

ANÁLISE DE QUANTIFICAÇÃO DE RECORRÊNCIA A CURTO E A LONGO INTERVALO DE TEMPO NA AVALIAÇÃO DE PATOLOGIAS LARÍNGEAS

V. J. D. Vieira*, S. C. Costa**, S. E. N. Correia**, W. C. de A. Costa** e F. M. de Assis***

*Doutorando do Programa de Pós-Graduação em Engenharia Elétrica, PPgEE/UFCG, Campina Grande, Brasil

** Instituto Federal de Educação, Ciência e Tecnologia da Paraíba, João Pessoa, Brasil

***Departamento de Engenharia Elétrica, Universidade Federal de Campina Grande, Campina Grande, Brasil

e-mail: vinicius.vieira@ee.ufcg.edu.br

Resumo: Este artigo trata da aplicação da análise de quantificação de recorrência a curto e a longo intervalo de tempo em sinais de voz de laringes saudáveis e sinais de voz de laringes patológicas (paralisia, edema e nódulos). O objetivo é identificar qual medida (ou conjunto de medidas) e qual das formas de análise apresenta melhor desempenho. São extraídas, a partir dos gráficos de recorrência, seis medidas de quantificação: determinismo, comprimento máximo das linhas diagonais, entropia de Shannon, laminaridade, tempo de permanência e comprimento máximo das linhas verticais. Os resultados obtidos, empregando um classificador baseado em análise discriminante quadrática, indicam que a análise a longo intervalo de tempo é mais promissora, principalmente com a medida do comprimento máximo das linhas diagonais, com a qual se atinge uma acurácia de até $96,27\% \pm 1,53\%$.

Palavras-chave: Análise de quantificação de recorrência, processamento digital de sinais de voz, patologias laringeas.

Abstract: *This paper deals with the application of recurrence quantification analysis in short and long time interval voice signals of healthy and pathological larynges (paralysis, edema and nodules). The goal is to identify which measure (or set of measures) and which way of the analysis reaches the best performance. From the recurrence plots six quantification measures are extracted: determinism, maximum length of the diagonal lines, Shannon's entropy, laminarity, trapping time and maximum length of the vertical lines. The obtained results, employing a classifier based on quadratic discriminant analysis suggest that the long time interval analysis is more promising, mostly with the maximum length of the diagonal lines measure, which reaches an accuracy of $96.27\% \pm 1.53\%$.*

Keywords: *Recurrence quantification analysis, digital speech signal processing, laryngeal pathologies.*

Introdução

Técnicas de processamento digital de sinais de voz aplicadas à avaliação de distúrbios vocais provocados por

patologias laringeas, em sua maioria, levam em consideração uma análise realizada a curto intervalo de tempo, em que os sinais são particionados em segmentos que mantenham sua estacionariedade, de forma que as medidas estatísticas não mudem no intervalo de tempo de análise. Entre essas técnicas estão aquelas baseadas no modelo linear de produção de voz (teoria fonte-filtro) [1, 2].

Nos anos mais recentes, as técnicas de análise dinâmica não linear apareceram como alternativas de avaliação da voz [3, 4], a exemplo da análise de quantificação de recorrência, a qual permite a análise de séries temporais que podem ou não ser estacionárias. Trabalhos recentes têm demonstrado o potencial deste método quando não realizado a curto intervalo de tempo, ou seja, quando não considera a estacionariedade do sinal de voz [5, 6].

Embora a análise de quantificação de recorrência realizada a longo intervalo de tempo tenha atingido resultados promissores na avaliação de patologias laringeas [5, 6], a investigação de seu potencial quando realizada a curto intervalo de tempo é relevante para indicar qual dos dois tipos de análise fornece a mais alta acurácia.

Neste trabalho, seis medidas de quantificação de recorrência são empregadas a curto e a longo intervalo de tempo: determinismo, comprimento máximo das linhas diagonais, entropia de Shannon, laminaridade, tempo de permanência e comprimento máximo das linhas verticais. A classificação, realizada com a análise discriminante quadrática, tem como objetivo identificar qual a forma mais eficaz de se empregar a análise de quantificação de recorrência na avaliação de patologias laringeas.

Análise de Quantificação de Recorrência

A análise de quantificação de recorrência é um método em que se extraem medidas dos chamados gráficos de recorrência [7]. Estes, por sua vez, são construídos a partir de séries temporais e refletem a dinâmica do sistema que produziu essas séries.

Um gráfico de recorrência é uma matriz quadrada de ordem N preenchida por pontos brancos e pretos. Um

ponto preto representa um estado recorrente, ou seja, um estado suficientemente próximo a outro em determinado instante de tempo. Um ponto branco indica estado não recorrente. Na Figura 1 são apresentados exemplos de gráficos de recorrência obtidos a partir de sinais de voz.

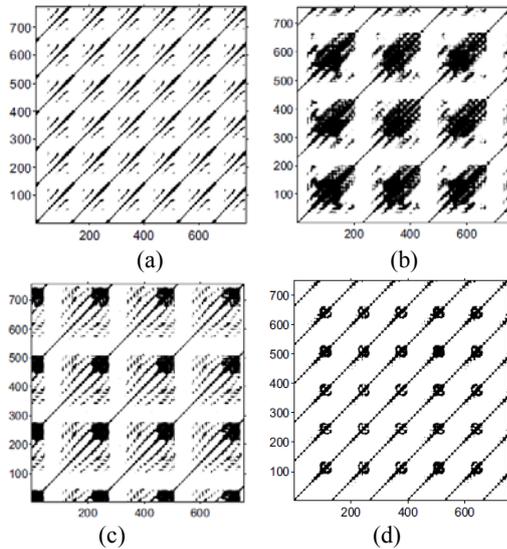


Figura 1: Gráficos de recorrência obtidos de sinais de voz: (a) laringe saudável; (b) laringe com paralisia; (c) laringe com edema; (d) laringe com nódulos.

Gráficos de recorrência para sinais de voz saudáveis (Figura 1a) apresentam a formação de estruturas diagonais pequenas e grandes, paralelas umas às outras. Isto representa a regularidade de um sistema de produção vocal saudável. Gráficos de recorrência de sinais de voz afetados por patologias laringeas (Figuras 1b, 1c e 1d) apresentam, além da formação de estruturas diagonais, estruturas verticais e horizontais espalhadas ao longo das diagonais maiores, indicando a possível presença de diferentes frequências resultantes da vibração irregular das pregas vocais [8].

As medidas de quantificação de recorrência dão informações objetivas a respeito das estruturas formadas no gráfico de recorrência, uma vez que a análise visual desses gráficos não deixa de ser subjetiva, podendo levar a diferentes interpretações. As principais medidas estão elencadas a seguir [9, 10].

Taxa de Recorrência (REC) – mede o percentual de pontos recorrentes no gráfico de recorrência:

$$REC = \frac{1}{N^2} \sum_{i,j=1}^N R_{i,j;i \neq j}^{m,\varepsilon}, \quad (1)$$

em que $R_{i,j;i \neq j}^{m,\varepsilon}$ é o elemento (i, j) da matriz de recorrência, formada com m graus de liberdade e raio de vizinhança ε . A taxa de recorrência pode ser utilizada em conjunto com o raio de vizinhança para a construção do gráfico de recorrência [5]. Dessa forma, o raio é escolhido de modo que a taxa de recorrência não ultrapasse determinado valor. Essa metodologia é empregada quando se deseja que outras medidas não tenham valores saturados [11], como é o caso da medida determinismo, descrita a seguir.

Determinismo (DET) – mede a quantidade dos pontos de recorrência presentes na formação das linhas diagonais em relação a todo o conjunto dos pontos de recorrência:

$$DET = \frac{\sum_{l=l_{min}}^N l \times P^{\varepsilon}(l)}{\sum_{i,j}^N R_{i,j}^{m,\varepsilon}}, \quad (2)$$

em que $P^{\varepsilon}(l) = \{l_i; i=1 \dots N_l\}$ representa a distribuição de frequência dos comprimentos l das estruturas diagonais no gráfico de recorrência e N_l é o número absoluto de linhas diagonais. l_{min} representa o número mínimo de pontos que são considerados para formar uma linha diagonal (geralmente, $l_{min}=2$). Esta medida está relacionada ao grau de previsibilidade do sistema.

Comprimento máximo das linhas diagonais (Lmax) – está relacionado ao tempo máximo de predição do sistema:

$$L_{max} = \max(\{l_i; i = 1, \dots, N_l\}). \quad (3)$$

Quanto mais bem comportado for o sistema, maior será o comprimento das estruturas diagonais presentes nos gráficos de recorrência [8].

Entropia de Shannon (ENTR) – mede a distribuição de frequência dos comprimentos das linhas diagonais. Ela reflete a complexidade da estrutura determinística no sistema:

$$ENTR = - \sum_{l=l_{min}}^N \frac{P^{\varepsilon}(l)}{\sum_{l=l_{min}}^N P^{\varepsilon}(l)} \ln \left(\frac{P^{\varepsilon}(l)}{\sum_{l=l_{min}}^N P^{\varepsilon}(l)} \right) \quad (4)$$

Laminaridade (LAM) – mede a quantidade dos pontos de recorrência presentes na formação das linhas verticais em relação a todo o conjunto dos pontos de recorrência:

$$LAM = \frac{\sum_{v=v_{min}}^N v \times P^{\varepsilon}(v)}{\sum_{v=1}^N v \times P^{\varepsilon}(v)} \quad (5)$$

em que $P^{\varepsilon}(v) = \{v_i; i=1 \dots N_v\}$ representa a distribuição de frequência dos comprimentos v das estruturas verticais no gráfico de recorrência e N_v é o número absoluto de linhas diagonais. v_{min} representa o número mínimo de pontos que são considerados para formar uma linha vertical (geralmente, $v_{min}=2$). A formação de estruturas verticais nos gráficos de recorrência indica a presença de estados laminares, ou seja, indica a presença de comportamento caótico no sistema [10].

Tempo de permanência (TT) – mede o comprimento médio das estruturas verticais:

$$TT = \frac{\sum_{v=v_{min}}^N v \times P^{\varepsilon}(v)}{\sum_{v=v_{min}}^N P^{\varepsilon}(v)} \quad (6)$$

Também conhecida como *Trapping Time*, esta medida contém informação acerca da quantidade e do comprimento das estruturas verticais no gráfico de recorrência, observando o tempo médio em que o sistema permanece em um estado específico [10].

Comprimento máximo das linhas verticais (Vmax) – mede o tempo máximo em que o sistema permanece em um estado específico, estando relacionada à duração máxima de um comportamento caótico:

$$V_{max} = \max(\{v_i; i = 1, \dots, N_v\}). \quad (7)$$

Materiais e métodos

Base de dados – São analisados 171 sinais de vozes da vogal sustentada /a/, oriundos de um banco de dados, *Disordered Voice Database, Model 4337*, da *Kay Elemetrics* [12]. Dos sinais avaliados, 53 são de vozes saudáveis e 118 são de vozes afetadas por patologias na laringe (55 com paralisia, 45 com edema de Reinke, e 18 com nódulo vocal). A taxa de amostragem trabalhada é de 25 kHz.

Análise a longo intervalo de tempo (LT) – São analisados trechos de 400 ms (10.000 amostras) de cada sinal de voz da base de dados. As medidas de quantificação de recorrência são extraídas por meio do software RQA (*Recurrence Quantification Analysis*) [13], considerando cinco patamares de taxa de recorrência: 1%, 2%, 3%, 4% e 5%.

Análise a curto intervalo de tempo (CT) – Os sinais de voz da base de dados são particionados em segmentos consecutivos de 32 ms (800 amostras). Ao todo, são 30 segmentos obtidos por sinal. As medidas de quantificação de recorrência são obtidas por meio da *toolbox CRP (Cross Recurrence Plot)* [14]. A extração dessas medidas considera os mesmos patamares de taxa de recorrência que na análise a longo intervalo de tempo.

Classificação – Em ambas as formas de análise, a classificação é realizada com a análise discriminante quadrática (QDA – *Quadratic Discriminant Analysis*), utilizando validação cruzada e 10 subconjuntos de forma estratificada. Com as seis medidas utilizadas (*DET, Lmax, ENTR, LAM, TT e Vmax*), a classificação é realizada para cada uma delas individualmente e com elas combinadas duas a duas, três a três, quatro a quatro, cinco a cinco e todas conjuntamente. O desempenho da classificação é medido por acurácia, sensibilidade e especificidade [5].

Resultados

Na Tabela 1 estão apresentados os principais resultados obtidos na análise a longo intervalo de tempo (LT) e a curto intervalo de tempo (CT), no caso de discriminação entre sinais de voz de laringes saudáveis e sinais de voz de laringes patológicas (paralisia, edema e nódulos). A longo intervalo de tempo, a medida *Lmax* foi suficiente para proporcionar o melhor desempenho do classificador, na taxa de recorrência de 2%. A curto intervalo de tempo, são necessárias duas medidas, *Lmax* e *Vmax*, mesmo assim sem atingir o desempenho da outra análise.

Tabela 1: Melhor desempenho na classificação saudável versus patológica nas análises a longo (LT) e a curto intervalo de tempo (CT).

Métrica	LT	CT
Medida(s)	<i>Lmax</i>	<i>Lmax</i> e <i>Vmax</i>
<i>REC</i> ≤	2%	1%
Acurácia (%)	94,15 ± 1,24	88,58 ± 1,94
Sensibilidade (%)	94,85 ± 1,95	89,25 ± 1,23
Especificidade (%)	92,33 ± 4,34	88,27 ± 0,64

Nas Tabelas 2, 3 e 4 são apresentados os principais resultados obtidos na classificação dos sinais de laringes saudáveis e dos sinais de laringes patológicas. Individualmente ou combinada com outras medidas, *Lmax* proporcionou o melhor desempenho do classificador, na análise a longo intervalo de tempo.

Tabela 2: Melhor desempenho na classificação saudável versus paralisia nas análises a longo (LT) e a curto (CT) intervalo de tempo.

Métrica	LT	CT
Medida(s)	<i>Lmax</i>	<i>DET, Lmax, ENTR</i> e <i>Vmax</i>
<i>REC</i> ≤	2%	1%
Acurácia (%)	96,27 ± 1,53	91,52 ± 0,48
Sensibilidade (%)	96,33 ± 2,46	94,34 ± 0,66
Especificidade (%)	96,00 ± 2,67	88,65 ± 0,72

Tabela 3: Melhor desempenho na classificação saudável versus edema nas análises a longo (LT) e a curto (CT) intervalo de tempo.

Métrica	LT	CT
Medida(s)	<i>Lmax, LAM</i> e <i>TT</i>	<i>Lmax</i> e <i>Vmax</i>
<i>REC</i> ≤	2%	1%
Acurácia (%)	90,78 ± 2,45	85,57 ± 0,72
Sensibilidade (%)	89,00 ± 3,71	89,37 ± 0,84
Especificidade (%)	92,33 ± 3,14	80,98 ± 1,36

Tabela 4: Melhor desempenho na classificação saudável versus nódulos nas análises a longo (LT) e a curto (CT) intervalo de tempo.

Métrica	LT	CT
Medida(s)	<i>DET</i> e <i>Lmax</i>	TODAS, menos <i>Vmax</i>
<i>REC</i> ≤	1%	1%
Acurácia (%)	95,71 ± 2,18	91,69 ± 0,89
Sensibilidade (%)	95,00 ± 5,00	90,82 ± 0,99
Especificidade (%)	96,00 ± 2,67	94,26 ± 0,94

Discussão

Nos resultados obtidos, a medida do comprimento máximo das linhas diagonais (*Lmax*) foi a que mais se destacou, refletindo a maior regularidade da vibração de pregas vocais saudáveis. A presença de *DET* e *ENTR* entre os principais resultados também indica que a formação das linhas diagonais é um fator discriminativo entre as classes consideradas. Apesar de os sinais provenientes de laringes patológicas apresentarem a formação de estruturas diagonais, com menor quantidade e menor comprimento que os sinais saudáveis.

A contribuição das medidas *LAM, TT* e *Vmax* indica que a presença de estados laminares (comportamento caótico), que aparece em maior número nos sinais de laringes patológicas e é refletido com a formação de linhas verticais [8], é um fator discriminante entre sinais

de voz laringes saudáveis e sinais de voz de laringes patológicas.

Na discriminação saudável *versus* patológica (Tabela 1) e saudável *versus* paralisia (Tabela 2), o classificador se mostrou mais sensível para identificar o estado patológico da laringe (maiores valores para a sensibilidade). No caso saudável *versus* edema (Tabela 3), o classificador obteve valores maiores para a especificidade na análise a longo intervalo de tempo e para a sensibilidade a curto intervalo de tempo. Na discriminação entre laringes saudáveis e laringes com nódulos (Tabela 4), o classificador obteve maior valor para especificidade em ambas as formas de análise.

A análise de quantificação de recorrência ainda tem sido pouco explorada em aplicações de processamento digital de sinais de voz voltado à avaliação de distúrbios vocais. Em trabalho anterior, este método foi aplicado na discriminação entre vozes saudáveis e patológicas, a longo intervalo de tempo, obtendo acurácia de $94,19\% \pm 0,87\%$, combinando a medida Entropia de Shannon com 12 coeficientes LPC no classificador QDA [15]. Em [16], 15 medidas de recorrência foram extraídas, a curto intervalo de tempo, com redes neurais MLP como classificador, tendo o melhor resultado na classificação entre saudável e nódulos, com acurácia de $99,11\%$.

Godino et al. [1] e Wang et al. [2], utilizaram vetores multidimensionais de características com análise a curto intervalo de tempo. No primeiro, 24 coeficientes mel-cepstrais e 6 modelos de misturas gaussianas foram utilizados para classificar sinais saudáveis e patológicos, chegando a uma acurácia de $94,07\% \pm 3,28\%$. No segundo, foram utilizados 16 coeficientes mel-cepstrais e 16 misturas gaussianas, aplicados a máquinas de vetor de suporte, obtendo acurácia de $96,1\% \pm 2,51\%$, na discriminação entre vozes saudáveis e patológicas.

Os resultados obtidos neste trabalho indicam que a análise de quantificação de recorrência é uma técnica promissora na discriminação entre sinais de vozes saudáveis e sinais afetados por patologias laringeas. Em relação ao que tem se encontrado na literatura, uma pequena quantidade de características foi suficiente para atingir resultados satisfatórios, confirmando uma melhor precisão da análise a longo intervalo de tempo.

Agradecimentos

À CAPES e ao CNPq pelo financiamento parcial desta pesquisa.

Referências

[1] Godino-Llorente JI, Gomes-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*. 2006; 53(10):1943-1953.

[2] Wang X, Zhang J, Yan Y. Discrimination between pathological and normal voices using GMM-SVM approach. *Journal of Voice*. 2011; 25(1):38-43.

[3] Henríquez P, Alonso JB, Ferrer MA, Travieso CM, Godino-Llorente, Díaz-de-Maria F. Characterization of healthy and pathological voice through measures based on nonlinear dynamics. *IEEE Transactions on Audio, Speech and Language Processing*. 2009; 17(6):1186-1195.

[4] Vaziri G, Almasganj F, Behroozmand R. Pathological assessment of patients' speech signals using nonlinear dynamical analysis. *Computers in Biology and Medicine*. 2010; 40(1):54-63.

[5] Costa WCA. Análise dinâmica não linear de sinais de voz para detecção de patologias laringeas [tese de Doutorado]. Campina Grande, Universidade Federal de Campina Grande; 2012.

[6] Lopes LW, Costa SLNC, Costa WCA, Correia SEN, Vieira VJD. Análise da dinâmica não linear de vozes infantis: nova proposta de avaliação e monitoramento vocal. *In: Pesquisas em Fonoaudiologia*. Sociedade Brasileira de Fonoaudiologia; 2013.

[7] Eckmann JP, Kamphorst SO, Ruelle D. Recurrence plots of dynamical systems. *Europhysics Letters*. 1987; 4:973-977.

[8] Vieira VJD. Avaliação de distúrbios da voz por meio de análise de quantificação de recorrência [dissertação]. João Pessoa: Instituto Federal de Educação, Ciência e Tecnologia da Paraíba; 2014.

[9] Webber-Jr. CL, Zbilut JP. Dynamical assessment of physiological systems and states using recurrence plot strategies. *Journal of Applied Physiology*. 1994; 76(2):965-973.

[10] Marwan N. Encounters with neighbours – current developments of concepts based on recurrence plots and their applications [PhD Thesis]. Potsdam: University of Potsdam; 2003.

[11] Webber-Jr CL, Zbilut JP. Recurrence quantification analysis of nonlinear dynamical systems. *In: Tutorials in Contemporary Nonlinear Methods for the Behavioral Sciences*. National Science Foundation; 2005.

[12] Kay Elemetrics Corp. Disordered Voice Database, Model 4337; 1994.

[13] Webber-Jr. C. L. Recurrence quantification analysis (RQA) software 13. 2009. Available from: <http://homepages.luc.edu/~cwebber/>.

[14] Marwan N. Cross Recurrence Plot toolbox for Matlab, version 5.17(r28.16). 2013. Available from: <http://tocsy.pik-potsdam.de/CRPtoolbox/index.html>.

[15] Costa WCA, Vieira VJD, Costa SC, Assis FM, Aguiar Neto BG. Avaliação do uso combinado de medidas de quantificação de recorrência e análise LPC na classificação de vozes patológicas. *In: Anais do XXIII Congresso Brasileiro de Engenharia Biomédica (CBEB 2012)*; Porto de Galinhas, Brasil. 2012.

[16] Vieira VJD, Costa SC, Costa WCA, Correia SEN, Assis FM. Discriminação de sinais de voz com análise de quantificação de recorrência e redes neurais MLP. *In: Anais do XXXI Simpósio Brasileiro de Telecomunicações (SBrT 2013)*; Fortaleza, Brasil. 2013.