados são 1.Palato mole 2.Osso hióide 3.Epiglote 4.Cartilagem cricóide 5.Cavidade nasal 6.Palato duro 7.Língua 8.Cartilagem tiróide 9.Pregas vocais 10.Traquéia 11.Pulmões 12.Esterno.

Especificamente, a laringe que contém as pregas vocais, está situada no pescoço ao nível da 3ª e 6ª vértebras cervicais, e visível na região onde se localiza o pomo de Adão. Ela consiste de 9 cartilagens denominadas tireóide, cricóide e epiglote, que são as cartilagens simples; e as cartilagens parelhadas denominadas aritenóide, corniculada e cuneiforme. Sua principal função é impedir que corpos estranhos atinjam os pulmões pela ação do fechamento da epiglote, e auxiliar na deglutição. A Figura 2 apresenta mais detalhes da laringe, mostrando as visões anterior e posterior.

A articulação das cartilagens aritenóides com a cartilagem cricoide permite que as pregas vocais se aproximem ou se afastem deixando passar o ar entre elas. De modo semelhante, as articulações entre a tireóide e cricóide permitem que as pregas se alonguem e se encurtem aumentando ou diminuindo sua tensão. A ação das cartilagens, ao aproximar as pregas vocais, permitem que estas vibrem durante a passagem do ar criando uma variada gama de sons que apresenta a fala humana através do trato vocal. A produção dos sons é feita pela vibração da pregas vocais localizadas na laringe sendo radiada e filtrada

Figura 2: Vista posterior e anterior dos principais componentes da laringe.

através desse trato quando o ar é expelido pelos pulmões. O fluxo de ar resultante é perturbado pelo estreitamento ou abertura da região do trato vocal, envolvendo a cavidade bucal. A cavidade bucal atua como um filtro ampliando ou atenuando as frequências produzidas durante a fala. Na Figura 3 tem-se as pregas vocais nas posições aberta e fechada.

Figura 3: Esquema das pregas vocais em posição respiratória (aberta) e em posição fonatória (fechadas). Figura modificada de *Mayo Foundation for medical education e Research*

A mecânica da produção da voz começa com os músculos das pregas vocais inicialmente em repouso. Uma pressão vinda diretamente dos pulmões age abaixo da glote forçando os músculos das pregas vocais a se afastarem um do outro num movimento acelerado. Conforme o fluxo de ar passa por entre as pregas vocais, a pressão local é reduzida conforme a equação de *Bernoulli*, e força as pregas vocais a voltarem à sua posição inicial.

Uma vez que as pregas estão novamente próximas uma da outra, o fluxo de ar diminui enquanto que a pressão subglotal estática aumenta. A partir daí o ciclo é novamente repetido obtendo-se a onda de mucosa. Esse movimento das pregas vocais determina a frequência do sinal de voz que é da ordem 125 Hz para o homem e 250 Hz para as mulheres (Behlau, 2001). O mecanismo de vibração das pregas vocais permite um pulso glotal quase periódico suficiente para excitar o sistema acústico acima das pregas vocais. A Figura 4 mostra uma sequência dos pulsos glóticos de um sinal vocal.

Figura 4: Forma de onda do pulso glotal

O pulso glótico pode ser verificado utilizando técnicas de filtragem inversa sobre o sinal de voz, onde todo o efeito da radiação labial e tato vocal são extraídos.

O espectro de potência do sinal glotal decresce com o aumento da frequência. Para produção da fala esse sinal é filtrado no trato vocal atenuando ou amplificando as componentes de frequência do sinal. Em seguida, o sinal é radiado dos lábios, que também filtra o sinal. A partir desse ponto tem-se a produção da fala do modo como a ouvimos. O pulso glotal na figura 4 é apenas ilustrativa mas pode ser extraído através da filtragem inversa de sinal de voz real conforme realizado por Rosa, Pereira, & Grellet (2000).

Alguns modelos foram propostos na literatura para se representar um sinal de voz. O modelo clássico, chamado modelo linear, foi proposto por Fant (1960) e Davis (1979), onde o sinal de voz é representado como o resultado de uma sequência de pulsos periódicos (excitação) filtrado pela glote, o trato vocal e os lábios. Esse modelo é representado na Figura 5.

O efeito da radiação labial, $R(Z)$, pode ser aproximado pela função dada por (Rabiner & Schafer, 1978)

$$R(Z) = R_0(1 - Z^{-1}) \quad (2.1)$$

A função de transferência $V(Z)$ define a estrutura de um modelo cujos parâmetros variam lentamente com o tempo. $G(Z)$ representa o modelo do pulso glotal. Esse modelo

Figura 5: Modelo para o trato vocal modificada

faz uma aproximação da produção da voz através da convolução do pulso glotal periódico, representado pela velocidade do fluxo de ar através da glote, com a resposta ao impulso. Por essa representação, pode-se verificar que o sinal de voz é o resultado de uma sequência de pulsos periódicos (excitação) filtrado pela glote, pelo trato vocal e pelos lábios.

Os modelos acima são representações para voz normal. Para vozes patológicas Deliyski (1993) desenvolveu um modelo acústico para descrever os efeitos que ocorrem na forma de onda nesse tipo de voz. Nesse trabalho foi feita uma avaliação quantitativa das características que usualmente são medidas como frequência média fundamental, relação ruído harmônico, componentes sub-harmônicas, dentre outras.

O método clássico para descrever a acústica da fala baseia-se na análise de perturbações de período e amplitude do sinal, respectivamente chamados de *Jitter* e *shimmer*, além da análise de ruídos. No entanto, a voz real contém componentes irregulares devido provavelmente à natureza caótica do mecanismo da Laringe (Deliyski, 1993).

O sistema vocálico também pode ser interpretado como um sistema fonte-filtro sendo a fonte o conjunto pulmão+ traquéia e o filtro, o trato vocal. A partir dessas considerações Flanagan (1972) propôs o modelo da Figura 6.

A motivação para se construir modelos para produção de voz é devido à possibilidade de controlar os parâmetros que envolvem o sinal de voz, permitindo a produção e estudo de sinais com irregularidades características de vozes patológicas (Jiang, Zhang, & McGilligan, 2006).

2.1 Classificação dos sons da fala

Os sinais vocais podem ser classificados dentro de 3 classes de acordo com seu modo de excitação, descritos a seguir.

Figura 6: Sistema Sistema fonte filtro para a formação da voz

- Sons vocálicos: são produzidos pela passagem do ar através da glote com a tensão das cordas vocais ajustadas para vibrar num modo "relaxado". Isso produz pulsos de ar quase periódicos que excitam o trato vocal. Exemplos desses tipos de sons são as vogais;
- Sons fricativos: são produzidos pela formação de uma constrição parcial e contínua em algum ponto no trato vocal, geralmente na parte frontal da boca, e forçando a passagem do ar através dessa constrição em alta velocidade, produzindo uma turbulência. Esse tipo de som produz uma fonte de ruído com uma faixa larga no espectro de frequência para estimular o trato vocal. Exemplos desses tipos de sons são os fonemas "f", "v", etc;
- Plosivos: Sons plosivos resultam do fechamento completo e momentâneo, da boca por exemplo, criando uma pressão por trás dos lábios e abertura abrupta. Exemplos desse tipo de sons são os fonemas "p", "b", etc.

2.2 Tipos de Vozes

Titze (1995) classificou os sinais de vozes em Tipo 1, Tipo 2 e Tipo 3 de acordo com as características em sua periodicidade intrínseca. Vozes do Tipo 1 são sinais quase periódicos; vozes do Tipo 2 são sinais que contêm modulações e subharmônicas e do Tipo 3 são sinais irregulares e aperiódicos. As formas de onda típica desses sinais estão indicadas na Figura 7.

Figura 7: Formas de onda típicas de acordo com a classificação de Ingo R. Titze

2.3 Voz normal e voz patológica

A maioria das doenças na laringe e nas pregas vocais causa mudanças significativas no processo de produção da fala. Infecções respiratórias, uso inadequado e excessivo da voz e falta de cuidados simples são algumas das causas que levam à formação de uma patologia. Fatores anatomofuncionais predisponentes, como proporção glótica, assimetria laríngea e o ângulo de abertura das pregas vocais também podem causar algum tipo de patologia.

Essas patologias podem causar problemas na voz devido a uma série de fatores como interferência no contato das pregas vocais durante a fase fechada do ciclo glótico, escape de ar, ou soprosidade, em consequência da patologia e alterações teciduais nas camadas das pregas vocais, que inibem ou suprimem a onda de mucosa (Harris, 1998).

Vários procedimentos clínicos para detecção de patologias na laringe são de natureza invasiva, que podem causar desconforto aos pacientes; mas a análise acústica se tornou uma boa ferramenta para análise não invasiva dos sinais de voz. Essas análises podem ser feitas no domínio do tempo, da frequência e sob o enfoque da dinâmica não linear, uma

vez que a voz apresenta não linearidades (Herzel & Wendler, 1991). Devido à essas não linearidades, uma voz patológica pode apresentar comportamento caótico.

Uma voz normal, sob o ponto de vista perceptivo-auditivo, pode ser entendida como uma voz agradável, melódica, audível e coerente, sendo essencial para a boa comunicação permitindo a cada indivíduo ajustar-se às situações sociais e diferentes tipos de ambiente. Já uma voz com alguma patologia, como aspereza ou rouquidão grave pode causar constrangimento à pessoa que fala.

Dentre as patologias mais comuns, pode se citar aquelas de origem organofuncional como os nódulos, pólipos, edema de *Reinke*, úlcera de contato, granuloma e leucoplasia. A seguir, uma breve descrição de nódulos vocais e edema de Reinke, que são as patologias estudadas nesse trabalho.

2.4 Nódulos vocais

Os nódulos são lesões de massa, benignas, bilaterais, de característica esbranquiçada ou levemente avermelhada. São lesões muito comuns em mulheres adultas jovens, entre 25 e 35 anos de idade, e em crianças de ambos os gêneros, devido às características peculiares da função vocal e dos órgãos que participam da produção vocal, especialmente a conformação laríngea (Gindri & Cielo, 2008).

Os nódulos vocais desenvolvem na região anterior das pregas vocais, na metade da área de maior vibração glótica, decorrentes essencialmente do abuso vocal. Acredita-se que o nódulo vocal seja a lesão laríngea mais comum (Herrington-Hall, Lee, Stemple, Niemi & Mchome, 1988). Esse tipo de lesão nas pregas vocais pode causar um movimento de comportamento caótico entre as pregas (Jiang et al, 2006). A Figura 8 mostra uma imagem típica de nódulos em adultos.

De modo geral, rouquidão e sopro são os principais sinais perceptivo-auditivos da presença de nódulos vocais. O grau de rouquidão ou sopro depende do tamanho do nódulo e da rigidez de seus tecidos. A rouquidão correlaciona-se ao grau de irregularidade de vibração das pregas vocais, enquanto que a sopro refere-se ao fechamento glótico incompleto. Sintomas típicos para portadores de nódulos são fadiga vocal, perda de potência da voz com o uso, dor na laringe ou no pescoço, podendo ainda causar dificuldades em produzir notas agudas.

Figura 8: Nódulos vocais

2.5 Edema de Reinke

O edema de *Reinke* é uma lesão difusa na camada superficial da prega vocal, de coloração rosada, caracterizada por acúmulo de fluido, de modo irregular em alguma região da porção membranosa ou em toda prega. Também é chamado de hipertrofia edematosa crônica, dentre outros nomes (Behlau, 2001).

A causa do edema é devido à uma reação natural do tecido ao trauma fonatório associado ao consumo de tabaco por longo tempo e se traduz por uma distensão do espaço de *Reinke*, pela presença de material mucóide, como resultado de uma irritação crônica (Shapsay & Aretz, 1984). Ocorrem em pessoas adultas de ambos os sexos, entre 45 e 65 anos de idade, que apresentam uma frequente associação de uso intensivo da voz, abusos vocais variados e tabagismo. A Figura 9 mostra uma imagem típica de edema de Reinke em adultos.

Figura 9: Edema de *Reinke*

Os sintomas típicos dessa patologia incluem voz grave para a idade e sexo do paciente e rouquidão (Colton & Casper, 1996), fato que pode tornar a voz feminina masculinizada. O edema produz um *pitch* grave e modulação restrita com extensão fonatória e dinâmica limitadas. Bennet, Bishop & Lumpkin (1987) obtiveram frequência fundamental média de 108 Hz para um grupo de 29 mulheres e uma média de 91 Hz para um grupo de 6 homens. Os valores de *Shimmer* apresentam-se geralmente bastante elevados, característica da lentificação da vibração das pregas vocais e do ruído no espectro, devido à rouquidão. Em casos severos da patologia pode haver incoordenação pneumofônica e queixa respiratória.

3 Análise acústica de sinais de voz

A Análise acústica de sinais de voz é um eficiente método não invasivo para avaliação e detecção de vozes patológicas (Boyanov & Hadjitodorov, 1997). Vários métodos e modelos baseados na dinâmica não linear foram desenvolvidos para analisar sinais de vozes normais e patológicas. Dentre alguns métodos pode-se citar cálculo da Dimensão de Hausdorff, Dimensão da informação e Dimensão de correlação; cálculo da entropia, como Entropia de *Kolmogrov* e entropia de segunda ordem, reconstrução do espaço de fase, expoente de *Lyapunov* e Seção de Poincaré.

A Reconstrução do espaço de fase, entropia de segunda ordem e dimensão de correlação foram utilizados por Yu Zhang (2004) para análise de vozes de pacientes com pólipos antes e depois da extração do pólipo. Nesse trabalho, foi verificada uma significativa redução na dimensão de correlação e na entropia de segunda ordem. Nesse mesmo estudo, o *Jitter* e o *Shimmer* também foram comparados e foi constatado que não houve diferença na medida do *Shimmer*. No entanto, o *Jitter* apresentou redução após a extração do pólipo. Entretanto, Zeitels et al (2002) não verificou redução do *Jitter*, mas verificou a redução do *Shimmer*.

De acordo com Titze (1995) e Karnell et al (1997) medidas de *Jitter* e *Shimmer* não são confiáveis para sinais fortemente aperiódicos, o que poderia ocorrer em sinais de vozes com pólipos. Titze & Liang (1993) e Titze (1995) sugerem que essas medidas sejam confiáveis somente para casos onde elas não ultrapassem 5% , e podem se tornar não confiáveis quando se verifica grandes variações da forma de onda do sinal vocal.

Herzel, Berry, Titze, & Saleh (1994) analisaram vozes patológicas de vários pacientes utilizando técnicas tradicionais (*Jitter*, *Shimmer*, *pitch*, avaliação perceptiva, etc) e métodos de análise não linear e concluíram que as mudanças abruptas na frequência do sinal de voz podem ser melhor estudadas e compreendidas sob o aspecto da dinâmica não linear. Deliyski (1993) também mostrou os efeitos de não linearidade presentes na forma de onda de vozes patológicas através de um modelo acústico onde as irregulari-

dades na amplitude e na frequência fundamental, as componentes sub-harmônicas, ruídos e voz crepitante foram expressos como resultado de uma influência randômica na função de excitação glotal.

Outros estudos também mostraram que turbulências e não linearidades de vozes patológicas são características de sinais de vogais sustentadas. Esses estudos mostram que vozes patológicas apresentam comportamento caótico cujo comportamento turbulento pode ser representado como um aumento da rouquidão (Boyanov & Hadjitodorov, 1997).

A vantagem de se utilizar técnicas de análise não linear tanto para vozes normais quanto vozes patológicas é que mesmo a representação dos modelos mais simples das pregas vocais apresentam não-linearidades (Ishizka & Flanagan, 1972). A partir desses modelos é possível analisar vozes patológicas e normais simuladas e comparar com vozes reais. Alguns dos métodos tradicionais de análise de um sinal de voz são descritos a seguir.

3.1 Análise Perceptivo auditiva

Essa avaliação se baseia no julgamento que um avaliador treinado faz a respeito das características de uma voz. Embora, questionada por ser uma avaliação subjetiva, a Avaliação Perceptiva Auditiva é uma ferramenta fundamental na avaliação vocal e faz parte da prática dos profissionais da área de fonoaudiologia. As escalas para avaliação de voz mais difundidas são o modelo fonético de descrição da qualidade vocal e o modelo de escala "GRBAS", criado pelo comitê para testes de função fonatória da sociedade japonesa de logopedia e foniatria em 1969.

3.2 Análise espectrográfica de um sinal de voz

A análise espectrográfica é útil para visualização das frequências presentes no sinal. Sons quase periódicos, como das vogais, são representados pela superposição de várias componentes senoidais, cada uma com sua frequência, amplitude e defasagem. As Figuras 10 e 11 mostram um sinal de voz normal e patológica (nódulo) e seus espectros.

Pode-se observar na Figura 11 a redução das harmônicas no espectro do sinal devido à imperfeição do pulso glotal desse tipo de patologia.

Figura 10: Sinal de voz normal e seu espectro

Figura 11: Sinal de voz com nódulo e seu espectro

3.3 Medição da Frequencia Fundamental

A frequência fundamental, F_0 , corresponde ao número de ciclos vibratórios das pregas vocais por segundo. Do ponto de vista perceptivo ela é referida como *pitch*. Na maioria das vozes normais, o *pitch* é calculado utilizando os pontos de máximo do sinal. No entanto, em vozes patológicas, os pulsos de máximo podem consistir de 2 ou 3 sucessivos pulsos (Boyanov & Hadjitodorov, 1997)].

A informação básica para a extração de medidas de F_0 é a demarcação de cada ciclo

glótico. Uma abordagem comum para determinar o valor de F0 de cada ciclo é através da comparação da forma de onda de dois segmentos adjacentes e de mesmo comprimento. Esse método utiliza o coeficiente de correlação entre os trechos analisados mas demonstrou-se eficiente apenas para sinais periódicos.

Boyanov & Hadjitodorov (1997) propuseram o cálculo do *pitch* utilizando o GCM (Centro de massa geométrico). O GCM leva em consideração a forma de onda do pico do sinal e concentração de energia no *pitch*. Também foram utilizados algoritmos baseados nos pontos de pico e cruzamento do sinal pelo zero (Vieira & Mc Innes FR, 1997). Outros trabalhos para detecção de F0 também foram realizados por Milenkovic (1987) e Titze & Liang (1993).

3.4 Cálculo do *Shimmer*

O *shimmer* é uma medida das variações ciclo a ciclo na intensidade da excitação glótica. De acordo com Tsuji I.D.; Inamura, R.; Ubirajara, LS (2006), o *shimmer* de dois ciclos consecutivos A(1) e A(2) pode ser calculado pela função de perturbação de primeira ordem como

$$S(\%) = 2 \frac{|A(1) - A(2)|}{|A(1) + A(2)|} * 100 \quad (3.1)$$

A obtenção do *shimmer* através dessa expressão é feita tomando-se a média de todos os pares de períodos glóticos vizinhos.

Uma outra medida para o *Shimmer* mais abrangente que a anterior foi utilizada por Zhan et al (2004). Nesse trabalho o *Shimmer* foi calculado usando *Mult Dimensional Voice Program* (MDVP, Kay Elemetrics) e a expressão a seguir

$$S(\%) = \frac{\frac{1}{k-1} \sum_{i=1}^{(k-1)} |A_0^i - A_0^{(i+1)}|}{\frac{1}{k} \sum_{i=1}^k A_0^{(i)}} \quad (3.2)$$

com $i=1,2,3,\dots,k$ e k igual ao número de período de *pitch*

Uma proposta para medição de *Shimmer* foi feita por Dajer, Sobrinho, & Pereira (2009) calculando o desvio padrão da dispersão dos pontos na Seção de Poincaré correspondentes a todos os pontos de máximo do sinal temporal. Posteriormente, Dajer,

Sobrinho, & Pereira (2010) desenvolveram um algoritmo que calcula o desvio padrão da dispersão da seção de Poincaré correspondente à vários trechos do sinal temporal em seu respectivo espaço de fase tridimensional. Esse algoritmo demonstrou ser mais promissor que o anterior porque utiliza vários trechos do sinal no tempo, e não apenas os pontos de máximo.

As alterações de *Shimmer* ocorrem em patologias que apresentam lesões de massa nas pregas vocais, como nódulos, pólipos, edemas e carcinomas. A característica acústica associada a esse parâmetro é a rouquidão (Andrade et al., 2002). A Figura 12 ilustra o *Shimmer* em um sinal de voz entre os 2 picos em destaque.

Figura 12: Perturbação em amplitude (Shimmer)

3.5 Cálculo do *Jitter*

O Jitter é uma medida que procura quantificar as variações ciclo a ciclo de cada período de um sinal de voz. Lieberman (1963) definiu Fator de Perturbação de Frequência (PFF - *Period Frequency Factor*) como sendo a frequência relativa das perturbações do período maiores do que 0,5 ms. Behlau (2001) propõem uma medida para o *Jitter* considerando dois períodos glóticos consecutivos de frequência $F0(1)$ e $F0(2)$, de acordo com a expressão abaixo

$$J(\%) = 2 \left| \frac{F0(1) - F0(2)}{F0(1) + F0(2)} \right| * 100 \quad (3.3)$$

A análise completa é feita tomando-se todos os pares de ciclos góticos e a medida final será a média. Devido à sensibilidade do *Jitter* à demarcação do ciclo glótico também pode-se utilizar mais de dois ciclos para diminuir os erros na demarcação dos períodos. Da mesma forma que o *Shimmer*, o *Jitter* também possui uma expressão mais abrangente utilizada por Zhan et al (2004) dada pela expressão a seguir

$$J(\%) = \frac{\frac{1}{k-1} \sum_{i=1}^{(k-1)} |T_0^i - T_0^{(i+1)}|}{\frac{1}{k} \sum_{i=1}^k T_0^{(i)}} \quad (3.4)$$

com $i=1,2,3,\dots,k$ e k igual ao número de período de *pitch*.

Apesar dos estudos que procuram associar o *Jitter* vocal com patologias na laringe e com o aumento na aspereza vocal, deve haver cuidado no uso desta medida uma vez ela perde a confiabilidade quando a aperiodicidade aumenta.

Ao *Jitter* estão relacionadas a presença de massa ou doenças neurológicas e psicológicas (Scalassara, 2009) e também está associado ao aspecto perceptivo de aspereza (Andrade et al.,2002). Pela Figura 13 pode-se analisar o *Jitter* como sendo a média das variações de cada do sinal estudado.

Figura 13: Perturbação em Período (Jitter)

3.6 Relação Sinal Ruído

Há vários métodos para quantificar o nível de ruído produzido pela passagem turbulenta do ar pelas pregas vocais. Essas medidas estão associadas à sopro-sidade da voz. A relação sinal ruído num sinal de voz mede a energia dos harmônicos, que estão associados à vibração glótica periódica, e a energia inter-harmônica devido à ruído glótico e sub-harmônicos.

Para realizar essas medidas, podem ser usados espectros de banda estreita ou processamento direto do registro temporal. Utilizando o registro temporal, faz-se inicialmente a demarcação de todos os ciclos glóticos; em seguida, obtém-se um ciclo de referência (sinal), que é a média de todos os ciclos da elocução. Nesta média o ruído tende a ser cancelado. Por fim, cada um dos outros ciclos demarcados é subtraído do ciclo médio e a diferença (ruído) é acumulada. A expressão para a relação sinal ruído então torna-se

$$SNR(dB) = 10\log_{10}\left[\frac{\textit{ciclo médio}}{\textit{ciclo ruidoso} - \textit{ciclo médio}}\right] \quad (3.5)$$

Onde o ciclo médio e o ciclo ruidoso referem-se às energias dos ciclos, isto é, a soma do quadrado de cada amostra do período.

Outros parâmetros no domínio do tempo também são propostos na literatura como o Fator de Perturbação Direcional (DPF - Directional Perturbation Factor) (Hecker & Kruel, 1970), Perturbação Média Relativa (RAP - *Relative Average Perturbation*) (Koike, Takahashi, & Cacaterre, 1977); Jitter Médio (JM - *Jiter Media*) e Razão de Jitter (JR - *Jitter Ratio*) (Horii, 1979); Fator de Jitter (JF - Jitter Factor) (Murry & Doherty, 1980) dentre outros.

4 Técnicas de Análise não-linear

Alguns métodos de análise de sinais de voz tem como objetivo separar e remover o ruído existente no sinal. Essas técnicas são importantes para melhorar a transmissão de voz através de um canal de telefonia e também para análise e reconhecimento de voz.

As técnicas convencionais utilizam as características do espectro do sinal e do ruído para separar e eliminar o ruído pelo processo de filtragem. Uma alternativa para essas análises é estudar o sinal de voz sob o contexto da dinâmica não linear. Os métodos não lineares para análise de sinais de voz no tempo utilizam o espaço de fase reconstruído, que é um vetor no domínio do tempo cuja dimensão são amostras do sinal original com um tempo de atraso pré determinado.

Essa técnica permite visualizar a dinâmica do sinal temporal, chamada de trajetórias ou atratores; fornecendo mais informações do que a representação do sinal no domínio da frequência (Kantz & Schreiber, 2004).

4.1 Reconstrução do espaço de fase

Os métodos de análise não linear em série temporais são baseados na teoria dos sistemas dinâmicos, onde a evolução da série no tempo é definida no espaço de fase. Sistemas dinâmicos são definidos por um conjunto de equações ordinárias de primeira ordem agindo sobre o espaço de fase. A teoria matemática das equações diferenciais garante a existência de uma única trajetória, dada uma condição inicial.

O espaço de fase, ou espaço de estado, pode ser considerado como um vetor de dimensão finita R_m . Um estado é especificado por um vetor $x \in R_m$. A dinâmica do sistema pode então ser descrita por um sistema de m equações diferenciais de primeira ordem contínuo (ou discreto) de acordo com as equações a seguir

$$\frac{dx(t)}{dt} = f(x(t)), \quad t \in R \quad (4.1)$$

Para o caso discreto tem-se:

$$x_{n+1} = F(x_n), \quad n \in Z \quad (4.2)$$

O vetor f na equação 4.1 por definição não depende do tempo. Nesse caso, o sistema é chamado autônomo. A existência e unicidade da solução da equação 4.1 depende do vetor ser contínuo (caso autônomo). A solução, $x(t)$, da equação 4.1 é chamada de trajetória do sistema dinâmico, com $x(0)$ sendo a condição inicial.

Através do espaço de fase reconstruído Lindgren, Johson, & Povinelli (2003) desenvolveram um método para reconhecimento de voz. Neves, Fernandes, & Restivo (2006) melhoraram processos de filtragem utilizando o espaço de fase de uma série temporal.

Para a construção do espaço de fase é necessário conhecer as variáveis dinâmicas que governam o sistema; mas nem sempre essas variáveis são conhecidas ou estão disponíveis, como é o caso da dinâmica vocal, que é um sistema dinâmico complexo (Alonso et al, sd). Num sinal de voz tem-se disponível apenas a série temporal; para esse caso, deve-se converter a série temporal em vetores de estado para reconstruir o espaço de fase utilizando a técnica de tempo de atraso (Packard, Crutchfield, Farmer, & Shaw, 1980).

O espaço de fase reconstruído é formado por vetores em R_m , cujos elementos são amostras atrasadas da série tempo original conforme indicado na expressão

$$X = \{x[n], x[n - \tau], x[n - 2\tau], \dots, x[n - (m - 1)\tau]\} \quad (4.3)$$

onde m é a dimensão de imersão e τ é o tempo de atraso.

As amostras na série de tempo devem possuir um tamanho ideal para que o espaço de fase reconstruído seja equivalente ao atrator original. A garantia para que o atrator reconstruído seja semelhante ao original é estabelecida quando $m > 2D + 1$, onde D é a dimensão de Hausdorff (Takens, 1980).

4.2 Tempo de atraso

Um parâmetro importante para reconstrução do espaço de fase é o tempo de atraso, τ . Inicialmente, imagina-se que o tempo de atraso τ , podia ser escolhido aleatoriamente (Takens, 1980). Contudo, experimentos mostraram que a qualidade do atrator reconstruído dependia da escolha correta de τ . Ainda nos estudos iniciais sobre a reconstrução

do espaço de fase, não havia na literatura nenhum critério para a escolha de τ .

De acordo com Fraser & Swinney (1986), se a escolha de τ para a reconstrução do espaço de fase for muito pequena, as trajetórias serão definidas por uma linha dada por $x_0(t) = x_1(t)$; nesse caso os dados estariam correlacionados. Isso é chamado de redundância em (Casdagli et al., 1991a) e (Gibson et al., 1992). A escolha de τ deve garantir que as trajetórias no espaço de fase sejam linearmente independentes. Se τ for muito grande, perde-se a correlação e os pontos no espaço de fase ficarão muito dispersos.

Uma outra escolha de τ para garantir a independência linear é o valor na qual a função de autocorrelação passa pelo zero. A desvantagem desse método é que ele é baseado na estatística linear, no entanto as pregas vocais possuem característica não linear (Herzel H., 1993); (Jiang; Zhang, 2001), (Awrejceicz, 1990 p. 151 - 156), (Berry; Herzel; Titze, Krisher, 1994 p.3595-3604), (Jiang; Xhang, 2002 p.2127-2133).

A determinação da dimensão de imersão usando decomposição em valores singulares também pode ser utilizada em sistemas reais, onde há presença de ruídos. Nessa técnica, o espaço de estado é dividido em 2 dois subespaços, um contendo o sinal e o outro, contendo o ruído. Eliminando o ruído, reduz-se a dimensão e remove-se os efeitos indesejados, obtendo-se o atrator reconstruído (Broomhead & King, 1986).

Um critério que pode ser utilizado é o que verifica a geometria do atrator. Nesse método, o atrator deve estar "desdobrado", ou seja, a extensão do atrator no espaço deve ser aproximadamente a mesma (Kantz & Schreiber, 2004).

Todos os métodos estudados apresentaram bons resultados, mas o melhor tempo de atraso foi determinado pelo método da informação mútua (Fraser & Swinney, 1986). Nesse estudo, foi reconstruído o espaço de fase da reação chamada "Belousov-Zhabotinskii" utilizando os dois critérios para escolha de τ : o primeiro mínimo da informação mútua do sistema e o primeiro zero da função de autocorrelação. Foi verificado que o critério do primeiro mínimo da informação mútua resultava um atrator mais próximo do original. O atrator original, chamado de atrator *Rössler*, pôde ser reconstruído pois suas variáveis dinâmicas são disponíveis e dadas por

$$\dot{x} = -y - z \quad (4.4)$$

$$\dot{y} = x + ax \quad (4.5)$$

$$\dot{z} = b + z(x - c) \quad (4.6)$$

Onde a , b e c são valores previamente estabelecidos.

4.3 Informação mútua

Se as duas variáveis estiverem defasadas de um tempo τ , a informação mútua é a informação que se tem sobre o valor $X(t + \tau)$ se o valor $X(t)$ for conhecido. O valor $X(t + \tau)$ é a máxima informação obtida a respeito de $X(t)$.

Uma expressão para determinação da Informação Mútua é atribuída à Shanon (1968). Dado um conjunto de dados, cria-se um histograma de resolução ϵ para a distribuição de probabilidade de dados. Indicando por p_i a probabilidade do sinal assumir um valor dentro do i -ésimo intervalo do histograma, e seja $p_{ij}(\tau)$ a probabilidade que $s(\tau)$ esteja no intervalo i e $s(t + \tau)$ esteja no intervalo j . A informação mútua para o tempo de atraso é dada por

$$I(\tau) = \sum_{ij} p_{ij}(\tau) \cdot \ln p_{ij}(\tau) - 2 \sum_i p_i \ln p_i \quad (4.7)$$

4.4 Seção de Poincaré

Uma técnica clássica para análise de sistemas dinâmicos é devido a Henry Poincaré. Essa técnica transforma um sistema dinâmico contínuo no tempo para um sistema discreto utilizando o mapa de Poincaré ou seção de Poincaré.

Considera-se a representação de um sistema dinâmico autônomo, n -dimensional em seu espaço de fase. A seção de Poincaré reduz a dimensão desse espaço de fase para um sistema $(n - 1)$ dimensional e é encontrada realizando um corte perpendicular às trajetórias do espaço de fase e tomando somente os pontos que interceptam essa seção.

Herzel & Wendler (1991) observaram que a presença de estruturas regulares na seção de Poincaré de sinais de voz era uma indicação da presença de uma dinâmica de baixa dimensão. Em sinais de vozes fortemente soprosas a seção de Poincaré apresentou um aspecto nebuloso, indicando um comportamento caótico. Pela seção de Poincaré é possível verificar que as trajetórias de alguns sistemas dinâmicos no espaço de fase nunca repetem o mesmo percurso, mas estão confinadas dentro de um feixe bem limitado.

Pela seção de Poincaré também é possível determinar o *pitch* de um sinal temporal (Mann & Laughlin, 1997). O principal argumento é que uma revolução no espaço de fase

reconstruído é igual a um período de *pitch*. Em outro estudo, Propes (1996) utilizou a seção de Poincaré para analisar o comportamento de longo termo de um sistema periódico forçado, no caso um circuito Diodo-R-L.

Roux, Simoyi, & Swinney (1983) estudaram o atrator reconstruído de uma reação química homogênea não equilibrada. No atrator resultante, extraíram a seção de Poincaré e constataram o alongamento e dobramento do atrator, característica típica dos atratores estranhos, que se caracterizam por ter dimensões não inteira (dimensão fractal), e dinâmica com comportamento caótico (Ruelle & Takens, 1971).

4.4.1 Determinação da seção de Poincaré

A seção de Poincaré de um sinal pode ser obtida utilizando o algoritmo proposto (Gernot Kubin, 1997), conforme o procedimento descrito abaixo.

1. Construir uma trajetória N-dimensional $X(n)$, ou espaço de fase, de um sinal de voz $x(n)$ observado numa janela estacionária dada por

$$X = \{x[n], x[n - \tau], x[n - 2\tau], \dots, x[n - (m - 1)\tau]\} \quad (4.8)$$

2. Selecionar uma vizinhança do espaço de fase $x(n_0)$, e a partir dessa vizinhança, selecionar todos os pontos $x(n)$ com um raio R ao redor de um dado ponto $x(n_0)$;
3. Estimar, com base nesses pontos, um vetor normal $f(n_0)$ à um plano definido pelos pontos, cuja orientação seja a da trajetória do espaço de fase;
4. Definir o plano de Poincaré que irá percorrer por toda trajetória do espaço de Fase e que seja perpendicular à trajetória;
5. Encontrar os pontos de intersecção da trajetória com o espaço de fase;
6. Aplicar uma transformação local de coordenadas dos pontos de intersecção. Usar vetores singulares como nova coordenada.

Por meio desse algoritmo, Kubin utilizou o sinal de voz representado na Figura 14.

Dessa série temporal foi reconstruído o espaço de fase e definido o plano de corte ao longo da trajetória. O plano de Poincaré percorre toda a trajetória com o ângulo de fase variando de 0° a 360° , conforme indicado na Figura 15.

Figura 14: Vogal o

Figura 15: Espaço de fase reconstruído obtido através de 10000 amostras da vogal sustentada /o/. Os cortes em vermelho é o plano de Poincaré (Kubin, 1997).

A Figura 16 representa as 20 seções de Poincaré resultante da reconstrução do espaço de fase da vogal /o/.

Figura 16: Seção de Poincaré obtida a partir da reconstrução do espaço da fase da vogal /o/ (Kubin, 1997).

Uma proposta para se traçar a seção de Poincaré foi feita por Dajer, Sobrinho, & Pereira (2009) onde foi utilizado os pontos correspondentes aos pontos de máximo do sinal temporal. Posteriormente, Dajer, Sobrinho, & Pereira (2010) desenvolveram um algoritmo

que traça a seção de Poincaré correspondente à vários trechos do sinal temporal em seu respectivo espaço de fase tridimensional. Esse algoritmo demonstrou ser mais promissor que o anterior porque utilizou vários trechos do sinal no tempo, e não apenas os pontos de máximo.

4.5 Análise das Componentes Principais - PCA

A análise por componentes principais, também chamada de Transformação *Karhunen-Loève*, decomposição em valor singular ou função ortogonal empírica, tem com princípio caracterizar a série temporal pelos seus componentes mais relevantes no espaço R_m , onde m é dimensão de imersão. Para o cálculo das componentes principais deve-se encontrar a matriz de covariância $m \times m$ dada pela expressão

$$C_{ij} = \langle (s)_i (s)_j \rangle = \frac{1}{n - m + 1} \sum_{n=1}^{N-m+1} s_{n-m+i} s_{n-m+j} \quad (4.9)$$

Como a matriz C_{ij} é real e simétrica, seus autovalores são reais e seus autovetores são ortogonais. Os autovalores c_{ij} são os comprimentos ao quadrado dos semi-eixos da elipsóide que melhor se encaixa na "nuvem" de pontos plotados; os correspondentes autovetores dão a direção dos eixos. Como resultado desses cálculos, as componentes principais fazem a translação da origem do sistema original de coordenadas de modo que o novo eixo passe pela nuvem de pontos nas direções de máxima variância.

Nesse trabalho, as componentes principais são utilizadas para descorrelacionar os dados da seção de Poincaré para se obter o desvio padrão da dispersão dos pontos seção. A Figura 17 mostra um exemplo de rotação do eixo original através das componentes principais.

4.6 Caos na produção de voz

Caos é um termo que descreve um comportamento pseudo-randômico gerado por um sistema que é ao mesmo tempo determinístico e não linear. A saída de um sistema caótico, apesar de ser regido por uma equação determinística, é extremamente sensível às condições iniciais, resultando em comportamentos imprevisíveis. O estudo de atividades caóticas em sistemas fisiológicos sugerem que as mudanças ocorridas em sistemas não lineares podem indicar presença de alguma patologia.

Figura 17: Rotação dos eixos originais através das componentes principais

Nas últimas duas décadas, os estudos em modelos computacionais de pregas vocais, experimentos com laringes extirpadas e análise dinâmica não linear de vozes humanas verificaram a existência de caos na produção de voz. De fato, muitos pesquisadores notaram as não linearidades intrínsecas das pregas vocais como relação pressão-fluxo na glote, colisão das pregas vocais e a curva não linear de deformação da prega vocal. Foi Titze et al (1993) quem originalmente sugeriu como melhorar a compreensão dos distúrbios das vozes através dos conceitos da dinâmica não linear.

Dentre alguns parâmetros propostos na literatura para verificar comportamento caótico pode-se citar a dimensão de correlação D_2 , proposto por Grassberger & Procaccia (1983). Essa é uma medida geométrica das trajetórias no espaço de fase que descreve o quanto 2 pontos na trajetória estão correlacionados. A dimensão de correlação pode assumir os valores 0, 1 e 2, caso em que suas respectivas trajetórias serão chamadas de ponto fixo, ciclo limite e "2-torus". Um trajetória com comportamento caótico irá apresentar uma dimensão fractal. Essa é uma medida bastante simples usada por pesquisadores devido à sua rápida convergência.

Através da dimensão de correlação pode-se distinguir ruído branco de sistemas caóticos, ou seja, D_2 para o ruído branco não converge conforme se aumenta a dimensão de imersão m ao passo que para um sistema caótico, D_2 converge para um valor finito.

Outro parâmetro para verificar a presença de caos é o expoente de *Lyapunov*. Um sistema contendo pelo menos um expoente positivo já é definido como caótico. Essa

medida descreve a taxa média da convergência ou divergência das órbitas vizinhas no espaço de fase.

A entropia de *Kolmogorov* quantifica a taxa de perda de informação sobre o estado de um sistema dinâmico conforme ele evolui no tempo. Para comportamentos regulares a entropia é igual a zero, para um sistema caótico com grau de liberdade finito o valor da entropia também é um valor finito. Como exemplo, pode-se citar o trabalho de Alonso et al (2006) que utilizou a dimensão de correlação e expoente de *Lyapunov* para quantificar a qualidade de um sinal vocal e monitorar a evolução de tratamentos na laringe.

5 Materiais e Métodos

5.1 Materiais

Os sinais de vozes analisados pertencem à base de dados do Laboratório de Processamento de Sinais - LPS - da Escola de Engenharia de São Carlos. Foram selecionadas 16 normais, 16 vozes com nódulo e 16 vozes com edema de Reinke. Foram analisados trechos de 500ms da vogal sustentada /a/ de pessoas de ambos os sexos na faixa etária de 19 a 39 anos. Todo procedimento de captura do sinal de voz foi feito por um profissional especializado na área de Fonoaudiologia.

Para reconstrução do espaço de fase foi utilizado o programa de análise de dinâmica não linear desenvolvido a partir do pacote *TISEAN*. Em seguida foi extraída a seção de Poincaré através de um algoritmo desenvolvido pelo próprio autor desse trabalho no software MATLAB 7.0 - versão acadêmica.

5.2 Validação da ferramenta utilizando senóides

Não há um algoritmo padrão para se traçar a seção de Poincaré de um sinal de voz. Alguns autores desenvolveram algoritmos próprios de acordo com aplicação desejada. A validação da ferramenta foi feita utilizando o espaço de fase reconstruído de um sinal senoidal inserindo *Jitter* e *Shimmer* gradativamente.

Para cada espaço de fase foram feitos 8 cortes regularmente espaçados das trajetórias formadas e de cada corte foi determinada a média da dispersão dos pontos da seção em relação aos eixos (x,y) e a media aritmética da dispersão em relação ao ponto médio da seção.

Considera-se Média = 0,0 (zero) para um sinal sem *Shimmer* e *Jitter*. Em seguida, fixa-se 3 valores para *Jitter* (0, 5 e 10%) e para cada valor de *Jitter* varia-se o *Shimmer* de 5 em 5%. Procedendo dessa forma foram gerados os resultados da Tabela 1 para dispersão

Tabela 1: Média da dispersão dos pontos da seção de Poincaré em relação aos eixos coordenados.

Shimmer (%)	Jitter 0%	Jitter 5%	Jitter 10%
5	0,1334	0,1296	0,1346
10	0,1771	0,1841	0,1836
15	0,2406	0,2167	0,2158
20	0,2624	0,2627	0,2439
25	0,2895	0,2915	0,2752
30	0,3158	0,2985	0,3379

dos pontos da seção em relação aos eixos coordenados.

A Tabela 1 gerou a Figura 18.

Figura 18: Média da dispersão dos pontos da seção em relação aos eixos (x; y)

Também foram calculadas as médias da dispersão em relação ao ponto médio da seção, resultando na Tabela 2. A Tabela 2 gerou a Figura 19.

Tabela 2: Média da dispersão dos pontos da seção de Poincaré em relação ao ponto médio da seção

Shimmer (%)	Jitter 0%	Jitter 5%	Jitter 10%
5	0,063	0,074	0,075
10	0,112	0,141	0,155
15	0,184	0,204	0,202
20	0,233	0,237	0,26
25	0,27	0,309	0,315
30	0,369	0,333	0,391

Figura 19: Média da dispersão dos pontos da seção em relação ao ponto médio

Pelo gráfico na figura 19, pode-se observar que para cada variação de *shimmer*, houve uma variação proporcional das médias da dispersão; mesmo na presença de *Jitter*. Conclui-se que a ferramenta rotacionou corretamente os pontos da seção através da reta tangente à seção de modo que os pontos ficassem paralelos ao plano $x=0$. O mesmo procedimento utilizado para a senoide foi adotado para determinar a seção de Poincaré dos sinais de voz e, em seguida, a dispersão dos pontos na seção.

A validação da ferramenta foi feita calculando-se a média dos pontos da seção em vez do cálculo do desvio padrão. Uma vez que a lei de formação do sinal senoidal através da inserção de *Jitter* e *Shimmer* é aleatório, o desvio padrão também será aleatório, não

sendo possível verificar a variação linear da dispersão, como indicado nos gráficos acima. Por isso, utilizou-se o cálculo da média.

5.3 Procedimento

O programa abre um sinal de voz completo, indicado na Figura 20.

Figura 20: Sinal de voz completo

Em seguida, o sinal é normalizado de acordo com a expressão abaixo

$$x_n(i) = \frac{x(i)}{x_{max}} \quad (5.1)$$

Onde:

x_n : Sinal normalizado

$x(i)$: Sinal original

x_{max} : Amplitude máxima do sinal original

Após esse procedimento, obtém-se a Figura 21.

Figura 21: Sinal de voz normalizado

No sinal de voz da figura 21, seleciona-se um trecho de 500ms numa região de maior estacionariedade (indicado pelo tracejado vermelho), sem considerar o início e o fim do sinal, para evitar os transientes. A Figura 22 mostra somente o trecho de 500ms.

Figura 22: Trecho de 500ms do sinal de voz normalizado

O tempo de 500ms foi escolhido porque permite trajetórias bem definidas no espaço de fase e uma boa quantidade de pontos na seção de Poincaré, além de diminuir a carga computacional necessária para trechos maiores. Depois da seleção do trecho, foi feita uma interpolação linear desse sinal aumentando em 10 vezes o número de pontos para melhorar a precisão das medidas. Procedendo dessa forma, obtém-se a Figura 23.

Figura 23: Trecho de um sinal de voz interpolado

As barras tracejadas nos pontos 2,0s e 2,5s limitam o trecho de 500ms. A distância entre dois picos na forma de onda representa uma trajetória completa no espaço de fase. O número aproximado de trajetórias pode ser calculado pela expressão

$$n = \frac{x(N) - x(1)}{x(i) - x(i-1)} \quad (5.2)$$

Onde:

n : é o número de trajetórias completas no espaço de fase

$x(N)$: é o último pico do sinal

$x(1)$: é o primeiro pico do sinal

$x(i)$: um ponto de máximo qualquer do sinal

Os 4 elementos para o cálculo de n são obtidos dando um *click* no *mouse* nos pontos indicados na Figura 22. O objetivo principal dessa expressão é determinar o número de pontos da seção de Poincaré, uma vez que cada trajetória corta o plano uma única vez. Esse valor nem sempre é exato porque depende do *Jitter* do sinal. Os testes demonstraram que esse erro está em torno 3%. Como o sinal é pequeno, uma inspeção visual mostra 60 picos, ou 60 trajetórias no espaço de fase. Em seguida obtém-se o gráfico da função de informação mútua, e dessa função seleciona-se o primeiro mínimo, conforme indicado pelo cruzamento das linhas vermelhas na Figura 24.

A partir desse sinal, gera-se o espaço de fase reconstruído utilizando a técnica de

Figura 24: Informação Mútua

tempo de atraso. O tempo de atraso τ , para reconstrução do espaço de fase é determinado tomando-se o primeiro mínimo da informação mútua (Fraser, A.M.; Swinney, H.L., 1986), que nesse caso foi de 8 amostras atrasadas. Portanto, o eixo horizontal do espaço de fase será o próprio sinal $x(n)$ e o eixo vertical será sinal atrasado de 8 amostras $x(n-8)$.

Figura 25: Espaço de fase

Na própria Figura 25, selecionam-se 2 pontos $P_1(x_1, y_1)$ e $P_2(x_2, y_2)$ de modo a formar uma reta tangente à trajetória. A partir desses 2 pontos faz-se uma rotação de todo o espaço de fase de um ângulo θ , de acordo com a expressão a seguir

$$\theta = \arctan\left(\frac{y_1 - y_2}{x_1 - x_2}\right) \quad (5.3)$$

Essa rotação é feita para se trabalhar com o eixo x como referência.

Após a rotação, faz-se uma ampliação do trecho em estudo resultando na Figura 26. Nessa figura, seleciona-se 3 pontos indicados por P_3 , P_4 e P_7 entre os pontos P_1 e P_2 para delimitar a região do corte de onde se vai extrair a seção de Poincaré; e mais 2 pontos, indicados por P_5 e P_6 .

Figura 26: Espaço de fase ampliado. As barras tracejadas delimitam a região do corte e o trecho de orientação da reta normal à região do corte

Numericamente, a Região do Corte (RC) é determinada de acordo com a expressão abaixo:

$$P_{3x} \leq (RC)_x \leq P_{4x} \quad (5.4)$$

$$P_{7y} \leq (RC)_y \leq \text{Min}(P_{3y}; P_{4y}) \quad (5.5)$$

As barras paralelas delimitam as duas regiões pra estudo e os pontos pretos em destaque são os pontos pertencentes à trajetória do espaço de fase. A região delimitada pelos pontos P_5 , P_6 e P_7 é necessária para orientação da trajetória e é determinada da mesma forma que a RC. O ponto médio do primeiro e segundo corte irá formar a reta tangente à trajetória, que é normal ao plano de Poincaré.

Fazendo-se uma ampliação da Figura 26 verifica-se que várias trajetórias contêm mais de um ponto. A elipse na Figura 27 mostra a ocorrência de 2 pontos. A seção, como já explicado, deve conter apenas um ponto.

Figura 27: Espaço de fase ampliado. Observa-se a ocorrência de mais de um ponto no trecho em estudo

Em seguida, determinam-se os pontos da terceira coordenada (Eixo z), tomando-se o sinal original com 2 vezes o tempo de atraso, nesse caso, $x(n-16)$, e calcula-se o produto escalar entre a reta tangente à trajetória e todas as retas formadas pelo ponto médio da região do corte e todos os outros pontos da região do corte. A Figura 28 ilustra esse procedimento.

Figura 28: Determinação dos pontos pertencentes ao plano de Poincaré através do produto escalar

A Figura 28 indica apenas os pontos da região do corte e a reta tangente à trajetória. Os pontos da seção de Poincaré serão selecionados da seguinte forma: dado um ponto $B(i)$, esse ponto será um ponto da seção se

$$\vec{OA} \cdot \vec{O(i)} \leq \epsilon \quad (5.6)$$

Onde:

$\vec{OA}$: Reta tangente ao plano de Poincaré

$\vec{O(i)}$: Reta formada pelo ponto médio da seção e os pontos da região do corte

ϵ : Resultado do produto escalar - Esse valor é calculado tendo como base o número de trajetórias no espaço de fase.

5.3.1 Cálculo de ϵ

Para o cálculo de ϵ , define-se um vetor bidimensional associando cada ponto da região do corte ao resultado do seu respectivo produto escalar. Em seguida, ordena-se esse vetor e toma-se os n primeiros elementos do vetor, onde n é o resultado da equação 5.2. O valor de ϵ será o resultado do produto escalar do e -ésimo elemento do vetor.

Uma ilustração desse procedimento pode ser feita utilizando um trecho de 1000 ms. Realizando os procedimentos anteriores obteve-se a Figura 29.

Figura 29: Trecho do espaço de fase de um sinal de 100ms. O resultado do produto escalar selecionou apenas os pontos pertencentes ao plano de Poincaré - indicado em vermelho

Na Figura 29 pode-se visualizar a reta tangente à trajetória (indicada pelos pontos em azul), a região do corte selecionada (formada pelos pontos pretos) e os pontos da seção de Poincaré, em vermelho, que obedecem à equação do produto escalar. Depois de selecionar todos os pontos da seção de Poincaré, indicados na Figura 29, faz-se uma rotação em 3 dimensões dos pontos da seção. A rotação dos pontos é feita utilizando as expressões abaixo e tomando-se a reta tangente à trajetória como referência.

$$R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix} \quad (5.7)$$

$$R_y(\theta) = \begin{bmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin(\theta) & 0 & \cos \theta \end{bmatrix} \quad (5.8)$$

$$R_z(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (5.9)$$

Em seguida, projetam-se os pontos no plano $x = 0$ para visualizar os pontos da seção no plano. Antes de realizar os cálculos do desvio padrão, é necessário extrair as componentes principais da seção para decorrelacionar os pontos da seção. Finalmente, obtém-se a seção de Poincaré e seus histogramas nas figuras a seguir.

Figura 30: Seção de Poincaré do trecho (2,0-2,5ms) do sinal de voz

Figura 31: Dispersão dos pontos em relação ao ponto médio da seção

Figura 32: Dispersão dos pontos em relação ao eixo y

Figura 33: Dispersão dos pontos em relação ao eixo z

5.4 Cálculo da Dispersão

Foram estabelecidos 2 parâmetros para a análise da dispersão do sinal de voz na seção de Poincaré:

- DPD: Desvio Padrão da Dispersão - desvio em relação ao ponto médio da seção.
- DPE: Desvio Padrão em relação aos eixos coordenados

Para o cálculo do DPD utiliza-se a expressão abaixo:

$$DPD = \sqrt[2]{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (5.10)$$

Onde:

n : é a quantidade de pontos da seção de Poincaré

$\bar{x}$: é o ponto médio da seção

$x(i)$: são os pontos da seção

Para o cálculo do DPE utiliza-se a expressão abaixo:

$$DPE = \sqrt[2]{\sigma_x^2 + \sigma_y^2} \quad (5.11)$$

Onde:

σ_x : é o desvio padrão em relação ao eixo x

σ_y : é o desvio padrão em relação ao eixo y

6 Resultados e discussões

6.1 Seção de Poincaré de Vozes Normais

6.1.1 Voz normal n° 6

Foi gerado o sinal temporal normalizado da voz intitulada de voz normal 6, indicado na Figura 34.

Figura 34: Voz normal 6

Desse sinal foi selecionado o trecho 2.2-2.7s, resultando na Figura 35.

Figura 35: Trecho do sinal de voz normal 6

A partir do trecho que vai de 2.2s até 2.7s foi gerada o espaço de fase indicado na Figura 36, contendo os 10 cortes para se extrair a seção de Poincaré.

Figura 36: Espaço de fase do sinal de voz normal 6 com 10 cortes

Os 10 cortes geraram as 10 seções indicadas nas figuras a seguir.

Figura 37: Seção de Poincaré do Corte 1 e seus histogramas

Figura 38: Seção de Poincaré do Corte 2 e seus histogramas

Figura 39: Seção de Poincaré do Corte 3 e seus histogramas

Figura 40: Seção de Poincaré do Corte 4 e seus histogramas

Figura 41: Seção de Poincaré do Corte 5 e seus histogramas

Figura 42: Seção de Poincaré do Corte 6 e seus histogramas

Figura 43: Seção de Poincaré do Corte 7 e seus histogramas

Figura 44: Seção de Poincaré do Corte 8 e seus histogramas

Figura 45: Seção de Poincaré do Corte 9 e seus histogramas

Figura 46: Seção de Poincaré do Corte 10 e seus histogramas

Os desvios DPE e DPD das 10 seções estão indicados na Tabela 3.

Tabela 3: Desvio padrão da voz normal n° 6

Resultados		
	DPE	DPD
Corte 1	0,1719	0,0093
Corte 2	0,3221	0,0330
Corte 3	0,1636	0,00860
Corte 4	0,1813	0,0116
Corte 5	0,1528	0,0075
Corte 6	0,1378	0,0067
Corte 7	0,1554	0,0074
Corte 8	0,1687	0,0092
Corte 9	0,1504	0,0069
Corte 10	0,1430	0,0077

Os desvios DPE e DPD foram calculados realizando-se 10 cortes regularmente espaçados no espaço de fase de 16 vozes normais. Foi calculado o DPE de todos os cortes e extraída a média dessas 10 medidas. Dessa forma, foram geradas as Tabelas 4 e 5.

Tabela 4: Desvio padrão da dispersão das vozes normais em relação aos eixos xy

Resultados	
Normal	Média DPE - 10 Seções
1	0,244
2	0,184
3	0,219
4	0,1897
5	0,1725
6	0,1561
7	0,2339
8	0,1976
9	0,2125
10	0,1681
11	0,21769
12	0,2031
13	0,1714
14	0,178
15	0,2175
16	0,1751

A média final do DPE de todas as medidas foi de 0,1962 com desvio padrão de 0,0258.

Tabela 5: Desvio padrão das vozes normais em relação ao ponto médio da seção

Resultados	
Normal	Média DPD - 10 Seções
1	0,0203
2	0,0122
3	0,0177
4	0,0137
5	0,0109
6	0,008
7	0,022
8	0,013
9	0,0159
10	0,0096
11	0,0179
12	0,0158
13	0,0116
14	0,012
15	0,0172
16	0,0104

A média final do DPD de todas as medidas foi de 0,0143 com desvio padrão de 0,004. Observou-se que a voz 7 apresentou valores maiores comparados com o resto das vozes. A média final do DPD de todas as medidas foi de 0,0143 com desvio padrão de 0,004.

6.2 Seção de Poincaré de Vozes com nódulo

6.2.1 Voz com nódulo nº2

Foi gerado o sinal temporal da voz normal intitulada de voz com nódulo nº 2, indicado na figura 47.

A partir do trecho que vai de 2.4s até 2.9s foi gerado o espaço de fase indicado na Figura 48, contendo os 10 cortes para se extrair a seção de Poincaré.

Figura 47: Voz com nódulo n° 2

Figura 48: Espaço de fase da voz com nódulo n° 2 com 10 cortes

Os 10 cortes geraram as 10 seções indicadas nas figuras a seguir.

Figura 49: Seção de Poincaré do Corte 1 e seus histogramas

Figura 50: Seção de Poincaré do Corte 2 e seus histogramas

Figura 51: Seção de Poincaré do Corte 3 e seus histogramas

Figura 52: Seção de Poincaré do Corte 4 e seus histogramas

Figura 53: Seção de Poincaré do Corte 5 e seus histogramas

Figura 54: Seção de Poincaré do Corte 6 e seus histogramas

Figura 55: Seção de Poincaré do Corte 7 e seus histogramas

Figura 56: Seção de Poincaré do Corte 8 e seus histogramas

Figura 57: Seção de Poincaré do Corte 9 e seus histogramas

Figura 58: Seção de Poincaré do Corte 10 e seus histogramas

Tabela 6: Desvio padrão da voz com nódulo nº 2

	DPE	DPD
Corte 1	0,3602	0,03779
Corte 2	0,3221	0,0330
Corte 3	0,3316	0,0360
Corte 4	0,4128	0,0513
Corte 5	0,3513	0,0517
Corte 6	0,3384	0,04123
Corte 7	0,2886	0,02313
Corte 8	0,3667	0,0434
Corte 9	0,3210	0,02968
Corte 10	0,3390	0,0367

Os desvios DPE e DPD das 10 seções estão indicados na Tabela 6.

Realizando o mesmo procedimento com 16 vozes com nódulo resultaram as Tabelas 7 e 8.

Tabela 7: Desvio padrão das vozes com nódulo em relação aos eixos cartesianos

Resultados	
Nódulo	Média DPE - 10 Seções
1	0,2929
2	0,3432
3	0,2138
4	0,2607
5	0,2823
6	0,3203
7	0,2163
8	0,2617
9	0,3918
10	0,2778
11	0,3428
12	0,2559
13	0,2555
14	0,3768
15	0,2581
16	0,348

A média final do DPE de todas as medidas foi de 0,2936 com desvio padrão de 0,0548. A média final do DPD de todas as medidas foi de 0,0319 com desvio padrão de 0,0126.

Comparando com as medidas realizadas com vozes normais observa-se que os desvios das vozes com nódulo são maiores que os das vozes normais.

6.3 Voz com edema de Reinke

6.3.1 Voz com edema de Reinke n° 14

Foi gerado o sinal temporal da voz normal intitulada de voz com edema de Reinke n° 14, indicado na Figura 59.

Tabela 8: Desvio padrão das vozes com nódulo em relação ao ponto médio da seção

Resultados	
Nódulo	Média DPD - 10 Seções
1	0,0287
2	0,0384
3	0,0154
4	0,0239
5	0,0287
6	0,0342
7	0,0187
8	0,023
9	0,0598
10	0,0246
11	0,0419
12	0,0242
13	0,0241
14	0,0602
15	0,0217
16	0,0423

Figura 59: Voz com edema de Reinke

A partir do trecho que vai de 2.0s até 2.5s foi gerado o espaço de fase indicado dicado na Figura 60, contendo os 10 cortes para se extrair a seção de Poincaré.

Os 10 cortes geraram as 10 seções indicadas nas figuras a seguir.

Figura 60: Espaço de fase do sinal de voz com edema de Reinke com 10 cortes

Figura 61: Seção de Poincaré do Corte 1 e seus histogramas

Figura 62: Seção de Poincaré do Corte 2 e seus histogramas

Figura 63: Seção de Poincaré do Corte 3 e seus histogramas

Figura 64: Seção de Poincaré do Corte 4 e seus histogramas

Figura 65: Seção de Poincaré do Corte 5 e seus histogramas

Figura 66: Seção de Poincaré do Corte 6 e seus histogramas

Figura 67: Seção de Poincaré do Corte 7 e seus histogramas

Figura 68: Seção de Poincaré do Corte 8 e seus histogramas

Figura 69: Seção de Poincaré do Corte 9 e seus histogramas

Figura 70: Seção de Poincaré do Corte 10 e seus histogramas

O desvio padrão de todas as seções estão indicados na Tabela 9.

Realizando o procedimento com as 16 vezes com edema de Reinke foram obtidas as Tabelas 10 e 11.

A média final do DPE de todas as medidas foi de 0,2721 com desvio padrão de 0,0583.

Tabela 9: Desvio padrão da voz com Edema de Reinke 14

Resultados		
	DPE	DPD
Corte 1	0,2653	0,0313
Corte 2	0,2152	0,0153
Corte 3	0,1976	0,0127
Corte 4	0,2958	0,0334
Corte 5	0,3003	0,0287
Corte 6	0,3029	0,0345
Corte 7	0,2750	0,0232
Corte 8	0,2693	0,0265
Corte 9	0,2074	0,0216
Corte 10	0,2376	0,0229

A média final do DPD de todas as medidas foi de 0,0279 com desvio padrão de 0,0115. Comparando-se as medidas extraídas para o edema de Reinke com as medidas obtidas para vozes normais e vozes com nódulo, observa-se um valor intermediário entre essas 2 últimas. A Tabela 12 compara as média obtidas dos 3 tipos de vozes.

Tabela 10: Desvio padrão das vozes com edema de Reinke em relação aos eixos cartesianos

Resultados	
Edema	Média DPE - 10 seções
1	0,2180
2	0,2245
3	0,2286
4	0,2727
5	0,3473
6	0,2878
7	0,2534
8	0,1700
9	0,3152
10	0,3447
11	0,35577
12	0,2726
13	0,2302
14	0,2566
15	0,3524
16	0,2240

Tabela 11: Desvio padrão das vozes com edema de Reinke em relação ao ponto médio da seção

Resultados	
Edema	Média DPD - 10 seções
1	0,0194
2	0,0174
3	0,0184
4	0,0249
5	0,0398
6	0,0284
7	0,0219
8	0,0100
9	0,0414
10	0,0439
11	0,0487
12	0,0243
13	0,0181
14	0,0245
15	0,0463
16	0,0184

Tabela 12: Tabela comparativa entre as medidas DPE e DPD dos 3 tipos de vozes estudadas

Resultados			
	DPE		DPD
Normal	0,1962	Normal	0,0143
Edema	0,2721	Edema	0,0278
Nódulo	0,2936	Nódulo	0,0319

7 Teste ANOVA

O teste ANOVA compara a média de 2 ou mais grupos de dados e verifica se as medias são iguais. O teste retorna o valor P para hipótese de nulidade mas não informa qual do grupos é diferente.

Foi utilizado o teste anova para verificar se as diferenças entre os desvios dos 3 grupos foram significativas. Para a análise do DPE foi utilizada a Tabela 13, contendo as médias das dispersões (DPE) de cada sinal de voz analisado.

Tabela 13: DPE das vozes normais, com nódulo e com edema de Reinke

DPE		
Normal (1)	Edema (2)	Nódulo (3)
0.24399	0.21796	0.29285
0.18396	0.22447	0.34317
0.21896	0.22856	0.21378
0.18898	0.27269	0.26071
0.17254	0.34731	0.28226
0.15605	0.28776	0.3203
0.23393	0.25343	0.2163
0.19763	0.16997	0.26171
0.21254	0.31524	0.39175
0.16807	0.34465	0.27781
0.21756	0.35566	0.342791
0.20305	0.27258	0.25591
0.17141	0.23017	0.255471
0.17799	0.2566	0.3768
0.21749	0.3524	0.25814
0.17506	0.22403	0.34795

A partir da Tabela 13 realizou-se o teste anova, resultando na Figura 71.

Figura 71: Teste anova para o DPE

O valor P foi $= 1,438.10^{-6}$, menor que o nível de significância de 0,05, assim podemos rejeitar a hipótese nula, ou seja, as médias dos DPE's são diferentes.

Para a análise do DPD foi utilizada a Tabela 14, contendo as médias das dispersões (DPD) de cada sinal de voz analisado.

Tabela 14: DPD das 16 vozes normais, com nódulo e com edema de Reinke

DPD		
Normal (1)	Edema (2)	Nódulo (3)
0.02034	0.01937	0.02874
0.01217	0.01737	0.03837
0.01774	0.01839	0.01538
0.01369	0.02486	0.02391
0.01087	0.03975	0.02872
0.00803	0.02836	0.03416
0.02197	0.02185	0.01873
0.01303	0.00996	0.02303
0.01589	0.04141	0.05982
0.00956	0.04385	0.02456
0.01791	0.04869	0.04186
0.01584	0.0243	0.02418
0.0116	0.01813	0.02412
0.01197	0.02457	0.06021
0.01721	0.04634	0.02167
0.01035	0.01843	0.04229

A partir da Tabela 14 realizou-se o teste anova, resultando na Figura 72.

Figura 72: Teste anova para o DPD

O valor P foi $= 7,2778.10^{-5}$, menor que o nível de significância de 0,05, assim podemos rejeitar a hipótese nula, ou seja, as médias dos DPD's são diferentes.

Realizando o teste em grupos de dois a dois obteve-se os resultados na tabela a seguir.

Tabela 15: Valores de P para o teste Anova em grupos de dois a dois

Valores de P		
Grupo	DPE	DPD
NormalxEdema	$3.58.10^{-7}$	$1.8592.10^{-4}$
NormalxNódulo	$3.5124.10^{-5}$	$2.2754.10^{-5}$
NóduloxEdema	0.2823	0.3830

Observa-se pela tabela 15 que não há diferença entre as médias dos desvios padrões entre os grupos Edema e Nódulo porque o valor P foi maior que o nível de significância 0,05. Mas entre os grupos Normal x Nódulo e Normal x Edema pode-se dizer que há diferença entre as médias porque o valor P foi significativamente menor que o nível de significância. Também foi realizado o teste Tukey para o DPE e o DPD onde foi verificado que há diferença estatística somente entre as médias dos desvios padrões entre voz com nódulo e voz com edema de Reinke. Entre voz normal e patológica o teste de Tukey demonstrou que não há diferença.

8 Conclusão

A utilização da ferramenta para traçar a seção de Poincaré e calcular os desvios das dispersões demonstrou-se eficaz para comparar os 3 grupos de vozes estudadas nesse trabalho. O teste Anova demonstrou que existe diferenças nas dispersões dos pontos da seção de Poincaré de vozes normais e patológicas. Na comparação entre vozes com nódulo e edema, o teste demonstrou que não houve diferença.

A dispersão espaço de fase fica caracterizada pela distância entre as trajetórias tomando a seção de Poincaré com referência. Uma consideração importante é que as vozes analisadas não possuem informações sobre o grau de severidade de cada patologia.

O maior valor médio das dispersões para os 2 tipos de patologias pode ter ocorrido devido a um excesso de soprosidade, ruído causado por uma massa nas pregas vocais, rigidez na mucosa das pregas, dentre outros fatores. Esses fatores podem ter provocado uma perturbação na amplitude e frequência e essa instabilidade pode ter sido refletida no aumento das médias dos desvios.

Apesar disso, a análise perceptivo-auditiva da voz, como instrumento básico de atuação fonoaudiológica clínica, é indispensável para uma avaliação vocal precisa. Mas pode ser associada às técnicas de análise não linear para complementar e acompanhar a evolução de tratamentos clínicos.

Sugestão para trabalhos futuros:

- Traçar a seção de Poincaré do sinal de voz completo e calcular os desvios padrões.
- Traçar a seção de Poincaré de um sinal de voz conhecendo-se o grau de severidade da patologia e calcular os desvios padrões.
- Aumentar o número de cortes na seção.
- Fazer o plano de Poincaré percorrer todo o espaço de fase dinamicamente e verificar o que ocorre nas regiões onde não foi possível especificar um corte.

Referências

- Adnene, C., Lamia, B. e Mounir, M, *Analysis of pathological voices by speech processing*, IEEE, pp. 365-368 , 1997
- Alonso, J. B., Díaz-de-María, F., Travieso, C. M. e Ferrer, M. A. *Using nonlinear features for voice disorder detection* (2006), Acesso em 2010, disponível em <http://www.gpds.ulpgc.es/index.htm>
- Banci, Monini, Falaschi e Sario, *Vocal fold disorder evaluation by digital speech analysis*, J. Phonetics , 14, pp 495-499
- Belhau, M., 2001 , *A voz. O livro do especialista (Vol. I)*
- Berg, J. v., *Myoelastic - Aerodynamic theory of voice production*
- Birgmeier, M., Bernhard, H.P. e Kubin, G., *Nolinear long-term prediction of speech signal*, IEEE , pp. 1283-1286
- Boukabou, A. e Mansouri, N., *Control of chaotic behavior in high order dynamical system*, IEEE , pp. 516-521
- Boyanov e Chollet, *Pathological voice analysis using spectra, bispectra and group delay functions*, In Proceedings of International Conference on Spoken Language Processing , 2, pp. 1039-1042
- Boyanov, B. e Hadjitodorov, S., *Acoustic analysis of pathological voices*, IEEE, Engineering And Biology , pp. 74-82
- Boyanov, Doskov, Ivanov e Hadjitodorov. *PC based system for analysis the voice of patient with laryngeal diseases*. ENT Journal , 70, pp. 767-772
- Broomhead, D. S. e King, G. P., *On the Qualitative analysis of experimental dynamical systems*. In: D. S. Broomhead, e G. P. King, *Nonlinear Phenomena and Chaos*, pp. 113 - 144, Bristol: Adam Hilger
- Buzug, T. M. *Characterisation of Time Series from Non-Linear system*, IEEE , pp. II-29 - II-234
- Casdagli, M., Eubank, S., Farmer, J. D., e Gibson, J. *State space reconstruction in the presence of noise*, Physica D , 51, 52
- Cesar, M., e Hugo, R. *Acoustic analysis of speech for detection of laryngeal pathologies*, In: IEEE (Ed.), *Proceedings of the 22th Annual EMBS International Conference*, pp. 2369-2372, Chicago IL

- Childers, D. G., *Speech Processing and Synthesis for Assessing Vocal Disorders*, IEEE ENGINEERING IN MEDICINE AND BIOLOGY MAGAZINE , 69-71
- Costa, S., Correia, S., Falcão, H., Almeida, N., e Assis, F., *O uso da entropia na discriminação de vozes patológicas*, II Congresso de Pesquisa e Inovação da Rede Norte Nordeste de Educação Tecnológica
- Dajer, M. E., *Padrões visuais de sinais de voz através através de técnicas de análise não linear*, Dissertação de mestrado, Programa de Pós - Graduação Interunidades em Bioengenharia. Escola de Engenharia de São Carlos (EESC), Faculdade de Medicina de Ribeirão preto (FMRP) e Insituto de Química de São Carlos (IQSC), Universidade de São Paulo (USP)
- Dajer, M. E., Sobrinho, F. A., e Pereira, J. C. *Tecnicas não lineares e convencionais para análise de vozes de crianças*, XVII Congresso Argentino de Bioengenharia SABI Rosario, Argentina
- Dager, M. E., Sobrinho, F. A., e Pereira, Dr. J. C., *Tecnicas no lineales y tecnicas tradicionales para analisis de la calidad vocal en ninos*, XI Congresso Argentino de Otorrinolaringologia Y Fonoudiologia Pediatrica. Salta, Argentina
- Daizhi, L., Xihai, L., Ke, Z., e Zhigang, L. (14-17 de 12 de 2003). Attractor analysis of feature pahse space. IEE Int. Conf. Neural Networks e Signal Processing , pp 783-785
- Dajer, M. E., Pereira, J. C., e Maciel, C. D. (ISM 2005). Nonlinear dynamical analysis of normal voices. In Proceedings of the Seventh IEEE Internatinal Symposium on Multimedia (ISM 2005) , pp. 765-771
- Dajer, M., Sobrinho, F., e Pereira, J. (2010). Qualitative and Quantittive Analysis of vocal dynmical visual patters. 28th World Congress of the International Association of Logopedics and Phoniatrics. Athenaeum Intercontinental - Athens, Greece
- Deliyski, D. D. (1993). Acoustic model and evaluation of pathological voice production. 3-rd Conference on Speech Communication and Technology EUROSPEECH'93 , pp 1969-1972
- Ding, W., e Kasuya, H. (s.d.). A novel approach to the estimation of voice source and vocal tract parameters from speech signals, 1257-1260
- E.R.Carson, H.J.Leicester, e Roudsari, A. (1996). Linear and non-invasive time series analysis in non-invasive physiological measurement. 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (pp. 970-971). Amsterdam: IEEE
- Fant, G. (1982). The voice source - acoustic modeling. STL-QPSR , 23 (4), pp. 028-048
- Flanagan, J. L. (1972). Speech analysis, synthesis and perception. New York: Springer - Verlag, 2a ed
- Fraser, A. M. (1989). Reconstruction attractors from scalar time series: A comparision of Singular System and Redundancy Criteria. Phisica D 34 , 391-404

- Fraser, A. M., e Swinney, a. H. (1986). Independent coordinates for strange attractors from mutual information. *Physical Review A* , 33 (2), 1134-1137
- Gibson, J. F., Farmer, J. D., Casdagli, M., e Eubank, S. (1992). An analytic approach to practical state space reconstruction. *Physica D* , 57 (1)
- Gindri, Gigiane; CIELO, Carla Aparecida. Disfonia por nódulos vocais na infância. *Salusvita, Bauru*, v. 27, n. 1, p. 91-110, 2008.
- Grassberger, P., e Procaccia, a. I. (1983). Measuring the strangeness of strange attractors. *Physica 9D* , 189-208
- Gunter, H. E. (February de 2003). A mechanical model of vocal-fold collision with high spatial and temporal resolution. *J. Acoust. Soc. Am.* 113 (2)
- Haddad, W. M., Nersesov, S. G., e Chellaboina, V. (2002). A lyapunov function proof of poincare theorem. *Proceedings of the American Control Conference*. Anchorage, AK
- Hagmuller, M., e Kubin, a. G. (s.d.). Poincaré sections for pitch mark determination. Fonte: <http://www.spsc.tugraz.at>
- Hecker, M. H., e Kruegel, E. J. (1970). Description of the speech of patients with cancer of the vocal folds. Part I: Measures of fundamental frequency. *Journal of the Acoustical Society of America* , 49 (4), 1275-1282
- Herzel, H. (1993). Bifurcation and chaos in voice signals. *Applied Mechanics Review* , 399-413
- Herzel, H., e Wendler, J. (1991). Evidence of chaos in phonatory samples. *EUROSPEECH 91*, (pp. 262-266). Genoa, Italy
- Herzel, H., Berry, D., Titze, I. R., e Saleh, M. (1994). Analysis of vocal disorders with methods from nonlinear dynamics. *Journal of Speech and Hearing Research* , 1008-1019
- Hong-guang, M., e Chong-Zhao, H. (2006). Selection of embedding dimension and delay time in phase space reconstruction. *Front. electr. Electron. Eng* , pp. 111-114
- Horii, Y. (1979). Fundamental frequency perturbation observed in sustained phonation. *Journal of Speech and Hearing Research* , 12-14
- Hu, T., R. Teel, A., e Lin, Z. (2004). A Lyapunov approach to frequency analysis. In: 2. AACC (Ed.), *Proceeding of the 2004 American Control Conference* FrA07.5, pp 4145-4150. Boston, Massachusetts
- Hussain, A. (1997). A new neural network structure for temporal signal processing. *IEEE* , pp. 3441-3444
- Ishizka, K., e Flanagan, J. L. (1972). Synthesis of voice sounds from a two-mass model of the vocal cords. *Bel Syst. Tech. J.* , 50 (6), 1233-1268
- Jackson, M. A., e Burnett, I. S. (1999). Phase space portraits of speech employing mutual information and perceptual masking. *IEEE* , pp. 61-63

- Jiang, J. J., Zhang, Y., e McGilligan, a. C. (2006). Chaos in voice, from modeling to measurement. *Journal of Voice* , 20 (12-17)
- Johson, M. T., Lindgren, A. C., Poinelli, R. J., e Yuan, X. (2003). Performance of nonlinear speech enhancement using phase space reconstruction. *IEEE* , pp. I-920 - I-923
- Kantz, H., e Schreiber, T. (2004). *Nonlinear time series analysis*. Cambridge: Cambridge University Press
- Karnell, M. P., chang, A., Smith, A., e Hoffman, H. (1997). Impact of signal type of validity of voice perturbation measures. *NCVS Status and Progress Report* , 11, 91-94
- Kember, G., e Fowler, A. (1993). A correlation function for choosing time delays in phase portrait reconstructions. *Physics letters A* 179 , 72-80
- Kennel, M. B., Brown, R., e Abarbanel, H. D. (1992). Determining embedding dimension for phase space reconstruction using a geometrical construction. *Physical Review A* , 45 (6)
- Koike, Y., Takahashi, H., e Cacattera, T. C. (1977). Acoustic measures for detecting laryngeal pathology. *Journal of Phonetics* , 14, 517-524
- Kubin, G. (1997). Poincaré sections Techniques for speech. *IEEE* , pp. 7-8
- Kumar, A., e Mullick, S. K. (1990). Atractor dimension, entropia and modeling of speech time series. *Electronics letters*, Vol. 26 (No. 21), pp. 1790-1792
- Lieberman, P. (1963). Some acoustic measures of he fundamental periodicity of nomal and patological larynges. *Journal of the Acoustic Society of America* , 35 (3), 344-353
- Lin, S., Qiao, j., Wang, G., Zhang, S., e Zhi, L. (2006). Phase space recontruction of nonlinear time series based on kernel method. *Proceeddings of the 6th world on Intelligent Control and Auomation* , pp. 4364-4368
- Lindgren, A. C., Johson, M. T., e Povinelli, R. J. (203). Speech recognition using reconstructed phase space features. *IEEE* , pp. I-60-I-63
- Little, M., McSharpy, P., e Roberts, I. M. (2006). Nolinear, biophysically - Informed speechpPathology detection. *IEEE* , pp. II-1080 - II-1083
- Mann, I., e Laughlin, a. S. (1997). Poincaré maps and pitch detection in speech. *IEEE* , pp. 1-5
- Mende, W., Herzel, H., e Wermke, K. (1990). Bifurcatins ans chaos in newbon infant cries. *Physics letters A* , 145 (8,9)
- Murry, T., e Doherty, E. T. (1980). Selected acoustics characteristics of pathologic and normal speakers. *Jurnal of Speech and Hearing Research* , 23, 361-369
- Neves, J. A., Fernandes, J. C., e Restivo, F. (2006). Phase space signal filtering. *IEEE* , pp. 1805-1809

- Noakes, L. (1991). The Takens embedding theorem. *International Journal of Bifurcation and Chaos* , 1 (4), pp. 867-872
- Packard, N. H., Crutchfield, J. P., Farmer, J., e Shaw, a. R. (1980). Geometry from a time series. *Physical Review Letters* , 45 (9), 712-716
- Parker, T. S., e Chua, L. O. (1987). Chaos: A tutorial for engineers. *Proceedings of the IEEE* , 75, pp. 982-1007
- Povinelli, R. J., e Lindgren, A. C. (JUNE de 2006). Statistical models of reconstructed phase spaces for signal classification. *IEEE Transactions on signal processing*, VOL. 54 (NO. 6), pp. 2178-2186
- Propes, N. C. (1996). Poincare sections and the R-L-Diode circuit. *IEEE* , pp. 195-197
- Rabiner, L. R., e Schafer, R. W. (1978). *Digital processing of speech signals*. New Jersey: Prentice Hall
- Rauf, F., e Ahmed, H. M. (1993). Adaptive signal processing techniques for chaotic system. *IEEE* , pp. III-563 - III-566
- Rosa, M. O. (1998). Análise acústica da voz para pré-diagnóstico de patologias de laringe. Dissertação de Mestrado, Escola de Engenharia de São Carlos (EESC), Universidade de São Paulo (USP), São Carlos
- Rosa, M. O. (2002). Laringe digital. Tese de doutorado, Escola de Engenharia de São Carlos (EESC), Universidade de São Paulo (USP), São Carlos. SP, Brasil
- Rosa, M. O., Pereira, J. C., e Grellet, M. (2000). Adaptive estimation of residue signal for voice pathology diagnosis. *IEEE Transactions on Biomedical Engineering* , 47 (1), pp. 96-104
- Roux, J. C., Simoyi, R. H., e Swinney, H. L. (1983). Observation of a strange attractor. *Physica D* , 257-266
- Scalassara, P. R. (2009). Utilização de medidas de previsibilidade em sinais de voz para discriminação de patologias de laringe. Tese de doutorado, Escola de engenharia de São Carlos (EESC), Universidade de São Paulo (USP), São Carlos
- Shanon, C. E. (1968). A mathematical theory of communication. *Bell system Tech J.* , 27, 623-656
- Story, B. H., Titze, I. R., e Wong, D. (1996). A simplified model for simulation and transformation of speech. *IEEE* , pp. 320-327
- Takens, F. (1980). Detecting strange attractors in turbulence
- Tishby, N. (s.d.). *A dynamical systems approach to speech processing*. Murray Hill
- Titze, I. R. (1995). Workshop on acoustic voice analysis; Summary statement. p. 36
- Titze, I. R., e Liang, H. (1993). Comparison of F0 extraction method for high precision voice perturbation measurements. *J. Speech Hear. Res.* , 36, 1120-1133

- Titze, I. R., e Strong, W. J. (1975). Normal modes in vocal cord tissue. *J. Acoust. Soc. Am.* , 57 (3), 736-744
- Titze, I., Baken, r., e Herzel, H. (1993). Evidence of chaos in vocal fold vibration. In: I. Titze, e I. titze (Ed.), *Vocal fold Physiology: New frontiers in basic Science* (pp. 143-188) San Diego: Singular Publishing Group
- Tsuji I.D.; Inamura, R. Ubirajara, LS. Fisiologia da laringe. In: Pinho, S; Tsuji, DH; Bohadana, SC. *Fundamentos em laringologia e voz*. Ed. Revinter, Rio de Janeiro, 2006
- Vieira, M. N., e Mc Innes FR, Jack, M.A. (1997). Comparative assessment of electroglottographic and acoustic measures of jitter in pathological voices. *J Speech Lang Hear Res* , 40, 170-182
- Yu Zhang, C. M. (2004). Nonlinear dynamic analysis of voices before and after surgical. *J. Acoustic. Soc. Am.* 115 (5) , 2270-2277
- Zeitels, S.M, Hillman R.E., Desloge, R., Mauri, M., and Doyle, P.B..Phonemicsurgery in singers and performing artists: treatment outcomes, management theories, and future directions, *Ann. Otol. Rhinol. Laryngol.* 111,21-40 (2002)