

# Classificação de disfonias por meio da análise de medidas não lineares e de quantificação de recorrência

Giulliana Karla Lacerda Pereira de Queiroz <sup>[1]</sup>, Silvana Luciene do Nascimento Cunha Costa <sup>[2]</sup>, Suzete Élide Nóbrega Correia <sup>[3]</sup>, Leonardo Wanderley Lopes <sup>[4]</sup>, Vinícius Jefferson Dias Vieira <sup>[5]</sup>

[1] giullianalacerda@gmail.com. [2] silvanacunhacosta@gmail.com. [3]suzete.correia@gmail.com. [4] lwlopes@hotmail.com. UFPB – Universidade Federal da Paraíba, Campus – João Pessoa. [5] viniciusjdv@gmail.com. UFCG – Universidade Federal de Campina Grande, Campus – Campina Grande.

## RESUMO

Nos últimos anos, técnicas de processamento digital de sinais de voz vêm sendo empregadas como ferramentas auxiliares na avaliação de desvios vocais, trazendo maior comodidade ao paciente, baixo custo e caráter objetivo, comparado às técnicas tradicionalmente empregadas, como a análise perceptivo-auditiva. A avaliação da qualidade vocal, por análise acústica de sinais de voz, se constitui cada vez mais em prática clínica para detecção de distúrbios vocais que, em alguns casos, podem ser ocasionados por lesões laringeas ou abuso vocal. Nesta pesquisa, foram empregadas algumas medidas não lineares tradicionais, combinadas com medidas de quantificação de recorrência, para a análise discriminativa dos desvios vocais rugosidade, soprosidade e tensão. As características da análise dinâmica não linear, abordadas para o processo de classificação, foram o Passo de Reconstrução ( $\tau$ ), o Primeiro Mínimo da Função de Informação Mútua ( $PM$ ) e a Dimensão de correlação ( $D2$ ). As medidas de quantificação empregadas foram: Determinismo ( $Det$ ), Entropia de Shannon ( $Entr$ ), Comprimento médio das linhas diagonais ( $Lmed$ ), Comprimento máximo das linhas verticais ( $Vmax$ ) e Transitividade ( $Trans$ ). Por meio de testes estatísticos, foi avaliado o potencial de cada característica em discriminar os tipos de sinais de voz considerados. No processo de classificação, foi utilizada a rede neural MLP (*Multilayer Perceptron*), com algoritmo de aprendizado supervisionado Scheduling Conjugate Graded ( $SCG$ ). Obteve-se uma acurácia média de 90% na discriminação entre vozes saudáveis e desviadas. Na classificação entre vozes saudáveis e tensas, obteve-se uma acurácia média de 76%, com as medidas combinadas  $Trans$ ,  $\tau$ ,  $Vmax$ ,  $Lmed$ ,  $Det$  e  $D2$ . Na detecção do desvio rugosidade, obteve-se uma acurácia média de 89% com as medidas  $Lmed$ ,  $Entr$ ,  $Trans$  e  $D2$  e, na distinção entre as vozes saudáveis e soprosas, obteve-se 91,17% com apenas duas medidas combinadas  $Trans$  e  $\tau$ , demonstrando o caráter promissor da técnica empregada.

**Palavras-chave:** Análise Não Linear. Análise de Quantificação de Recorrência. Análise Acústica. Processamento Digital de Sinais de Voz. Avaliação da Qualidade Vocal.

## ABSTRACT

*In recent years techniques of digital processing of speech signals have been used as an auxiliary tool in the evaluation of vocal deviations, providing the patient with greater comfort low cost and objectivity when compared to the techniques traditionally employed, such as perceptual-auditory analysis. The evaluation of vocal quality, through acoustic analysis of voice signals, is becoming a very popular clinical practice for the detection of vocal disorders that in some cases can be caused by laryngeal lesions or vocal abuse. In this research, we used some traditional non-linear measures combined with measures of recurrence quantification for the discriminative analysis of vocal deviations, breathiness, roughness and strain. The characteristics of the non-linear dynamic analysis, used in the classification process, were the Reconstruction Step ( $\tau$ ), the First Minimum of the Mutual Information Function ( $PM$ ) and the Correlation Dimension ( $D2$ ). The quantification measures employed were: Determinism ( $Det$ ), Shannon entropy ( $Entr$ ), Mean length of diagonal lines ( $Lmed$ ), Maximum length of vertical lines ( $Vmax$ ) and Transitivity ( $Trans$ ). Through these statistical tests, the potential of each characteristic to discriminate the types of voice signals was evaluated. In the classification process, the neural network MLP (*Multilayer Perceptron*) was used, with supervised learning algorithm Graded Conjugate Gradient ( $SCG$ ). There was an average accuracy of 90% in the discrimination between healthy and deviant voices. In the classification between healthy and strained voices, an average accuracy of 76% was obtained with the combined measures  $Trans$ ,  $\tau$ ,  $Vmax$ ,  $Lmed$ ,  $Det$  and  $D2$ . In the detection of the roughness deviation, an average accuracy of 89% was obtained with the  $Lmed$ ,  $Entr$ ,  $Trans$  and  $D2$  measures and in the distinction between healthy and breathy voices, 91.17% of accuracy was obtained with only two combined measures,  $Trans$  and  $\tau$ , showing the promising character of the used technique.*

**Keywords:** *Non-linear Analysis. Recurrence Quantification Analysis. Acoustic Analysis. Digital Processing of Speech Signals. Vocal Quality Assessment.*

## 1 Introdução

A voz é uma ferramenta de grande complexidade e possui inúmeros fatores envolvidos na sua produção. As estruturas anatômicas e os processos fisiológicos envolvidos no sistema de produção da fala são responsáveis pelos diversos fonemas emitidos e, conseqüentemente, carregam informações sobre suas formas, fisiologia e dinâmica (BEHLAU, 2001).

A partir da análise da voz, é possível extrair informações sobre a saúde geral e bem-estar de uma pessoa. Além de revelar quem somos e como nos sentimos, a voz fornece consideráveis percepções sobre a estrutura e função de certas partes do corpo. A presença de algum tipo de disфония causa mudanças significativas em seus padrões vibratórios, afetando assim a qualidade da produção vocal.

A qualidade vocal é a impressão global que se relaciona com o timbre, a frequência, a amplitude e a composição dos harmônicos da onda sonora produzida bem como com o modo de vibração das pregas vocais, o tamanho e o formato do trato vocal, a tensão e os tónus das paredes faríngeas, podendo ser modificada por fatores anatômicos ou patológicos (FANT, 2004).

Uma das formas mais utilizadas na avaliação da voz é a análise perceptivo-auditiva, que detecta os distúrbios e caracteriza o tipo de voz, tais como: ressonância, ataque vocal, tessitura vocal, qualidade vocal, velocidade de fala, respiração, pitch e tensão geral do aparelho fonador. Deve ser ressaltada, no entanto, a importância da utilização de instrumentos complementares para a avaliação fonoaudiológica, possibilitando uma avaliação mais precisa da qualidade vocal (CRUZ; OLIVEIRA, 2014).

A análise acústica permite um registro e oferece a possibilidade de detalhamento do processo de geração do sinal sonoro, o qual corresponde a eventos das porções glótica e supraglótica do trato vocal (CAMPISI *et al.*, 2000). Nesse contexto, o interesse pela análise acústica como uma ferramenta para o auxílio ao diagnóstico e tratamento de doenças na laringe vem aumentando cada vez mais nos últimos anos (COSTA, 2012; VIEIRA, 2014). A escolha de características relevantes, que pos-

sam representar bem as variações existentes no sinal de voz introduzidas pela presença de disfonias e/ou de alterações na laringe, pode aumentar a eficiência no processo de detecção de desvios vocais como também na classificação do tipo e do grau do desvio vocal.

Nos últimos anos, medidas não lineares, como aquelas baseadas na Teoria do Caos, e as medidas de quantificação de recorrência vêm ganhando destaque no estudo da produção da voz, considerando que o sistema de produção vocal seja não linear, levando em conta a complexidade inerente à vibração das pregas vocais (JIANG, ZHANG, MCGILLIGAN, 2006; COSTA, 2012; VIEIRA, 2014; SANTOS, 2015; LOPES *et al.*, 2016).

As teorias não lineares têm sido utilizadas cada vez mais para a interpretação, explicação e previsão dos comportamentos de fenômenos físicos e biológicos. Os modelos não lineares ainda não são capazes de explicar toda complexidade presente nos sistemas humanos, e é fundamental que estudos surjam de modo a buscar o refinamento desta ciência. Estes modelos estão ajudando a explicar certos tipos de sistemas que os modelos tradicionais lineares não explicam (COSTA, 2012; VIEIRA, 2014).

A análise de quantificação de recorrência surgiu como forma de potencializar a avaliação do comportamento de séries temporais, proporcionando análise de sinais tanto na forma visual (gráficos de recorrência) quanto paramétrica, a partir das medidas de quantificação de recorrência (WEBBER, ZBILUT, 1994 *apud* VIEIRA, 2014; MARWAN, 2003).

Neste trabalho, o comportamento do sinal de voz será avaliado utilizando as medidas não lineares (Passo de reconstrução, Primeiro mínimo da função de informação mútua e Dimensão de correlação) e medidas de quantificação de recorrência (Determinismo, Entropia de Shannon, Comprimento médio das linhas diagonais, Comprimento máximo das linhas verticais e Transitividade) para a classificação de desvios (rugosidade, soprosidade e tensão) em vozes adultas. As medidas serão avaliadas de forma individual e combinada. A escolha destas medidas se deu pelo fato de terem apresentado os melhores resultados em trabalhos já existentes (COSTA, 2012; LOPES *et al.*, 2016). Os sinais foram avaliados de forma a indicar a presença ou não do desvio (Saudável x Desviada) e qual o desvio presente (Saudável x Soprosa, Saudável x Rugosa, Saudável x Tensão).

O processo de classificação é realizado empregando redes neurais artificiais, com validação cruzada.

O artigo está organizado da seguinte forma: na Seção 2, são apresentadas as pesquisas correlacionadas; na Seção 3, está contida a fundamentação teórica; na Seção 4, é apresentada a metodologia empregada; e na Seção 5, estão apresentados os resultados obtidos, seguidos das considerações finais, na Seção 6.

## 2 Pesquisas correlatas

Diversas pesquisas têm sido realizadas, nos últimos anos, utilizando a análise acústica, para avaliação de desordens vocais em sinais de vozes afetados por patologias na laringe bem como para a avaliação da qualidade vocal, identificando a presença de desvios vocais como rugosidade, soprosidade e tensão.

A seguir, são apresentados os resultados de algumas pesquisas encontradas na literatura que buscam desenvolver novos métodos para dar mais precisão à análise de distúrbios da voz.

Zhang e Jiang (2008) pesquisaram as características de vogais faladas de forma contínua e de forma sustentada por locutores com vozes saudáveis e pacientes com patologias da laringe. Foram usadas as medidas de perturbação (incluindo Jitter e Shimmer), a Relação Sinal-Ruído (SNR) e os métodos de dinâmica não linear (dimensão de correlação e entropia de segunda ordem). Os resultados obtidos mostraram que, para as vogais sustentadas, a análise de Jitter, Shimmer, a dimensão de correlação e a entropia de segunda ordem revelaram diferenças significativas entre vozes saudáveis e patológicas. Por outro lado, para voz contínua, Jitter e Shimmer não discriminaram, estatisticamente, vozes saudáveis de patológicas, porém uma diferença significativa foi encontrada para a SNR (Relação Sinal-Ruído), dimensão de correlação e entropia de segunda ordem.

Henríquez *et al.* (2009) pesquisaram a possibilidade de se quantificar a qualidade da voz (saudável ou patológica) por meio do uso de seis medidas não lineares: as entropias de Rényi de primeira e segunda ordem, a entropia de correlação (ou entropia de segunda ordem), a dimensão de correlação, a entropia de Shannon e o valor do primeiro mínimo da função de informação mútua. Foram empregados dois bancos de dados: um não comercial, composto de quatro diferentes tipos de qualidade da voz (voz saudável e três níveis de vozes patológicas – alta, média e baixa severidade), de acordo com o parâ-

metro do grau de rouquidão (*Grade of hoarseness*), o (G) da escala GRBAS; e um comercial, “MEEI Voice Disorders Database”, composto de vozes saudáveis e vozes patológicas. Utilizando um classificador baseado em redes neurais MLP e todas as características combinadas, a taxa global de sucesso do sistema de classificação empregado foi de 82,47% (banco de dados não comercial) e de 99,69% (banco de dados MEEI). Segundo os autores, a diferença entre as taxas de classificação obtidas a partir dos dois bancos de dados deve-se ao fato da não existência, no banco de dados MEEI, de vozes patológicas de baixa severidade.

Dajer (2010), em sua Tese, avaliou vozes saudáveis e com alterações patológicas aplicando a análise de Padrões Visuais da Dinâmica Vocal (PVDV) em conjunto com a análise acústica e a análise perceptivo-auditiva. Foram utilizadas as medidas de Jitter, Shimmer e Frequência Fundamental. Na análise perceptivo-auditiva foram avaliadas rugosidade, soprosidade, tensão e instabilidade. Foi utilizada a técnica de reconstrução de espaço de fase e, qualitativamente, foram usados parâmetros de laços, regularidade e convergência de traçados para analisar a dinâmica dos sinais de voz dos PVDV. Os resultados demonstram que o Jitter apresentou uma correlação negativa com laços, regularidade e convergência de laços.

Costa (2012), em sua Tese, trata da caracterização e da classificação de sinais de vozes saudáveis e vozes afetadas por diferentes patologias laríngeas (edema, paralisia e nódulos nas pregas vocais), por meio da análise dinâmica não linear (e teoria do caos) como também por meio da análise de quantificação de recorrência. Na classificação, foi empregada a técnica de análise discriminante com função linear ou quadrática, com validação cruzada. A partir da combinação de características dos conjuntos das medidas de análise não linear (MNL) e das medidas de quantificação de recorrência (MQR), as médias da taxa de acurácia obtidas variaram nos intervalos de confiança: [95,44%; 100%] para a classificação entre vozes saudáveis e patológicas; [94,75%; 100%] entre vozes saudáveis e afetadas por edema, e entre saudáveis e nódulos. Para a classificação entre saudável e paralisia, obteve-se uma acurácia de 100%. Também são avaliados os efeitos do uso de vetores híbridos formados por características MNL, MQR e coeficientes extraídos da análise preditiva linear (LPC). Neste caso, as taxas de acurácia variaram nos intervalos de confiança: [95,02%; 97,62%] na discri-

minação entre vozes afetadas por paralisia e edema; [98,29%; 99,93%] para paralisia versus nódulos e [97,98%; 99,84%] para edema versus nódulos.

Vieira (2014), em seu trabalho de dissertação, empregou a análise de quantificação de recorrência para a avaliação dos distúrbios da voz. Foi realizada a análise de sinais de vozes de locutores afetados por patologias laríngeas, com a base da *Kay Elemetrics*, e análise de sinais de vozes infantis diagnosticadas de acordo com o grau de intensidade do desvio vocal. As medidas empregadas na análise sem segmentação foram: determinismo, comprimento máximo das linhas diagonais, entropia de Shannon, tendência, laminaridade, tempo de permanência e comprimento máximo das linhas verticais. Na análise dos sinais segmentados, foram acrescentadas as medidas: transitividade, relação entre determinismo e taxa de recorrência, divergência, comprimento médio das linhas diagonais, tempo de recorrência do tipo 1, tempo de recorrência do tipo 2, entropia do tempo de recorrência do tipo 1 e o raio de vizinhança. Na separação entre vozes saudáveis e vozes patológicas, a medida de comprimento máximo da diagonal principal foi a mais eficiente, com taxa de recorrência de 2% e análise discriminante quadrática (QDA) na classificação. Na classificação com redes neurais artificiais e medidas extraídas com segmentação, a acurácia máxima foi  $99,11\% \pm 0,90\%$ , discriminando vozes saudáveis e vozes afetadas por nódulos. Na discriminação entre o grau geral leve e o moderado dos desvios vocais, obteve-se uma acurácia de  $88,47\% \pm 3,71\%$ . Os resultados encontrados, nesta pesquisa, indicam que a análise de quantificação pode ser empregada em métodos de avaliação, triagem e monitoramento vocal.

Santos (2015), em sua dissertação, utilizou as medidas de energia e entropia dos coeficientes de detalhe da transformada wavelet para a avaliação da qualidade vocal em crianças. Dois estudos de casos foram abordados em sua pesquisa: 1) análise acústica do grau da intensidade do desvio vocal e 2) análise acústica da qualidade vocal predominante (rugosidade e sopro). Para o primeiro caso, foi obtida uma acurácia acima de 95% e, para o segundo caso, as medidas de acurácia foram superiores a 90%, com a utilização de um vetor híbrido de medidas combinadas para ambos os casos. Os resultados da classificação mostraram que a técnica é promissora e pode ser considerada para ser empregada como uma ferramenta para a análise acústica.

Souza (2015), em sua dissertação, analisou a discriminação entre sinais de vozes saudáveis e patológicos, por meio de duas abordagens: extração dos descritores de textura diretamente das sub-bandas da transformada wavelet; e extração dos descritores de Haralick, a partir da matriz de coocorrência. Paralisia – edema de Reinke e nódulos nas pregas vocais foram as patologias laríngeas consideradas na pesquisa. Os melhores resultados foram obtidos com os descritores de Haralick, empregando redes neurais MLP (*Multilayer Perceptron*) na classificação, em conjunto com o algoritmo de otimização por enxame de partículas PSO (*Particle Swarm Optimization*), empregado na seleção das características mais representativas, atingindo uma acurácia de 91,79% na discriminação das vozes saudáveis e com nódulos.

Baravieira (2016), em sua tese, utilizou uma rede neural artificial para a avaliação de vozes rugosas e soprosas. A avaliação perceptivo-auditiva foi feita pela escala visual analógica de 100 mm e pela escala numérica de quatro pontos. A extração de características do sinal de voz foi por meio da Transformada Wavelet Packet e dos parâmetros acústicos: Jitter, Shimmer, amplitude da derivada e amplitude do Pitch; e validação do classificador por meio da parametrização, treino, teste e avaliação das redes neurais artificiais. O melhor desempenho, com a utilização da rede neural, foi encontrado na discriminação da sopro e rugosidade, em que se obteve uma taxa de acerto de 74%.

Lopes *et al.* (2016), em sua pesquisa, analisaram a precisão das medidas de quantificação de recorrência na discriminação de pacientes com e sem problema na voz. Doze medidas de quantificação de recorrência foram extraídas usando 1% de taxa de recorrência: determinismo, comprimento médio e máximo das linhas diagonais, entropia de Shannon, laminaridade, tempo de permanência, comprimento máximo das linhas verticais, tempo de recorrência do tipo 1 e do tipo 2, entropia do tempo de recorrência do tipo 1, transitividade e divergência. Parâmetros relacionados com a topologia recorrente do sistema vocal também foram analisados (dimensão de imersão, passo de reconstrução e raio vizinhança), totalizando um conjunto de 15 medidas. Os valores médios das medidas foram empregados no processo de classificação por meio de análise discriminante quadrática. Os resultados mostram que as medidas de recorrência, isoladas ou combinadas, apresentaram um bom desempenho na discriminação de indivíduos com e

sem problema na voz, atingindo uma acurácia de 83,27%.

### 3 Análise não linear e análise de quantificação de recorrência

A análise de séries temporais considera uma série temporal escalar, em geral, associada a uma aquisição experimental para entender o comportamento dinâmico de um sistema. O ponto essencial dessa análise é que uma série temporal contém informações sobre variáveis não observáveis do sistema, o que permite a reconstrução do espaço de estados (COSTA, 2012). No contexto da pesquisa, os sinais de vozes em análise constituem, cada um, uma série temporal que será avaliada nos métodos da análise dinâmica não linear e teoria do caos.

A principal característica ligada a um sistema que possui dinâmica caótica é a sua sensibilidade e variação das condições iniciais, ou seja, uma pequena variação nas condições iniciais faz com que o sistema evolua de forma imprevisível (LORENZ, 1963).

Segundo Kantz e Schreiber (2004), a ligação mais direta entre a teoria do caos e o mundo real é a análise de dados de séries temporais, nos termos da dinâmica não linear. É devido, principalmente, ao desenvolvimento de uma base matemática coerente para a descrição do caos em dinâmicas determinísticas não lineares que se tem dado uma significativa atenção à análise dinâmica não linear e ao modelamento determinístico dos dados de uma série temporal (COSTA, 2012; KUMAR, MULLICK, 1996).

Uma série temporal é um conjunto de observações ordenadas no tempo, não necessariamente igualmente espaçadas, que apresentam dependência serial, isto é, dependência entre instantes de tempo. A notação usada aqui para denotar uma série temporal é  $S_1, S_2, S_3, \dots, S_T$  que indica uma série de tamanho  $T$ .

Quando se deseja determinar se a dinâmica que deu origem a uma série temporal é determinística (ou não), há de se lançar mão de outros métodos. Se for determinística, é possível ter condições de desenvolver um modelo capaz de descrever a dinâmica.

No caso de comportamento estocástico, ao contrário, trata-se de um processo a muitos graus de liberdade, do qual se pode esperar uma descrição estatística do sistema, não sendo possível descrevê-lo por meio de equações diferenciais (FIEDLER-FERRARA; PRADO, 1994).

### 3.1 Medidas de análise dinâmica não linear

A avaliação da análise dinâmica de sinais pode ser feita de duas formas: por meio de um modelo matemático associado ao sistema de produção desses sinais ou por meio da análise de séries temporais. As séries temporais contêm informações sobre variáveis não observáveis do sistema, o que permite a reconstrução do espaço de estados. Cada sinal de voz a ser analisado corresponde a uma série temporal (KENT; READ, 1992).

- **Passo de reconstrução ( $\tau$ )**

Esta medida é baseada no teorema da imersão de Takens (1981). Ele demonstrou que, com o uso da técnica dos tempos de retardo ou método das coordenadas defasadas, é possível reconstruir certas propriedades topológicas do espaço de estados (atrator) a partir da série temporal,  $\{x_t\}$ , em que vetores  $m$ -dimensionais são reconstruídos, de forma que:

$$\vec{\xi}_i = \{x(t_i), x(t_i + \tau), \dots, x(t_i + (m - 1)\tau)\} \quad (1)$$

Neste trabalho, foi utilizado o método de informação mútua, em que o valor do tempo de defasagem ideal corresponde ao primeiro mínimo local, quando este existir, da função de informação mútua média  $I(\tau)$  (FRASER; SWIMMER, 1986).

- **Primeiro mínimo da função de informação mútua (PM)**

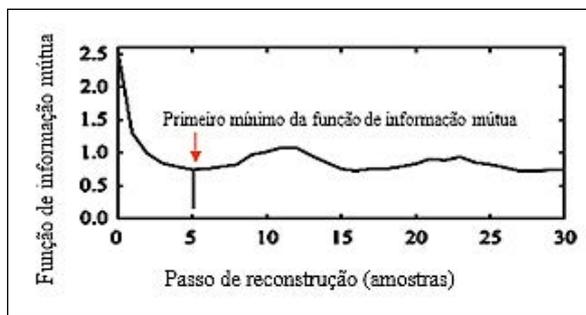
A mesma informação, em sistemas lineares, fornecida pela função de correlação é também fornecida pela função de informação mútua média, sendo que, na verdade, é um tipo de generalização para sistemas não lineares (FANT, 2004). Tem-se que:

$$I(\tau) = -\sum_{i=1}^b \sum_{j=1}^b P_{ij}(x(t), x(t+\tau)) \log_2 \left[ \frac{P_{i,j}(x(t), x(t+\tau))}{P_i(x(t)) P_j(x(t+\tau))} \right] \quad (2)$$

, em que  $P_i(x(t))$  é a probabilidade de que o sinal  $x(t)$  assuma um valor dentro do  $i$ -ésimo intervalo do histograma,  $P_j(x(t+\tau))$  é a probabilidade de que  $x(t+\tau)$  esteja no  $j$ -ésimo intervalo e  $P_{i,j}(x(t), x(t+\tau))$  é a probabilidade de que  $x(t)$  esteja, simultaneamente, no  $i$ -ésimo intervalo e  $x(t+\tau)$  no  $j$ -ésimo intervalo. Sendo  $I(\tau) \geq 0$ . Porém, se  $x(t)$  e  $x(t+\tau)$  são iguais, então  $I(\tau)$  será máximo. Por outro lado, se  $x(t)$  e  $x(t+\tau)$  são completamente independentes, então o

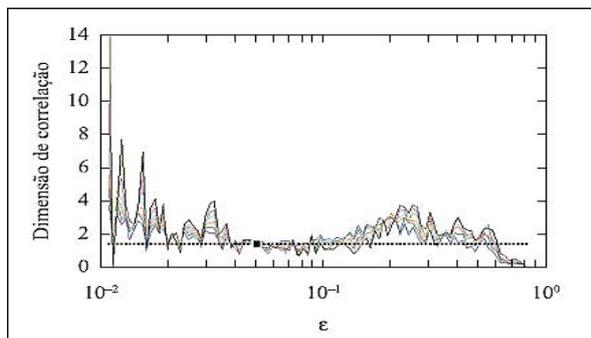
argumento do termo logarítmico é a unidade e  $I(t) = 0$  (FRASER; SWINNEY, 1986 *apud* VIEIRA, 2014).

Na Figura 1, está ilustrado o comportamento da  $I(\tau)$  em função de  $\tau$ . A linha tracejada indica o ponto onde ocorre o primeiro mínimo de  $I(\tau)$  (FMMI – *First Minimum of Mutual Information*) e, assim, a determinação do passo de reconstrução “ótimo” para este segmento ( $\tau = 5$ ).



Fonte: Adaptado de Costa (2012).

**Figura 1** – Função de Informação Mútua e determinação do passo de reconstrução ótimo



Fonte: Adaptado de Costa (2012).

• **Dimensão de correlação ( $D_2$ )**

Esta característica tem por objetivo medir a probabilidade de se encontrar um par aleatório de pontos dentro de um determinado volume elementar e é definida da seguinte forma:

$$D_2 = \lim_{\epsilon \rightarrow 0} \frac{\log \sum_{i=1}^{N(\epsilon)} (f_i)^2}{\log(\epsilon)} \quad (3)$$

em que o termo  $f_i$  representa a frequência relativa com que dois pontos de um atrator estejam dentro da  $i$ -ésima caixa de tamanho  $\epsilon$ .

Na Figura 2, é apresentado o comportamento da dimensão de correlação, em função de  $m$  e de  $\epsilon$ , para um segmento de 32 ms de um sinal de voz saudável. A fim de se estimar o valor de  $D_2$  em cada segmento do sinal são consideradas as inclinações da integral de correlação para diferentes dimensões de imersão, com  $m$  variando de 4 a 1.

**Figura 2** – Comportamento da dimensão de correlação para um segmento de 32 ms de um sinal de voz saudável

**3.2 Gráficos de recorrência**

Os Gráficos de Recorrência (ECKMANN; KAMPHORST; RUELLE, 1987) têm surgido como mais uma técnica de análise não linear de sistemas dinâmicos. A análise de quantificação de recorrência surgiu como forma de potencializar as avaliações, a partir do desenvolvimento das medidas de quantificação de recorrência (WEBBER, ZBILUT, 1994; MARWAN, 2003). A principal vantagem dos Gráficos de Recorrência é a possibilidade do uso dessa técnica para a análise de séries curtas e não estacionárias (COSTA, 2012).

Geralmente, um experimento não mede todas as variáveis de estado do sistema e tem-se disponível a evolução no tempo de apenas uma variável de estado, representada por uma série temporal,  $x(t)$ , que representa a trajetória de um sistema dinâmico.

O gráfico de recorrência é definido pela seguinte expressão matemática (MARWAN, 2003):

$$\mathcal{R}_{i,j}^{m,\epsilon} = \Theta(\epsilon - \|\tilde{\xi}_i - \tilde{\xi}_j\|), \tilde{\xi}_i \in \mathcal{R}^m, i, j = 1 \dots N \quad (4)$$

em que:

- »  $N$  é o número de estados considerados;
- »  $\epsilon$  é o raio vizinhança (*threshold*) no ponto ;
- »  $\|\cdot\|$  é a norma da vizinhança, comumente a norma euclidiana;
- »  $\Theta(\cdot)$  é a função de degrau unitário;
- »  $m$  é a dimensão de imersão (graus de liberdade).

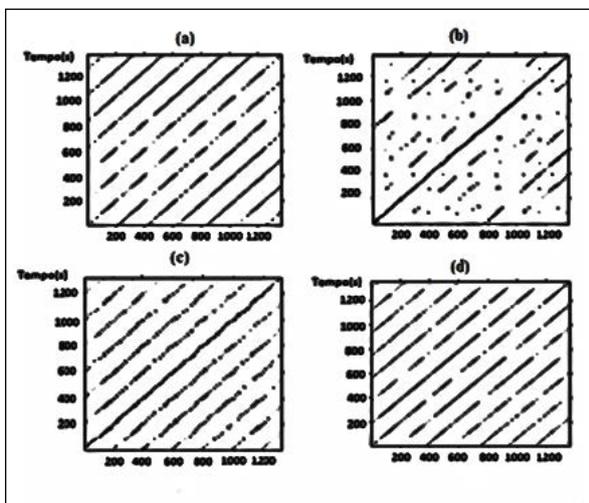
Os gráficos de recorrência são formados por uma matriz quadrada de ordem  $N$ , em que  $N$  é o número de vetores (estados do sistema) de dimensão  $m$ , preenchida por pontos brancos e pretos. O ponto

preto, chamado de ponto recorrente, é colocado na matriz de recorrência com coordenadas  $i$  e  $j$  somente se o estado  $i$  for suficientemente próximo ao estado  $j$ , ou seja, se a distância entre o estado corrente do sistema e o estado a ser comparado for menor que certa distância (raio)  $r$ , fixada no centro do estado corrente (ECKMANN; KAMPHORST; RUELLE, 1987 *apud* VIEIRA, 2014).

Nos gráficos de recorrência, é possível identificar pontos isolados e linhas tanto diagonais como verticais e horizontais. A combinação destas pode formar estruturas variadas (MARWAN; KURTHS, 2005)

Na Figura 3, são apresentados exemplos de quatro gráficos de recorrência, obtidos a partir de sinais de vozes saudáveis e desviadas (rugosa, soprosa e tensa), em que cada gráfico corresponde a 32 ms do sinal.

**Figura 3** – Gráficos de recorrência de sinais de voz: (a) saudável; (b) rugosa; (c) soprosa e (d) tensa.



Fonte: Autoria própria.

Pode-se observar que o gráfico de recorrência do sinal de voz saudável (3.a) apresenta uma formação periódica nas linhas diagonais. No sinal vocal desviado e com predomínio de rugosidade (3.b), pode-se perceber uma menor formação das linhas diagonais (o que parece ter uma relação com a presença de maior aperiodicidade no sinal). A qualidade vocal desviada e predominantemente soprosa (3.c) apresenta menor número de pontos de recorrência (no gráfico, pode-se observar praticamente cada ponto individualmente, embora dê a impressão que as linhas diagonais têm maior comprimento). A configuração da qualidade vocal tensa (3.d) é semelhante à confi-

guração da voz considerada saudável, semelhante ao que foi encontrado em outros tipos de análise, como o “diagrama de desvio fonatório”, em que as vozes tensas se situam muito próximas às vozes saudáveis (LOPES *et al.*, 2016).

A análise visual dos gráficos de recorrência, porém, fornece um resultado subjetivo. Para trazer objetividade na quantificação da dinâmica de um sistema não linear, Webber e Zbilut (1994) desenvolveram medidas de complexidade, as chamadas Medidas de Quantificação de Recorrência. As medidas inicialmente propostas – taxa de recorrência, determinismo, comprimento máximo das linhas diagonais, Entropia de Shannon da distribuição de frequências dos comprimentos das linhas diagonais e tendência – estão relacionadas à densidade dos pontos de recorrência e às estruturas diagonais presentes nos gráficos de recorrência. Posteriormente, Marwan (2003) propôs outras três medidas de recorrência; desta vez, baseadas nas estruturas verticais ou horizontais, são elas: a laminaridade, o tempo de permanência e o comprimento máximo das estruturas verticais.

### 3.3 Medidas de Quantificação de Recorrência

Esta análise é um método que tem por objetivo extrair medidas quantitativas dos gráficos de recorrência. De modo geral, a distribuição dos pontos recorrentes pode ser analisada por meio de inspeção visual a partir de três elementos: as linhas diagonais e as estruturas verticais/horizontais (MARWAN; KURTHS, 2005 *apud* VEIRA, 2014). Neste trabalho, serão descritas apenas as medidas utilizadas na pesquisa.

- **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 (WEBBER; ZBILUT, 2007 *apud* COSTA, 2012):

$$Det = \frac{\sum_{l=lm_{in}}^N l \times P^E(l)}{\sum_{i,j}^N R_{i,j}^{m,E}} \quad (5)$$

em que  $P^E(l) = \{l; i=1...N\}$  representa a distribuição de frequência dos comprimentos  $l$  das estruturas diagonais no gráfico de recorrência, e  $Nl$  é o número absoluto de linhas diagonais. O termo  $l_{min}$  representa o número mínimo de pontos que são considerados para formar uma linha diagonal.

• **Transitividade (*Trans*)**

A medida Transitividade (*Trans*) é uma espécie de taxa de recorrência local, que pode ser utilizada em conjunto com o raio vizinhança para a construção do gráfico de recorrência.

$$Trans = \frac{\sum_{i,j,k=1}^N R_{i,j}^{m,\varepsilon} R_{i,k}^{m,\varepsilon} R_{k,i}^{m,\varepsilon}}{\sum_{i,j,k=1}^N R_{i,j}^{m,\varepsilon} R_{k,i}^{m,\varepsilon}} \quad (6)$$

• **Comprimento médio das linhas diagonais (*Lmed*)**

Está relacionado ao tempo médio que dois segmentos de uma trajetória estão próximos um do outro e pode ser interpretado como o tempo médio de predição. É calculado da seguinte maneira:

$$Lmed = \frac{\sum_{l=lmin}^N l \cdot P^E(l)}{\sum_{i,j}^N P^E(l)} \quad (7)$$

• **Entropia de Shannon (*Entr*)**

Refere-se à entropia de Shannon da distribuição de frequência dos comprimentos das linhas diagonais. Ela reflete a complexidade da estrutura determinística no sistema:

$$Entr = - \sum_{l=lmin}^N p(l) \cdot \ln p(l)$$

, em que

$$p(l) = \frac{P^E(l)}{\sum_{l=lmin}^N P^E(l)} \quad (8)$$

• **Comprimento máximo das linhas verticais (*Vmax*)**

Esta medida mede o tempo máximo em que o sistema permanece em um estado específico, estando relacionada à duração máxima de um comportamento caótico, dada por:

$$Vmax = \max(\{Vl: i = 1 \dots Nv\}) \quad (9)$$

em que *Nv* é o número total de linhas verticais e *Vl* o comprimento da linha vertical.

## 4 Material e métodos

### 4.1 Base de dados

O banco de dados dos sinais utilizados na pesquisa é proveniente do Laboratório Integrado de Estudos da Voz (LIEV) da Universidade Federal da Paraíba. A base faz parte de um projeto intitulado “Integração das medidas acústicas, perceptivas, fisiológicas e de autoavaliação em pacientes com e sem alteração laríngea”, avaliado e aprovado pelo Comitê de Ética em Pesquisa do Centro de Ciências da Saúde/UFPB, com o parecer número 52492/12 (LOPES *et al.*, 2016). Foram incluídos no estudo pacientes com idade entre 18 e 65 anos, que apresentassem queixa vocal e que já tivessem realizado avaliação laringológica, com laudo otorrinolaringológico por escrito. Todos os participantes assinaram o termo de consentimento livre e esclarecido.

Os pacientes foram classificados em relação ao grau de intensidade do desvio vocal (Grau 1, Grau 2 e Grau 3), a partir da análise perceptivo-auditiva, por três avaliadores fonoaudiólogos. Esta análise foi realizada utilizando a Escala Analógica Visual (EAV). A escala constitui um intervalo de 100 milímetros, dentro do qual há três pontos de corte definidos a partir de estudos clínicos realizados no Brasil (YAMASAKI *et al.*, 2008). Caso o profissional treinado marque na escala algum valor entre 0 e 35,5 mm, a voz é considerada normal; valores entre 35,5 e 50,5 mm indicam uma voz com desvio leve; valores entre 50,5 e 90,5 mm indicam uma voz com desvio moderado; e valores acima de 90,5, desvio intenso. Os sinais foram digitalizados a uma taxa de amostragem de 44.100 amostras/s e 16 bits/amostra. Nesta pesquisa, foram selecionados 120 sinais, sendo 30 sinais para cada tipo de voz.

### 4.2 Metodologia

Para a extração das medidas, foram empregados os seguintes *software*: *Time Series Analysis* (TISEAN) v.3.01, do qual foram extraídas as medidas de Passo de reconstrução ( $\tau$ ), Primeiro mínimo da função de informação mútua (*PM*) e Dimensão de correlação (*D2*), e o *Cross Recurrence Plot* (CRP), do qual foram extraídas as medidas de Determinismo (*Det*), Comprimento médio das linhas diagonais (*Lmed*), Entropia de Shannon (*Entr*), Comprimento máximo das linhas verticais (*Vmax*) e Transitividade (*Trans*), além de ela-

boração de rotinas e uso de funções predefinidas no ambiente Matlab®.

Após a extração das medidas consideradas nesta pesquisa, foram realizados os testes de hipóteses estatísticas utilizando o *software Statistical Package for the Social Sciences (SPSS) Statistics®*, versão 19.0, com o objetivo de validar as conclusões a respeito do potencial discriminativo de cada medida. O nível de significância considerado em todos os testes estatísticos realizados foi de 5%. Para o Teste de Normalidade, foi feito o Teste de Shapiro-Wilk, pois o número de amostras analisadas foi menor que 50 ( $N < 50$ ). Em seguida, foram aplicados os Testes paramétricos (Teste de Levene), quando os testes de normalidade sugeriram que as distribuições dos resultados fossem normais, e Testes não paramétricos (Teste de Mann-Whitney por se comparar dois grupos), quando a variável de observação não apresenta distribuição normal. Após a etapa da análise estatística, foi realizada a classificação dos sinais.

No processo de classificação, foi utilizado um classificador MLP com o algoritmo de aprendizado supervisionado do *Scheduling Conjugate Graded (SCG)* (SIMSEK; TANSEL; HÜSEYİN, 2016 *apud* COURAS, 2017). Este algoritmo, definido por Moller (1993), inicia a discussão sobre algoritmos de ordem mais elevada e utiliza informações mais complexas com o objetivo de melhorar o desempenho da rede. Além disso, ele conseguiu convergir rapidamente e com um bom desempenho. A arquitetura utilizada nesta pesquisa foi a rede *FeedForward* de Múltiplas Camadas, por possuir uma camada oculta ou intermediária.

No processo de classificação, foram realizadas duas fases: treinamento e teste, para verificar o desempenho do sistema. Em seguida, foram analisados utilizando as características de forma individual e combinada. Na fase de treinamento do sistema de classificação, foram utilizados 09 neurônios na camada oculta, pois foi observado que seria a quantidade mais eficiente para esta classificação, e 02 na camada de saída. Na camada de entrada, a quantidade de neurônios varia de acordo com o número de características empregadas, podendo variar de 01 a 08.

A validação cruzada foi empregada, de forma que os sinais de treino e os sinais de teste são separados aleatoriamente em 2/3 e 1/3, respectivamente, por 10 vezes consecutivas para que o sistema não fique especialista em um único subconjunto.

Para mensurar a precisão do classificador, três medidas comumente empregadas foram utilizadas: acurácia, sensibilidade e especificidade. Essas medidas estão relacionadas à capacidade de um classificador em diagnosticar uma doença em um paciente que realmente tenha a doença (Verdadeiro Positivo – VP) ou em um paciente saudável (Falso Positivo – FP), ou, ainda, diagnosticar um estado saudável em um paciente saudável (Verdadeiro Negativo – VN) ou em um paciente que tenha a doença (Falso Negativo – FN) (COSTA, 2012). No presente estudo o estado patológico se refere à presença de distúrbio vocal.

A Acurácia (*ACUR*) mede a taxa de classificação correta global, refletindo a capacidade do classificador de identificar corretamente quando há e quando não há a presença do desvio vocal.

$$ACUR = \frac{VP+VN}{VP+VN+FP+FN} \quad (10)$$

A Sensibilidade (*SENS*) mede a capacidade do classificador em identificar a presença do distúrbio quando ele de fato existe.

$$SENS = \frac{VP}{VP+FN} \quad (11)$$

A Especificidade (*ESP*) mede a capacidade do classificador em identificar corretamente a ausência do distúrbio quando de fato ele não existe.

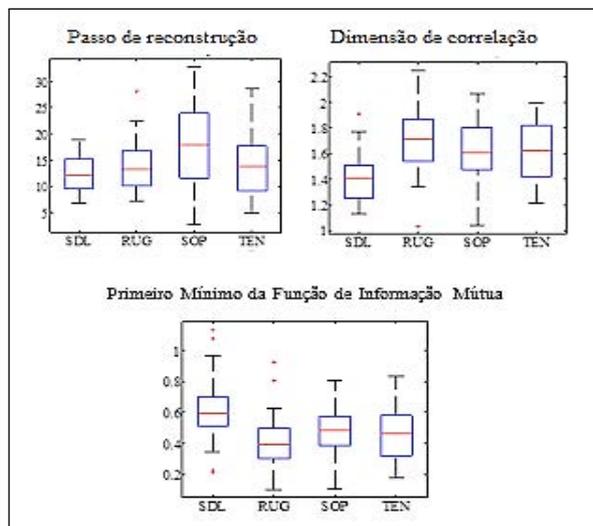
$$ESP = \frac{VN}{VN+FP} \quad (12)$$

## 5 Resultados

Foram aplicados, nesta pesquisa, testes estatísticos, com o objetivo de avaliar o potencial discriminativo de cada uma das características.

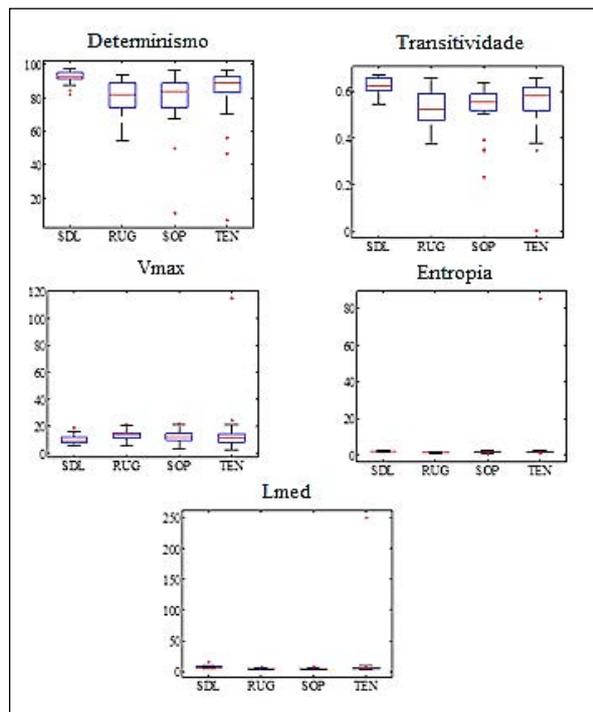
Nas Figuras 4 e 5, respectivamente, é representada a distribuição dos valores médios das características empregadas com as medidas não lineares tradicionais e as medidas de quantificação de recorrência, por meio de gráficos de caixas, os boxplots.

**Figura 4** – Distribuição dos valores médios das medidas associadas à teoria não linear



Fonte: Autoria própria.

**Figura 5** – Distribuição dos valores médios das medidas de quantificação de recorrência



Fonte: Autoria própria.

Nos gráficos de caixas (boxplots), apresentados nas Figuras 4 e 5, observa-se que a medida de Passo de Reconstrução para as vozes soprosas, que são consideradas sinais mais desviados, apresenta mediana superior às demais classes, assim como a dispersão. No caso da Dimensão de correlação, as vozes saudáveis apresentam menor mediana, sendo

que as vozes alteradas apresentam valores dentro do mesmo patamar. As medidas de Lmed e Entr apresentaram uma pequena dispersão, impossibilitando até a visualização de suas medianas. Para o grupo de vozes saudáveis, a medida de Determinismo obteve valores concentrados entre 80% e 100%, enquanto nos sinais com desvios vocais os valores são mais dispersos. Isso pode ser explicado pelo fato desses sinais estarem compreendidos com diferentes tipos de graus de severidade em um mesmo desvio de um mesmo grupo. Além disso, nesta pesquisa não houve separação dos sinais por gênero.

Nas Tabelas 1 e 2, estão os resultados do potencial discriminativo destas medidas, respectivamente, em que os pontos marcados com  $\checkmark$  representam os casos nos quais foram verificadas diferenças estatísticas significativas, considerando  $p\text{-value} < 5\%$  entre os grupos analisados, e os pontos marcados por  $\emptyset$  foram os que não obtiveram diferenças estatísticas. O objetivo destes testes é evidenciar o potencial discriminativo de cada uma das características.

**Tabela 1** – Potencial discriminativo das medidas não lineares tradicionais

Medidas não lineares tradicionais			
Discriminação	$\tau$	PM	D2
SDL x DESV	$\emptyset$	$\checkmark$	$\checkmark$
SDL x SOP		$\checkmark$	$\checkmark$
SDL x RUG		$\emptyset$	$\checkmark$
SDL x TEN		$\emptyset$	$\checkmark$

SDL: Saudável; DESV: Desviada; SOP: Soprosa; RUG: Rugosa e TEN: Tensa.

Fonte: Autoria própria.

**Tabela 2** – Potencial discriminativo das medidas de quantificação de recorrência

Medidas de quantificação de recorrência					
Discriminação	Det	Trans	Lmed	Entr	Vmax
SDL x DESV	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
SDL x SOP	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
SDL x RUG	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
SDL x TEN	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\emptyset$

SDL: Saudável; DESV: Desviada; SOP: Soprosa; RUG: Rugosa e TEN: Tensa.

Fonte: Autoria própria.

No geral, as medidas obtiveram uma boa diferença estatística nos casos de discriminações. A classificação entre as vozes saudáveis e tensas (SDLxTEN) foi a que teve menos medidas com diferenças estatísticas significativas, assim como a medida de Passo de reconstrução ( $\tau$ ) não obteve diferença estatística significativa em nenhum caso de discriminação.

Na Tabela 3, são apresentados os melhores resultados, obtidos com as características individuais e combinadas, entre a discriminação das vozes saudáveis e desviadas.

**Tabela 3** – Avaliação do desempenho – Discriminação entre as vozes saudáveis e desviadas

Avaliação de desempenho		
SDL x DESV	Individuais	Combinadas
Características	Entr	PM + Lmed + Det + Entr + Trans + Vmax
Acurácia (%)	84,67 ± 0,65	90,33 ± 0,6
Sensibilidade (%)	79,00 ± 2,05	91,33 ± 1,02
Especificidade (%)	90,33 ± 0,92	89,33 ± 1,2

SDL: Saudável; DESV: Desviada.

Fonte: Autoria própria.

A medida de Entropia obteve a maior acurácia, atingindo um resultado de 84,67% na classificação individual. Porém, quando combinada com outras medidas (*PM*, *Lmed*, *Det*, *Entr*, *Trans* e *Vmax*), obteve um aumento em sua acurácia de 6%.

Na Tabela 4, são apresentados os melhores resultados, obtidos com as características individuais e combinadas, entre a discriminação das vozes saudáveis e soprosas.

**Tabela 4** – Avaliação do desempenho – Discriminação entre as vozes saudáveis e soprosas

Avaliação de desempenho		
SDL x SOP	Individuais	Combinadas
Características	Entr	$\tau$ + Trans
Acurácia (%)	80,33 ± 1,61	91,17 ± 0,43
Sensibilidade (%)	82,33 ± 5,44	91,67 ± 1,2
Especificidade (%)	78,33 ± 2,64	86,67 ± 2,11

SDL: Saudável; SOP: Soprosa.

Fonte: Autoria própria.

Nos resultados das medidas individuais, a medida de Entropia (*Entr*) se destacou mais uma vez no melhor resultado da classificação individual. Já na classificação com as medidas combinadas, a maior acurácia obtida foi com as medidas de Transitividade (*Trans*) e o Passo de reconstrução ( $\tau$ ), com um aumento de 10%, aproximadamente, na acurácia.

Na Tabela 5, são apresentados os melhores resultados, obtidos com as características individuais e combinadas, entre a discriminação das vozes saudáveis e rugosas.

**Tabela 5** – Avaliação do desempenho – Discriminação entre as vozes saudáveis e rugosas

Avaliação de desempenho		
SDL x RUG	Individuais	Combinadas
Características	Entr	Lmed+Entr+ Trans+D2
Acurácia (%)	84,50±0,5	89,50±1,32
Sensibilidade (%)	89,33±1,63	91,00±1,22
Especificidade (%)	79,67±1,75	88,00±2,4

SDL: Saudável; RUG: Rugosa.

Fonte: Autoria própria.

Nessa classificação, a maior acurácia individual foi com a medida de Entropia (*Entr*). As medidas de Comprimento médio das linhas diagonais (*Lmed*), Transitividade (*Trans*), Entropia (*Entr*) e Dimensão de correlação ( $D_2$ ), quando combinadas, obtiveram a melhor acurácia na classificação nesta discriminação.

Na Tabela 6, são apresentados os melhores resultados, obtidos com as características individuais e combinadas, entre a discriminação das vozes saudáveis e tensas.

**Tabela 6** – Avaliação do desempenho – Discriminação entre as vozes saudáveis e tensas

Avaliação de desempenho		
SDL x TEN	Individuais	Combinadas
Características	Trans	Trans + $\tau$ + Vmax + Lmed + Det + D2
Acurácia (%)	68,00 ± 1,66	76,50 ± 1,77
Sensibilidade (%)	77,67 ± 4,1	83,33 ± 2,28
Especificidade (%)	58,33 ± 5,65	69,67 ± 2,83

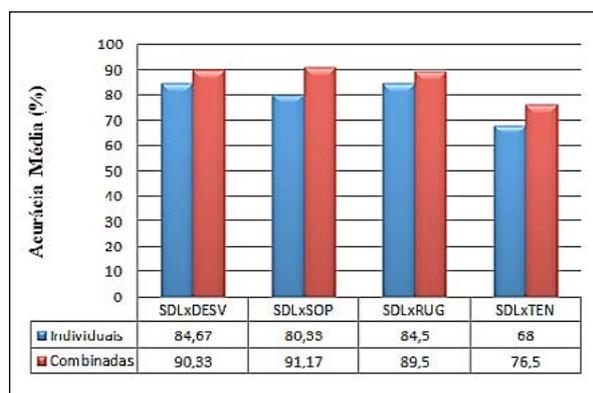
SDL: Saudável; TEN: Tensa.

Fonte: Autoria própria.

A medida de Transitividade (*Trans*) se destacou nos resultados da acurácia, estando presente também na classificação com as medidas combinadas. Quando esta medida foi combinada com as medidas de Passo de reconstrução ( $\tau$ ), Comprimento máximo das linhas verticais ( $V_{max}$ ), Comprimento médio das linhas diagonais ( $L_{med}$ ), Determinismo ( $Det$ ) e Dimensão de correlação ( $D_2$ ), obteve um aumento significativo na acurácia, atingindo 76,5%. Esse aumento corresponde a 8%, aproximadamente, a mais do que o resultado da melhor acurácia individual.

Na Figura 6, é apresentado um resumo com os principais resultados obtidos nas classificações realizadas com as medidas extraídas, tanto individuais como combinadas, indicando a acurácia máxima, para cada caso discriminativo desta análise.

**Figura 6** – Comparação das médias das acurácias dos resultados individuais e combinados



SDL: Saudável; DESV: Desviada; SOP: Soprosa; RUG: Rugosa e TEN:Tensa.

Fonte: Autoria própria.

Observa-se que houve um aumento nos resultados das classificações, quando as medidas foram combinadas, destacando-se um aumento de cerca de 10% na discriminação entre SDLxSOP.

## 6 Considerações finais

O uso da análise acústica, na fonoaudiologia, como ferramenta adicional para diagnóstico de sinais de voz na discriminação entre sinais saudáveis e sinais com alguma presença de patologia ou desvio vocal, vem sendo cada vez mais estudado. As principais motivações para esta pesquisa estão no fato de não ter sido encontrado nenhum trabalho que relacionasse a análise de medidas não lineares tradicionais

com as medidas de quantificação de recorrência à avaliação de distúrbios da voz, assim como a intensidade destes desvios. Além disso, também não foi encontrado nenhum trabalho relacionado à análise visual dos gráficos de recorrência com relação aos desvios vocais (rugosidade, sopro e tensão). Estas medidas vêm se destacando cada vez mais em seus resultados, mostrando uma técnica bastante promissora, atingindo uma acurácia de 91,17% na discriminação entre as vozes saudáveis e soprosas, apenas com duas medidas. Essa taxa é considerada excelente, segundo o critério de Hosmer-Lemeshow (HOSMER; LEMESHOW, 2000). O grande desafio, no entanto, é encontrar características que melhor representem o comportamento caótico de um sinal de voz, pois cada patologia ou desvio vocal possuem características particulares que melhor os definem.

## REFERÊNCIAS

BARAVIEIRA, P. B. **Aplicação de uma rede neural artificial para a avaliação da rugosidade e sopro vocal**. 2016. 101 f. Tese (Doutorado em Ciências)–Programa de Pós-Graduação Interunidades em Bioengenharia EESC/FMRP/IQSC, Universidade de São Paulo, 2016.

BEHLAU, M. **Voz, O Livro do Especialista**. Revinter, 2001.

CAMPISI, P. *et al.* MultiDimensional Voice Program analysis in children with vocal cord nodules. **J Otolaryngol**, v. 29, n. 5, p. 302-308, 2000.

COSTA, W. C. de A. **Análise dinâmica não linear de sinais de voz para detecção de patologias laringeas**. 2012. 176 f. Tese (Doutorado em Engenharia Elétrica) – Universidade Federal de Campina Grande Campina Grande, 2012.

COURAS, M. F. K. B. **Classificação de desvios vocais utilizando características baseadas no modelo linear de produção da fala**. 2017. 115 f. Dissertação (Mestrado em Engenharia Elétrica)– Instituto Federal da Paraíba, João Pessoa, 2017.

CRUZ, B. S. C.; OLIVEIRA, I. B. Disfonias nas Laringectomias Parciais: revisão sistemática de literatura em procedimentos avaliativos. In: ENCONTRO DE INICIAÇÃO CIENTÍFICA, 19., 2014, Campinas. **Anais...** Campinas: Pontifícia Universidade Católica, 2014.

DAJER, M. E. **Análise de sinais de voz por padrões visuais de dinâmica vocal**. 2010. 154 f. Tese (Doutorado em Ciências)— Universidade de São Paulo, São Carlos, 2010.

ECKMANN, J. P.; KAMPHORST S. O.; RUELLE, D. Recurrence plots of dynamical systems. **Europhys. Lett**, v. 56, n. 5, p. 973-977, 1987.

FANT, G. **Speech acoustics and phonetics**. Springer, 2004.

FIEDLER-FERRARA, N.; PRADO, C. P. C. **Caos - uma Introdução**. Editora Edgard Blücher, 1994.

FRASER, A. M.; SWINNEY, H. L. Independent coordinates for strange attractors from mutual information. **Physical Review A**, v. 33, n. 2, p. 1134-1140, 1986.

HENRIQUEZ, P. *et al.* Characterization of healthy ant pathological voice through measures based on nonlinear dynamics. **IEEE Transactions on Audio, Speech and Language Processing**, v. 17, n. 6, p. 1186-1195, 2009.

HOSMER, D. W. JR.; LEMESHOW, S. **Applied logistic regression**. New York: John Wiley & Sons, 2000.

JIANG, J.; ZHANG, Y.; MCGILLIGAN, C. Chaos in voice, from modeling to measurement. **Journal of Voice**, v. 20, n. 1, p. 2-17, jan. 2006.

KANTZ, H.; SCHREIBER, T. **Nonlinear time series analysis**. Cambridge University Press, 2004.

KENT, R. D.; READ, C. **The acoustic analysis of speech**. San Diego: Singular Pub. Group, 1992.

KUMAR, A.; MULLICK, S. K. Nonlinear dynamical analysis of speech. **Journal of Acoustical Society of America**, v. 100, n. 1, p. 615-629, 1996.

LOPES, L. W. *et al.* Effectiveness of recurrence quantification measures in discriminating patients with and without voice disorders. In: INTERNATIONAL CONFERENCE ON VOICE PHYSIOLOGY AND BIOMECHANICS, 10., 2016, Viña del Mar, Chile. **Anais...** Viña del Mar, 2016. p. 14-17.

LORENZ, E. N. Deterministic non-periodic flow. **Journal of Atmospheric Science**, v. 26, p. 130-136, 1963.

MARWAN, N. **Encounters With Neighbours - Current Developments of Concepts**

**Based on Recurrence Plots and Their Applications**. University of Potsdam, 2003.

MARWAN, N.; KURTHS, J. Line structures in recurrence plots. **Physics Letters A**, v. 336, n. 4, p. 349-357, 2005.

MOLLER, M. F. A scaled conjugate gradient algorithm for fast supervised learning. **Neural Networks**, v. 6, n. 4, p. 525-533, 1993.

SANTOS, M. O. **Análise Acústica de Desvios Vocais Infantís utilizando a Transformada Wavelet**. 2015. 79 f. Dissertação (Mestrado em Engenharia Elétrica)—Instituto Federal de Educação, Ciência e Tecnologia da Paraíba, João Pessoa, 2015.

SOUZA, T. A. **Aplicação da Transformada Wavelet em Análise de Texturas de Gráficos de Recorrência para Detecção de Patologias Laringeas**. 2015. 79 f. Dissertação (Mestrado em Engenharia Elétrica)—Instituto Federal de Educação, Ciência e Tecnologia da Paraíba, João Pessoa, 2015.

TAKENS, F. Detecting strange attractors in turbulence. **Lecture notes in mathematics**, v. 898, n. 1, p. 366-381, 1981.

VIEIRA, V. J. D. **Avaliação de Distúrbios da Voz por meio de Análise de Quantificação de Recorrência**. 2014. 2017 f. Dissertação (Mestrado em Engenharia Elétrica)—Instituto Federal de Educação, Ciência e Tecnologia da Paraíba, João Pessoa, 2014.

WEBBER, C. Jr.; ZBILUT, J. Recurrence quantifications: Feature extractions from recurrence plots. **Int. J. Bifurcation and Chaos**, v. 17, p. 3467-3475, 2007.

WEBBER, C. L.; ZBILUT, J. P. D. Dynamical assessment of physiological systems and states using recurrence plot strategies. **J. Appl. Physiol**, v. 76, p. 965-973, 1994.

YAMASAKI, R. *et al.* Correspondência entre escala analógico-visual e a escala numérica na avaliação perceptivo-auditiva de vozes. In: CONGRESSO BRASILEIRO DE FONOAUDIOLOGIA, 16., Campos de Jordão, 2008. **Anais...** Campos de Jordão: Sociedade Brasileira de Fonoaudiologia Campos de Jordão, 2008. p. 24-27.

ZHANG, Y.; JIANG, J. J. Acoustic analyses of sustained and running voices from patients with laryngeal pathologies. **Journal of Voice**, v. 22, n. 1, p. 1-9, jan. 2008.