

Análise acústica de sinais de vozes patológicas por meio de características temporais

Maria do Carmo de Luna Malheiros Frazão ^[1], Silvana Luciene do Nascimento Cunha Costa ^[2], Suzete Élide Nóbrega Correia ^[3]

[1] marafrazao@gmail.com ; [2] silvana@ifpb.edu.br ; [3] suzete.correia@gmail.com. Instituto Federal de Educação, Ciência e Tecnologia da Paraíba. Av. 1° de Maio, 720, Jaguaribe, João Pessoa - PB, Brasil, CEP: 58015-430

RESUMO

Este trabalho trata da avaliação de desempenho de características temporais de sinais de vozes no pré-diagnóstico de patologias laringeas (edema de Reinke, paralisia e nódulos vocais). As medidas temporais empregadas são: período de pitch (frequência fundamental), jitter, shimmer, energia segmental, autocorrelação e taxa de cruzamentos por zero. Na etapa de classificação é empregada a análise discriminante com as funções linear (LDA) e quadrática (QDA). A avaliação de desempenho do classificador é realizada com as medidas empregadas de forma individual e de forma combinada, ressaltando a eficácia da análise de discriminante linear na classificação entre vozes saudáveis e patológicas. A medida temporal que apresentou os melhores resultados na classificação foi Shimmer percentual, chegando a uma precisão de 100% entre sinais saudáveis ou com paralisia, como também para sinais com edema de Reinke ou com nódulo vocal.

Palavras-chave: Análise acústica. Vozes patológicas. Características temporais. Diagnóstico de patologias laringeas.

ABSTRACT

This paper deals with performance evaluation of time characteristics of speech signals in pre-diagnosis of laryngeal diseases (Reinke's edema, paralysis and vocal nodules). The measures used are: pitch period (fundamental frequency), jitter, shimmer, segmental energy, autocorrelation and zero-crossing rate. In classification stage is employed discriminant analysis with linear (LDA) and quadratic (QDA) functions. The classifier performance evaluation is carried out with the measures employed individually and combined, highlighting the effectiveness of linear discriminant analysis in the classification between healthy and pathological voices. A temporary measure that showed the best results in the classification was Shimmer percentage, reaching an accuracy of 100% among healthy signs or paralysis, as well as for signals with Reinke's edema or vocal nodules.

Keywords: Acoustic analysis. Pathological voices. Temporal characteristics. Diagnosis of laryngeal diseases.

1 Introdução

A análise acústica usando técnicas de processamento digital de sinais, para auxiliar diagnósticos de médicos especializados em doenças da laringe, é um procedimento de medição simples, de baixo custo e não invasivo para os pacientes. Pode ser empregada como ferramenta auxiliar em pré-diagnósticos e acompanhamento de tratamentos, diminuindo a necessidade do uso de exames mais invasivos como os de videolaringoscopia (GODINO-LLORENTE et al., 2006; BOYANOV; HADJITODOROV, 1997).

A chave para a modelagem acústica de vozes desordenadas é o entendimento das mudanças referentes às medidas acústicas, produzidas pelos efeitos da fonte de excitação e do trato vocal (COSTA et al., 2008a). Para que se possa obter um bom modelamento acústico da patologia em estudo, as características ou parâmetros devem ser escolhidos de tal forma que representem o sinal, de forma a diferenciar um sinal saudável de um patológico, para uma discriminação eficaz da patologia. O pré-diagnóstico para fornecer uma detecção precoce de patologias laringeas aumenta significativamente a eficácia do tratamento (VIEIRA, 2011).

Um dos principais desafios da modelagem acústica é capturar a variabilidade presente no sinal de voz. Assim, a voz é dinâmica ou variante no tempo e a modelagem precisa considerar dois aspectos: 1) as dependências temporais explícitas do sinal de voz e 2) a estimação das características. Esses aspectos têm de estar baseados na análise estatística a curto intervalo de tempo, além disso o modelo deve representar o comportamento das irregularidades introduzidas pela própria patologia (COSTA, 2008b).

As patologias na laringe podem ser de origem orgânica como nódulos, pólipos, cistos, cânceres, edemas de Reinke, entre outras, ou neurológica como paralisia nas pregas vocais (GODINO-LLORENTE et al., 2006). Neste trabalho, são focalizadas as patologias edemas de Reinke, nódulos e paralisia nas pregas vocais. Essas patologias na laringe provocam mudanças significativas na voz causadas por uma modificação na morfologia de excitação, produzindo um padrão de vibração irregular, pois a distribuição de massa na prega vocal e sua rigidez são maiores (DIBAZAR et al., 2006).

Nesse sentido, as mudanças na morfologia das pregas vocais podem provocar modificações

significativas no sinal de voz. As alterações nas vozes desordenadas podem ser observadas tanto nas modificações da frequência fundamental quanto na envoltória espectral do sinal (GODINO-LLORENTE et al., 2006).

Neste trabalho são empregadas as características temporais de frequência fundamental (F_0) bem como de duas medidas de perturbação obtidas a partir desta: *jitter* percentual e *shimmer* percentual, além de energia segmental, autocorrelação e taxa de cruzamentos por zero (*TCZ*). A classificação dos sinais em saudável ou patológico é feito por meio de um classificador baseado em análise discriminante linear ou quadrática. É realizada a análise de desempenho da classificação para verificar a viabilidade de implementação da técnica empregada em discriminar vozes patológicas (afetadas por paralisia, nódulos ou edema nas pregas vocais) de vozes saudáveis.

O objetivo da análise discriminante, com as funções linear e quadrática, inclui identificar a contribuição relativa das variáveis empregadas na classificação para a separação dos grupos. Assim, é avaliado o desempenho dos classificadores com as características individuais e, posteriormente, é realizada uma combinação das medidas, com o propósito de melhorar o desempenho na classificação.

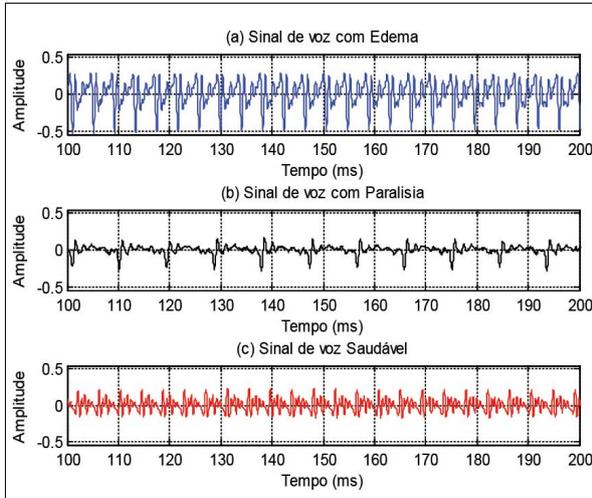
2 Material e Métodos

Para o desenvolvimento deste trabalho, foram utilizados os sinais de voz extraídos da base de dados desenvolvida pelo Massachusetts *Eye and Ear Infirmary* (MEEI) *Voice and Speech Lab* (KayElementrics, 1994a), modelo 4337, construída com o objetivo de auxiliar a análise perceptual de vozes desordenadas para aplicações clínicas e de pesquisa. Os arquivos utilizados são provenientes da emissão da vogal /ah/ sustentada com duração de 1 a 3 segundos, digitalizados com frequência de amostragem de 25 kHz e 16 bits por amostra, coletados em um ambiente controlado com baixo nível de ruído e distância fixa do microfone.

A escolha da vogal sustentada deve-se ao fato de que durante a emissão da vogal sustentada, as dobras vocais vibram, permitindo a observação de seu comportamento na presença da patologia (VIEIRA, 2011). Para este trabalho são selecionados os seguintes sinais: 53 arquivos de vozes saudáveis, além de 43 sinais com edema, 53 afetados por paralisia e 18 com nódulos vocais.

Na Figura 1, estão representados trechos de 100 ms de sinais com as patologias edema de Reinke (Fig. 1a), paralisia nas pregas vocais (Fig. 1b) e para um sinal saudável (Fig. 1c) com o nome padrão da base de dados utilizada no trabalho.

Figura 1 – (a) Edema -JXF11AN.NSP; (b) Paralisia - JFN11AN.NSP; e (c) Saudável -JEG1NAL.NSP.



As características temporais extraídas dos sinais de vozes empregadas neste trabalho são: frequência fundamental e suas variações como *jitter* percentual (perturbações na frequência), *shimmer* percentual (perturbações na amplitude); energia segmental, autocorrelação e taxa de cruzamentos por zero. A extração das medidas temporais de energia segmental, da taxa de cruzamentos por zero e da autocorrelação foi realizada por meio de programação desenvolvida no ambiente Matlab. E para garantir a estacionariedade do sinal de voz (intervalos de 16 a 32 ms), ele foi dividido em segmentos de 20 ms.

A energia segmental, E_{seg} , é definida como:

$$E_{seg} = N_A \cdot E\{[s(n)]^2\} = \sum_{n=0}^{N_A-1} [s(n)]^2 \quad (1)$$

Em que N_A é o tamanho da janela ou o número de amostras do segmento.

E pode ser calculada também como:

$$E_{seg} = 10 \log [E_{seg}] \quad (2)$$

O parâmetro TCZ, a curto intervalo de tempo, é definido, segundo Deller, Proakis e Hansen (1993), como:

$$TCZ = \frac{1}{N_A} \sum_{n=m-N_A+1}^m \frac{|sgn[s(n)] - sgn[s(n-1)]|}{2} \cdot w(m-n) \quad (3)$$

em que:

$$sgn[s(n)] = \begin{cases} +1, & s(n) \geq 0 \\ -1, & s(n) < 0 \end{cases} \quad (4)$$

e $w(m)$ é a janela, *frame* ou segmento do sinal

em análise e N_A é o tamanho da janela ou o número de amostras do segmento.

A expressão geral que define a função de autocorrelação de um sinal $x(n)$ é dada por (RABINER; SCHAFFER, 1978):

$$R_{xx}(k) = \frac{1}{N} \sum_{n=0}^{N-k-1} x(n) \cdot x(n+k), k \geq 0 \quad (5)$$

Os valores das medidas temporais de frequência fundamental, *jitter* percentual e *shimmer* percentual, foram adquiridos pelo *software* MULTI-SPEECH da Kay Elemetrics (KAY ELEMETRICS, 1994b). As medidas foram separadas pelo gênero e submetidas ao *software* Matlab, v.7.11.

A frequência média dos pulsos é denominada frequência fundamental de excitação, F_0 , e o período fundamental (*pitch*), P , é dado por:

$$P = \frac{1}{F_0} \quad (6)$$

O *Jitter* percentual, *Jitt*, é calculado a partir dos valores de *pitch* extraídos período a período do sinal de voz por:

$$Jitt = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |T_0^{(i)} - T_0^{(i+1)}|}{\frac{1}{N} \sum_{i=1}^N T_0^{(i)}} \quad (7)$$

Em que $T_0^{(i)}$ na equação (7) é o período de *pitch* extraído, com $i = 1, 2, \dots, N$, e N = número de

períodos de *pitch* extraídos do sinal de voz em análise (BRANDT, 2011).

O *Shimmer* percentual, *Shimm*, é calculado a partir dos dados de amplitude picoapico do sinal de voz, como:

$$Shimm = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |A^{(i)} - A^{(i+1)}|}{\frac{1}{N} \sum_{i=1}^N A^{(i)}} \quad (8)$$

Em que $A^{(i)}$ na equação (3) é o valor picoapico extraído da amplitude, com $i = 1, 2, \dots, N$, e N = número de impulsos extraídos (BRANDT, 2011).

A frequência fundamental média no Português brasileiro está em torno de 105 Hz para os homens e 213 Hz para as mulheres (RUSSO; BEHLAU, 1993). Logo, a separação pelo gênero é importante para a eficiência dos classificadores (FRAZÃO et al., 2012). Foram utilizados neste trabalho 53 sinais de vozes saudáveis, sendo: (21 masculinos) e (32 femininos) e 108 sinais patológicos, distribuídos em: para o gênero masculino (11 com edemas de Reinke, 1 com nódulo e 22 com paralisia nas pregas vocais) e para o gênero feminino (32 com edemas de Reinke, 17 com nódulos e 25 com paralisia nas pregas vocais).

O classificador utilizado neste trabalho utiliza a função *classify* do Matlab, empregando a análise discriminante linear e quadrática. A análise discriminante linear (LDA – *Linear Discriminant Analysis*) é uma técnica comumente usada para a classificação de dados e redução de dimensionalidade (BALAKRISHNAMA; GANAPATHIRAJU,). Essa técnica clássica em reconhecimento de padrões permite uma previsão direta do grupo que a variável pertence (SAMPAIO, 2006).

A técnica de LDA classifica um objeto da amostra em uma de duas categorias baseadas em determinadas propriedades do objeto (SAMPAIO, 2006). A análise discriminante é uma técnica da estatística multivariada utilizada para discriminar e classificar objetos em duas ou mais classes (KHATTREE; NAIK, 2000). A proposta consiste na obtenção de uma combinação linear das características observadas que apresente maior poder de discriminação entre populações. Esta combinação linear é denominada função discriminante. Quando a regra de classificação assume que as variâncias das populações são iguais, as funções discriminantes são ditas lineares e quando não, são funções discriminantes quadráticas (VARELLA, 2004).

O processo de classificação dos sinais é realizado em duas fases: treinamento e teste/classificação. Na etapa de treinamento, o sinal de voz é pré-processado, segmentado em 20 ms e logo após as características dos sinais são extraídas. Uma parte dos sinais é utilizada para treinamento do sistema (11 sinais saudáveis e 20 sinais patológicos masculinos; 22 saudáveis e 48 patológicos femininos), constituindo os padrões de referência. Os outros sinais (10 saudáveis e 14 patológicos masculinos; 10 saudáveis e 26 patológicos femininos) são utilizados na fase de teste.

Na etapa de teste ou classificação, os sinais são também pré-processados e as características são extraídas e comparadas com os padrões de referência. São empregados classificadores de análise discriminante: linear (LDA) e quadrática (QDA).

Numa primeira classificação, os sinais são classificados como saudável ou patológico, sem especificar qual a patologia presente no sinal de voz. Dois classificadores distintos são empregados, um para vozes femininas e outro para vozes masculinas, empregando a análise discriminante linear e quadrática. O resultado final da classificação é obtido a partir da média entre os resultados obtidos para cada gênero. Na sequência, é feita a classificação especificando a patologia: paralisia, nódulo ou edema nas pregas vocais. Assim, foram realizadas combinações com todas as características temporais empregadas, buscando um resultado mais preciso que consiste na taxa de classificação correta, ou seja, é o número de sinais classificados corretamente dividido pelo número total de sinais.

3 Resultados e Discussões

3.1 Classificação com as medidas F_0 , *Jitter* e *Shimmer*

A separação da classificação foi realizada em dois grupos, sendo o primeiro grupo dado por: frequência fundamental, *jitter* percentual (*Jitt*) e *shimmer* percentual (*Shimm*). Na Tabela 1, estão ilustrados os resultados obtidos na classificação individual com essas medidas temporais. Os sinais são denominados de: sinais patológicos (PAT), sinais afetados por edema de Reinke (EDM), paralisia (PRL) e sinais de vozes saudáveis (SDL). As medidas foram separadas pelo gênero e foi extraído o melhor resultado do

classificador LDA ou QDA. Observa-se que a medida temporal *Shimmer* apresentou os melhores resultados para as medidas de desempenho, chegando a obter uma taxa de correta classificação de 100% entre vozes saudáveis e vozes afetadas por paralisia, como também para vozes com edema de Reinke e vozes com nódulos vocais.

Tabela 1 – Classificação individual das medidas F_0 , *Jitt* e *Shimm*.

	F_0 (%)	<i>Jitt</i> (%)	<i>Shimm</i> (%)
SDLxPAT	79,17	70,14	91,67
SDLxPRL	70,00	82,50	100,00
SDLxEDM	71,07	76,79	90,36
SDLxNDL	87,50	75,00	93,75
PRLxEDM	41,79	64,64	68,57
PRLxNDL	79,17	70,14	91,67
EDMxNDL	70,00	82,50	100,00

Para melhorar o desempenho, as medidas foram combinadas duas a duas, cujos resultados são apresentados na Tabela 2.

Tabela 2 – Classificação das medidas F_0 , *Jitt* e *Shimm* combinadas 2 a 2.

	F_0 e <i>Jitt</i> (%)	F_0 e <i>Shimm</i> (%)	<i>Jitt</i> e <i>Shimm</i> (%)
SDLxPAT	81,25	95,14	86,11
SDLxPRL	85,00	97,50	92,50
SDLxEDM	88,93	92,86	90,36
SDLxNDL	93,75	93,75	81,25
PRLxEDM	57,50	58,93	69,64
PRLxNDL	62,50	62,50	43,75
EDMxNDL	62,50	87,50	87,50

Os melhores resultados obtidos por meio da combinação de duas medidas simultâneas, para a classificação entre sinais saudáveis e patológicos (Tabela 2), representam uma taxa de correta classificação de 95,14%, conseguidos pela combinação das medidas entre F_0 e *Shimm*. A combinação aumentou em cerca de 3,47% a precisão, comparada ao melhor resultado individual (*Shimmer*, 91,67%).

Para a classificação entre os sinais saudáveis e com edema de Reinke (Tabela 2), as melhores medidas de desempenho correspondem a uma taxa

de correta classificação de 92,86%, conseguidos pela combinação das medidas entre F_0 e *Shimmer*. Portanto, houve um aumento de aproximadamente 2,5% em relação ao melhor resultado individual (*Shimmer*, 90,36%).

Para a classificação entre vozes saudáveis e com nódulos, os melhores resultados foram obtidos através do parâmetro *Shimmer* individual, assim como também pela combinação entre as medidas F_0 e *Jitter* e F_0 e *Shimmer*, todos com o valor correspondente de 93,75%.

Na classificação de sinais de vozes com paralisia ou com edema de Reinke, a melhor taxa de correta classificação foi de 69,64%, obtida através da combinação das medidas *Jitter* e *Shimmer* (Tabela 2). Para os sinais de vozes afetados por paralisia ou por nódulo vocal, a melhor taxa de correta classificação foi conseguida pela medida *Shimmer* (91,67%).

3.2 Classificação com as medidas Energia Segmental, TCZ e Autocorrelação

Para o segundo grupo foram selecionadas as medidas de Energia Segmental $E(dB)$, Taxa de Cruzamentos por Zero (TCZ) e Autocorrelação. A Tabela 3 apresenta a classificação individual dessas medidas temporais.

Tabela 3 – Classificação individual das medidas $E(dB)$, TCZ e Autocorrelação (Autocor).

	$E(dB)$ (%)	TCZ (%)	Autocor (%)
SDLxPAT	61,29	54,84	62,90
SDLxPRL	54,76	66,67	73,81
SDLxEDM	58,82	61,76	67,65
SDLxNDL	50,00	65,38	65,38
PRLxEDM	50,00	58,33	58,33
PRLxNDL	46,43	57,14	57,14
EDMxNDL	25,00	45,00	45,00

Devido às desordens vocais provocadas por diferentes patologias (edema de Reinke, nódulo e paralisia) e com diferentes graus de severidade da patologia, algumas medidas apresentaram resultados pouco satisfatórios, como é o caso da energia segmental $E(dB)$. Entretanto, a combinação de classificadores pode contribuir através dos resultados dessas medidas temporais, para efetivamente distinguir vozes saudáveis de vozes patológicas.

Para aumentar o desempenho da classificação, as medidas foram combinadas duas a duas. Os resultados estão ilustrados na Tabela 4.

Tabela 4 – Classificação das medidas E(dB), TCZ e Autocor combinadas 2 a 2.

	E(dB) e TCZ (%)	E(dB) e Autocor (%)	TCZ e Autocor (%)
SDLxPAT	59,68	64,52	74,19
SDLxPRL	66,67	73,81	88,10
SDLxEDM	61,76	61,76	70,59
SDLxNDL	61,54	65,38	65,38
PRLxEDM	50,00	61,11	69,44
PRLxNDL	57,14	57,14	67,86
EDMxNDL	45,00	45,00	50,00

Pela Tabela 4, observa-se que houve um aumento significativo na classificação para os tipos de sinais saudáveis e com paralisia, em relação à Tabela 3, resultando em um acréscimo de 14,29%. Entretanto, não superou o valor obtido pela medida *Shimm* (100%).

3.3 Classificação pela combinação das medidas temporais

Para melhorar os resultados obtidos, foi realizada a combinação para as quatro melhores medidas temporais, F_0 , *Jitt*, *Shimm* e Autocorrelação. A Tabela 5 mostra os resultados encontrados.

Tabela 5 – Classificação pela combinação das medidas F_0 , *Jitt*, *Shimm* e Autocorrelação (%).

SDLxPAT	90,97
SDLxPRL	92,86
SDLxEDM	95,00
SDLxNDL	87,50
PRLx EDM	66,07
PRLxNDL	75,00
EDMxNDL	37,50

Observa-se na Tabela 5 que o melhor resultado da classificação corresponde aos sinais de vozes saudáveis e com edema de Reinke (95%), melhorando em 2,14% a taxa de correta classificação obtida pela Tabela 2.

Na busca por melhorar ainda mais a precisão dos resultados obtidos, foram realizadas combinações com as cinco melhores medidas, F_0 , *Jitter*, *Shimmer*, Autocorrelação e TCZ. A Tabela 6 ilustra os resultados obtidos. Para os tipos de sinais saudáveis e patológicos houve um acréscimo de 4,17% em relação ao resultado obtido (90,97%) na Tabela 5.

Tabela 6 – Classificação através da combinação das medidas F_0 , *Jitter*, *Shimmer*, Autocorrelação e TCZ (%).

SDLxPAT	95,14
SDLxPRL	90,00
SDLxEDM	85,36
SDLxNDL	87,50
PRLx EDM	76,79
PRLxNDL	87,50
EDMxNDL	43,75

4 Conclusões

Os resultados obtidos apontam a medida *Shimmer* percentual (*Shimm*) como a que, individualmente, consegue melhor capturar as desordens vocais provocadas pelas patologias: edema de Reinke, paralisia e nódulo. As medidas temporais frequência fundamental e *Shimmer*, quando combinadas, conseguem com maior eficiência discriminar vozes saudáveis de vozes patológicas, melhorando os resultados obtidos com as medidas individuais do conjunto selecionado. Quando as medidas foram combinadas entre si aconteceram também melhorias nos resultados obtidos, demonstrando a eficiência das características temporais.

Os resultados são bastante promissores e apontam para o uso dessas medidas em sistema de classificação de sinais como ferramenta auxiliar a pré-diagnósticos de patologias laringeas.

Como pesquisa futura, sugere-se a aplicação da técnica de validação cruzada aliada ao uso de classificadores considerados mais robustos, a exemplo de Máquinas de Vetor de Suporte.

REFERÊNCIAS

BALAKRISHNAMA, S.; GANAPATHIRAJU, A. **Linear discriminant analysis – a Brief Tutorial**. Institute for Signal and Information

Processing. Department of Electrical and Computer Engineering, Mississippi State University.

BRANDT, R. R. **Modelagem de vozes patológicas baseadas na estimação espectral do ruído glotal**. Dissertação (Doutorado em Engenharia Elétrica) – Universidade Federal de Campina Grande, Campina Grande, 2011. 88f

BOYANOV, B.; HADJITODOROV, S. Acoustic analysis of pathological voices. **A voice analysis system for the screening of laryngeal diseases**. Proceedings of IEEE Engineering in Medicine And Biology. p. 74-82, Jul./Aug. 1997.

COSTA, S. C.; AGUIAR NETO, B. G.; FECHINE, J. F.; CORREIA, S. **Parametric cepstral analysis for pathological voice assessment**. Proceedings of The 23rd ACM Symposium on Applied Computing 2008 (ACM SAC' 2008). Computer Applications in Health Care Track, Pages 1410-1414, Fortaleza,-Ceará, Brazil, Mar.2008a.

COSTA, S. C. **Análise acústica, baseada no modelo linear de produção da fala, para discriminação de vozes patológicas**. Tese de Doutorado. Universidade Federal de Campina Grande, Campina Grande, 2008b.

COSTA, S. C.; COSTA, W. C. A.; CORREIA, S. Análise Dinâmica não linear aplicada ao diagnóstico de patologias orgânicas e neurológicas da laringe. In: **VI Congresso de Pesquisa e Inovação da Rede Norte e Nordeste de Educação Tecnológica – VI CONNEPI**. . Anais do VI CONNEPI, 2011.

DELLER Jr. R.; PROAKIS, J. G.; HANSEN, J. H. L. **Discrete-time processing of speech signals**. Macmillan Publishing Co., 1993.

DIBAZAR, A. A.; BERGER, T.W.; NARAYANAN, S. S. **Pathological voice assessment**. Proceedings of the 28th IEEE EMBS Annual International Conference, New York- USA, p. 1669-1673, Ago., 2006.

FRAZÃO, M. C. L. M.; COSTA S. C.; CORREIA S.; COSTA, W. C. A. O uso de características temporais de sinais de voz na detecção de patologias laríngeas. In: **VII Congresso de Pesquisa e Inovação da Rede Norte e Nordeste de Educação Tecnológica – VII CONNEPI, 2012**. Anais do VII CONNEPI, 2012.

GODINO-LLORENTE, J. I.; GÓMEZ-VILDA, P.; BLANCO VELASCO, M. **Dimensionality reduction of a pathological voice quality assessment system based on gaussian mixture models and short-**

term cepstral parameters. IEEE Transactions on Biomedical Engineering, v. 53 (10), p.1943-1953, 2006.

KAY ELEMETRICS, **Kay Elemetrics Corp. Disordered Voice Database**, model 4337, 03 ed, 1994a.

KAY ELEMETRICS, **Kay Elemetrics Corp. Multi-Speech**, model 3700, 1994b.

KHATTREE, R.; NAIK, D.N. **Multivariate data reduction and discrimination with SAS software**. 558 p. Cary, NC, USA: SAS Institute Inc., 2000.

RABINER.R.; SCHAFER R. W. **Digital processing of speech signals**. Nova Jersey: Prentice-Hall, 1978.

RUSSO, I. C. P.; BEHLAU, M. **Percepção da fala: análise acústica do português brasileiro**. São Paulo: Lovise, 1993.

SAMPAIO, O. M. **Diferenciação entre rums cubanos e não cubanos**. Instituto de Química de São Carlos- Universidade de São Paulo. São Paulo, 2006.

VARELLA, C.A.A. **Estimativa da produtividade e do estresse nutricional da cultura do milho usando imagens digitais**. Tese (Doutorado em Engenharia Agrícola) – Universidade Federal de Viçosa- Viçosa, 2004. 92f.

VIEIRA, R. T. **Processamento digital de sinais de voz para detecção de patologias na laringe**. . Monografia (Bacharel em Engenharia Elétrica) – Instituto Federal de Educação, Ciência e Tecnologia da Paraíba, João Pessoa, 2011. 66f.